Practical Quantum Fingerprinting and Appointment Scheduling

by

Benjamin Lovitz

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

Waterloo, Ontario, Canada, 2017

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

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

Statement of Contributions

The work on quantum fingerprinting was done under the supervision of Norbert Lütkenhaus, and the work on appointment scheduling was done in collaboration with NL and Dave Touchette.

Abstract

Quantum protocols for many communication tasks have been found which significantly improve on their classical counterparts. However, many of these protocols are beyond the reach of current technology. In this work, we find more readily implementable protocols for the tasks of quantum fingerprinting and appointment scheduling. Our protocols maintain a quantum advantage even under realistic experimental imperfections.

In the task of quantum fingerprinting, two parties wish to evaluate the equality function on two n-bit strings in the simultaneous message passing model. The original quantum fingerprinting protocol uses a tensor product of a small number of $\mathcal{O}(\log n)$-qubit highly entangled signals [14], whereas a recently-proposed optical protocol uses a tensor product of $\mathcal{O}(n)$ single-qubit signals, while maintaining the $\mathcal{O}(\log n)$ information leakage of the original protocol [4]. The low-dimensionality of each signal in the recently proposed optical protocol makes it more amenable to experimental implementation [68, 32], but due to limited coherence times the large number of signals remains a significant barrier to observing a quantum advantage in information leakage. In contrast, the original protocol sends few signals, but the dimension of each signal is prohibitively high. We find a family of protocols which interpolate between the original and optical protocols while maintaining the $\mathcal{O}(\log n)$ information leakage, thus demonstrating a trade-off between the number of signals sent and the dimension of each signal, and opening the door for experimental implementations to find a "sweet spot" for which the number of signals sent and the dimension of each signal are both amenable to current technology.

In [68, 32] the recently proposed optical protocol is implemented using coherent states. We develop a coherent state protocol which reduces the number of signals by a factor 1/2 from the recently proposed optical protocol, while also reducing the information leakage. We consider several natural generalizations of this protocol to other coherent state protocols which further reduce the number of signals, but find numerical evidence that they have greater information leakage in the ideal setting and also under realistic experimental imperfections. Using a similar technique, we improve a recently proposed coherent state protocol for evaluating the Euclidean distance between two real unit vectors [41] by reducing the number of signals by a factor 1/2 while also reducing the information leakage. We also extend this protocol to handle complex unit vectors. Along the way, we find a simple beamsplitter measurement to perform optimal unambiguous state comparison between two coherent states.

In the task of appointment scheduling, two parties each have n-bit strings, and they wish to find a common intersection in the interactive communication model. The known quantum appointment scheduling protocol of [15] performs this task with $\mathcal{O}(\sqrt{n}\log n)$ qubits of communication, a nearly quadratic improvement over the classical lower bound of $\Omega(n)$ bits [39]. However, this protocol requires quantum states of high dimension and global unitary operations. We find appointment scheduling protocols which are more feasible for implementation and maintain a quantum advantage over the classical lower bound in terms of information cost, even under experimental imperfections. Our main protocols require the generation of coherent states of a fixed set of amplitudes, along with phase shifters and beamsplitters on two modes with relatively low splitting angle. They also require the parties to transfer two modes back-and-forth multiple times with relatively low loss. Although our protocols make progress towards the experimental implementation of quantum appointment scheduling, we expect that they still remain outside the scope of current technology.

Acknowledgements

First and foremost, thank you Norbert Lütkenhaus for providing me the opportunity to work on this project, and for keeping your door so often open for discussion.

Thank you Judy McDonnell, Michele Roche, Melissa Floyd, and the folks at QuTask for your secratarial, administrative, and technical support.

Thank you all of my fellow Optical Quantum Communication Theory group members for your support and fruitful discussions: Juan Miguel Arrazola, Patrick Coles, Electra Eleftheriadou, Michael Epping, Sumeet Khatri, Jie Lin, Christian Mastrommattei, Filippo Miatto, Ryo Namiki, Dave Touchette, and Yanbao Zhang. In particular, thank you Filippo Miatto and Dave Touchette for your invaluable help and collaboration on many occasions.

Thank you Daniel Puzzuoli for your patient teaching.

Thank you Konrad Banaszek for your insight on the $q-$ary encoding for quantum fingerprinting described in Appendix A.0.2, which helped me improve the presentation of these results. Thank you Dave Touchette for your helpful insight on the information leakage bound of Appendix B.0.2. Thank you Juan Miguel Arrazola, Richard Cleve, David Lovitz, and John Watrous for helpful insight and discussions on quantum fingerprinting.

Thank you Filippo Miatto, Dave Touchette, and Christian Mastromattei for helpful comments on this thesis.

I acknowledge the traditional territory of the Attawandaron (Neutral), Anishnawbe, and Haudenosaunee peoples on which I have lived and worked as a settler for the past two years. This land lies in the Haldimand tract, which was promised to the Haudenosaunee peoples in the Haldimand Treaty of 1784.

Table of Contents

List of Figures

Chapter 1

Quantum communication

Quantum communication is a field of quantum information concerned with the processing and storage of information using quantum states and channels under various constraints of nonlocality. A plethora of quantum protocols for communication tasks have been found which drastically outperform their classical counterparts. For example, quantum key distribution protocols have been found which produce provably secure secret keys for private communication, a task that is classically impossible. There are also many communication tasks for which there exist quantum protocols that use exponentially fewer resources than any classical protocol.

In recent years, experimental technology has progressed to the point where realization of some quantum protocols which exhibit such an advantage is possible. However, many existing quantum protocols remain unrealized because the resources required are not within reach of current technology. In this thesis we modify existing quantum protocols and develop new quantum protocols which are easier to implement using current technology and maintain a quantum advantage over classical protocols.

We develop practical quantum communication protocols for the tasks of quantum fingerprinting and appointment scheduling. Quantum fingerprinting is a task in which two parties each have some classical data, and they communicate with a single referee to determine whether or not their data is the same. We find practical quantum protocols to perform this task which maintain an advantage over the classical lower bound in terms of information leakage. Appointment scheduling is a task in which two parties each have calendars, and they communicate between each other in order to find a date in which they are both free. We find practical quantum protocols to perform this task while maintaining an advantage over the classical lower bound in terms of information cost.

This thesis is organized as follows. In the remainder of this chapter we introduce some mathematical objects of quantum information that we will use in this thesis. In Chapter 2 we introduce coherent states and linear optics, which are the primary quantum states and channels that we use in our protocols. In Chapter 3 we introduce the fields of quantum communication complexity and quantum information complexity, and the notions of information that we use to analyze our protocols. In Chapter 4 we present our results in quantum fingerprinting, and in Chapter 5 we present our results in appointment scheduling.

1.1 Mathematical objects of quantum information

Here we review the mathematical objects of quantum information that we will use in this work. We avoid explicit definitions of several standard objects, for which we refer the reader to the manuscript of John Watrous' book [67]. We have made an effort to refer to most objects by similar names to those given in [67].

We primarily consider finite-dimensional Hilbert spaces, which we call complex Euclidean spaces. We refer to the abstract physical container which holds states acting on some complex Euclidean space $\mathcal{A}$ as a *register*, which we denote by A, and define the *classical state set* A of A as the (finite and non-empty) set for which $\mathcal{A} = \mathbb{C}^{\mathrm{A}}$. We refer the reader to [67] for formal definitions of these objects.

We primarily use the symbols A, B, C, X, Y, Z to refer to registers with associated complex Euclidean spaces $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$ and classical state sets $\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{X}, \mathrm{Y}, \mathrm{Z}$. For brevity, when introducing these objects we frequently only introduce the registers and implicitly introduce the associated classical state sets and complex Euclidean spaces with the above notation. Registers referred to with the symbols X, Y, Z are frequently (but not always) classical registers.

We define the Euclidean inner product function $\langle \cdot, \cdot \rangle$ with the conjugate-transpose taken on the first argument. We define $(\cdot)^\dagger$ as the conjugate-transpose map, $(\cdot)^T$ as the transpose map, and $\overline{(\cdot)}$ as the complex conjugation map.

We frequently use Dirac notation: $|v\rangle \in \mathcal{A}$ denotes a column vector, and $\langle v| = (|v\rangle)^\dagger$. We define the standard basis elements of $\mathcal{A}$ as $|a\rangle$ for $a \in \mathrm{A}$. For two vectors $|v\rangle, |w\rangle \in \mathcal{A}$, we sometimes use the shorthand $\langle v, w\rangle := \langle |v\rangle, |w\rangle \rangle$ to denote their inner product.

We define $\log(\cdot)$ as the base-two logarithm.

We now list some standard objects that we will use.

Some standard objects we will use

- $\mathcal{S}(\mathcal{A})$: The set of unit vectors on $\mathcal{A}$ ($\mathcal{S}$ is for unit sphere).
- $\mathcal{L}(\mathcal{A},\mathcal{B})$: The set of linear operators from $\mathcal{A}$ to $\mathcal{B}$, and $\mathcal{L}(\mathcal{A}) := \mathcal{L}(\mathcal{A},\mathcal{A})$: The set of linear operators on $\mathcal{A}$.
- $\mathrm{Herm}(\mathcal{A})$: The set of Hermitian operators on $\mathcal{A}$.
- $\mathrm{Pos}(\mathcal{A})$: the set of positive semi-definite operators on $\mathcal{A}$.
- $\mathcal{D}(\mathcal{A})$: The set of quantum states (or simply, states) on $\mathcal{A}$ ($\mathcal{D}$ is for density operators).
- $\mathrm{SepD}(\mathcal{A}:\mathcal{B})$: The set of separable states on $\mathcal{A}\otimes\mathcal{B}$.
- $\mathcal{U}(\mathcal{A},\mathcal{B})$: The set of isometries from $\mathcal{A}$ to $\mathcal{B}$, and $\mathcal{U}(\mathcal{A}) := \mathcal{U}(\mathcal{A},\mathcal{A})$: The set of unitary operators on $\mathcal{A}$.
- $\mathcal{C}(\mathcal{A},\mathcal{B})$: The set of quantum channels from $\mathcal{L}(\mathcal{A})$ to $\mathcal{L}(\mathcal{B})$.
- $\mathrm{GL}(\mathcal{A})$: The general linear group of invertible linear operators on $\mathcal{A}$, under standard matrix multiplication.
- $L^2(\mathbb{R})$: The Hilbert space of square-integrable functions from $\mathbb{R}$ to $\mathbb{C}$.

Chapter 2

Coherent states and linear optics

The coherent states of position and momentum are pure quantum states which well-approximate the states of light produced by a laser [28], and therefore are practical states for optical implementation. Linear optics transformations are a class of unitary operators on Hilbert space which are also amenable to implementation. As such, we make frequent use of these objects in this thesis. In this chapter, we introduce these objects and use them to describe a general coherent state mapping (developed in [3]) from a broad class of quantum protocols to coherent state protocols. We use this mapping in Chapter 5 to develop coherent state versions of existing appointment scheduling protocols.

In Section 2.1 we construct the coherent states from the position and momentum operators, and show that they form a resolution of the identity operator and saturate Heisenberg's uncertainty principle. This treatment is largely of theoretical interest to the author, and will not be needed in the remainder of the thesis. In Section 2.2 we introduce linear optics transformations and show that they are unitary operators on Hilbert space. We also review beamsplitters and phase shifters, two fundamental building blocks of linear optics transformations that we frequently use in this thesis. The reader comfortable with linear optics transformations can safely skip this section. In Section 2.3 we review the coherent state mapping of [3].

Coherent states and linear optics transformations are operators on an infinite dimensional Hilbert space (usually $L^2(\mathbb{R})$). In the following treatment, we omit some technicalities which arise from this fact, and make an effort to inform the reader when we have omitted an element of rigor. For example, we frequently refer to the *creation and annihilation operators* $\hat{a}, \hat{a}^\dagger$, which are unbounded operators on infinite dimensional Hilbert space. As such, it must be verified that every object acted on by $\hat{a}, \hat{a}^\dagger$ is in the domain of $\hat{a}, \hat{a}^\dagger$. We will omit such verification, and refer the interested reader to, e.g., [34].

2.1 Coherent states

Here we construct the coherent states from any linear operator $\hat{a} \in \mathcal{L}(\mathcal{H})$ acting on some complex Hilbert space $\mathcal{H}$ residing in some register A which obeys the commutation relation

$$[\hat{a}, \hat{a}^\dagger] = \mathbb{1}_{\mathcal{H}} \tag{2.1}$$

and satisfies $\hat{a}\,|0\rangle = 0$ for some unit vector $|0\rangle$. Following the standard language of quantum optics, we refer to the register A as a *mode*. We can already infer that $\mathcal{H}$ is infinite-dimensional, as $\mathrm{Tr}([\hat{a}, \hat{a}^\dagger]) = 0$ and $\mathrm{Tr}(\mathbb{1}_{\mathcal{H}}) = \dim(\mathcal{H})$ in the finite case.

We will write the coherent states as a linear combination of a countably infinite set of orthogonal vectors in $\mathcal{H}$, which are constructed as follows: Define $\hat{N} = \hat{a}^\dagger \hat{a}$. By the commutation relation (2.1) it follows that

$$\hat{N}\hat{a}^{\dagger n}\,|0\rangle = n\hat{a}^{\dagger n}\,|0\rangle \tag{2.2}$$

for all $n \in \mathbb{N}$. As $\hat{N}$ is a normal matrix and the vectors $\hat{a}^{\dagger n}\,|0\rangle$ have distinct eigenvalues, they are all orthogonal. Now we normalize each $\hat{a}^{\dagger n}\,|0\rangle$. Define unit vectors $|n\rangle$ such that $a^{\dagger n}\,|0\rangle = c_n\,|n\rangle$ for some $c_n \in \mathbb{R}$. Then

$$\begin{aligned} c_n^2 &= \langle 0|\,\hat{a}^n \hat{a}^{\dagger n}\,|0\rangle \\ &= \langle 0|\,\hat{a}^{n-1}(\mathbb{1}_{\mathcal{H}} + \hat{N})\hat{a}^{\dagger n-1}\,|0\rangle \\ &= nc_{n-1}^2 \end{aligned}$$

for all $n = 1, 2, \ldots$, where the second equality follows from the commutation relation (2.1). As $c_0 = 1$, then $c_n = \sqrt{n!}$, so

$$a^{\dagger n}\,|0\rangle = \sqrt{n!}\,|n\rangle, \tag{2.3}$$

which implies

$$\begin{aligned} \hat{a}^\dagger\,|n\rangle &= \sqrt{n+1}\,|n+1\rangle \\ \hat{a}\,|n\rangle &= \sqrt{n}\,|n-1\rangle\,. \end{aligned} \tag{2.4}$$

For any $\alpha \in \mathbb{C}$, define the *coherent state* $|\alpha\rangle$ as

$$|\alpha\rangle := e^{\frac{-|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}\,|n\rangle\,. \tag{2.5}$$

Using equation (2.4) it can be shown that the coherent states are eigenvectors of the operator $\hat{a}$:

$$\hat{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle\,. \tag{2.6}$$

We find it convenient to write the coherent states as

$$|\alpha\rangle = \mathscr{D}_0(\alpha)\,|0\rangle \tag{2.7}$$

for a unitary $\mathscr{D}_0(\alpha) \in \mathcal{U}(\mathcal{H})$ (written as the output of a function $\mathscr{D}_0 : \mathbb{C} \to \mathcal{U}(\mathcal{H})$), which we now define. Recall the Baker-Hausdorff lemma:

Lemma 1 (Baker-Hausdorff)**.** *For any Hilbert space $\mathcal{H}$, and any linear operators $\hat{A}, \hat{B} \in \mathcal{L}(\mathcal{H})$ satisfying $[[\hat{A},\hat{B}],\hat{A}] = [[\hat{A},\hat{B}],\hat{B}] = 0$, it holds that*

$$e^{\hat{A}}e^{\hat{B}} = e^{\frac{1}{2}[\hat{A},\hat{B}]}e^{\hat{A}+\hat{B}} = e^{[\hat{A},\hat{B}]}e^{\hat{B}}e^{\hat{A}}. \tag{2.8}$$

By the Baker-Hausdorff lemma, the commutation relation (2.1) and equation (2.4) it follows that

$$\mathscr{D}_0(\alpha) := e^{\alpha\hat{a}^\dagger - \overline{\alpha}\hat{a}} \tag{2.9}$$

satisfies (2.7). We refer to the unitary operators $\mathscr{D}_0(\alpha)$ as the *displacement operators*.

Now we observe that the function $\mathscr{D} : \mathbb{R}\times\mathbb{C} \to \mathcal{U}(\mathcal{H})$ defined as

$$\mathscr{D}_\lambda(\alpha) := e^{i\lambda}\mathscr{D}_0(\alpha) \tag{2.10}$$

for each $\lambda \in \mathbb{R}$ and $\alpha \in \mathbb{C}$ forms a representation of the Heisenberg group. We will use this observation in the next section to prove that the coherent states of position and momentum form a resolution of the identity operator.

Definition 2. *We define the* ***Heisenberg group*** $(H,\cdot)$ *as the set of matrices*

$$\left\{ H(\alpha,\lambda) = \begin{bmatrix} 1 & \mathrm{Re}(\alpha) & \lambda \\ 0 & 1 & \mathrm{Im}(\alpha) \\ 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^3 : \alpha\in\mathbb{C}, \lambda\in\mathbb{R} \right\} \tag{2.11}$$

equipped with standard matrix multiplication.

Definition 3. *For a Hilbert space $\mathcal{H}$ and group $(G,\cdot)$, we say a function $\mathscr{A} : G \to \mathrm{GL}(\mathcal{H})$ forms a* ***representation*** *of G if*

$$\mathscr{A}(g_1\cdot g_2) = \mathscr{A}(g_1)\mathscr{A}(g_2) \tag{2.12}$$

for all $g_1, g_2 \in G$ (i.e. $\mathscr{A}$ is a homomorphism).

Definition 4. *We say a representation $\mathscr{A}$ of a group $(G,\cdot)$ is* ***irreducible*** *if the only subspaces of $\mathcal{A}$ which are invariant under $\mathscr{A}(g)$ for all $g \in G$ are the null space and $\mathcal{A}$ itself.*

Observing that $\mathscr{D}$ forms a representation of the Heisenberg group is straightforward. Indeed, the action of the Heisenberg group is given by

$$H(\alpha,\lambda)H(\beta,\nu) = H(\alpha+\beta,\lambda+\nu+\mathrm{Im}(\overline{\alpha}\beta)) \tag{2.13}$$

and by the Baker-Hausdorff lemma,

$$\begin{aligned}\mathscr{D}_\lambda(\alpha)\mathscr{D}_\nu(\beta) &= e^{\mathrm{Im}(\overline{\alpha}\beta)}\mathscr{D}_\nu(\beta)\mathscr{D}_\lambda(\alpha)\\ &= D_{\lambda+\nu+\mathrm{Im}(\overline{\alpha}\beta)}(\alpha+\beta).\end{aligned} \tag{2.14}$$

This representation is known as the Schrödinger representation. In the following section we will show that this representation is irreducible.

2.1.1 Coherent states of position and momentum

Now we focus our interest on a particular set of coherent states residing in the Hilbert space $L^2(\mathbb{R})$, for which $\hat{a}$ is given by

$$\hat{a} = \frac{1}{\sqrt{2}}(\hat{X}-\hat{D}) \tag{2.15}$$

where $\hat{X},\hat{D} \in \mathcal{L}(L^2(\mathbb{R}))$, defined as

$$(\hat{X}f)(x) = xf(x) \tag{2.16}$$

$$\hat{D}f = \frac{df}{dx}, \tag{2.17}$$

are known as the *position* and *momentum* operators, respectively. Under this choice it can be shown that

$$\hat{a}^\dagger = \frac{1}{\sqrt{2}}(\hat{X}+\hat{D}). \tag{2.18}$$

For the remainder of this text, when we refer to coherent states, we refer to the "coherent states of position and momentum" arising from $\hat{a}$ and $\mathcal{H}$ defined as above.

In what follows, we describe some notable properties of coherent states. Namely, we show that the coherent states form a resolution of the identity operator, and that they are states of minimum uncertainty.

Resolution of the identity operator

Here we show that integration over all coherent states produces a scalar multiple of the identity operator

$$\int_{\alpha\in\mathbb{C}} |\alpha\rangle\langle\alpha| \mathrm{d}^2\alpha = \pi\mathbb{1}_\infty, \tag{2.19}$$

where $\mathrm{d}^2\alpha = \mathrm{dRe}(\alpha)\mathrm{dIm}(\alpha)$. We leave calculation of the numerical factor π to the reader, and only show that the lefthand side of (2.19) is some scalar multiple of the identity. To prove the statement, we use Schur's lemma:

Theorem 5 (Schur's lemma)**.** *Let $\mathscr{A}$ be an irreducible representation of a group G on a complex Hilbert space $\mathcal{A}$ and let $\hat{B} \in \mathcal{L}(\mathcal{A})$ be a linear operator satisfying*

$$\hat{B}\mathscr{A}(g)v = \mathscr{A}(g)\hat{B}v \tag{2.20}$$

for all $g \in G$ and all $v \in \mathcal{A}$ (i.e. $\hat{B}$ is an intertwining map). Then $\hat{B} = \lambda\mathbb{1}_{\mathcal{H}}$ for some $\lambda \in \mathbb{C}$.

It is straightforward to show that $\mathscr{D}_\lambda(\alpha)$ commutes with the lefthand side of (2.19) for all $\alpha \in \mathbb{C}$, $\lambda \in \mathbb{R}$, so by Schur's lemma it remains only to show that $\mathscr{D}$ forms an irreducible representation of the Heisenberg group. First, we pause to prove Schur's lemma.

Proof sketch of Schur's Lemma [33]. As $\hat{B}$ is a linear operator on a vector space over an algebraically closed field ($\mathbb{C}$), then it must have at least one eigenvalue $\lambda \in \mathbb{C}$. Let $\mathcal{W} \subseteq \mathcal{A}$ denote the corresponding eigenspace for $\hat{B}$. Then,

$$\hat{B}\mathscr{A}(g)w = \lambda\mathscr{A}(g)w \tag{2.21}$$

for all $g \in G$, $w \in \mathcal{W}$, so the subspace $\mathcal{W}$ is invariant under $\mathscr{A}$. Since λ is an eigenvalue, $\mathcal{W} \neq 0$ which implies $\mathcal{W} = \mathcal{A}$ by irreducibility of $\mathscr{A}$, so $\hat{B} = \lambda\mathbb{1}_{\mathcal{H}}$ on all of $\mathcal{A}$. □

Now we show that $\mathscr{D}$ forms an irreducible representation of the Heisenberg group, which by Schur's lemma will complete the proof that the lefthand side of (2.19) is a scalar multiple of the identity. For each $f, g \in L^2(\mathbb{R})$ define a function (the "Fourier-Wigner Transform" [26]) $M(f,g) : \mathbb{C} \to \mathbb{C}$ as $M(f,g)(\alpha) = \langle\mathscr{D}_0(\alpha)f, g\rangle$. It can be shown [26] that for all $f_1, g_1, f_2, g_2 \in L^2(\mathbb{R})$,

$$\langle M(f_1,g_1), M(f_2,g_2)\rangle = \overline{\langle f_1, f_2\rangle}\,\langle g_1, g_2\rangle\,. \tag{2.22}$$

Irreducibility of $\mathscr{D}$ follows:

Proposition 6 ([26]). *The representation $\mathscr{D}$ is irreducible.*

Proof sketch. Supose $\mathcal{W} \subset L^2(\mathbb{R})$ is a nonzero subspace invariant under $\mathscr{D}_\lambda(\alpha)$ for all $(\lambda, \alpha) \in \mathbb{R} \times \mathbb{C}$, and $f \neq 0 \in \mathcal{W}$. If $g \perp \mathcal{W}$ then $M(f,g) = 0$, which by (2.22) implies $\|f\|_2 \|g\|_2 = 0$, so $g = 0$, which implies $\mathcal{W} = L^2(\mathbb{R})$. □

This completes the proof that the lefthand side of (2.19) is a scalar multiple of the identity.

Saturation of Heisenberg's uncertainty relation

In this section we state and prove Heisenberg's uncetainty relation, and then show that it is saturated by the coherent states. For any Hermitian operator $\hat{A} \in \mathrm{Herm}(\mathcal{H})$, define the *variance* $(\Delta_f \hat{A})^2$ of $\hat{A}$ with respect to $f \in \mathcal{H}$ as

$$(\Delta_f \hat{A})^2 := \langle f, \hat{A}^2 f \rangle - \langle f, \hat{A} f \rangle^2. \tag{2.23}$$

Corresponding to any Hermitian operator $\hat{A} \in \mathrm{Herm}(\mathcal{H})$ is a projective measurement onto the eigenvectors of $\hat{A}$ which assigns to each outcome a real number given by the eigenvalue of each eigenvector. The variance $(\Delta_f \hat{A})^2$ is a commonly studied value which informally quantifies how far each numerical outcome deviates from the average numerical outcome obtained by the measurement, or (also informally) how much uncertainty one has about the numerical outcome that will be obtained in the measurement corresponding to $\hat{A}$ on the state f before performing the measurement.

The measurements corresponding to the Hermitian operators $\hat{X}$ and $i\hat{D}$ translate physically to measurements of the position and momentum of the input state, respectively (with some constant prefactors which depend on one's choice of units) [28]. Heisenberg's uncertainty relation stipulates that for any state, the uncertainty (variance) of its position and momentum cannot both be low. If the state has low uncertainty in one quantity then it necessarily has high uncertainty in the other.

Lemma 7 (Heisenberg's uncertainty relation). $(\Delta_f \hat{X})(\Delta_f i\hat{D}) \geq 1/2$ *for any state* $f \in \mathcal{S}(L^2(\mathbb{R}))$ *in the domain of both* $\hat{X} i\hat{D}$ *and* $i\hat{D}\hat{X}$.

Proof sketch. It is straightforward to verify that

$$(\Delta_f \hat{X})^2 = \langle g, g \rangle, \tag{2.24}$$

where

$$g = (\hat{X} - \langle f, \hat{X} f \rangle \mathbb{1}_\infty) f, \tag{2.25}$$

and

$$(\Delta_f i\hat{D})^2 = \langle h, h \rangle, \tag{2.26}$$

where

$$h = (i\hat{D} - \langle f, i\hat{D} f \rangle \mathbb{1}_\infty) f. \tag{2.27}$$

By the Cauchy-Schwartz inequality,

$$\begin{align} (\Delta_f \hat{X})(\Delta_f i\hat{D}) &\geq |\langle g, h \rangle| \tag{2.28} \\ &= \frac{1}{2i} \langle f, [\hat{X}, i\hat{D}] f \rangle \tag{2.29} \\ &= \frac{1}{2}, \tag{2.30} \end{align}$$

where the first equality follows from straightforward calculation and the second equality follows from $[\hat{X}, i\hat{D}] = i\mathbb{1}_\infty$, which is easily verified from (2.31) and (2.1). □

Using (2.1), (2.6), and the relations

$$\begin{align} \hat{X} &= \frac{1}{\sqrt{2}} (\hat{a} + \hat{a}^\dagger) \\ \hat{D} &= \frac{1}{\sqrt{2}} (\hat{a}^\dagger - \hat{a}) \tag{2.31} \end{align}$$

it is straightforward to verify that $(\Delta_\alpha \hat{X}) = (\Delta_\alpha i\hat{D}) = \sqrt{1/2}$ for any coherent state $|\alpha\rangle \in \mathcal{S}(L^2(\mathbb{R}))$. Thus, $(\Delta_\alpha \hat{X})(\Delta_\alpha i\hat{D}) = 1/2$, so coherent states give rise to the minimum possible value of $(\Delta_f \hat{X})(\Delta_f i\hat{D})$ over all states $f \in \mathcal{S}(L^2(\mathbb{R}))$.

Heisenberg's uncertainty relation seems counterintuitive in comparison to our experience of the world. To the naked eye it seems the position and momentum of objects are both readily measurable without uncertainty. Because of this intuition, many use $(\Delta_f \hat{X})(\Delta_f i\hat{D})$ to measure the level of non-classicality of f. As the coherent states give rise to the minimum possible value of $(\Delta_f \hat{X})(\Delta_f i\hat{D})$, they are interpreted as the most classical states.

2.2 Linear optics

Here we introduce a family of unitary operators on $L^2(\mathbb{R})^{\otimes n}$ for finite $n \in \mathbb{N}$ that are known as *linear optics transformations*. We first define the operators in this family and then show that they are unitary. We then review beamsplitters and phase shifters, two fundamental linear optics transformations that we frequently use in this thesis.

We define a linear optics transformation by its action on each element of the following orthonormal basis of $L^2(\mathbb{R})^{\otimes n}$ (known as the Fock basis):

$$F = \{\frac{\hat{a}_1^{\dagger j_1} \ldots \hat{a}_n^{\dagger j_n}}{\sqrt{j_1! \cdots j_n!}} |0\rangle^{\otimes n} \in L^2(\mathbb{R})^{\otimes n} : j_1, \ldots, j_n \in \mathbb{N}\}, \tag{2.32}$$

where

$$\hat{a}_i^\dagger := \mathbb{1}^{\otimes i-1} \otimes \hat{a}^\dagger \otimes \mathbb{1}^{\otimes n-i} \in \mathcal{L}(L^2(\mathbb{R})^{\otimes n}) \tag{2.33}$$

for each $i = 1, \ldots, n$ (see, e.g., [34]). We often omit the normalization factors $\sqrt{j_1! \cdots j_n!}$ for brevity.

For any linear operator $U \in \mathcal{L}(\mathbb{C}^n)$ define $V_U \in \mathcal{L}(L^2(\mathbb{R})^{\otimes n})$ as

$$\begin{aligned} &V_U a_1^{\dagger j_1} \ldots a_n^{\dagger j_n} |0\rangle^{\otimes n} \\ &= \left(\sum_{i=1}^n U_{1,i} a_i^\dagger\right)^{j_1} \ldots \left(\sum_{i=1}^n U_{n,i} a_i^\dagger\right)^{j_n} |0\rangle^{\otimes n} \end{aligned} \tag{2.34}$$

for each $j_1, \ldots, j_n \in \mathbb{N}$. As shorthand, we can define the action of V_U by

$$a_j^\dagger \to \sum_{i=1}^n U_{j,i} a_i^\dagger \tag{2.35}$$

for each $j = 1, \ldots, n$.

Definition 8. *We say an operator $V \in \mathcal{L}(L^2(\mathbb{R})^{\otimes n})$ is a **linear optics transformation** if $V = V_U$ for some unitary $U \in \mathcal{U}(\mathbb{C}^n)$.*

Note that every unitary operator $U \in \mathcal{U}(\mathbb{C}^n)$ can be written as $U = e^{-iH}$ for some Hermitian operator $H \in \mathrm{Herm}(\mathbb{C}^n)$. It can be shown that the linear optics transformation $V_U \in \mathcal{U}(L^2(\mathbb{R})^{\otimes n})$ is then given by [55],[45]

$$V_U = \exp\left[-i \sum_{j,k=1}^n a_j^\dagger H_{j,k}^T a_k\right]. \tag{2.36}$$

We omit a proof of this fact, and simply prove from the definition (2.34) that V_U is unitary.

Lemma 9. *Every linear optics transformation is unitary.*

Proof sketch. To prove the lemma, we first compute the adjoint map of V_U for any linear map $U \in \mathcal{L}(\mathbb{C}^n)$. We then show that for any unitary U, $V_U V_U^\dagger = V_U^\dagger V_U = \mathbb{1}_\infty^{\otimes n}$.

The adjoint map $V_U^\dagger$ is defined by $\langle f, V_U g\rangle = \left\langle V_U^\dagger f, g \right\rangle$ for all $f, g \in L^2(\mathbb{R})^{\otimes n}$. Since F forms a basis of $L^2(\mathbb{R})^{\otimes n}$, this is equivalent to

$$\begin{aligned}&\left\langle a_1^{\dagger k_1} \dots a_n^{\dagger k_n} |0\rangle^{\otimes n}, \left(\sum_{i=1}^n U_{1,i} a_i^\dagger\right)^{j_1} \dots \left(\sum_{i=1}^n U_{n,i} a_i^\dagger\right)^{j_n} |0\rangle^{\otimes n} \right\rangle \\ &= \left\langle V_U^\dagger a_1^{\dagger k_1} \dots a_n^{\dagger k_n} |0\rangle^{\otimes n}, a_1^{\dagger j_1} \dots a_n^{\dagger j_n} |0\rangle^{\otimes n} \right\rangle \end{aligned} \tag{2.37}$$

for all $j_1, \dots, j_n, k_1, \dots, k_n \in \mathbb{N}$. We prove that the action of $V_U^\dagger$ is given by

$$a_j^\dagger \to \sum_{i=1}^n U_{j,i}^\dagger a_i^\dagger \tag{2.38}$$

by showing that (2.37) holds under this choice. Indeed,

$$\begin{aligned}&\left\langle a_1^{\dagger k_1} \dots a_n^{\dagger k_n} |0\rangle^{\otimes n}, \left(\sum_{i=1}^n U_{1,i} a_i^\dagger\right)^{j_1} \dots \left(\sum_{i=1}^n U_{n,i} a_i^\dagger\right)^{j_n} |0\rangle^{\otimes n} \right\rangle \\ &= k_1! \dots k_n! \sum_{\substack{J \in M^n \text{ such that} \\ J_{i,l} \in \mathbb{N} \quad \forall i,l=1,\dots,n \\ \sum_{l=1}^n J_{i,l} = j_i \quad \forall i=1,\dots,n \\ \sum_{i=1}^n J_{i,l} = k_l \quad \forall l=1,\dots,n}} \prod_{i=1}^n \binom{j_i}{J_{1,i}, \dots, J_{n,i}} \prod_{p,q=1}^n U_{p,q}^{J_{p,q}} \\ &= j_1! \dots j_n! \sum_{\substack{J \in M^n \text{ such that} \\ J_{i,l} \in \mathbb{N} \quad \forall i,l=1,\dots,n \\ \sum_{l=1}^n J_{i,l} = j_i \quad \forall i=1,\dots,n \\ \sum_{i=1}^n J_{i,l} = k_l \quad \forall l=1,\dots,n}} \prod_{i=1}^n \binom{k_i}{J_{i,1}, \dots, J_{i,n}} \prod_{p,q=1}^n {U_{p,q}^T}^{J_{q,p}} \\ &= j_1! \dots j_n! \sum_{\substack{K \in M^n \text{ such that} \\ K(i,l) \in \mathbb{N} \quad \forall i,l=1,\dots,n \\ \sum_{l=1}^n K_{i,l} = k_i \quad \forall i=1,\dots,n \\ \sum_{i=1}^n K_{i,l} = j_l \quad \forall l=1,\dots,n}} \prod_{i=1}^n \binom{k_i}{K_{1,i}, \dots, K_{n,i}} \prod_{p,q=1}^n {U_{p,q}^T}^{K_{p,q}} \\ &= \left\langle \left(\sum_{i=1}^n U_{1,i}^\dagger a_i^\dagger\right)^{k_1} \dots \left(\sum_{i=1}^n U_{n,i}^\dagger a_i^\dagger\right)^{k_n} |0\rangle^{\otimes n}, a_1^{\dagger j_1} \dots a_n^{\dagger j_n} |0\rangle^{\otimes n} \right\rangle, \end{aligned} \tag{2.39}$$

which proves the claim.

Now we show that for U unitary, $V_U V_U^\dagger = V_U^\dagger V_U = \mathbb{1}_\infty^{\otimes n}$. Indeed for U unitary, for any basis vector $a_1^{\dagger j_1} \dots a_n^{\dagger j_n} |0\rangle^{\otimes n}$,

$$\begin{aligned}
&V_U^\dagger V_U a_1^{\dagger j_1} \dots a_n^{\dagger j_n} |0\rangle^{\otimes n} \\
&= V_U^\dagger \left(\sum_{i=1}^n U_{1,i} a_i^\dagger \right)^{j_1} \dots \left(\sum_{i=1}^n U_{n,i} a_i^\dagger \right)^{j_n} |0\rangle^{\otimes n} \\
&= \left(\sum_{i=1}^n U_{1,i} \sum_{l=1}^n U_{i,l}^\dagger a_l^\dagger \right)^{j_1} \dots \left(\sum_{i=1}^n U_{n,i} \sum_{l=1}^n U_{i,l}^\dagger a_l^\dagger \right)^{j_n} |0\rangle^{\otimes n} \\
&= a_1^{\dagger j_1} \dots a_n^{\dagger j_n} |0\rangle^{\otimes n},
\end{aligned} \tag{2.40}$$

and similarly

$$V_U V_U^\dagger a_1^{\dagger j_1} \dots a_n^{\dagger j_n} |0\rangle^{\otimes n} = a_1^{\dagger j_1} \dots a_n^{\dagger j_n} |0\rangle^{\otimes n}, \tag{2.41}$$

so $V_U V_U^\dagger = V_U^\dagger V_U = \mathbb{1}_\infty^{\otimes n}$. This completes the proof. □

For a basis vector $a_1^{\dagger j_1} \dots a_n^{\dagger j_n} |0\rangle^{\otimes n}$ we refer to $j_1 + \dots + j_n$ as the *total photon number*. Note that the inner product (2.37) is zero for any two basis vectors with different total photon numbers. Thus, the unitary V_U can be decomposed as a direct sum of unitary operators on invariant subspaces of fixed total photon number.

Beamsplitters and phase shifters

Now we introduce the linear optics transformations beamsplitters and phase-shifters. A *beamsplitter* across modes $1 \leq j,k \leq n$ with $j \neq k$ specified by splitting angle $\theta \in [0, 2\pi)$ and phase $\phi \in [0, 2\pi)$ is the linear optics transformation corresponding to the unitary

$$\hat{B}_{j,k}(\theta, \phi) = \mathbb{1}_n - |j\rangle\langle j| - |k\rangle\langle k| + \begin{pmatrix} |j\rangle & |k\rangle \end{pmatrix} \begin{pmatrix} \cos(\theta) & -\sin(\theta)e^{-i\phi} \\ \sin(\theta)e^{i\phi} & \cos(\theta) \end{pmatrix} \begin{pmatrix} \langle j| \\ \langle k| \end{pmatrix}. \tag{2.42}$$

We define the 50/50 beamsplitter as the linear optics transformation corresponding to $\hat{B}_{j,k}(\pi/4, 0)$ (the Hadamard gate). We define the *light port* and the *dark port* of the 50/50 beamsplitter as the modes j and k, respectively, immediately after they have passed through the beamsplitter.

A *phase-shifter* on mode $1 \leq j \leq n$ with phase $\phi \in [0,2\pi)$ is the linear optics transformation corresponding to the unitary

$$\hat{P}_j(\phi) = \mathbb{1}_n - |j\rangle\langle j| + e^{i\phi}|j\rangle\langle j|. \tag{2.43}$$

Using Lemma 1 (Baker-Hausdorff) it can be shown that the action of beamsplitters on coherent states is given by

$$V_{\hat{B}_{j,k}}(\theta,\phi)\,|\alpha\rangle_j\,|\beta\rangle_k = \left|\cos(\theta)\alpha + e^{i\phi}\sin(\theta)\beta\right\rangle_j \left|-\sin(\theta)e^{-i\phi}\alpha + \cos(\theta)\beta\right\rangle_k, \tag{2.44}$$

where

$$|\psi\rangle_j := |0\rangle^{\otimes(j-1)}\,|\psi\rangle\,|0\rangle^{\otimes n-j}. \tag{2.45}$$

This action gives some intuition for why we assign the names light port and dark port to the modes j and k, respectively. If $\alpha = \beta$, mode k will be in the vacuum state (dark), and mode j will be in state $|\sqrt{2}\alpha\rangle$ (light). Note that if $\alpha = -\beta$ these roles are reversed. In Section 4.4 we use this fact to show that the beamsplitter measurement with single photon threshold detectors placed at both the dark and light ports can be used to perform optimal unambiguous state comparison on $\{|\alpha\rangle, |-\alpha\rangle\}$ when the states are given with equal a priori probability.

The action of phase shifters on coherent states is given by

$$V_{\hat{P}_j}(\phi)\,|\alpha\rangle_j = \left|e^{i\phi}\alpha\right\rangle_j. \tag{2.46}$$

Note that $\hat{P}_j(\pi)\hat{B}_{j,k}(0,0)$ gives a control-NOT gate. Any unitary operator can be implemented with control-NOT gates, Hadamard gates $\hat{B}_{j,k}(\pi/4,0)$, and phase gates $\hat{P}_j(\pi/4)$ [48]. Thus, any linear optics transformation can be implemented with beamsplitters and phase shifters [54].

2.3 Coherent state mapping

Arrazola and Lütkenhaus recently proposed a mapping from any quantum protocol which uses pure quantum states, unitary operations, and standard basis measurements to a corresponding protocol which uses coherent states, linear optics transformations, and single photon detectors [3]. Here we review this mapping, and in Chapter 5 we apply it to the protocol of appointment scheduling.

For any quantum protocol which uses pure states $|\psi\rangle \in \mathcal{S}(\mathbb{C}^n)$ for some finite $n \in \mathbb{N}$ and unitary operations $U \in \mathcal{U}(\mathbb{C}^n)$, the coherent state mapping proceeds as follows: For fixed $\alpha \in \mathbb{C}$, each pure state $|\psi\rangle$ is mapped to a tensor product of coherent states according to the function

$$f_\alpha : \mathbb{C}^n \to L^2(\mathbb{R})^{\otimes n} \tag{2.47}$$

defined as

$$f_\alpha\left(\sum_{i=1}^{n} \lambda_i |i\rangle\right) = \bigotimes_{i=1}^{n} |\lambda_i \alpha\rangle_i, \tag{2.48}$$

and each unitary $U \in \mathcal{U}(\mathbb{C}^n)$ is mapped to the linear optics transformation $V_U \in \mathcal{U}(L^2(\mathbb{R})^{\otimes n})$. The unitary V_U can be shown to satisfy $V_U f_\alpha(|\psi\rangle) = f_\alpha(U|\psi\rangle)$ for all $|\psi\rangle \in \mathbb{C}^n$, which ensures that at a given stage in the protocols, the state of the mapped protocol is equal to f_α applied to the state of the original protocol.

Standard basis measurement in the original protocol is mapped to photon number counting measurement performed on each mode. Arrazola and Lütkenhaus show that the probability distribution of the number of photons measured in each mode is equal to that obtained from repeated canonical basis measurements of the state of the original quantum protocol, where the number of repetitions is drawn from a Poisson distribution with mean $|\alpha|^2$. Alternatively, standard basis measurement can be mapped to single-photon threshold detection (described by the two-outcome measurement $\{|0\rangle\langle 0|^{\otimes n}, \mathbb{1} - |0\rangle\langle 0|^{\otimes n}\}$), which is easier to implement at the cost of not having such a direct connection with the original measurement.

A notable property of the mapped states $f_\alpha(|\psi\rangle)$ derived in [3] is that they are close in trace distance to states of $\mathcal{O}(\log n)$ qubits, just like the original protocol:

Theorem 10 (Adapted from [3]). *For any $\alpha \in \mathbb{C}$ the following holds: for any $\varepsilon > 0$ there exists a sequence of subspaces $\mathcal{H}_{\alpha,n} \subset L^2(\mathbb{R})^{\otimes n}$ such that*

$$\dim(\mathcal{H}_{\alpha,n}) = \mathcal{O}(\log n) \tag{2.49}$$

and for all $|\psi\rangle \in \mathcal{S}(\mathbb{C}^n)$ the states

$$g_\alpha(|\psi\rangle) = \frac{\Pi_{\mathcal{H}_{\alpha,n}} f_\alpha(|\psi\rangle)}{\|\Pi_{\mathcal{H}_{\alpha,n}} f_\alpha(|\psi\rangle)\|_2} \tag{2.50}$$

satisfy

$$\| |g_\alpha(|\psi\rangle)\rangle\langle g_\alpha(|\psi\rangle)| - |f_\alpha(|\psi\rangle)\rangle\langle f_\alpha(|\psi\rangle)| \|_1 \leq \varepsilon. \tag{2.51}$$

In Chapters 4 and 5 we extend this result to show that the asymptotic behaviour of the information leakage and information cost of many quantum communication protocols are preserved under the coherent state mapping. The apparent usefulness of this mapping leads us to ask whether other mappings might exist from general quantum protocols to protocols which are more realistic for implementation, while retaining certain figures of merit. We ask, for example, whether one might find such a mapping to protocols using tensor products of qubits or other easily implementable states.

Chapter 3

Communication complexity and information complexity

Here we review two subfields of quantum communication: communication complexity and information complexity. Communication complexity is concerned with the amount of resources (e.g. bits, qubits, and entanglement) exchanged during communication protocols. Figured more prominently in this work is information complexity, which is concerned with the amount of information exchanged during communication protocols, and which has recently received heightened interest in part because the quantities studied can be easier to work with than those of communication complexity. One commonly studied quantity of information complexity is the amount of resources needed per task to execute asymptotically many copies of the same task. These quantities are useful for proving lower bounds on the amount of resources used in the single shot setting. Information complexity also studies various quantifiers of privacy, i.e. how much one party learns about another's private information as a result of their communication.

Despite an early result of Holevo [36], stating that no fewer than n qubits must be used to transmit n bits from one party to another, there have been many examples in quantum communication in which quantum resources outperform their classical counterparts. We review a few of these results, borrowing ideas from the useful 2010 review [13].

We begin with a review of some results in communication complexity. Two of the earliest examples of a quantum improvement were found by Buhrman et al. [15] in the interactive communication setting (formally defined in Section 3.3), in which two parties (Alice and Bob) communicate back and forth to execute a task. In the first task, Alice and Bob each receive n-bit strings $x, y \in \{0,1\}^n$ which are guaranteed to be either equal or differ by $n/2$ bits, and they wish to evaluate the equality function on x and y with zero error. Buhrman et al. found a quantum protocol using $\log(n)$ qubits of communication, whereas any classical protocol must use at least $0.007n$ bits of communication [27]. In the second task (known as appointment scheduling), Alice and Bob again receive $x, y \in \{0,1\}^n$ (which can now be arbitary), and find a common index i for which $x_i = y_i = 1$. Buhrman et al. exhibited a quantum protocol in this case which uses nearly quadratically fewer qubits than the classical lower bound.

Another communication setting that has been studied is known as the simultaneous message passing model (formally defined in Section 3.2), in which Alice and Bob communicate one-way to a referee, who helps them peform some task. The first task with a quantum advantage in this model was found by Buhrman et al. [14]. Here, Alice and Bob each receive n-bit strings (which can now be arbitrary, in contrast to the equality task described above), and they communicate to a referee to evaluate the equality function. Buhrman et al. again found an exponential improvement over the classical lower bound in this setting. Since these early protocols, many more protocols have been found which exhibit quantum advantages in communication complexity under various resource and locality restrictions.

We now shift our focus to information complexity, which will be the primary framework we use to analyze communication protocols in this thesis. We first review some fundamental results in information complexity which quantify the asymptotic amount of resources needed to perform the primitive communication tasks of source coding and channel coding. We then use these primitives to form notions of information for tasks in the more complicated settings of the simultaneous message passing model and the interactive communication model, which we use in Chapters 4 and 5 to analyze the information content of practical quantum fingerprinting and appointment scheduling protocols that we have developed. See the 2012 review [11] for further reading on classical information complexity, and see [64] for further reading on quantum information complexity.

3.1 Background

In this section we motivate several entropic quantities which we will use to quantify information in this work through the settings of source coding and channel coding. For a positive operator $\rho \in \mathrm{Pos}(\mathcal{ABC})$ with eigenvalues λ_a, $a \in \mathrm{A} \times \mathrm{B} \times \mathrm{C}$, define the *entropy* of ρ as

$$\mathrm{H}(\rho) = \sum_{a \in \mathrm{A} \times \mathrm{B} \times \mathrm{C}} \lambda_a \log\left(\frac{1}{\lambda_a}\right) \tag{3.1}$$

(with the definition $0\log(1/0) := 0$) and the *mutual information* between registers A and B as

$$\mathrm{I}(\rho^A : \rho^B) = \mathrm{H}(\rho^A) + \mathrm{H}(\rho^B) - \mathrm{H}(\rho^{AB}). \tag{3.2}$$

We will frequently write $\mathrm{H}(A) = \mathrm{H}(\rho^A)$ and $\mathrm{I}(A : B) = \mathrm{I}(\rho^A : \rho^B)$ as shorthand. Define the *conditional entropy* as

$$\mathrm{H}(A|B) = \mathrm{H}(AB) - \mathrm{H}(B) \tag{3.3}$$

and the *conditional mutual information* as

$$\mathrm{I}(A : B|C) = \mathrm{H}(A|C) + \mathrm{H}(B|C) - \mathrm{H}(AB|C). \tag{3.4}$$

When $\rho \in \mathcal{D}(\mathcal{ABC})$ is a state, the mutual information somehow characterizes the amount of information that register A contains about register B. We will see several motivations for this intuition, but as a first motivation note the limiting behaviour: the mutual information is minimized to zero when ρ is a product state between the two input registers, and it is maximized when ρ is maximally entangled between the two input registers.

In Section 3.1.1 (source coding) we motivate the entropy and mutual information as asymptotic quantifiers of the minimum size to which the input state can be compressed and then successfully decompressed. In Section 3.1.2 (channel coding) we motivate the mutual information as an asymptotic quantifier of the maximum size of classical message that can be encoded into register A and decoded from register B.

3.1.1 Source coding

Here we state source coding theorems, which motivate both the entropy and mutual information functions as optimal asymptotic communication rates. First we consider a classical scenario in which Alice records the outcomes of some stochastic process, which she wishes to transmit to Bob through a perfect channel using the least number of bits possible. To avoid introducing more notation, we use the (somewhat bloated) language of quantum information for this scenario. Instead of Alice recording the outcomes of some classical random variable, we say she measures some classical state $P \in \mathcal{D}(\mathcal{C})$ contained in some classical register C in the standard basis. She then sends the outcome through a classical encoding channel, after which Bob applies a classical decoding channel and measures in the standard basis in hopes of recovering the outcome. A classical register is a register whose state is always diagonal with respect to the standard basis. A classical channel is any channel which can be written as $\Delta\Phi\Delta$, where Φ is some quantum channel and Δ is the completely dephasing channel with respect to the standard basis. An achievable rate is then defined as follows:

Definition 11. *We say a positive number* S *is an* ***achievable rate*** *for source coding of a classical state* $P \in \mathcal{D}(\mathcal{C})$ *if there exists a sequence of classical registers* X_N *satisfying*

$$\lim_{N\to\infty} \frac{N}{\log(|\mathrm{X}_N|)} = \mathrm{S} \tag{3.5}$$

and corresponding sequences of classical encoding channels $\Psi_{E,N} \in C(\mathcal{C}^{\otimes N}, \mathcal{X}_N)$*, and classical decoding channels* $\Psi_{D,N} \in C(\mathcal{X}_N, \mathcal{C}^{\otimes N})$ *such that*

$$\lim_{N\to\infty} \sum_{c_N \in \mathrm{C}^{\otimes N}} \langle c_N | P^{\otimes N} | c_N \rangle \langle c_N | \Psi_{D,N} \Psi_{E,N}(|c_N\rangle\langle c_N|) | c_N \rangle = 1, \tag{3.6}$$

i.e. the average probability of successful decoding goes to one in the limit $N \to \infty$*.*

Shannon's source coding theorem states that the optimal rate is exactly given by the entropy function:

Theorem 12 (Shannon's source coding theorem [59])**.** *The infimum over all achievable rates for source coding of a classical state* $P \in \mathcal{D}(\mathcal{C})$ *is equal to* $\mathrm{H}(P)$*. Conversely, for every sequence of classical registers* X_N *satisfying*

$$\lim_{N\to\infty} \frac{N}{\log(|\mathrm{X}_N|)} < \mathrm{H}(P), \tag{3.7}$$

every sequence of classical encoding channels $\Psi_{E,N} \in C(\mathcal{C}^{\otimes N}, \mathcal{X}_N)$ *and classical decoding channels* $\Psi_{D,N} \in C(\mathcal{X}_N, \mathcal{C}^{\otimes N})$ *satisfies*

$$\lim_{N\to\infty} \sum_{c_N \in \mathrm{C}^{\otimes N}} \langle c_N | P^{\otimes N} | c_N \rangle \langle c_N | \Psi_{D,N} \Psi_{E,N}(|c_N\rangle\langle c_N|) | c_N \rangle = 0, \tag{3.8}$$

i.e. the average success probability goes to zero in the limit $N \to \infty$*.*

Now we state a generalization of Shannon's source coding theorem in which C is a quantum register and P is an arbitrary quantum state (which we now denote ρ). Here, Alice sends ρ^C to Bob while preserving correlations with a purifying register R, i.e. the purification ρ^{CR} is effectively unchanged by the encoding and decoding maps. Accordingly, we now generalize the above definition of an achievable rate to quantum states.

Definition 13. *We say a positive number* S *is an* ***achievable rate*** *for a state* $\rho \in \mathcal{D}(\mathcal{C})$ *if there exists a sequence of registers* X_N *satisfying*

$$\lim_{N\to\infty} \frac{N}{\log(|\mathrm{X}_N|)} = \mathrm{S} \tag{3.9}$$

and corresponding sequences of encoding channels $\Psi_{E,N} \in C(\mathcal{C}^{\otimes N}, \mathcal{X}_N)$, and decoding channels $\Psi_{D,N} \in C(\mathcal{X}_N, \mathcal{C}^{\otimes N})$ such that

$$\lim_{N\to\infty} \mathrm{F}\left((\Psi_{D,N}\Psi_{E,N} \otimes \mathbb{1}_R)(\rho_{CR}^{\otimes N}), \rho_{CR}^{\otimes N}\right) = 1 \tag{3.10}$$

for any purification register R.

Schumacher's generalization of Shannon's source coding theorem states that the optimal rate is again given by the entropy function.

Theorem 14 (Schumacher's source coding theorem [56])**.** *The infimum over all achievable rates for source coding of any quantum state $\rho \in \mathcal{D}(\mathcal{C})$ is given by* $\mathrm{H}(\rho)$*. Conversely, for every sequence of registers X_N satisfying*

$$\lim_{N\to\infty} \frac{N}{\log(|\mathrm{X}_N|)} < \mathrm{H}(\rho), \tag{3.11}$$

every sequence of encoding channels $\Psi_{E,N} \in C(\mathcal{C}^{\otimes N}, \mathcal{X}_N)$ and decoding channels $\Psi_{D,N} \in C(\mathcal{X}_N, \mathcal{C}^{\otimes N})$ satisfies

$$\lim_{N\to\infty} \mathrm{F}((\Psi_{D,N}\Psi_{E,N} \otimes \mathbb{1}_R)(\rho_{CR}^{\otimes N}), \rho_{CR}^{\otimes N}) = 0. \tag{3.12}$$

for every purification register R.

Now we note that Schumacher's theorem also provides an interpretation of the mutual information which further motivates its use as a quantifier of information. Recall the definition $\mathrm{I}(A:B) = \mathrm{H}(A) + \mathrm{H}(B) - \mathrm{H}(AB)$. By Schumacher's theorem, the mutual information is exactly the difference between the following two quantities: one, the asymptotic rate of communication needed to transmit ρ^A and then ρ^B separately; and two, that needed to transmit the entire state ρ^{AB} at once. In this way, the mutual information somehow quantifies the amount of information that register A contains about register B (and vice versa).

Now we consider a yet more general task known as state redistribution [69][23] and motivate the conditional mutual information as the optimal rate for this task. In this task, two parties (Alice and Bob) share a state ρ^{ABC} with some purification ρ^{ABCR}. The state ρ^{ABCR} is known to (but not held by) both Alice and Bob. In the beginning, Alice holds register AC and Bob holds register B. Alice wishes to transmit the register C to Bob (using the minimum number of qubits) through a perfect channel. We will see that Alice and Bob can substitute some of their quantum communication for entanglement in this task, and they can even generate entanglement in some circumstances.

Define $|\Phi^{+}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ the canonical maximally entangled state, and for two registers X and X' with $|\mathrm{X}| = |\mathrm{X}'|$ a power of two, define $|\Phi^{+}\rangle\langle\Phi^{+}|^{XX'} = |\Phi^{+}\rangle\langle\Phi^{+}|^{\otimes \log|\mathrm{X}|} \in \mathcal{D}(\mathcal{X}\mathcal{X}')$, where the first qubit of each maximally entangled state is contained in register X, and the second is contained in register X'.

Definition 15. *We say a two-tuple* (S,E) *of positive numbers is an* ***achievable rate tuple*** *for state redistribution from Alice to Bob for the state* ρ^{ABC} *if there exists a sequence of registers* $X_N, T^{\text{in}}_{A,N}, T^{\text{in}}_{B,N}, T^{\text{out}}_{A,N}, T^{\text{out}}_{B,N}$ *satisfying* $|\mathrm{T}^{\text{in}}_{A,N}| = |\mathrm{T}^{\text{in}}_{B,N}|$, $|\mathrm{T}^{\text{out}}_{A,N}| = |\mathrm{T}^{\text{out}}_{B,N}|$,

$$\lim_{N\to\infty} \frac{N}{\log(|\mathrm{X}_N|)} = \mathrm{S}, \tag{3.13}$$

and

$$\lim_{N\to\infty} \frac{N}{\log(|\mathrm{T}^{\text{in}}_{A,N}|) - \log(|\mathrm{T}^{\text{out}}_{A,N}|)} = \mathrm{E}, \tag{3.14}$$

and corresponding sequences of encoding channels $\Psi_{E,N} \in C(\mathcal{A}^{\otimes N}\mathcal{C}^{\otimes N}\mathcal{T}^{\text{in}}_{A}, \mathcal{A}^{\otimes N}\mathcal{X}_N\mathcal{T}^{\text{out}}_{A})$ *and decoding channels* $\Psi_{D,N} \in C(\mathcal{B}^{\otimes N}\mathcal{X}_N\mathcal{T}^{\text{in}}_{B}, \mathcal{B}^{\otimes N}\mathcal{C}^{\otimes N}\mathcal{T}^{\text{out}}_{B})$ *satisfying*

$$\lim_{N\to\infty} \mathrm{F}\left(\mathrm{Tr}_{T^{\text{out}}_{A,N}T^{\text{out}}_{B,N}} \Psi_{D,N}\Psi_{E,N}\left(\rho^{\otimes N}_{ABCR} \otimes |\Phi^{+}\rangle\langle\Phi^{+}|^{T^{\text{in}}_{A,N}T^{\text{in}}_{B,N}}\right), \rho^{\otimes N}_{ABCR}\right) = 1 \tag{3.15}$$

and

$$\lim_{N\to\infty} \mathrm{F}\left(\mathrm{Tr}_{\neg T^{\text{out}}_{A,N}T^{\text{out}}_{B,N}} \Psi_{D,N}\Psi_{E,N}\left(\rho^{\otimes N}_{ABCR} \otimes |\Phi^{+}\rangle\langle\Phi^{+}|^{T^{\text{in}}_{A,N}T^{\text{in}}_{B,N}}\right), |\Phi^{+}\rangle\langle\Phi^{+}|^{T^{\text{out}}_{A,N}T^{\text{out}}_{B,N}}\right) = 1, \tag{3.16}$$

where we have implicitly tensored the encoding and decoding channels with $\mathbb{1}^{\otimes N}_{\mathcal{B}} \otimes \mathbb{1}_{\mathcal{R}}$ *and* $\mathbb{1}^{\otimes N}_{\mathcal{A}} \otimes \mathbb{1}_{\mathcal{R}}$*, respectively. We say* S *is* ***the communication rate*** *and* E *is the* ***entanglement cost*** *of the rate tuple* (S,E).

Note that the quantity E can be negative, in which case entanglement is generated rather than consumed. The following theorem equates the optimal rate communication rate with the coniditonal mutual information.

Theorem 16 (State redistribution [69][23])**.** *For any state* ρ^{ABC}*, any rate tuple* (S,E) *satisfying* $\mathrm{S} > \frac{1}{2}\mathrm{I}(C:R|B)$ *and* $\mathrm{S}+\mathrm{E} > \mathrm{H}(C|B)$ *is achievable. Conversely, no tuple* (S,E) *with* $\mathrm{S} < \frac{1}{2}\mathrm{I}(C:R|B)$ *or* $\mathrm{S}+\mathrm{E} < \mathrm{H}(C|B)$ *is achievable.*

A strong converse to this theorem also holds:

Theorem 17 (Strong converse to state redistribution [43][9])**.** *For any two-tuple* (S,E) *of positive numbers for which* $\mathrm{S} < \frac{1}{2}\mathrm{I}(C:R|B)$ *or* $\mathrm{S}+\mathrm{E} < \mathrm{H}(C|B)$*, the following holds: for every sequence of registers* $X_N, T^{\text{in}}_{A,N}, T^{\text{in}}_{B,N}, T^{\text{out}}_{A,N}, T^{\text{out}}_{B,N}$ *satisfying* $|\mathrm{T}^{\text{in}}_{A,N}| = |\mathrm{T}^{\text{in}}_{B,N}|$, $|\mathrm{T}^{\text{out}}_{A,N}| = |\mathrm{T}^{\text{out}}_{B,N}|$,

$$\lim_{N\to\infty} \frac{N}{\log(|\mathrm{X}_N|)} = \mathrm{S}, \tag{3.17}$$

and

$$\lim_{N\to\infty} \frac{N}{\log(|\mathrm{T}^{\text{in}}_{A,N}|) - \log(|\mathrm{T}^{\text{out}}_{A,N}|)} = \mathrm{E}, \tag{3.18}$$

and corresponding sequences of encoding channels $\Psi_{E,N} \in C(\mathcal{A}^{\otimes N}\mathcal{C}^{\otimes N}\mathcal{T}^{\text{in}}_A, \mathcal{A}^{\otimes N}\mathcal{X}_N\mathcal{T}^{\text{out}}_A)$ *and decoding channels* $\Psi_{D,N} \in C(\mathcal{B}^{\otimes N}\mathcal{X}_N\mathcal{T}^{\text{in}}_B, \mathcal{B}^{\otimes N}\mathcal{C}^{\otimes N}\mathcal{T}^{\text{out}}_B)$*,*

$$\lim_{N\to\infty} \mathrm{F}\left(\mathrm{Tr}_{T^{\text{out}}_{A,N}T^{\text{out}}_{B,N}} \Psi_{D,N}\Psi_{E,N}\left(\rho^{\otimes N}_{ABCR} \otimes |\Phi^+\rangle\langle\Phi^+|^{T^{\text{in}}_{A,N}T^{\text{in}}_{B,N}}\right), \rho^{\otimes N}_{ABCR}\right) = 0. \tag{3.19}$$

In [43] and [9] these state redistribution theorems are also proven in the scenario when Alice and Bob are allowed interactive (two-way) communication to redistribute the state ρ (instead of the strictly forward communication scenario we have presented here), and the same results are obtained.

3.1.2 Channel coding

Here we further motivate the mutual information as a quantifier of information through results in channel coding. In the previous section we determined, for a fixed state held by Alice (and known to Bob), the asymptotic number of (qu)bits she needs to send through a perfect channel to transmit the state to Bob while maintaining correlations with other registers. For channel coding, we instead fix a channel, and determine the asymptotic number of arbitrary (qu)bits per channel use that can be sent.

Unfortunately, both the source coding and channel coding quantities of interest are frequently referred to as the *rate* of the state and channel, respectively. Note that we wish to minimize the rate for source coding, while we wish to maximize the rate for channel coding. To avoid ambiguity, we use the symbol R to denote an achievable rate for channel coding.

Classical capacity of channels

First we consider a classical communication scenario in which one party (Alice) wishes to communicate a classical message $x \in \mathrm{X}$ for some classical state set X to another party (Bob) with high probability using a channel $\Phi \in C(\mathcal{A}, \mathcal{B})$. To do so, Alice prepares $|x\rangle\langle x| \in \mathcal{L}(\mathcal{X})$ in a classical register X and applies an encoding channel $\Psi_E \in C(\mathcal{X}, \mathcal{A})$, the output of which she sends through Φ. Bob then applies a decoding channel $\Psi_D \in C(\mathcal{B}, \mathcal{X})$ in hopes of recovering $|x\rangle\langle x|$. We define various error probabilities associated with the channel $\Psi_D \Phi \Psi_E \in C(\mathcal{X})$ as follows:

$$\begin{aligned} \mathrm{Pr}_e(x) &= 1 - \langle x|\, \Psi_D \Phi \Psi_E(|x\rangle\langle x|)\, |x\rangle \\ \mathrm{Pr}_e(\Psi_D \Phi \Psi_E) &= \max_{x \in \mathrm{X}} \mathrm{Pr}_e(x) \\ \overline{\mathrm{Pr}}_e(\Psi_D \Phi \Psi_E) &= \frac{1}{|\mathrm{X}|} \sum_{x \in \mathrm{X}} \mathrm{Pr}_e(x), \end{aligned}$$

the probability that Bob fails to recover x on input x, the worst case such probability over all $x \in \mathrm{X}$, and the average such probability over all $x \in \mathrm{X}$, respectively.

We now define an achievable rate for classical channel coding and the classical capacity of a channel:

Definition 18. *For a channel $\Phi \in C(\mathcal{A}, \mathcal{B})$, we say a positive number R is an **achievable rate** for classical channel coding if there exists a sequence of classical registers X_N satisfying*

$$\lim_{N \to \infty} \frac{\log |\mathrm{X}_N|}{N} = \mathrm{R} \tag{3.20}$$

and a corresponding sequence of encoding channels $\Psi_{E,N} \in C(\mathcal{X}_N, \mathcal{A}^{\otimes N})$, and decoding channels $\Psi_{D,N} \in C(\mathcal{B}^{\otimes N}, \mathcal{X}_N)$ such that

$$\lim_{N \to \infty} \mathrm{Pr}_e(\Psi_{D,N} \Phi^{\otimes N} \Psi_{E,N}) = \lim_{N \to \infty} \overline{\mathrm{Pr}}_e(\Psi_{D,N} \Phi^{\otimes N} \Psi_{E,N}) = 0. \tag{3.21}$$

*We define the **classical capacity** $\mathrm{C}(\Phi)$ as the supremum over all achievable rates.*

Now we present the Holevo-Schumacher-Westmoreland theorem for the classical capacity of a channel. This theorem strengthens our intuition for the mutual information as an asymptotic quantifier of the information shared between two registers. First, we need the following definition.

Definition 19. *For some registers A and B we define the set of* ***classical-quantum*** *states* $\mathrm{CQ}(\mathcal{X}:\mathcal{A}) \subset \mathcal{D}(\mathcal{X}\mathcal{A})$ *as those states which can be written as*

$$\sum_{x\in \mathrm{X}} P(x)|x\rangle\langle x| \otimes \sigma_x^A \tag{3.22}$$

for some probability distribution $P \in \Pr(\mathrm{X})$ *and states* $\sigma_x^A \in \mathcal{D}(\mathcal{A})$.

Now we state the theorem.

Theorem 20 (Classical capacity of a channel [35][57])**.** *For any channel* $\Phi \in C(\mathcal{A},\mathcal{B})$*,*

$$\mathrm{C}(\Phi) = \lim_{l\to\infty} \frac{1}{l} \sup_{Y} \sup_{\sigma\in \mathrm{CQ}(\mathcal{Y}:\mathcal{A}^{\otimes l})} \mathrm{I}(Y:B^{\otimes l})_{(\mathbb{1}_{\mathcal{Y}}\otimes\Phi^{\otimes l})(\sigma)} \tag{3.23}$$

where the left supremum is taken over all classical registers Y.

Theorem 20 equates the classical capacity of a channel with the limiting behaviour of the mutual information. We now restrict our attention to a set of channels for which the channel capacity is given by a single expression of the mutual information (rather than some limiting behaviour). This will solidify our intuition for expressions involving the mutual information (without limiting behaviour) as quantifiers of information.

We first consider the set of entanglement breaking channels, i.e. channels $\Phi \in C(\mathcal{A},\mathcal{B})$ for which $(\mathbb{1}_{\mathcal{Y}}\otimes\Phi)(\sigma) \in \mathrm{SepD}(\mathcal{Y}:\mathcal{B})$ for all states $\sigma \in \mathcal{D}(\mathcal{Y}\mathcal{A})$. We then present a corollary for classical-quantum channels (a subset of entanglement breaking channels), which we will use in Section 3.2 as motivation for the definition of the information leakage of simultaneous message passing model protocols.

Corollary 21 ([61])**.** *Let* $\Phi \in C(\mathcal{A},\mathcal{B})$ *be an entanglement breaking channel. Then*

$$\mathrm{C}(\Phi) = \sup_{Y} \sup_{\sigma\in \mathrm{CQ}(\mathcal{Y}:\mathcal{A})} \mathrm{I}(Y:B)_{(\mathbb{1}_{\mathcal{Y}}\otimes\Phi)(\sigma)}, \tag{3.24}$$

where the left supremum is taken over all classical registers Y.

Proof. The inequality

$$\frac{1}{l} \sup_{Y'} \sup_{\sigma\in \mathrm{CQ}(\mathcal{Y}':\mathcal{A}^{\otimes l})} \mathrm{I}(Y':B^{\otimes l})_{(\mathbb{1}_{\mathcal{Y}'}\otimes\Phi^{\otimes l})(\sigma)} \geq \sup_{Y} \sup_{\sigma\in \mathrm{CQ}(\mathcal{Y}:\mathcal{A})} \mathrm{I}(Y:B)_{(\mathbb{1}_{\mathcal{Y}}\otimes\Phi)(\sigma)} \tag{3.25}$$

for any channel $\Phi \in C(\mathcal{A},\mathcal{B})$ and any $l \in \mathbb{N}$ follows from the easily-ascertained fact that for any classical-quantum state $\sigma \in \mathrm{CQ}(\mathcal{Y} : \mathcal{A})$,

$$\frac{1}{l}\mathrm{I}(Y^{\otimes l} : B^{\otimes l})_{(\mathbb{1}_{\mathcal{Y}}^{\otimes l}\otimes\Phi^{\otimes l})(\rho)} = \mathrm{I}(Y : B)_{(\mathbb{1}_{\mathcal{Y}}\otimes\Phi)(\sigma)} \tag{3.26}$$

for $\rho = W\sigma^{\otimes l}W^{\dagger} \in \mathcal{D}(\mathcal{Y}^{\otimes l}\mathcal{B}^{\otimes l})$, where $W \in \mathcal{U}((\mathcal{Y}\mathcal{B})^{\otimes l}, \mathcal{Y}^{\otimes l}\mathcal{B}^{\otimes l})$ is defined as

$$W\,|y_1 b_1 \ldots y_l b_l\rangle = |y_1 \ldots y_l b_1 \ldots b_l\rangle \tag{3.27}$$

for each $y_1,\ldots,y_l \in \mathrm{Y}$ and $b_1,\ldots,b_l \in \mathrm{B}$.

Now we prove the reverse inequality when Φ is entanglement breaking. As Φ is entanglement breaking, then for any classical register Y' and state $\sigma \in \mathrm{CQ}(\mathcal{Y}' : \mathcal{A}^{\otimes l})$,

$$(\mathbb{1}_{\mathcal{Y}'}\otimes\Phi^{\otimes l})(\sigma) = \sum_{y\in\mathrm{Y}'} P(y)|y\rangle\langle y| \otimes \rho_{y,1}\otimes\cdots\otimes\rho_{y,l} \tag{3.28}$$

for some probability vector $P \in \mathrm{Pr}(\mathrm{Y}')$ and states $\rho_{y,j} \in \mathcal{D}(\mathcal{B})$. Thus,

$$\begin{aligned}
\mathrm{I}(Y' : B^{\otimes l})_{(\mathbb{1}_{\mathcal{Y}'}\otimes\Phi^{\otimes l})(\sigma)} &= \mathrm{H}\left(\sum_{y\in\mathrm{Y}'} P(y)\rho_{y,1}\otimes\cdots\otimes\rho_{y,l}\right) - \sum_{y\in\mathrm{Y}'}\sum_{j=1}^{l} P(y)\mathrm{H}(\rho_{y,j}) \\
&\le \sum_{j=1}^{l}\left[\mathrm{H}\left(\sum_{y\in\mathrm{Y}'} P(y)\rho_{y,j}\right) - \sum_{y\in\mathrm{Y}'} P(y)\mathrm{H}(\rho_{y,j})\right] \\
&= \sum_{j=1}^{l} \mathrm{I}(Y' : B)_{(\mathbb{1}_{\mathcal{Y}'}\otimes\Phi)(\sigma_j)} \\
&\le l \max_{j\in[l]} \mathrm{I}(Y' : B)_{(\mathbb{1}_{\mathcal{Y}'}\otimes\Phi)(\sigma_j)}
\end{aligned}$$

for states $\sigma_j = \sum_{y\in\mathrm{Y}'} P(y)\rho_{y,j} \in \mathcal{D}(\mathcal{A})$, $j = 1,\ldots,l$. The first equality follows from additivity of entropy under tensor product. The first inequality follows from Lemma 44 (subadditivity), and the rest of the above (in)equalities are straightforward. The result follows. □

In summary, Corollary 21 tells us that for entanglement breaking channels, the classical capacity is simply the supremum (taken over all inputs to the channel) of the mutual information between the purifying register of the input and the output of the channel.

Now we present a corollary for the special case of classical-quantum channels. A channel $\Phi \in C(\mathcal{A},\mathcal{B})$ is a *classical-quantum channel* if it can be written as $\Phi = \Psi\Delta$, where Δ is the completely dephasing channel with respect to the register A. Note that such channels are uniquely defined by their action on classical inputs. The following is an easy corollary to Corollary 21 for classical-quantum channels, and demonstrates an even more direct connection between the classical capacity and the mutual information.

Corollary 22 (Classical capacity of a CQ channel)**.** *Let $\Phi \in C(\mathcal{A}, \mathcal{B})$ be a classical-quantum channel defined as $\Phi(|a\rangle\langle a|) = \sigma_a^B$. For any probability vector $Q \in \mathrm{Pr}(\mathrm{A})$, define*

$$\sigma_Q^{AB} = \sum_{a \in \mathrm{A}} Q(a)|a\rangle\langle a| \otimes \sigma_a^B. \tag{3.29}$$

Then,

$$\mathrm{C}(\Phi) = \sup_{Q \in \mathrm{Pr}(\mathrm{A})} \mathrm{I}(A:B)_{\sigma_Q^{AB}}. \tag{3.30}$$

For classical-quantum channels, a strong converse theorem also holds:

Theorem 23 (Strong converse to classical capacity of a CQ channel [49])**.** *Let Φ be a classical-quantum channel defined as above, and for any probability vector $Q \in \mathrm{Pr}(\mathrm{A})$ let σ_Q^{AB} be defined as above. Then for any sequence of registers X_N and corresponding sequences of encoding and decoding channels $\Psi_{E,N} \in C(\mathcal{X}_N, \mathcal{A}^{\otimes N})$ and $\Psi_{D,N} \in C(\mathcal{B}^{\otimes N}, \mathcal{X}_N)$, if*

$$\lim_{N \to \infty} \frac{\log |\mathrm{X}_N|}{N} > \sup_{Q \in \mathrm{Pr}(\mathrm{A})} \mathrm{I}(A:B)_{\sigma_Q^{AB}}, \tag{3.31}$$

then

$$\lim_{N \to \infty} \mathrm{Pr}_e(\Psi_{D,N}\Phi^{\otimes N}\Psi_{E,N}) = \lim_{N \to \infty} \overline{\mathrm{Pr}}_e(\Psi_{D,N}\Phi^{\otimes N}\Psi_{E,N}) = 1. \tag{3.32}$$

In Section 3.2 we use the above result for classical-quantum channels to motivate the definition of the information leakage of protocols in the simultaneous message passing model.

Quantum capacity of channels

Now we consider a scenario in which Alice wishes to communicate arbitrary quantum states $\rho \in \mathcal{D}(\mathcal{X})$ to Bob using a quantum channel $\Phi \in C(\mathcal{A}, \mathcal{B})$. One motivation for this scenario is that Alice wishes to communicate ρ through a communication channel which introduces loss/noise which can be described by Φ. For example, we could have $\mathcal{X} = \mathbb{C}^2$ and a channel Φ that implements a bit flip or phase flip with some probability. We include quantum channel coding because it is instructive to see how the classical capacity theorems generalize to the quantum case, but the results presented in this section will not be used in the remainder of this thesis.

Now we define an achievable rate for quantum channel coding and the quantum capacity of a channel.

Definition 24. *For a quantum channel $\Phi \in C(\mathcal{A}, \mathcal{B})$, we say a positive number* R *is an **achievable rate** for quantum channel coding if there exists a sequence of registers* X_N *satisfying*

$$\lim_{N\to\infty} \frac{\log|\mathrm{X}_N|}{N} = \mathrm{R} \tag{3.33}$$

and a corresponding sequence of encoding and decoding channels $\Psi_{E,N} \in C(\mathcal{X}_N, \mathcal{A}^{\otimes N})$, $\Psi_{D,N} \in C(\mathcal{B}^{\otimes N}, \mathcal{X}_N)$ such that

$$\lim_{N\to\infty} \inf_{R} \inf_{\rho\in\mathcal{D}(\mathcal{X}_N\mathcal{R})} \mathrm{F}\left(\left(\Psi_{D,N}\Phi^{\otimes N}\Psi_{E,N} \otimes \mathbb{1}_{\mathcal{R}}\right)(\rho), \rho\right) = 1 \tag{3.34}$$

where the left infimum is taken over all registers R. We define the quantum capacity $\mathrm{Q}(\Phi)$ *as the supremum over all achievable rates.*

To state the quantum capacity theorem, we need the following definition.

Definition 25. *For any bipartite state $\rho \in \mathcal{D}(\mathcal{X}\mathcal{Y})$, define the **coherent information** as*

$$\mathrm{I}_c(X > Y)_\rho := \mathrm{H}(Y) - \mathrm{H}(XY). \tag{3.35}$$

Now we state the quantum channel capacity theorem.

Theorem 26 (Quantum capacity of a channel [44][60][22]). *For any channel $\Phi \in C(\mathcal{A}, \mathcal{B})$,*

$$\mathrm{Q}(\Phi) = \lim_{l\to\infty} \frac{1}{l} \sup_{R} \sup_{\sigma\in\mathcal{D}(\mathcal{R}\mathcal{A}^{\otimes l})} \mathrm{I}_c(R > B^{\otimes l})_{(\mathbb{1}_{\mathcal{R}}\otimes\Phi^{\otimes l})(\sigma)} \tag{3.36}$$

where the left supremum is taken over all registers R.

Now we state a corollary of this theorem for degradable channels, proven in [21]. A channel Φ is *degradable* if there exists a channel Ψ such that $\Phi^C = \Psi\Phi$, where Φ^C is the complement of the channel Φ. The quantum capacity of a degradable channel takes a similar form to the classical capacity of an entanglement breaking channel.

Corollary 27 (Quantum capacity of a degradable channel [21])**.** *For any degradable channel $\Phi \in C(\mathcal{A}, \mathcal{B})$,*

$$Q(\Phi) = \sup_{R} \sup_{\sigma \in \mathcal{D}(\mathcal{R}\mathcal{A})} \mathrm{I}_c(R > B)_{(\mathbb{1}_{\mathcal{R}} \otimes \Phi)(\sigma)} \tag{3.37}$$

where the left supremum is taken over all registers R.

There has been work on strong converse theorems for quantum channel coding (along with some negative results). A “pretty strong” converse theorem for the quantum capacity of a degradable channel was proven in [46].

3.2 Simultaneous message passing model

The simultaneous message passing model is a communication setting consisting of three parties: Alice, Bob, and the referee. For some classical state sets X and Y, Alice and Bob receive inputs $x \in \mathrm{X}$ and $y \in \mathrm{Y}$, respectively, chosen according to some probability distribution $P \in \mathrm{Pr}(\mathrm{X} \times \mathrm{Y})$. Conditioned on x and y, Alice and Bob send some (quantum or classical) states $\sigma_x \in \mathcal{D}(\mathcal{A}), \sigma'_y \in \mathcal{D}(\mathcal{B})$ to the referee. The registers held by Alice, Bob, and the referee are constricted to be uncorrelated at the start of the protocol (besides the correlations introduced by P). Also disallowed are back communication from the referee to Alice and Bob, and any communication between Alice and Bob.

We frequently consider simultaneous message passing model protocols in which the referee performs a measurement on register AB to attempt to determine some relationship between x and y. In Chapter 4 we consider simultaneous message passing model protocols for which the referee determines the equality function on x and y with high probability. We also consider the case in which x and y are unit vectors and the referee determines the Euclidean distance between them with high probability.

Denoting the above general protocol by the symbol Π, we define the information leakage of Π as follows.

Definition 28. *For some input distribution P and protocol Π defined as above, let*

$$\rho_P = \sum_{x \in \mathrm{X}, y \in \mathrm{Y}} P(x,y) |xy\rangle\langle xy| \otimes \sigma_x \otimes \sigma'_y, \tag{3.38}$$

*and define the **information leakage** of Π on input P as*

$$\mathrm{QIL}(\Pi, P) = \mathrm{I}(XY : AB)_{\rho_P} \tag{3.39}$$

and the (prior free) information leakage of Π *as*

$$\mathrm{QIL}(\Pi) = \sup_{P \in \Pr(\mathrm{X}\times\mathrm{Y})} \mathrm{I}(XY : AB)_{\rho_P}. \tag{3.40}$$

Using our intuition for the mutual information as a quantifier of the amount of information one register contains about the other, the information leakage quantifies the amount of information Alice and Bob leak to the referee about their inputs by sending them the registers A and B. Furthermore, note that $\mathrm{QIL}(\Pi)$ is equal to the classical capacity of the classical-quantum channel $\Phi \in C(\mathcal{X}\mathcal{Y}, \mathcal{A}\mathcal{B})$ defined as $\Phi(|xy\rangle\langle xy|) = \sigma_x \otimes \sigma'_y$. As a corollary, in the limit of infinitely many repetitions of the protocol, the information leakage upper bounds the number of bits of Alice and Bob's joint input that the referee can obtain per protocol repetition, without error probability going to one by Theorem 23 (strong converse to classical capacity of CQ channels).

For protocols in which σ_x, σ'_y are pure states for all $x \in \mathrm{X}$ and $y \in \mathrm{Y}$, the information leakage also has a source coding interpretation. Indeed, since the entropy of pure states is zero, then $\mathrm{H}(AB|XY) = 0$ for any input distribution P, so $\mathrm{QIL}(\Pi) = \sup_P \mathrm{H}(AB)$. Thus, by Schumacher's theorem, the information leakage gives the optimum rate at which register AB can be compressed for communication to the referee (maximized over all input distributions P). The states σ_x, σ'_y are pure for every simultaneous message passing model protocol that we consider in this thesis.

3.3 Interactive model

The interactive model is a communication setting consisting of two parties (Alice and Bob) who use quantum communication, local operations, and fixed pre-shared entanglement $\psi^{T_A^{\mathrm{in}} T_B^{\mathrm{in}}}$ to implement some channel $\Phi \in C(\mathcal{A}_{\mathrm{in}}\mathcal{B}_{\mathrm{in}}, \mathcal{A}_{\mathrm{out}}\mathcal{B}_{\mathrm{out}})$ using an interactive protocol Π, as shown in Figure 3.1. Although we avoided generalizing the simultaneous message passing model to quantum inputs ρ, we find it instructive to present this generalization in the interactive model.

In reference to Figure 3.1, we have chosen

$$\begin{aligned}
A'_0 &= A_{\mathrm{in}} T_A^{\mathrm{in}} \\
B'_0 &= B_{\mathrm{in}} T_B^{\mathrm{in}} \\
B'_M &= B'_{M+1} = B_{\mathrm{out}} B'_{\mathrm{left}} \\
A'_{M+1} &= A_{\mathrm{out}} A'_{\mathrm{left}} \\
B'_i &= B'_{i-1} \text{ for odd } i \text{ with } 1 \le i < M \\
A'_i &= A'_{i-1} \text{ for even } i \text{ with } 1 < i \le M.
\end{aligned}$$

In this way, for odd i, $U_i \in \mathcal{U}(\mathcal{A}'_{i-1}\mathcal{C}_{i-1}, \mathcal{A}'_i\mathcal{C}_i)$ and for even i, $U_i \in \mathcal{U}(\mathcal{B}'_{i-1}\mathcal{C}_{i-1}, \mathcal{B}'_i\mathcal{C}_i)$, where C_0 is the trivial register. These definitions implicitly assume M is even (as in Figure 3.1) but they can be easily adapted for M odd.

For any such protocol Π acting on input state $\rho \in \mathcal{D}(\mathcal{A}_{\text{in}}\mathcal{B}_{\text{in}})$, we define

$$\rho_i = U_i \dots U_1(\rho \otimes \psi)U_1^\dagger \dots U_i^\dagger \tag{3.41}$$

for all $i = 1, \dots, M+1$, where we have implicitly extended each isometry to the appropriate space by tensoring with the identity matrix. We let R be any register which purifies ρ_i for all $i = 1, \dots, M+1$.

Definition 29 ([63]). *For an interactive communication protocol* Π *defined as above, define the* ***information cost*** *of* Π *on input* $\rho^{A_{\text{in}}B_{\text{in}}}$ *as*

$$\mathrm{QIC}(\Pi, \rho) = \sum_{\substack{i=1\\ \text{odd}}}^{M} \mathrm{I}(C_i : R | B'_{i-1})_{\rho_i} + \sum_{\substack{i=2\\ \text{even}}}^{M} \mathrm{I}(C_i : R | A'_{i-1})_{\rho_i},$$

and the (input-independent) information cost of Π *as*

$$\mathrm{QIC}(\Pi) = \sup_{\rho \in \mathcal{D}(\mathcal{A}_{\text{in}}\mathcal{B}_{\text{in}})} \mathrm{QIC}(\Pi, \rho).$$

For certain protocols we will find a convenient form for the information cost which intuitively represents the flow of information in the protocol. For the general form above, it is shown in [63] using Theorem 16 (state redistribution) that the information cost is equal to the asymptotic rate of communication needed to execute the protocol $\Pi(\rho)$ using preshared entanglement when both parties have knowledge of ρ and M is finite.

3.3.1 Safe classical-quantum interactive protocols

In this section we restrict our attention to protocols Π which implement classical-quantum channels Φ. This set includes any protocol which performs a classical task (for example, appointment scheduling). For classical inputs $\rho \in \mathcal{D}(\mathcal{A}_{\text{in}}\mathcal{B}_{\text{in}})$ we instead use the notation $P \in \mathcal{D}(\mathcal{X}\mathcal{Y})$, replacing the symbol ρ with P, register A_{in} with X, and register B_{in} with Y. At times we abuse notation and use the symbol P to refer to both the classical state as well as the associated probability distribution on $\mathrm{X} \times \mathrm{Y}$, and trust that the intended use will be clear from the context.

We further restrict that the protocol Π is safe:

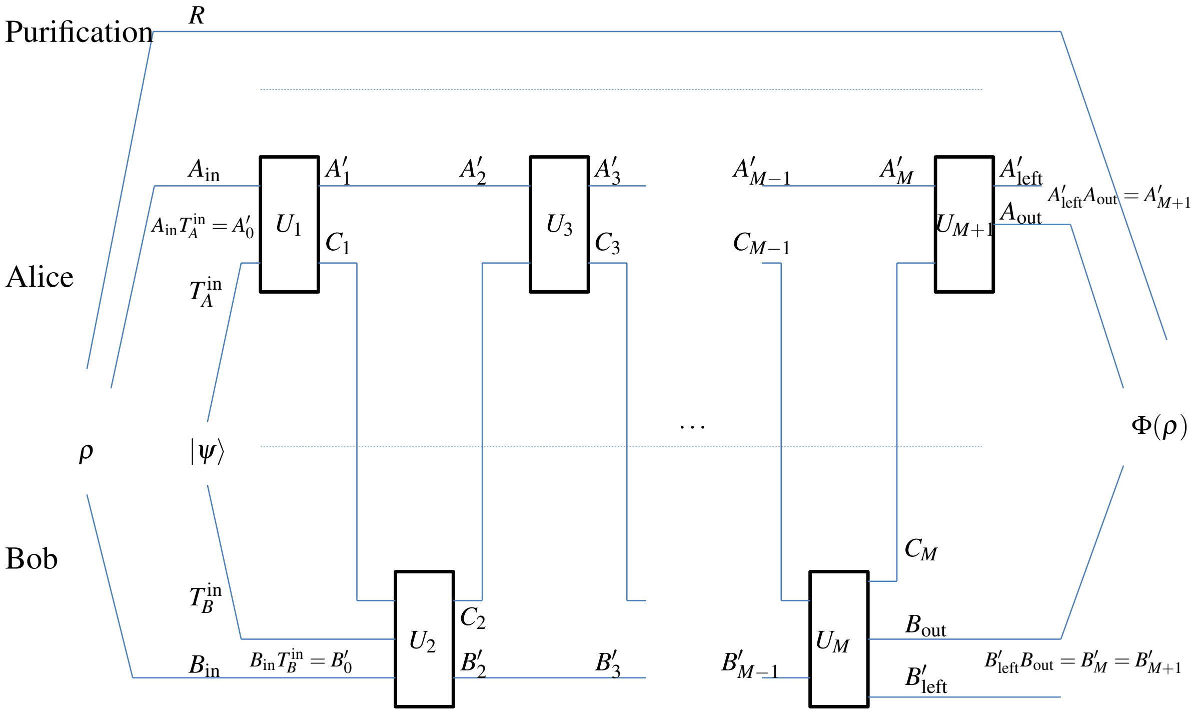

Figure 3.1: Interactive quantum communication protocol Π to implement the channel Φ. Borrowed from [42].

Definition 30. *We say a classical-quantum protocol* Π *is* ***safe*** *if Alice and Bob only use their respective input registers as control registers. Equivalently, the state of register* XY *is equal to* P^{XY} *throughout the protocol.*

The set of safe protocols contains any protocol which executes a classical task in which Alice and Bob receive classical inputs that they are not allowed to manipulate in any way. This set includes every interactive protocol that we consider in this thesis. The term "safe" was coined in [38].

In what follows, we will show that for safe protocols on classical inputs the information cost can be interpreted as a quantifier of the amount of information Alice and Bob learn about eachother's inputs as a result of the protocol.

We will frequently consider safe classical-quantum interactive protocols in which Alice and Bob perform measurements on the registers A_{out}, B_{out}, respectively, to attempt to determine some relationship between their classical inputs. In Chapter 5 (appointment scheduling) we consider interactive protocols for which $\mathrm{X} = \mathrm{Y} = \{0,1\}^n$, and on inputs $x, y \in \{0,1\}^n$ Alice and Bob wish to determine some index i such that $x_i = y_i = 1$, or determine with high probability that no such index exists.

For any safe protocol, we define registers A_i and B_i such that $A_i' = XA_i$ and $B_i' = YB_i$, respectively for each $i = 1, \ldots, M$ and A_{left} and B_{left} such that $A_{\text{left}}' = XA_{\text{left}}$ and $B_{\text{left}}' = YB_{\text{left}}$. Under this definition, the information cost of any safe classical-quantum interactive protocol Π takes the following form, which follows from the fact that the purification register R is now given by $X'Y'$, where X' and Y' are copies of the registers X and Y, respectively.

Lemma 31 (QIC: Safe classical-quantum protocols [42])**.** *The quantum information cost of any safe classical-quantum interactive protocol* Π *on input* $P \in \mathrm{Pr}(\mathrm{X} \times \mathrm{Y})$ *is given by*

$$\mathrm{QIC}(\Pi, P) = \sum_{i=1}^{M} \mathrm{QIC}_i(\Pi, P),$$

where

$$\mathrm{QIC}_i(\Pi, P) = \mathrm{I}(C_i : X | YB_i)_{\rho_i} + \mathrm{I}(C_i : Y | XA_i)_{\rho_i}.$$

Using our intuition for the mutual information as the amount of information that one register contains about another, the information cost of each message quantifies the amount of new information contained in the communication register about Alice's register given Bob's memory, plus the corresponding quantity with the roles of Alice and Bob reversed.

Another information-theoretic quantity of interest for safe classical-quantum protocols is the Holevo information cost:

Definition 32. *Define the* ***Holevo information cost*** *of a safe classical-quantum interactive protocol* Π *on input P as*

$$\mathrm{HIC}(\Pi,P) = \mathrm{I}(X : B_{\mathrm{out}}B_{\mathrm{left}}|Y)_{\rho_{M+1}} + \mathrm{I}(Y : A_{\mathrm{out}}A_{\mathrm{left}}|X)_{\rho_{M+1}} \tag{3.42}$$

and the (input-independent) Holevo information cost of Π *as*

$$\sup_{P\in\mathrm{Pr}(\mathrm{X}\times\mathrm{Y})} \mathrm{HIC}(\Pi,P). \tag{3.43}$$

The HIC quantifies the amount of information left over from the protocol: the amount of new information Bob's output contains about Alice's input as a result of Π, plus the corresponding quantity with the roles of Alice and Bob reversed. For safe classical-quantum protocols, the information cost upper bounds the HIC:

Lemma 33 ([42])**.** *For any safe classical-quantum protocol* Π *on input P,*

$$\mathrm{QIC}(\Pi,P) \geq \mathrm{HIC}(\Pi,P), \tag{3.44}$$

and thus

$$\mathrm{QIC}(\Pi) \geq \mathrm{HIC}(\Pi). \tag{3.45}$$

This lemma thus gives another interpretation of the information cost as an upper bound on the amount of information each party learns about the other party's input as a result of the protocol.

Note that for classical protocols the information cost is equal to the HIC [42].

Chapter 4

Quantum fingerprinting

Quantum fingerprinting is a task in the simultaneous message passing model (reviewed in Chapter 3) in which Alice and Bob receive inputs $x \in \mathrm{X}$ and $y \in \mathrm{Y}$ for $\mathrm{X} = \mathrm{Y} = \{0,1\}^n$, chosen according to some prior distribution $P \in \mathrm{Pr}(\mathrm{X} \times \mathrm{Y})$, and they wish to evaluate the quality function on x and y. Using classical states, the information leakage is lower bounded by $\Omega(\sqrt{n})$ [5]. In contrast, there exist protocols using quantum states with information leakage $\mathcal{O}(\log n)$ [14, 4].

The original quantum fingerprinting protocol uses $\mathcal{O}(\log n)$-qubit highly entangled signals and a controlled-swap measurement [14]. A more recent and experimentally realizable “optical” protocol uses a tensor product of $\mathcal{O}(n)$ single-qubit signals and a beamsplitter comparison measurement on each signal [4]. In this work, we find a family of protocols which interpolate between these two, thus demonstrating a trade-off between the number of signals sent and the dimension of each signal. We show that this family of protocols has information leakage $\mathcal{O}(\log n)$.

There has been interest in experimental realizations of the optical protocol of [4] using coherent states [68, 32], but for large n the number of signals required is a significant barrier to experimental demonstration of a quantum advantage in information leakage due to limited coherence times. We introduce several families of optical coherent state protocols which reduce the number of signals below that of the existing optical protocol. We improve on the existing optical protocol by utilizing the imaginary component of the phase space representation of coherent states, which reduces the number of signals by a factor 1/2 while also reducing the information leakage. We introduce several natural generalizations of this protocol which further reduce the number of signals, but find numerical evidence that the information leakage of these protocols is higher in both the ideal and experimental settings, even under the optimal measurement performed by the referee.

Using a similar technique, we also reduce the number of signals and information leakage of a recently-proposed optical protocol for evaluating the Euclidean distance between two real unit vectors [41], and find a similar protocol which evaluates the Euclidean distance between complex unit vectors.

Along the way, we find a simple beamsplitter measurement which can be used to perform optimal unambiguous state comparison (USC) between any two coherent states of equal amplitude and opposite phase when the states are given with equal a priori probabilities. Optimal USC of two unknown quantum states given with equal a priori probabilities was first solved in [8] and generalized to arbitrary a priori probabilities in [40]. A method to realize the optimal USC of two single-photon states prepared with arbitrary a priori probabilities is proposed in [51], but to our knowledge optimal USC has not yet been experimentally realized. Our scheme has the advantage of being more experimentally realizable, with the drawback of being sub-optimal for not-equal a priori probabilities. A related comparison task on coherent states using a beamsplitter has been proposed in [2].

This chapter is organized as follows. In Section 4.1 we interpolate between the original and existing optical equality protocols. In Section 4.2 we introduce our improvements to the existing optical equality and Euclidean distance protocols, and review our numerical evidence that several natural generalizations of our equality protocol have higher information leakage. In Section 4.3 we derive a bound on the information leakage of the protocols considered. In Section 4.4 we find a simple beamsplitter measurement which performs optimal unambiguous state comparison of two coherent states.

4.1 Interpolation

In this section we interpolate between the original equality protocol, which uses a small number of $\mathcal{O}(\log n)$-qubit signals, and the existing optical equality protocol, which uses $\mathcal{O}(n)$ single-qubit signals; thus demonstrating a trade-off between the number of signals sent and the dimension of each signal. In Sections 4.1.1 and 4.1.2 we introduce slight adaptations to the existing protocols which are more natural candidates for the interpolation, and in Section 4.1.3 we interpolate between these adaptations.

Before proceeding, we outline a general protocol framework which holds for all equality protocols that we consider in this work. First, Alice and Bob receive inputs $x, y \in \{0,1\}^n$ respectively, conditioned on which they send pure states $|\psi_x\rangle, |\psi_y\rangle$ to the referee which are sufficiently distinguishable when $x \neq y$. The referee then performs a comparison measurement on $|\psi_x\rangle|\psi_y\rangle$ and outputs either Equal or NotEqual. We define the error probability of the protocol as the worst case error probability over all $x, y \in \{0,1\}^n$. In the ideal setting, the error probability of every protocol is one-sided: if the inputs are equal the referee will always output Equal. In every protocol, the states $|\psi_x\rangle$ are product vectors. We refer to individual tensor factors of $|\psi_x\rangle$ as *signals*. For many protocols that we consider, the state $|\psi_x\rangle$ will be a tensor product of multiple copies of identical states, each composed of signals.

To make $|\psi_x\rangle, |\psi_y\rangle$ sufficiently distinguishable to the referee when $x \neq y$, inputs x, y are mapped to codewords $\mathrm{E}(x), \mathrm{E}(y) \in \{0,1\}^m$ of an error-correcting code characterized by some minimum distance. The codewords are then encoded into states whose overlap is a decreasing function of the distance between codewords, which ensures that they are sufficiently distinguishable to the referee. The code E is chosen to have constant minimum distance and constant rate, which we will see ensures the $\mathcal{O}(\log n)$ information leakage of all protocols.

4.1.1 Adaptation of existing optical equality protocol

Here we review the existing optical equality protocol (in the ideal setting) and propose a slight adaptation which is a more natural candidate for the interpolation. In the existing protocol, each signal consists of one of two qubits, which the referee measures with a beamsplitter setup. The desired error probability is attained by adjusting the inner product between the two possible qubit signals to make them sufficiently distinguishable. In our adapted protocol, we instead fix the inner product between the two possible qubit signals, and attain the desired error probability by sending multiple identical copies of each signal.

We show that when the desired error probability is attained with equality, the states used in the existing optical protocol and our adapted protocol are equal up to a change of basis, i.e. there exists an isometry mapping the states of the existing optical protocol to the states of our adapted protocol. We will use Property 3 of [17] that two sets of pure states $\{|v_a\rangle \in \mathcal{H}_v\}_{a \in \mathrm{Z}}$, $\{|w_a\rangle \in \mathcal{H}_w\}_{a \in \mathrm{Z}}$ are equal up to change of basis if and only if there exist real numbers $\theta_a, a \in \mathrm{Z}$ such that $\langle v_a, v_b\rangle = e^{i(\theta_a - \theta_b)} \langle w_a, w_b\rangle$ for all $a, b \in \mathrm{Z}$, and an easy corollary that equality up to change of basis is transitive. By the invariance of entropy under isometries, the information leakage (defined in Section 4.3) is equal for protocols using states that are equal up to change of basis.

In the original formulation of the existing optical protocol, the qubit signals are written in a basis as coherent states. We begin by introducing the existing optical protocol in this basis before converting to the qubit picture and introducing our adapted protocol. In the existing optical protocol, the j-th signal is one of two coherent states depending on the j-th codeletter of the codeword $\mathrm{E}(x) \in \{0,1\}^m$:

$$|\alpha_x\rangle_{\mathrm{EQ},1} = \bigotimes_{j=1}^{m} \left| (-1)^{\mathrm{E}(x)_j} \frac{\alpha}{\sqrt{m}} \right\rangle_j . \tag{4.1}$$

For each index j, the referee interferes the $j-$th pair of signals received from Alice and Bob in a beamsplitter, and measures the dark port with a single photon threshold detector, obtaining one of two outcomes: "dark port detection" or "no dark port detection". The referee outputs NotEqual if at least one outcome "dark port detection" occurs. On input $|\beta_a\rangle\,|\beta_b\rangle$, outcome "no dark port detection" occurs with probability

$$|\langle\beta_a,\beta_b\rangle| = e^{-\frac{1}{2}|\beta_a-\beta_b|^2}. \tag{4.2}$$

It follows that the error probability given different inputs $x \neq y$ is equal to $|\,\langle\alpha_x,\alpha_y\rangle\,|$, and the error probability given equal inputs is zero. The worst case error probability occurs when the codewords differ by minimum distance δm bits, and is equal to $\exp[-2\,|\alpha|^2\,\delta]$, which is brought to within any $\varepsilon > 0$ through appropriate choice of α.

In the existing optical protocol, the set of two possible coherent states for each signal span a two-dimensional space, and thus can be written in a basis as two qubits $|q_0^\varepsilon\rangle, |q_1^\varepsilon\rangle$ with inner product determined by ε. In our adaptated protocol, instead of using qubit signals with inner product determined by ε, the parties instead fix qubits $|q_0'\rangle, |q_1'\rangle$ independent of ε, and send the minimum number r of identical copies of each qubit signal needed to attain ε (we refer to each individual copy as a signal). The referee uses the same beamsplitter measurement (converted to the qubit basis) on each signal, and outputs NotEqual if any outcome "dark port detection" occurs. See Section 4.4 for an explicit description of the beamsplitter measurement in the qubit basis.

Now we show that if ε is attained with equality in both the existing protocol and our adapted protocol, then the states used in each protocol are equal up to a change of basis. Specifically, we show that the set of signals $\{|\frac{\alpha}{\sqrt{m}}\rangle, |\frac{-\alpha}{\sqrt{m}}\rangle\}$ used in the existing optical protocol are equal up to change of basis to the set of states $\{|q_0'\rangle^{\otimes r}, |q_1'\rangle^{\otimes r}\}$ containing the r copies of each signal used in our adapted protocol. Indeed, the error probability is given by $\varepsilon = |\,\langle q_0', q_1'\rangle\,|^{\delta mr} = |\,\langle\frac{-\alpha}{\sqrt{m}}, \frac{\alpha}{\sqrt{m}}\rangle\,|^{\delta m}$, so $|\,\langle q_0', q_1'\rangle\,|^r = |\,\langle\frac{-\alpha}{\sqrt{m}}, \frac{\alpha}{\sqrt{m}}\rangle\,|$, which completes the proof by Property 3 of [17] (reviewed above).

4.1.2 Adaptation of original equality protocol

Here we describe the original quantum fingerprinting protocol, and propose a slight adaptation that is a more natural candidate for the interpolation.

Similar to the adapted optical protocol, Alice and Bob send the minimum number r of identical signals

$$|\psi_x^{(m)}\rangle = \frac{1}{\sqrt{m}}\sum_{i=1}^{m}|i\rangle\,|\mathrm{E}(x)_i\rangle \tag{4.3}$$

needed to attain worst case error probability ε. We rewrite a single copy of the combined signal $|\psi_{xy}^{(m)}\rangle := |\psi_x^{(m)}\rangle|\psi_y^{(m)}\rangle$ as

$$|\psi_{xy}^{(m)}\rangle = \frac{1}{m}\sum_{i=1}^{m}|i\rangle\,|\mathrm{E}(x)_i\rangle\,|i\rangle\,|\mathrm{E}(y)_i\rangle + \frac{1}{m}\sum_{\substack{l,h=1\\ l\neq h}}^{m}|l\rangle\,|\mathrm{E}(x)_l\rangle\,|h\rangle\,|\mathrm{E}(y)_h\rangle\,. \tag{4.4}$$

In the original protocol, the referee performs the controlled-swap measurement (a projective measurement onto the symmetric and anti-symmetric subspaces) on each pair of signals $|\psi_{xy}^{(m)}\rangle$, and outputs NotEqual if any outcome "anti-symmetric" occurs. The worst case error probability occurs when the codewords differ by minimum distance δm bits, and is given by $(1-\delta(1-\frac{\delta}{2}))^r$ [14].

In our adapted protocol, the referee performs the beamsplitter measurement described in Section 4.1.1 on each pair of qubits in the first term of the decomposition (4.4), and performs the controlled-swap measurement on the second term. The referee outputs NotEqual if any outcome "dark port detection" or "anti-symmetric" occur. In Appendix A.0.1 we show that this protocol has worst case error probability

$$\mathrm{Pr}_m^{\mathrm{I}}(\mathrm{Err}) = \left[1-\delta\left(1-\frac{\delta}{2}+\frac{1}{2m}\right)\right]^r, \tag{4.5}$$

a minor improvement over the original protocol.

In Section 4.1.1 we show that the beamsplitter measurement in the qubit basis can be decomposed as a direct product of a controlled-swap measurement with an unambiguous state comparison measurement. Thus, the full adapted measurement can also be decomposed into these two measurements.

4.1.3 Interpolation between adapted protocols

Here we interpolate between the adapted protocols described in Sections 4.1.1 and 4.1.2, demonstrating a trade-off between the number of signals sent and the dimension of each signal.

For the interpolation protocol with block-size k, blocks of k bits of $\mathrm{E}(x)$ are encoded into each signal:

$$|\psi_x^{(k)}\rangle = \bigotimes_{j=1}^{\lceil m/k \rceil} \left[\frac{1}{\sqrt{k}} \sum_{i \in \mathrm{I}[j,k]} |i\rangle \left| q^{(k)}_{\mathrm{E}(x)_i} \right\rangle \right] \in (\mathbb{C}^k \otimes \mathbb{C}^2)^{\otimes \lceil \frac{m}{k} \rceil} \tag{4.6}$$

where $|q_0^{(k)}\rangle, |q_1^{(k)}\rangle$ are qubits. The set $\mathrm{I}[j,k]$ indexes the j-th block of k bits, i.e. it is the set of integers in the range $[(j-1)k+1, jk]$. If k does not divide m the remaining qubits in the final signal are set to $|q_0^{(k)}\rangle$. As before, Alice and Bob send the minimum number r of identical copies of the state $|\psi_x^{(k)}\rangle$ needed to attain the desired error probability ε.

The referee's measurement proceeds similarly to that of our adapted original protocol. We rewrite the j-th pair of signals contained in the combined state $|\psi_{xy}^{(k)}\rangle := |\psi_x^{(k)}\rangle|\psi_y^{(k)}\rangle$ as

$$\frac{1}{k} \sum_{i \in \mathrm{I}[j,k]} \left[|i\rangle \left| q^{(k)}_{\mathrm{E}(x)_i} \right\rangle |i\rangle \left| q^{(k)}_{\mathrm{E}(y)_i} \right\rangle \right] + \frac{1}{k} \sum_{\substack{l,h \in \mathrm{I}[j,k] \\ l \neq h}} \left[|l\rangle \left| q^{(k)}_{\mathrm{E}(x)_l} \right\rangle |h\rangle \left| q^{(k)}_{\mathrm{E}(y)_h} \right\rangle \right]. \tag{4.7}$$

For each index j, the referee performs the beamsplitter measurement on each pair of qubits in the first term of this decomposition, and performs the controlled-swap measurement on the second term. If outcome "dark port detection" or "anti-symmetric" occur for any index j the referee decides NotEqual. The worst case error probability is derived in Appendix A.0.1. As noted in the previous section, this measurement can equivalently be described as a direct product of a controlled-swap measurement with an unambiguous state discrimination measurement.

As evidenced by (4.6), each signal encodes k qubits, each chosen from the set $\{|q_0^{(k)}\rangle, |q_1^{(k)}\rangle\}$. We choose these qubits to satisfy $|\langle q_0^{(k)}, q_1^{(k)}\rangle| = 1 - k/m$, which ensures that the interpolation converges to the adapted orginal and optical protocols for $k = m$ and $k = 1$ respectively. Furthermore, for fixed error ε, this choice gives rise to fixed repetition number r, which allows us in Section 4.3.1 to bound the information leakage as $\mathcal{O}(\log n)$ for every block-size k.

In summary, this family of protocols interpolates between the adapted original and optical protocols while maintaining information leakage $\mathcal{O}(\log n)$. For block-size $k = 1, \dots, m$, our interpolation uses $r\lceil m/k \rceil$ signals, each of dimension $2k$, thus demonstrating a trade-off between the number of signals sent and the dimension of each signal.

4.2 Optical protocols

In this section we consider several families of optical coherent state simultaneous message passing model protocols which reduce the number of signals from that of the existing protocols. In Section 4.2.1 we introduce optical protocols for equality and Euclidean distance which reduce the number of signals by a factor 1/2 and reduce the information leakage from the existing optical protocols of [4] and [41] for equality and Euclidean distance, respectively. In Section 4.2.2 we introduce two families of optical equality protocols which further reduce the number of signals, but find numerical evidence that these families increase the information leakage in both the ideal and experimental settings, even under the optimal measurement.

4.2.1 Improved optical protocols for equality and euclidean distance

In our improved optical equality protocol, two bits of $\mathrm{E}(x)$ are encoded into each signal by utilizing the imaginary component of the phase space representation of coherent states, which reduces the number of signals by a factor 1/2. Codeletters $01/10$ are encoded into phases $\pm i$, and codeletters $00/11$ are encoded into phases ± 1, as shown in Figure 4.1. Explicitly, the parties send the states

$$|\alpha_x\rangle_{\mathrm{EQ},2} = \bigotimes_{\substack{j=1 \\ \text{odd}}}^{m} \left| (-1)^{\mathrm{E}(x)_j} \times (i)^{\mathrm{E}(x)_j \oplus \mathrm{E}(x)_{j+1}} \frac{\alpha}{\sqrt{m/2}} \right\rangle_j . \tag{4.8}$$

The referee uses the same beamsplitter measurement as in the existing optical protocol: she interferes pairs of signals in a beamsplitter, measures the dark port with a single photon threshold detector, and decides NotEqual if at least one outcome “dark port detection” occurs.

The above states have the same total mean photon number $|\alpha|^2$ and give rise to the same error probability as the existing optical protocol. The second statement follows from the expression (4.2) for the probability of “no dark port detection” in terms of the squared distance between the complex amplitudes of the incoming coherent state signals, along with the fact that for pairs of codeletters $(\mathrm{E}(x)_j, \mathrm{E}(x)_{j+1})$ and $(\mathrm{E}(y)_j, \mathrm{E}(y)_{j+1})$ which differ by one bit this distance is given by $w^2 = |\alpha|^2/m$ as in the existing optical protocol (see Figure 4.1), and for such pairs which differ by two bits this distance is $2w^2$, which gives rise to probability of “no dark port detection” equal to the probability of “no dark port detection” occuring for both signals j and $j+1$ in the existing optical protocol. More details are given in Appendix A.0.2. By the information leakage bound $\sim \mathcal{O}(|\alpha|^2 \log m_k)$ (where m_k is the number of signals) derived in Appendix B, our improved protocol has lower information leakage than the existing protocol. It can be shown that this statement also holds under the stronger bound derived in Section 4.3.2 for $|\alpha|^2 \ll m_k$ using standard approximation techniques. Below we will refer to this protocol, including its use of the beamsplitter measurement, as the two-bit protocol.

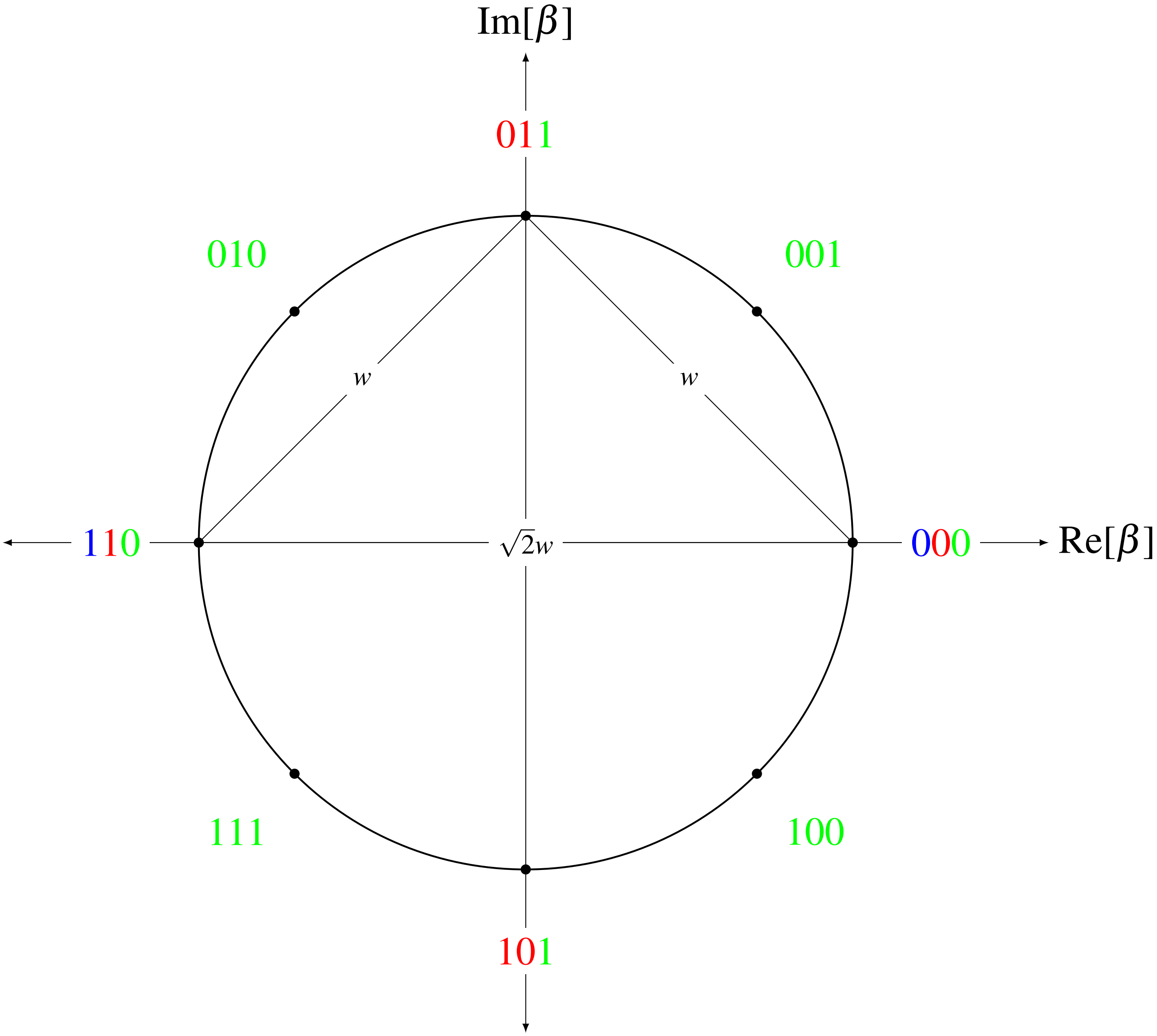

Figure 4.1: Gray coding of $k-$bit blocks into a ring of coherent states in phase space for $k = 1$ (blue), $k = 2$ (blue and red combined), and $k = 3$ (blue, red and green combined).

We improve the existing optical Euclidean distance protocol of [41] in similar fashion. In the existing protocol, Alice and Bob receive real unit vectors $u, v \in \mathbb{R}^s$ respectively and prepare the states

$$|\alpha_u\rangle_{\mathrm{ED},1} := \bigotimes_{j=1}^{s} |u_j \alpha\rangle_j. \tag{4.9}$$

The referee interferes each pair of signals received from Alice and Bob in a beamsplitter and measures both output ports with single photon threshold detectors. The quantity $\|u-v\|^2$ is a function of $|\alpha|^2$ and the expected difference between the number of detections observed in the two output ports, so using Chernoff bounds the referee can estimate $\|u-v\|^2$ to within an additive constant ε with probability at least $1-\delta$ by repeating this process $\mathcal{O}(\log(1/\delta)/\varepsilon^2)$ times.

As in our improved equality protocol, our improved Euclidean distance protocol utilizes the imaginary component of the phase space representation of coherent states to reduce the number of signals by a factor 1/2. Alice and Bob prepare the states

$$|\alpha_u\rangle_{\mathrm{ED},2} := \bigotimes_{\substack{j=1 \\ \text{odd}}}^{s} |(u_j + iu_{j+1})\alpha\rangle_j, \tag{4.10}$$

and the referee uses the same measurement as before. These states have the same total mean photon number $|\alpha|^2$ as before, and using nearly identical analysis to that of [41] it can be shown that this protocol attains the same error probability as the existing protocol. Thus, as before, our protocol has lower information leakage than the existing protocol under the information leakage bound of Appendix B. Alternatively, one can adapt the existing protocol to evaluate the Euclidean distance between two complex unit vectors $u, v \in \mathbb{C}^s$ using the same measurement and the states (4.9).

4.2.2 Families of optical equality protocols

In this section we introduce two families of optical coherent state equality protocols which further reduce the number of signals. In Section 4.2.2 we find numerical evidence that these protocols have higher information leakage than the two-bit protocol in the ideal setting, even under the optimal measurement. In Section 4.2.2 we find numerical evidence of this behaviour under realistic experimental imperfections.

Ideal setting

We first describe our families of optical equality protocols in the ideal setting. In the ring (lattice) protocol with block size k, blocks of k bits of $\mathrm{E}(x)$ are encoded into one of 2^k coherent state signals arranged in a ring (lattice) in phase space using a Gray code, as shown in Figure 4.1. The size of the ring (lattice) is determined by the referee's measurement and the desired worst case error probability ε, and is held constant for all signals.

The ring (lattice) Gray code is a mapping from k-bit strings to a ring (lattice) of coherent state signals which satisfies the property that all nearest neighbour signals differ in exactly one bit [30, 20, 16]. Note that in the ring encoding, each coherent state signal has two nearest neighbours, while in the lattice encoding a given coherent state signal can have as many as four nearest neighbours. We have chosen this code so that a greater number of bit differences between two k-bit blocks of $\mathrm{E}(x)$ and $\mathrm{E}(y)$ corresponds to greater distinguishability between the two coherent state signals. In Appendix A.0.2 we prove that this is the optimal encoding of k-bit blocks of binary codewords for all $k = 2,3,4$, and that it outperforms an analogous encoding of $q-$ary codewords.

Now we introduce the two measurements performed by the referee that we consider. The beamsplitter measurement proceeds identically to that of Section 4.1.1: the referee interferes pairs of signals in a beamsplitter and measures the dark port with a single photon threshold detector. She decides NotEqual if at least one outcome "dark port detection" occurs. Recall that the error probability under the beamsplitter measurement is one-sided, i.e. there is zero error for equal inputs. We also consider the optimal one-sided error measurement, which is described in Appendix A.0.3, and is shown to have error probability lower bounded by the square of the error probability of the beamsplitter measurement.

In Figure 4.2 we plot the information leakage of the ring encoding under the bound of Section 4.3.2, compared to the classical information leakage lower bound of [5]. We have optimized over δ and assumed the code saturates the Gilbert-Varshamov bound [29, 66, 65]

$$\frac{n}{m} = 1 - \mathrm{h}(\delta). \tag{4.11}$$

For $k = 1,2,3$ we plot the information leakage under the beamsplitter measurement, and for $k = 4,5,6$ we lower bound the information leakage under the optimal measurement using the quadratic bound on the error probability derived in Appendix A.0.3. We see that the one-bit and two-bit protocols have the lowest information leakage (they are numerically indistinguishable). We have observed the same result for ε in the range $[10^{-10}, 10^{-2}]$ in both the ring and lattice protocols.

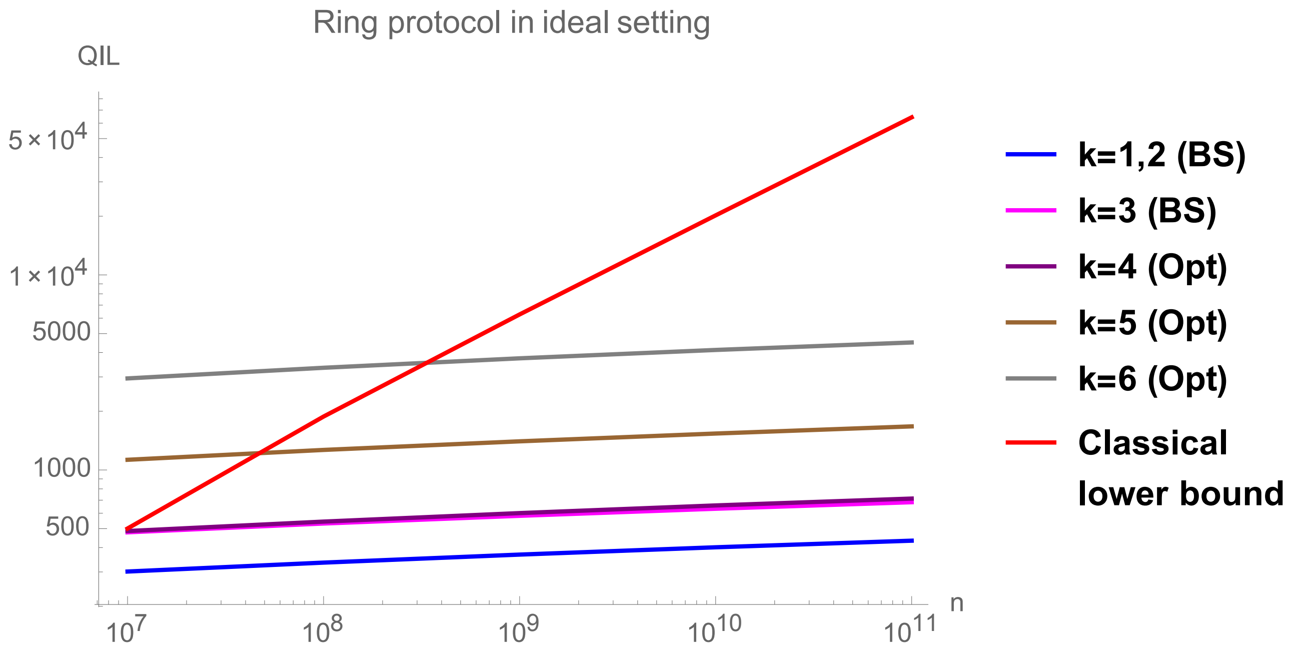

Figure 4.2: Information leakage of ring protocols in the ideal setting for $\varepsilon = 0.01$.

Before continuing on to consider experimental imperfections, we argue that for fixed block-size k, both the ring and lattice protocols have asymptotic information leakage $\mathcal{O}(\log n)$ in the ideal setting. In Appendix B we show that any simultaneous message passing model protocol which uses m_k coherent state signals and maximum total mean photon number μ_{max} has information leakage $\sim \mathcal{O}(\mu_{\text{max}} \log m_k)$. Note that $m_k = m/k = \mathcal{O}(n)$ by the fact that E is a constant rate code. Furthermore, in Appendix A.0.2 we show that any fixed error probability is attained with μ_{max} constant in n. Together, these results imply that the information leakage is $\mathcal{O}(\log n)$.

Experimental Setting

For larger block-sizes in the experimental setting, the states have fewer signals, so dark counts have less effect on the error probability. In this section we find numerical evidence that this effect does not change the observed ideal behaviour, and speculate the same result holds for the lattice protocol.

The experimental ring protocol uses the same states and beamsplitter measurement as in the ideal setting. However, to account for transmittivity η the initial total mean photon number is rescaled to μ/η, and to account for dark count probability p_{dark} the referee uses a different criteria to decide Equal or NotEqual. The referee decides NotEqual if the number of outcomes "dark port detection" exceeds a threshold value $\mathcal{T}_k$ which is chosen to minimize the worst case error probability over all inputs $x, y \in \{0,1\}^n$ (which is no longer one-sided when $p_{\text{dark}} > 0$).

Now we determine the optimal threshold value $\mathcal{T}_k$. Define a random variable $D_{E,k}$ for the number of outcomes "dark port detection" given equal inputs. Define $D_{D,k}$ identically, but for different inputs which have the lowest expected number of outcomes "dark port detection". For a given threshold value $\mathcal{T}_k'$, the worst case error probability is then given by $\max\{\Pr(D_{E,k} \geq \mathcal{T}_k'), \Pr(D_{D,k} < \mathcal{T}_k')\}$. As the first (second) element is monotonically decreasing (increasing) with $\mathcal{T}_k'$, it follows that the optimal threshold value $\mathcal{T}_k$ satsifies

$$\Pr(D_{E,k} \geq \mathcal{T}_k) = \Pr(D_{D,k} < \mathcal{T}_k), \tag{4.12}$$

which is also the worst case error probability of the protocol under this choice. Actually, as both probabilities are step functions, it is possible that (4.12) can only be attained with approximate equality.

For $m \gg k$ it can be shown that the number of clicks are well-approximated by binomial distributions $D_{E,k} \sim \text{Bin}(m/k, p_{E,k})$ and $D_{D,k} \sim \text{Bin}(m/k, p_{D,k})$, where

$$\begin{aligned} p_{D,k} &= (1-(\delta k - \lfloor \delta k \rfloor)) \left(1 - e^{-\frac{\mu}{m/k}\left[1-\cos\left(\frac{2\pi}{2^k}\lfloor \delta k \rfloor\right)\right]}\right) \\ &\quad + (\delta k - \lfloor \delta k \rfloor) \left(1 - e^{-\frac{\mu}{m/k}\left[1-\cos\left(\frac{2\pi}{2^k}(\lfloor \delta k \rfloor + 1)\right)\right]}\right) + p_{\text{dark}} \\ p_{E,k} &= p_{\text{dark}} \end{aligned} \tag{4.13}$$

for all $k = 1, \ldots, 6$ under the Gray code (see Appendix A.0.2). We have used this approximation to calculate the optimal value of $\mathcal{T}_k$ and the corresponding worst case error probability (4.12).

In Figure 4.3 we plot the information leakage of the ring protocols under realistic experimental imperfections, compared to the classical information leakage lower bound of [5]. As before, we have optimized over δ and assumed the code satisfies the Gilbert-Varshamov bound (4.11). We have chosen worst case error probability $\varepsilon = 0.01$, transmittivity $\eta = 0.3$, and dark count probability $p_{\text{dark}} = 7.3 \times 10^{-11}$. We include a plot of of the existing optical protocol with interferometric visibility 99% for reference. We observe the same hierarchy as the ideal setting, but as dark counts have less effect for protocols sending fewer signals, the $k = 2$ (two-bit) protocol now has visibly lower information leakage than the $k = 1$ protocol. We have also observed the same hierarchy for p_{dark} in the range $[0, 10^{-6}]$ and ε in the range $[10^{-10}, 10^{-2}]$. We speculate the same behaviour holds for the lattice protocol under experimental imperfections.

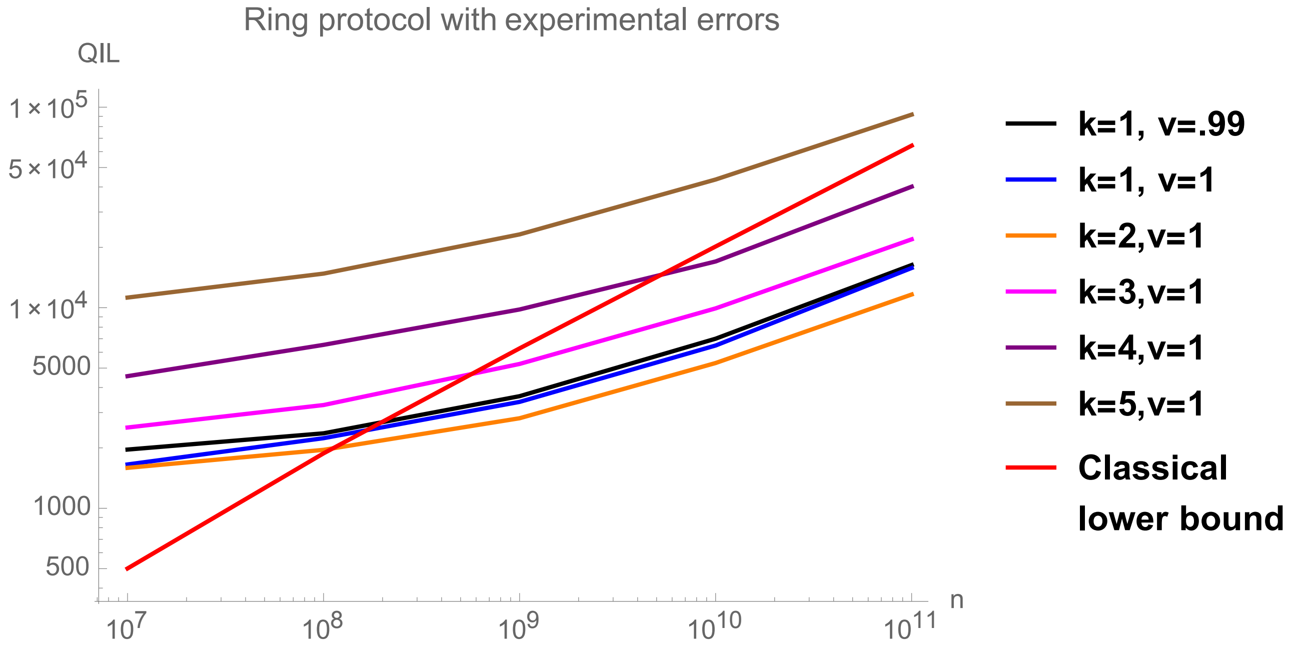

Figure 4.3: Information leakage of ring protocols to attain error probability $\varepsilon = 0.01$ under transmittivity $\eta = 0.3$ and dark count probability $p_{\text{dark}} = 7.3 \times 10^{-11}$. Existing optical protocol with interferometric visibility 99% included for reference.

4.3 Information leakage

In this section we bound the information leakage of any simultaneous message passing model protocol Π in which conditioned on $x \in \mathrm{X}$ and $y \in \mathrm{Y}$ Alice and Bob send pure states $|\psi_x\rangle\langle\psi_x|$, $|\psi'_y\rangle\langle\psi'_y|$ which can be written in a basis as states $|\phi_x\rangle\langle\phi_x|$, $|\phi'_y\rangle\langle\phi'_y|$ with diagonal entries equal to some fixed probability vector Λ for all $x \in \mathrm{X}$, $y \in \mathrm{Y}$ as $2\mathrm{H}(\Lambda)$. We find such a basis in both the interpolation and optical ring protocol families, and bound the information leakage of every protocol in either family as $\mathcal{O}(\log n)$. In Appendix B we bound the information leakage of the optical lattice protocol family as $\mathcal{O}(\log n)$. The bound of Appendix B also holds for the optical ring protocols, but we have found numerically that the bound of Section 4.3.2 is $15-40\%$ lower in the considered parameter regime. The intuition for this improvement is that the bound of Section 4.3.2 takes into account the particular structure of the coherent states used in the protocol, whereas the bound of Appendix B only uses the fact that every state has the same total mean photon number.

Recall from Section 3.2 that the information leakage of any pure state protocol is given by

$$\mathrm{QIL}(\Pi) = \max_{P \in \Pr(X\times Y)} \mathrm{I}(XY:AB)_{\rho_P} = \max_{P \in \Pr(X\times Y)} \mathrm{H}(AB), \tag{4.14}$$

where

$$\rho_P = \sum_{x\in\mathrm{X}, y\in\mathrm{Y}} P(x,y)|xy\rangle\langle xy| \otimes |\psi_x\rangle\langle\psi_x| \otimes |\psi'_y\rangle\langle\psi'_y|.$$

Define a state σ_P^{AB} as the state ρ_P^{AB} with each pure state $|\psi_x\rangle\langle\psi_x|$ and $|\psi'_y\rangle\langle\psi'_y|$ replaced by $|\phi_x\rangle\langle\phi_x|$ and $|\phi'_y\rangle\langle\phi'_y|$ respectively. By the definition of change of basis (see Section 4.1.1), there exists an isometry between ρ_P^{AB} and σ_P^{AB}, so by the invariance of quantum entropy under isometries, $\mathrm{H}(\rho_P^{AB}) = \mathrm{H}(\sigma_P^{AB})$. As the diagonal entries of each pure state in σ_P^{AB} are equal, then $\mathrm{Diag}(\sigma_P^{AB}) = \Lambda^{\otimes 2}$ for all $P \in \Pr(\mathrm{X}\times\mathrm{Y})$. Thus, by the Schur-Horn theorem [58, 37], $\sigma_P^{AB} \succ \Lambda^{\otimes 2}$ for all $P \in \Pr(\mathrm{X}\times\mathrm{Y})$, which implies $\mathrm{H}(\sigma_P^{AB}) \leq \mathrm{H}(\Lambda^{\otimes 2})$ for all $P \in \Pr(\mathrm{X}\times\mathrm{Y})$ (see, e.g. [47]). By additivity of entropy under tensor product,

$$\mathrm{QIL}(\Pi) \leq 2\mathrm{H}(\Lambda). \tag{4.15}$$

In the what follows, we apply this bound to the interpolation and optical ring protocol families.

4.3.1 Information leakage: Interpolation

Here we bound the information leakage of the interpolation under the choice of qubit overlap $|\langle q_0^{(k)}, q_1^{(k)}\rangle| = 1 - k/m$ as $\mathcal{O}(\log n)$:

Proposition 34. *For any fixed error probability $\varepsilon > 0$, there exists a constant $C \geq 0$ such that for all $n \geq 1$ the following holds: for all $k = 1, \ldots, m$ (where m is the length of the codewords $\mathrm{E}(x) \in \{0,1\}^m$), the information leakage of the interpolation protocol with block-size k and qubit overlap $|\langle q_0^{(k)}, q_1^{(k)} \rangle| = 1 - k/m$ is no greater than $C \log(n)$.*

As a corollary, for block-size k any function of n such that $k(n) \in [m]$ for each n, the family of protocols which uses the interpolation family with block-size $k(n)$ for each n will have information leakage $\mathcal{O}(\log n)$. For example, k could be held to a fixed constant as in the existing optical protocol, or the ratio k/m could be held fixed as in the original protocol.

Proof. We prove that the information leakage is upper bounded by $C' r \log n$ for some $C' > 0$. In Appendix A.0.1 we show that for fixed error probability ε, the repetition number r is fixed, completing the proof.

We first write the states $|\psi_x^{(k)}\rangle^{\otimes r}$ in a basis with the desired form. First, by Property 3 of [17] (reviewed in Section 4.1.1), the states $|\psi_x^{(k)}\rangle^{\otimes r}$ can be written in a basis as

$$|\phi_x^{(k,r)}\rangle = \bigotimes_{j=1}^{\lceil m/k \rceil} \left[\sqrt{1-p_k}\,|00\rangle + \sqrt{\frac{p_k}{2k}} \sum_{i \in \mathrm{I}[j,k]} |i\rangle \left(|0\rangle + (-1)^{\mathrm{E}(x)_i} |1\rangle \right) \right]^{\otimes r},$$

because the states $|\phi_x^{(k,r)}\rangle$ preserve the overlap structure. The diagonal entries of $|\phi_x^{(k,r)}\rangle\langle\phi_x^{(k,r)}|$ are equal for all $x \in \{0,1\}^n$, and are given by the probability vector

$$\Lambda_{\mathrm{I},k}^{\otimes \lceil m/k \rceil r} = \left(1 - p_k, \frac{p_k}{2k}, \ldots, \frac{p_k}{2k}\right)^{\otimes \lceil m/k \rceil r}. \tag{4.16}$$

Thus, by (4.15) the information leakage of the interpolation protocol with block-size k under the choice $p_k = k/m$ is upper bounded by

$$\begin{aligned}
\mathrm{QIL}(\Pi_k^{\mathrm{I}}) &\leq 2\lceil m/k \rceil r \mathrm{H}(\Lambda_{\mathrm{I},k}) \\
&= 2\lceil m/k \rceil r \left[(1 - k/m) \log\left(\frac{1}{1 - k/m} \right) + (k/m) \log(2m) \right] \\
&\leq 2r(2 + (1 + k/m) \log(2m)). \qquad (4.17)
\end{aligned}$$

As E is a constant rate code, then m is linear in n, so this quantity is upper bounded by $C' r \log n$ for some fixed constant $C' > 0$. In Appendix A.0.1 we show that for fixed error probability ε the repetition number r is also fixed, which completes the proof. □

4.3.2 Information leakage: Optical ring protocols

Here we bound the information leakage of the optical ring protocol family described in Section 4.2.2. In Appendix B we use a method similar to that used in Theorem 1 of [4] to bound the information leakage of both the ring and lattice protocol families as $\mathcal{O}(\log n)$ for fixed block-size k. We have found numerically that for the ring protocols the bound of this section is lower than that of Appendix B by $15-40\%$.

We use a similar technique as in the interpolation, and write the states in a basis for which the diagonal entries are equal for all $x \in \{0,1\}^n$. For block-size k and total mean photon number μ_k, each signal consists of one of 2^k coherent states equally spaced on a ring with amplitude $\beta_k = \mu_k/(m/k)$. Formally, each signal is contained in the set

$$S_k = \left\{ \left|\omega^j \beta_k\right\rangle \left\langle \omega^j \beta_k\right|, j = 0, \ldots, 2^k - 1 \right\}, \tag{4.18}$$

where

$$\omega = \exp\left[\frac{2\pi i}{2^k}\right]. \tag{4.19}$$

Now we show that the set S_k can be written in a basis such that the diagonal entries are equal for all $j = 0, \ldots, 2^k - 1$. As $\omega^{jl} = \omega^{jn}$ for all $n \equiv l \bmod 2^k$, then the states in S_k are equal to

$$\left|\omega^j \beta_k\right\rangle = e^{\frac{-|\beta_k|^2}{2}} \sum_{l=0}^{2^k-1} \omega^{jl} \sum_{n \equiv l \bmod 2^k} \frac{\beta_k^n}{\sqrt{n!}} |n\rangle \tag{4.20}$$

for $j = 0, \ldots, 2^k - 1$, which can be written in a basis as

$$\sum_{l=0}^{2^k-1} \omega^{jl} \lambda_l |l\rangle \tag{4.21}$$

for

$$\lambda_l = e^{\frac{-|\beta|^2}{2}} \sqrt{\sum_{n \equiv l \bmod 2^k} \frac{|\beta|^{2n}}{n!}}. \tag{4.22}$$

In this basis, the diagonal entries of each state in S_k form the probability vector $\Lambda_{\mathrm{O},k} = (\lambda_0^2, \ldots, \lambda_{2^k-1}^2)$. Writing each signal of the states used in the optical ring protocol in this basis, the diagonal entries of the transformed states are given by $\Lambda_{\mathrm{O},k}^{\otimes m/k}$. By (4.15) and additivity of entropy under tensor product, the information leakage of the optical ring protocol Π_k^{O} with block-size k is upper bounded by

$$\mathrm{QIL}(\Pi_k^{\mathrm{O}}) \leq 2m/k\mathrm{H}(\Lambda_{\mathrm{O},k}). \tag{4.23}$$

The righthand side is straightforward to calculate for practical values of n. It can also be used to bound the limiting behaviour as $\mathcal{O}(\log n)$ using standard asymptotic techniques. We have found that this bound gives an advantage of 15-40% over the bound of Appendix B.

4.4 Optimal unambiguous state comparison with a beamsplitter

Here we explicitly write any set of two coherent states $\{|\beta_0\rangle, |\beta_1\rangle\}$ in a basis as qubits, and construct a measurement in this basis (based on a measurement introduced in [8]) which reproduces the outcome probabilities of the beamsplitter measurement described in Section 4.1.1. Using a result of [8] we then find that a beamsplitter measurement with single photon threshold detectors placed at both output ports performs optimal unambiguous state comparison on $\{|\beta\rangle, |-\beta\rangle\}$ for any complex number β when the states are given with equal a priori probabilities.

We begin by writing $\{|\beta_0\rangle, |\beta_1\rangle\}$ in a basis as qubits. Let $\beta = \frac{1}{2}(\beta_0 - \beta_1)$ and

$$\begin{aligned} |q_a\rangle &= e^{\frac{-|\beta|^2}{2}} \left[\sqrt{\cosh(|\beta|^2)}\,|0\rangle + (-1)^a \sqrt{\sinh(|\beta|^2)}\,|1\rangle \right] \\ &= \sqrt{1-p}\,|0\rangle + (-1)^a \sqrt{p}\,|1\rangle \end{aligned} \tag{4.24}$$

for $p = \exp[-|\beta|^2]\sinh(|\beta|^2)$. It is straightforward to verify that $|\langle q_0, q_1\rangle| = |\langle \beta_0, \beta_1\rangle|$, so by Property 3 of [17] (reviewed in Section 4.1.1), the sets $\{|q_0\rangle, |q_1\rangle\}$ and $\{|\beta_0\rangle, |\beta_1\rangle\}$ are equal up to a change of basis.

Now we review a measurement (introduced in [8]) in the qubit basis which reproduces the outcome probabilities of the beamsplitter measurement described in Section 4.1.1, i.e. on input

$$\begin{aligned} |q_a\rangle\,|q_b\rangle =& (1-p)\,|00\rangle + (-1)^{a+b} p\,|11\rangle \\ &+ \sqrt{p(1-p)}[(-1)^b\,|01\rangle + (-1)^a\,|10\rangle] \end{aligned} \tag{4.25}$$

for $a, b \in \{0,1\}$, it outputs “dark port detection” with probability $1-|\langle q_a, q_b\rangle|$, and outputs “no dark port detection” with probability $|\langle q_a, q_b\rangle|$. To avoid confusion with the quantum fingerprinting protocols when we transition to describe the unrelated setting of unambiguous state comparison, we now change notation and rename the outcome “dark port detection” to “different” and outcome “no dark port detection” to “no difference detection”.

We describe the qubit measurement as a direct sum of measurements on the first and second terms of the above decomposition. The first term takes one of two forms depending on the equality of a and b. On this term, the measurement performs unambiguous state discrimination (USD) between these two forms. In particular, it uses the USD measurement which is optimal for the case in which each form is given with equal a priori probability [50]. Of course, this choice is sub-optimal for different a priori probabilities, but we make it because it gives rise to a measurement which reproduces the outcome probabilities of the beamsplitter measurement. We refer to the two unambiguous outcomes of this measurement as “plus” and “minus” corresponding to the two forms of the first term of (4.25), and the inconclusive outcome as “?”. On the second term of (4.25), a controlled-swap measurement is performed (which projects onto the symmetric and anti-symmetric subspaces). Outcomes “minus” and “anti-symmetric” are mapped to a single outcome “different” (previously “dark port detection”, see above) which unambiguously determines $a \neq b$, and occurs with probability $1-|\langle q_a, q_b\rangle|$ [8], thus reproducing the outcome probabilities of the beamsplitter measurement.

Following [8], we map outcomes “plus” and “symmetric” to a single outcome “same” which unambiguously determines $a = b$ with probability $1-\left|\langle q_a, q_{\overline{b}}\rangle\right|$ for $\overline{b} = 1 \oplus b$ [8]. In [8] it is shown that when $\Pr(a,b) = 1/4$ for all $a, b \in \{0,1\}$ this measurement performs optimal unambiguous comparison between the cases $a = b$ and $a \neq b$, i.e. it minimizes the probability of an inconclusive outcome “?”.

On input $|(-1)^a\beta\rangle\left|(-1)^b\beta\right\rangle$ to a beamsplitter with single photon threshold detectors placed at both the light and dark ports, a dark port detection occurs with probability $1-|\langle(-1)^a\beta, (-1)^b\beta\rangle|$ and a light port detection occurs with probability $1-|\langle(-1)^a\beta, (-1)^{\overline{b}}\beta\rangle|$, which are identical to the outcome probabilities of the above optimal unambiguous state comparison measurement in the qubit basis. Thus, this beamsplitter measurement performs optimal unambiguous state comparison on $\{|\beta\rangle, |-\beta\rangle\}$ when the states are given with equal a priori probabilities.

4.5 Discussion

In this chapter we developed several families of practical quantum fingerprinting protocols. One family demonstrated a trade-off between the number of signals sent and the dimension of each signal, thus opening the possibility for experimental implementations to find a "sweet spot" for which the number of signals and dimension of each signal are both technologically feasible. The other two families use coherent state signals arranged in a ring and lattice in phase space, respectively. We found that one such protocol reduced the number of signals from the existing coherent state protocol by a factor 1/2, while also reducing the information leakage. Although the other protocols in the ring and lattice families use even fewer signals, we found convincing numerical evidence that they have higher information leakage under the bounds we have used. We ask whether any other family of coherent state protocols might be used to reduce the number of signals and information leakage, and whether our information leakage bounds could be improved for these protocols.

Along the way, we discovered a simple beampslitter measurement to perform optimal unambiguous state comparison between two coherent states of equal amplitude and opposite phase when the states are given with equal a priori probability. We next ask whether a similar measurement could be used to perform optimal unambiguous state comparison of coherent states for arbitrary a priori probabilities.

Chapter 5

Appointment scheduling

Appointment scheduling is a task in the interactive communication setting (reviewed in Chapter 3) in which Alice and Bob receive inputs $x \in \mathrm{X}$ and $y \in \mathrm{Y}$ for $\mathrm{X} = \mathrm{Y} = \{0,1\}^n$, chosen according to some prior distribution $P \in \mathrm{Pr}(\mathrm{X} \times \mathrm{Y})$, and they wish to find an index i for which $x_i = y_i = 1$. The original appointment scheduling protocol was developed in [15] and uses $\mathcal{O}(\sqrt{n}\log n)$ qubits of communication, but only works when x and y intersect in exactly zero or one index i. This protocol was extended in [10], [18] to a protocol which works on any two inputs with again $\mathcal{O}(\sqrt{n}\log n)$ qubits of communication. This protocol has a nearly quadratic improvement over the classical communication complexity lower bound $\Omega(n)$ of appointment scheduling [53, 39]. In essense, this protocol uses a distributed version of the Grover search algorithm [31]. Accordingly, we refer to it as the distributed Grover search protocol. A lower bound of $\Omega(\sqrt{n})$ for the quantum communication cost of appointment scheduling was proven in [52], and a protocol attaining this bound was developed in [1].

In this chapter we present several appointment scheduling protocols which use coherent states and linear optics. We first present a protocol which implements the coherent state mapping described in Section 2.3 of the distributed Grover search protocol of [10], [18], and in Appendix C.3 we prove that it has information cost $\mathcal{O}(\sqrt{n}\log n)$. This is again a nearly quadratic improvement over the classical information cost lower bound of $\Omega(n)$ proven in [7] and [19] for the zero-error and nonzero-error cases, respectively. The distributed Grover search protocol consists of two unitaries repeatedly applied in succession: oracle calls and inversion about the mean. We find a method to implement the linear optics transformation corresponding to oracle calls which uses only local phase-shifters and swapping of two modes. Unfortunately, the linear optics transformation corresponding to inversion about the mean requires a global transformation on many modes, which poses a potential barrier to implementation of this protocol.

We circumvent this issue with appointment scheduling protocols which run a two-bit AND protocol subroutine on each pair of bits of x, y, a method introduced in [12]. We develop two subroutine AND protocols. Our first subroutine AND protocol is the direct coherent state mapping described in Section 2.3 applied to a qubit subroutine AND protocol used in [12] and originally developed in [38]. This protocol requires the exchange of coherent states in two modes, beamsplitters of high transmittivity, phase shifters, and single photon threshold detectors. Our second subroutine AND protocol is newly developed in this work, and it only requires the exchange of coherent states in two modes, beamsplitters of high transmittivity, and coherent state preparation. We design both protocols to handle experimental imperfections of loss and dark counts.

Although the asymptotic information cost of these protocols is $\mathcal{O}(n)$ (matching the classical lower bound), in Appendix C we find that the second protocol reduces the QIC in the finite setting by a factor of nearly 1/2 below the classical lower bound in the low loss regime of 99% transmittivity and dark count probability 4×10^{-8}. For the first protocol, we find that the factor of nearly 1/2 improvement comes for 99.9% transmittivity and dark count probability 4×10^{-8}. Due to the low loss required for a quantum advantage, we believe these protocols are still outside the scope of current technology.

5.1 Distributed Grover search protocol

In this section we describe the distributed Grover search protocol of [15, 10, 18], and then proceed to describe our implementation of the protocol's coherent state mapping (defined in Section 2.3).

5.1.1 Original distributed Grover search protocol

In the distributed Grover search protocol of [15, 10, 18] Alice and Bob receive $x, y \in \{0,1\}^n$ distributed according to some probability distribution $P \in \Pr(\mathrm{X} \times \mathrm{Y})$, which either have no intersection or intersect in k unknown indices $a_1, \ldots a_k$. These works describe a protocol which uses $\mathcal{O}(\sqrt{n} \log n)$ qubits of communication to either find a common intersection or determine with high probability that x and y do not intersect. In this section we review the protocol of [10][18] under the simplifying assumption $k \ll n$ (see [18] for an analysis without this assumption). We first consider the case in which k is known, and briefly discuss the extension to unknown k at the end of this section.

Now we describe the distributed Grover search protocol, which always outputs No Intersection on non-intersecting inputs, and outputs No Intersection on intersecting inputs with probability at most ε (with probability $1-\varepsilon$ it finds an intersection).

Distributed Grover search appointment scheduling protocol
First, Alice prepares the state

$$|s\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |i\rangle \,. \tag{5.1}$$

Choose iteration number $r = \lfloor \pi/(4\theta) \rfloor$, for θ satisfying $\sin^2\theta = k/n$. Then the following is iterated r times:

1. Alice and Bob jointly perform the oracle call unitary

$$U_A = \mathbb{1} - 2\sum_{j=1}^{k} |a_j\rangle\langle a_j| \tag{5.2}$$

   using the protocol outlined below.

2. Alice performs the inversion about the mean unitary

$$U_S = 2|s\rangle\langle s| - \mathbb{1}. \tag{5.3}$$

Then, Alice measures the state in the canonical basis, obtaining some outcome $i \in [n]$, and sends (i, x_i) to Bob. Bob then sends y_i to Alice. If they find that $x_i = y_i = 1$, they output this index. Otherwise, they repeat the protocol. If they repeat the protocol $K = \lceil \log(1/\varepsilon)/\log(n/k) \rceil$ times without finding an intersection, they output No Intersection.

The iteration number r is chosen as above because if x and y intersect in k indices, then the probability that Alice's measurement produces a non-intersecting index i is no greater than k/n under this choice (as discussed further below). The repetition number K is chosen to attain error probability ε.

In more details, the protocol evolves as follows. Let

$$|t\rangle = \frac{1}{\sqrt{1-(k/n)}}\left(|s\rangle - \frac{1}{\sqrt{n}}\sum_{j=1}^{k}|a_i\rangle\right) \tag{5.4}$$

and

$$|\tilde{a}\rangle = \frac{1}{\sqrt{k}}\sum_{j=1}^{k}|a_j\rangle, \tag{5.5}$$

then as shown in [10], after l applications of $U_S U_A$ the state is given by

$$(U_S U_A)^l |s\rangle = \sin((2l+1)\theta)|\tilde{a}\rangle + \cos((2l+1)\theta)|t\rangle, \tag{5.6}$$

for θ defined as above. In [10] it is shown that for $r = \lfloor \pi/(4\theta) \rfloor$, the probability $\cos^2((2r+1)\theta)$ that Alice's measurement does not output an intersecting index i satisfies $\cos^2((2r+1)\theta) \leq k/n$.

Now we detail how Alice and Bob jointly perform the oracle call unitary U_A.

Procedure to implement oracle call unitary U_A

Alice prepares auxilliary qubits $|0\rangle|-\rangle$, so the state of her entire register is $|\psi\rangle|0\rangle|-\rangle$, where $|\psi\rangle \in \mathcal{S}(\mathbb{C}^n)$ is the resultant state from the previous step in the appointment scheduling protocol.

1. Alice applies $(U_x \otimes \mathbb{1}_2)$ and sends the entire state to Bob.
2. Bob applies $(\mathbb{1}_n \otimes W)(V)(\mathbb{1}_n \otimes W)$ and sends the entire state back to Alice.
3. Alice applies $(U_x \otimes \mathbb{1}_2)$, and discards the qubits $|0\rangle|-\rangle$.

Where $U_x, U_y \in \mathcal{U}(\mathbb{C}^n \otimes \mathbb{C}^2)$ act as

$$U_x|i\rangle|z\rangle = |i\rangle|x_i \oplus z\rangle \text{ for all } i = 1,\ldots,n \tag{5.7}$$

$$U_y|i\rangle|z\rangle = |i\rangle|y_i \oplus z\rangle \text{ for all } i = 1,\ldots,n, \tag{5.8}$$

(which Alice and Bob can implement, respectively), $W \in \mathcal{U}(\mathbb{C}^2 \otimes \mathbb{C}^2)$ is the swap operator which acts as

$$W|i\rangle|j\rangle = |j\rangle|i\rangle \text{ for all } i,j = 1,2, \tag{5.9}$$

and V is the control-U_y gate, where U_y acts on the first two systems, and the state of the third system is the control.

It is straightforward to show that

$$(U_x \otimes \mathbb{1}_2)(\mathbb{1}_n \otimes W)(V)(\mathbb{1}_n \otimes W)(U_x \otimes \mathbb{1}_2) |i\rangle |0\rangle |-\rangle = (U_A \otimes \mathbb{1}_2 \otimes \mathbb{1}_2) |i\rangle |0\rangle |-\rangle \tag{5.10}$$

for all $i \in [n]$, so the above procedure implements U_A.

Thus, for each application of U_A, Alice and Bob exchange $2(\log(n)+2)$ qubits. For $k \ll n$, U_A must be implemented at most $Kr = \mathcal{O}(\sqrt{n/k})$ times. For each repetition of the protocol, Alice sends Bob her measurement outcome i (which is $\log n$ bits) along with x_i (which is one bit), and Bob sends Alice y_i (which is one bit). Thus, the amount of communication in these stages is upper bounded by $K(\log n + 2) = \mathcal{O}(\log n)$ bits. Thus, the protocol uses a total of $\mathcal{O}(\sqrt{n/k}\log(n))$ qubits of communication.

Now we consider the case in which k is unknown to either party, but is known to be much less than n (see [18] for an analysis without this assumption). The implementation of the unitaries U_S and U_A is independent of k, so they can still be applied, but the iteration number r is a function of k (and n), and must now be chosen in a different manner. The protocol proposed in [10], [18] uses a randomized algorithm to choose the iteration number r, and finds a common intersection (or determines no intersection with high probability) while maintaining the $\mathcal{O}(\sqrt{n}\log(n))$ behaviour.

5.1.2 Coherent state distributed Grover search with practical oracle calls

We proceed to describe the coherent-state mapping of Section 2.3 applied to the distributed Grover search protocol. We find a protocol for which the linear optics transformation V_A corresponding to the oracle call U_A uses only local phase shifters and the swapping of two modes. Unfortunately, the linear optics transformation V_S corresponding to the inversion about the mean U_S still requires a global transformation of the state. In Appendix C.3 we prove that this protocol has information cost $\mathcal{O}(\sqrt{n}\log n)$, a nearly quadratic improvement over the classical information cost lower bound of $\Omega(n)$ proven in [7] and [19] for the zero-error and nonzero-error cases, respectively.

We again consider only the case in which x and y either have no intersection or intersect in k unknown indices $a_1, \ldots a_k$ for $k \ll n$. We suggest that this protocol could be adapted in similar fashion to [10], [18] if this is not the case.

We first describe the coherent state mapping of the distributed Grover search protocol in terms of V_S and V_A. Let $|\psi\rangle = \sum_{i=1}^{n} \lambda_i |i\rangle \in \mathcal{S}(\mathbb{C}^n)$ be an arbitrary pure state, which will help us describe the action of V_S and V_A. The following coherent state mapping of the distributed Grover search protocol always outputs No Intersection on non-intersecting inputs, and outputs No Intersection on intersecting inputs with probability at most ε (with probability $1-\varepsilon$ it finds an intersection).

Coherent state mapping of distributed Grover search protocol
For some constant $\alpha \in \mathbb{C}$ (which can be optimized over), Alice prepares the state

$$\bigotimes_{i=1}^{n} \left|\alpha/\sqrt{n}\right\rangle_i \tag{5.11}$$

Choose iteration number $r = \lfloor \pi/(4\theta) \rfloor$, for θ satisfying $\sin^2\theta = k/n$. Then the following is repeated r times:

1. Alice and Bob jointly perform the linear optics transformation V_A corresponding to the oracle call U_A, which acts as

$$V_A f_\alpha(|\psi\rangle) = \bigotimes_{i=1}^{n} \left|(-1)^{x_i \wedge y_i} \lambda_i \alpha\right\rangle_i, \tag{5.12}$$

   using the protocol outlined below.

2. Alice performs the linear optics transformation V_S corresponding to the inversion about the mean U_S, which acts as

$$V_S f_\alpha(|\psi\rangle) = \bigotimes_{i=1}^{n} |(2v-\lambda_i)\alpha\rangle_i, \tag{5.13}$$

   for $\nu = (\lambda_1 + \cdots + \lambda_n)/n$.

Alice measures each mode with single photon threshold detectors. If no detectors click, she announces this and the parties repeat the protocol. Otherwise, she chooses a random index i for which she received a click, and sends (i, x_i) to Bob. Bob then sends y_i to Alice. If $x_i = y_i = 1$, the parties output this index. Otherwise, they repeat the protocol. If they repeat the protocol

$$K = \left\lceil \log(1/\varepsilon) \Big/ \log\left(\frac{1}{1 - e^{-|\alpha|^2 \frac{k}{n}} + e^{-|\alpha|^2}}\right) \right\rceil \tag{5.14}$$

times without finding an intersection, they output No Intersection.

The iteration number r is chosen as above because if x and y intersect in k indices, then the probability that Alice's measurement produces a non-intersecting index i is no greater than $(1 - e^{-|\alpha|^2 \frac{k}{n}})$ under this choice (as discussed further below). The repetition number K is chosen to attain error probability ε. The extra term $e^{-|\alpha|^2}$ is the probability that no clicks occur.

In more details, the protocol evolves as follows. After l applications of $V_S V_A$, coherent states in intersecting modes will have amplitude $\sin((2l+1)\theta)\frac{\alpha}{\sqrt{k}}$, and coherent states in non-intersecting modes will have amplitude $\cos((2l+1)\theta)\frac{\alpha}{\sqrt{n-k}}$. This follows directly from (5.6) and the coherent state mapping. Thus, after r iterations of $V_S V_A$, coherent states in intersecting modes will have mean photon number $\sin^2((2r+1)\theta)\frac{|\alpha|^2}{k} \geq \frac{1-k/n}{k}|\alpha|^2$ and coherent states in non-intersecting modes will have mean photon number $\cos^2((2r+1)\theta)\frac{|\alpha|^2}{n-k} \leq \frac{k}{n(n-k)}|\alpha|^2$. Thus, at least one of the $n-k$ non-intersecting modes i will click with probability no greater than

$$1 - e^{-|\alpha|^2 \frac{k}{n(n-k)}(n-k)} = 1 - e^{-|\alpha|^2 \frac{k}{n}}. \tag{5.15}$$

No clicks occur with probability

$$e^{-|\alpha|^2}. \tag{5.16}$$

Thus, when x and y intersect in k locations, the probability that Alice sends Bob a non-intersecting index i or that no clicks occur is upper bounded by $1 - e^{-|\alpha|^2 \frac{k}{n}} + e^{-|\alpha|^2}$, which justifies the above choice of repetition number K.

Now we describe Alice and Bob's procedure to implement V_A.

Procedure to implement linear optics transformation V_A corresponding to oracle call U_A

First, Alice prepares n auxilliary modes initialized to $|0\rangle$, so the state of her entire register is

$$\bigotimes_{i=1}^{n}(|\lambda_i\alpha\rangle\,|0\rangle), \tag{5.17}$$

where $\bigotimes_{i=1}^{n}|\lambda_i\alpha\rangle$ is the resultant state from the previous step in the coherent state protocol. Then,

1. For each i in which $x_i = 1$, Alice swaps the i-th pair of modes $|\lambda_i\alpha\rangle\,|0\rangle \rightarrow |0\rangle\,|\lambda_i\alpha\rangle$ (and otherwise applies the identity map), and sends the entire state to Bob.
2. For each i in which $y_i = 1$, Bob flips the sign of the second mode corresponding to index i using a phase shifter, and sends the entire state back to Alice.
3. Alice repeats the first step: For each i in which $x_i = 1$, she swaps the i-th pair of modes $|\lambda_i\alpha\rangle\,|0\rangle \rightarrow |0\rangle\,|\lambda_i\alpha\rangle$ (and otherwise applies the identity map). Alice then discards the n auxilliary modes.

It is straightforward to show that this procedure implements V_A exactly.

5.2 Bitwise-AND protocols: Ideal setting

Here we present a protocol for appointment scheduling using a two-bit AND protocol as a subroutine, a method introduced in [12]. We then develop two subroutine AND protocols. Our first subroutine AND protocol is the direct coherent state mapping described in Section 2.3 applied to a qubit subroutine AND protocol used in [12] and originally developed in [38]. Our second subroutine AND protocol is newly developed in this work. In this section we present these protocols in the ideal setting, and in Section 5.3 we extend these protocols to handle the experimental imperfections of loss and dark counts.

As we will see in Appendix C.2, both of our coherent state two-bit AND protocols have high information cost when the inputs $x_i, y_i \in \{0,1\}$ have high probability of intersection. This follows intuitively from the fact that when $\text{AND}(x_i, y_i) = 1$, Alice and Bob both gain full knowledge of the other's input. To avoid this effect in the case when the input strings $x, y \in \{0,1\}^n$ intersect in many coordinates with high probability, the parties first classically randomly sample some number s of the n bits and check whether any intersect. If any of these bits intersect, the protocol terminates, and if not the protocol proceeds with a bitwise-AND quantum protocol. If no intersection is found in the s randomly chosen bits, then the remaining bits have low probability to intersect in many locations, which reduces the information cost of the protocol. This classical subsampling is an adaptation of the protocol of [12], which instead uses pre-shared entanglement for this stage to attain lower information cost at the expense of being more difficult to implement.

In the ideal setting, both the two-bit AND protocols we consider are zero-error protocols, i.e. they always evaluate the AND function correctly. We now construct a zero-error appointment scheduling protocol from any such two-bit AND protocol. Note that this protocol differs from the Grover search protocol in that the Grover search protocol had some probability of error, whereas this protocol has zero error in the ideal setting.

Zero-error appointment scheduling protocol Π_D running zero-error protocol Π_A for two-bit AND as subroutine

1. Alice randomly samples with replacement s coordinates, and sends coordinate set S to Bob.
2. Alice sends x_i to Bob for each $i \in S$.
3. If Bob finds any $i \in S$ with $x_i = y_i = 1$, he sends smallest such i to Alice, and both parties output this index, else:
4. Alice and Bob run Π_A on remaining $n - s$ coordinates x_i, y_i.
5. If they find any index $i \in [n]$ such that $\text{AND}(x_i, y_i) = 1$, both output smallest such i
6. If they find $\text{AND}(x_i, y_i) = 0$ for all $i \in [n]$, they output No Intersection.

In the following sections we describe our two coherent state two-bit AND protocols.

5.2.1 Coherent state mapping of qubit two-bit AND protocol

In this section we describe the zero-error two-bit AND protocol used in [12] and originally developed in [38], and its coherent state mapping (as defined in Section 2.3), which will be the first of two coherent state two-bit AND protocols that we consider.

Qubit Protocol

We first recall the zero-error qubit two-bit AND protocol of used in [12] and originally developed in [38]. On inputs $x_i, y_i \in \{0,1\}$ given to Alice and Bob respectively, this protocol computes $\text{AND}(x,y)$ in r rounds for any even positive integer r.

First, let $\theta = \frac{\pi}{4r}$ and $|v\rangle = \cos(\theta)|0\rangle + \sin(\theta)|1\rangle$. Let U_v be the unitary operator reflecting about the vector $|v\rangle$, i.e. $U_v|0\rangle = \cos(2\theta)|0\rangle + \sin(2\theta)|1\rangle$ and $U_v|1\rangle = \sin(2\theta)|0\rangle - \cos(2\theta)|1\rangle$. Let U_0 be the operator reflecting about $|0\rangle$, i.e. $U_0|0\rangle = |0\rangle$ and $U_0|1\rangle = -|1\rangle$.

The unambiguous qubit AND protocol of [38] proceeds as follows.

Qubit AND protocol of [38]

First, Alice prepares a qubit-register C initialized to the state $|0\rangle$. Then, on each round, Alice and Bob do the following:

1. On $x_i = 0$ ($x_i = 1$), Alice performs the identity map (U_v map) on the register C and sends it to Bob.

2. On $y_i = 0$ ($y_i = 1$), Bob performs the identity map (U_0 map) on the register C and sends it to Alice.

After r rounds the state of register C will be $|0\rangle$ $(-|1\rangle)$ if $\text{AND}(x_i, y_i) = 0\ (1)$. Alice measures C in the standard basis to determine the result, which she communicates to Bob. Clearly this is an unambiguous two-bit AND protocol with zero probability of an inconclusive outcome.

Coherent state version

The first coherent state AND protocol $\widetilde{\Pi}_A^1$ that we consider is the direct coherent-state mapping of Section 2.3 applied to the above qubit protocol. As we will see, in contrast to the qubit protocol, $\widetilde{\Pi}_A^1$ has some probability p of an inconclusive outcome. We therefore modify protocol $\widetilde{\Pi}_A^1$ to a protocol $\Pi_A^1(\widetilde{\Pi}_A^1)$ with zero-error as follows:

Definition 35. *Given a protocol $\widetilde{\Pi}$ which can be inconclusive, and such that Alice and Bob always agree on whether a run was inconclusive or not, we recursively define $\Pi(\widetilde{\Pi})$ as follows:*

Protocol $\Pi(\widetilde{\Pi})$:

1. Run Protocol $\widetilde{\Pi}$.
2. If $\widetilde{\Pi}$ returns an output, return this output.
3. Else, if $\widetilde{\Pi}$ is inconclusive, rerun $\Pi(\widetilde{\Pi})$.

Now we explicitly define $\widetilde{\Pi}_A^1$. Define a unitary V_0 as

$$V_0 \left|\alpha\right\rangle \left|\beta\right\rangle = \left|\alpha\right\rangle \left|-\beta\right\rangle, \tag{5.18}$$

which acts as a phase flip on the second mode. Clearly, $V_0 f_\alpha(|\psi\rangle) = f_\alpha(U_0|\psi\rangle)$ for every state $|\psi\rangle$ used in the qubit protocol.

Define a unitary R_θ as

$$R_\theta \left|\alpha\right\rangle \left|\beta\right\rangle = \left|\cos(\theta)\alpha - \sin(\theta)\beta\right\rangle \left|\sin(\theta)\alpha + \cos(\theta)\beta\right\rangle \tag{5.19}$$

which acts as a beamsplitter specified by angle θ. Define a unitary $V_v = R_\theta V_0 R_\theta^\dagger$, where $R_\theta^\dagger = R_{-\theta}$. It can be shown that $V_v f_\alpha(|\psi\rangle) = f_\alpha(U_v|\psi\rangle)$ for every state $|\psi\rangle$ used in the qubit protocol.

The protocol $\widetilde{\Pi}_A^1$ then proceeds as follows.

Coherent state mapping $\widetilde{\Pi}_A^1$ of qubit AND protocol

First, Alice prepares a two-mode register C in state $|\alpha,0\rangle$, for some $\alpha > 0$. On each of the r rounds, Alice and Bob do the following:

1. On $x_i = 0$ ($x_i = 1$), Alice performs the identity map (V_v map) on the register C and sends it to Bob.
2. On $y_i = 0$ ($y_i = 1$), Bob performs the identity map (V_0 map) on the register C and sends it to Alice.

After r rounds, Alice measures each mode of C with single photon threshold detectors and communicates the result to Bob.

In ideal implementations, after all unitaries are performed Alice ends up with $|\alpha, 0\rangle$ on inputs $(0,0)$, $(0,1)$ and $(1,0)$, and with $|0,-\alpha\rangle$ on input $(1,1)$. Thus, she might detect a photon in the first mode only if the output to AND is 0 and she might detect a photon in the second mode only if the output to AND is 1. If she does not detect any photon, she tells Bob that the run was inconclusive. Note that Alice obtains a click with probability $1-e^{-|\alpha|^2}$ for any input. Thus, this protocol never outputs a wrong answer, and has some uniform probability $p = e^{-|\alpha|^2}$ of outcome "Inconclusive". For clarity, we explicitly write down how the protocol evolves for different inputs:

Evolution of $\widetilde{\Pi}_A^1$ for different inputs:

On (0, 0): $|\alpha,0\rangle \quad \to_A |\alpha,0\rangle \quad \to_B |\alpha,0\rangle \quad \to_A \cdots$

On (0, 1): $|\alpha,0\rangle \quad \to_A |\alpha,0\rangle \quad \to_B |\alpha,0\rangle \quad \to_A \cdots$

On (1, 0):

$$\begin{array}{lll} |\alpha,0\rangle & \to_A |\cos(2\theta)\alpha, \sin(2\theta)\alpha\rangle & \to_B |\cos(2\theta)\alpha, \sin(2\theta)\alpha\rangle \\ & \to_A |\alpha,0\rangle & \to_B |\alpha,0\rangle \\ & \to_A |\cos(2\theta)\alpha, \sin(2\theta)\alpha\rangle & \to_B |\cos(2\theta)\alpha, \sin(2\theta)\alpha\rangle \\ & \to_A |\alpha,0\rangle & \to_B |\alpha,0\rangle \\ & \vdots & \vdots \end{array}$$

On (1, 1):

$$\begin{array}{lll} |\alpha,0\rangle & \to_A |\cos(2\theta)\alpha, \sin(2\theta)\alpha\rangle & \to_B |\cos(2\theta)\alpha, -\sin(2\theta)\alpha\rangle \\ & \to_A |\cos(4\theta)\alpha, \sin(4\theta)\alpha\rangle & \to_B |\cos(4\theta)\alpha, -\sin(4\theta)\alpha\rangle \\ & \vdots & \vdots \\ & \to_A |\cos(2r\theta)\alpha, \sin(2r\theta)\alpha\rangle & \to_B \left|\cos(\frac{\pi}{2})\alpha, -\sin(\frac{\pi}{2})\alpha\right\rangle \\ & =|0,-\alpha\rangle & \end{array}$$

On (0,0) and (0,1) Alice and Bob's manipulations leave the state unchanged. On (1,0) Alice performs V_v and Bob does nothing. Since V_v is its own inverse, the state oscillates between two forms in this case. On (1,1) Alice and Bob's manipulations bring the state to $|0,-\alpha\rangle$ after r rounds.

Having described the first coherent state two-bit AND protocol, we now proceed to describe the second protocol.

5.2.2 Zero-state injection two-bit AND protocol

Now we propose a second coherent state two-bit AND protocol $\widetilde{\Pi}_A^2$ which only requires Alice to apply beamsplitters and Bob to prepare coherent states. The second protocol has the advantage of being potentially more easily implementable than the first, and it also has lower information cost under the bounds derived in Appendix C.2. We call this protocol the zero-state injection protocol because it is the coherent state mapping (see Section 2.3) of a qubit AND protocol in which Bob replaces the state of the communication register with $|0\rangle$ in some cases. For brevity, we avoid explicit description of this qubit protocol and proceed directly to the coherent state protocol.

As in protocol $\widetilde{\Pi}_A^1$, protocol $\widetilde{\Pi}_A^2$ computes AND with some probability p of an inconclusive outcome. We recursively define the zero-error AND protocol $\Pi_A^2(\widetilde{\Pi}_A^2)$ from $\widetilde{\Pi}_A^2$ as in Definition 35. In contrast to protocol $\widetilde{\Pi}_A^1$, any positive integer number of rounds r can be chosen for protocol $\widetilde{\Pi}_A^2$.

Zero-state injection coherent state AND protocol $\widetilde{\Pi}_A^2$

First, Alice prepares a two-mode register C in state $|\alpha,0\rangle$ and Bob prepares a $2r$-mode register B in state $|\alpha,0\rangle^{\otimes r}$. On each of the r rounds, Alice and Bob do the following:

1. On $x_i = 0$ ($x_i = 1$), Alice applies the identity map (R_θ map with $\theta = \frac{\pi}{2r}$) to the register C and sends the transformed state to Bob.
2. On $y_i = 0$, for round j, Bob swaps the state of register C with his j-th copy of $|\alpha,0\rangle$ and sends it to Alice. On $y_i = 1$, Bob applies the identity map to register C and returns it to Alice.

After r rounds, Alice measures each mode of C with single photon threshold detectors and communicates the result to Bob.

In ideal implementations, after all unitaries are performed, Alice ends up with $|\alpha,0\rangle$ on inputs $(0,0)$, $(0,1)$ and $(1,0)$, and with $|0,\alpha\rangle$ on input $(1,1)$. Thus, (as in protocol $\widetilde{\Pi}_A^1$) she might detect a photon in the first mode only if the output to AND is 0 and she might detect a photon in the second mode only if the output to AND is 1. If she does not detect any photon, she tells Bob that the run was inconclusive. This protocol never errs, and is inconclusive with probability $e^{-|\alpha|^2}$. For clarity, we explicitly write down how the protocol evolves on different inputs. The terms in parentheses denote Bob's memory.

Evolution of $\widetilde{\Pi}_A^2$ for different inputs:

On (0, 0): $|\alpha,0\rangle(|\alpha,0\rangle^{\otimes r}) \quad \to_A |\alpha,0\rangle(|\alpha,0\rangle^{\otimes r}) \quad \to_B |\alpha,0\rangle(|\alpha,0\rangle^{\otimes r}) \quad \to_A \cdots$

On (0,1): $|\alpha,0\rangle(|\alpha,0\rangle^{\otimes r}) \quad \to_A |\alpha,0\rangle(|\alpha,0\rangle^{\otimes r}) \quad \to_B |\alpha,0\rangle(|\alpha,0\rangle^{\otimes r}) \quad \to_A \cdots$

On (1, 0):

$$
\begin{aligned}
&|\alpha,0\rangle(|\alpha,0\rangle^{\otimes r}) \\
&\to_A |\cos(\theta)\alpha,\sin(\theta)\alpha\rangle(|\alpha,0\rangle^{\otimes r}) && \to_B |\alpha,0\rangle(|\cos(\theta)\alpha,\sin(\theta)\alpha\rangle\,|\alpha,0\rangle^{\otimes r-1}) \\
&\to_A |\cos(\theta)\alpha,\sin(\theta)\alpha\rangle(|\cos(\theta)\alpha,\sin(\theta)\alpha\rangle\,|\alpha,0\rangle^{\otimes r-1}) \\
& && \to_B |\alpha,0\rangle(|\cos(\theta)\alpha,\sin(\theta)\alpha\rangle^{\otimes 2}\,|\alpha,0\rangle^{\otimes r-2}) \\
&\to_A |\cos(\theta)\alpha,\sin(\theta)\alpha\rangle(|\cos(\theta)\alpha,\sin(\theta)\alpha\rangle^{\otimes 2}\,|\alpha,0\rangle^{\otimes r-2}) \\
& && \to_B |\alpha,0\rangle(|\cos(\theta)\alpha,\sin(\theta)\alpha\rangle^{\otimes 3}\,|\alpha,0\rangle^{\otimes r-3}) \\
&\quad\vdots && \vdots \\
&\to_A |\cos(\theta)\alpha,\sin(\theta)\alpha\rangle(|\cos(\theta)\alpha,\sin(\theta)\alpha\rangle^{\otimes r-1}\,|\alpha,0\rangle^{\otimes 1}) \\
& && \to_B |\alpha,0\rangle(|\cos(\theta)\alpha,\sin(\theta)\alpha\rangle^{\otimes r}) \\
&= |\alpha,0\rangle(|\cos(\theta)\alpha,\sin(\theta)\alpha\rangle^{\otimes r})
\end{aligned}
$$

On (1, 1):

$$
\begin{aligned}
&|\alpha,0\rangle(|\alpha,0\rangle^{\otimes r}) \\
&\to_A |\cos(\theta)\alpha,\sin(\theta)\alpha\rangle(|\alpha,0\rangle^{\otimes r}) && \to_B |\cos(\theta)\alpha,\sin(\theta)\alpha\rangle(|\alpha,0\rangle^{\otimes r}) \\
&\to_A |\cos(2\theta)\alpha,\sin(2\theta)\alpha\rangle(|\alpha,0\rangle^{\otimes r}) && \to_B |\cos(2\theta)\alpha,\sin(2\theta)\alpha\rangle(|\alpha,0\rangle^{\otimes r}) \\
&\vdots && \vdots \\
&\to_A |\cos(r\theta)\alpha,\sin(r\theta)\alpha\rangle(|\alpha,0\rangle^{\otimes r}) && \to_B \left|\cos(\tfrac{\pi}{2})\alpha,\sin(\tfrac{\pi}{2})\alpha\right\rangle(|\alpha,0\rangle^{\otimes r}) \\
&= |0,\alpha\rangle(|\alpha,0\rangle^{\otimes r}).
\end{aligned}
$$

On (0,0) and (0,1) Alice and Bob's manipulations leave the state unchanged. On (1,0) Alice rotates the state and then Bob replaces it with $|\alpha,0\rangle$ on each round. On (1,1) Alice and Bob's manipulations bring the state to $|0,\alpha\rangle$ after r rounds.

This concludes the description of the second coherent state two-bit AND protocol.

5.3 Bitwise-AND protocols: Experimental imperfections

In this section we extend the zero-error protocol Π_D for appointment scheduling described in Section 5.2 to a protocol which handles channel transmitivity η and noisy detectors characterized by dark count probability p_d (the probability that a detector will click even when no incident photons from the signal are present). We define the extended appointment scheduling protocol $\Pi_D(\eta, p_d, \varepsilon)$ based on any extended subroutine $\Pi_A(\eta, p_d, \varepsilon)$ which evaluates two-bit AND with one-sided error $\varepsilon \geq 0$ (which can be set to zero only if $p_d = 0$). That is, if $\mathrm{AND}(x_i, y_i) = 1$, Π_A returns "1" with probability at least $1 - \varepsilon$ and returns "0" with probability at most ε, and if $\mathrm{AND}(x_i, y_i) = 0$, Π_A always returns "0". We then describe extensions of the two coherent state two-bit AND protocols described above to protocols $\Pi_A(\eta, p_d, \varepsilon)$ which attain one-sided error ε.

The extended protocol $\Pi_D(\eta, p_d, \varepsilon)$ for Appointment Scheduling (running protocol $\Pi_A(\eta, p_d, \varepsilon)$ for AND as subroutine) is exactly the same as the protocol Π_D described in Section 5.2 with Π_A extended to $\Pi_A(\eta, p_d, \varepsilon)$. If x and y do not intersect, this protocol will always output No Intersection. If x and y do intersect in, say, k positions, then this protocol will erroneously output No Intersection with probability no greater than $\varepsilon^k \leq \varepsilon$.

Now we extend the protocols Π_A^l, $l = 1, 2$ described in Section 5.2 to $\Pi_A^l(\eta, p_d, \varepsilon)$. We define both protocols simultaneously. In similar spirit to Definition 35, we recursively define $\Pi_A^l(\eta, p_d, \varepsilon) = \Pi_A^l(\widetilde{\Pi}_A^l(\eta), \varepsilon)$ from a subroutine $\widetilde{\Pi}_A^l(\eta)$, defined as follows:

Subroutine protocol $\widetilde{\Pi}_A^l(\eta)$, for $l = 1, 2$:

1. In the initialization phase, for $\widetilde{\Pi}_A^1(\eta)$ Alice prepares $|\alpha/\eta^r, 0\rangle$. For $\widetilde{\Pi}_A^2(\eta)$ Alice prepares $|\alpha/\eta^r, 0\rangle$ and Bob prepares $\left|\alpha/\eta^{r-1/2}, 0\right\rangle \left|\alpha/\eta^{r-3/2}, 0\right\rangle \ldots \left|\alpha/\eta^{1/2}, 0\right\rangle$.

2. The parties then perform the same operations as in $\widetilde{\Pi}_A^l(1)$ (the original $\widetilde{\Pi}_A^l$ of Section 5.2) and Alice measures both modes with single photon threshold detectors. Note that if $p_d = 0$, the first (second) mode might click only if $\mathrm{AND}(x, y) = 0(1)$, and if $p_d > 0$, any input could give rise to a click in either mode (or both).

3. If only the second mode clicks, Alice sends x_i to Bob and Bob sends y_i to Alice. They then output "0" or "1" corresponding to the value of $\mathrm{AND}(x_i, y_i)$.

4. If neither mode clicks or both modes click, Alice sends the result to Bob and both parties output "Inconclusive".

5. If only the first mode clicks, Alice sends the result to Bob and both parties output "Zero?".

The two-bit AND protocol $\Pi^l_A\left(\widetilde{\Pi}^l_A(\eta), p_d, \varepsilon\right)$ is defined for any $\varepsilon \geq 0$ in the case $p_d = 0$ and any $\varepsilon > 0$ in the case $p_d > 0$, and proceeds as follows:

Two-bit AND protocol $\Pi^l_A\left(\widetilde{\Pi}^j_A(\eta), p_d, \varepsilon\right)$ for $l = 1,2$:

Initialize $K' = 0$.

1. Run Protocol $\widetilde{\Pi}^l_A(\eta)$.
2. If $\widetilde{\Pi}^l_A(\eta)$ returns “0” or “1”, return this output.
3. If $\widetilde{\Pi}^l_A(\eta)$ returns “Inconclusive,” restart $\Pi^l_A\left(\widetilde{\Pi}^l_A(\eta), \varepsilon\right)$.
4. If $\widetilde{\Pi}^l_A(\eta)$ returns “Zero?” and $p_d = 0$, return “0”.
5. If $\widetilde{\Pi}^l_A(\eta)$ returns “Zero?” and $p_d > 0$, set $K' = K' + 1$.
6. If $K' \geq \frac{\log(1/\varepsilon)}{\log(1/p_s)}$, return “0.” Else, restart $\Pi^l_A\left(\widetilde{\Pi}^l_A(\eta), p_d, \varepsilon\right)$.

Where

$$p_s = e^{-|\alpha|^2}(1 - p_d)p_d \tag{5.20}$$

is the probability that only the second mode clicks for non-intersecting inputs.

It is straightforward to show that the above protocol computes AND with one-sided error ε. This concludes the description of the two coherent state two-bit AND subroutine protocols.

5.4 Numerical results

In this section we present our numerical results on our bitwise-AND coherent state appointment scheduling protocols.

In Figure 5.1 we plot QIC$/n$ of the bitwise-AND protocol running $\widetilde{\Pi}^2_A$ as subroutine in the ideal setting for $\varepsilon = 0$, and under experimental imperfections $p_d = 4 \times 10^{-8}$ and $\eta = 0.99$ for $\varepsilon = 4 \times 10^{-8}$, using the bounds derived in Appendix C. We compare these results to the classical

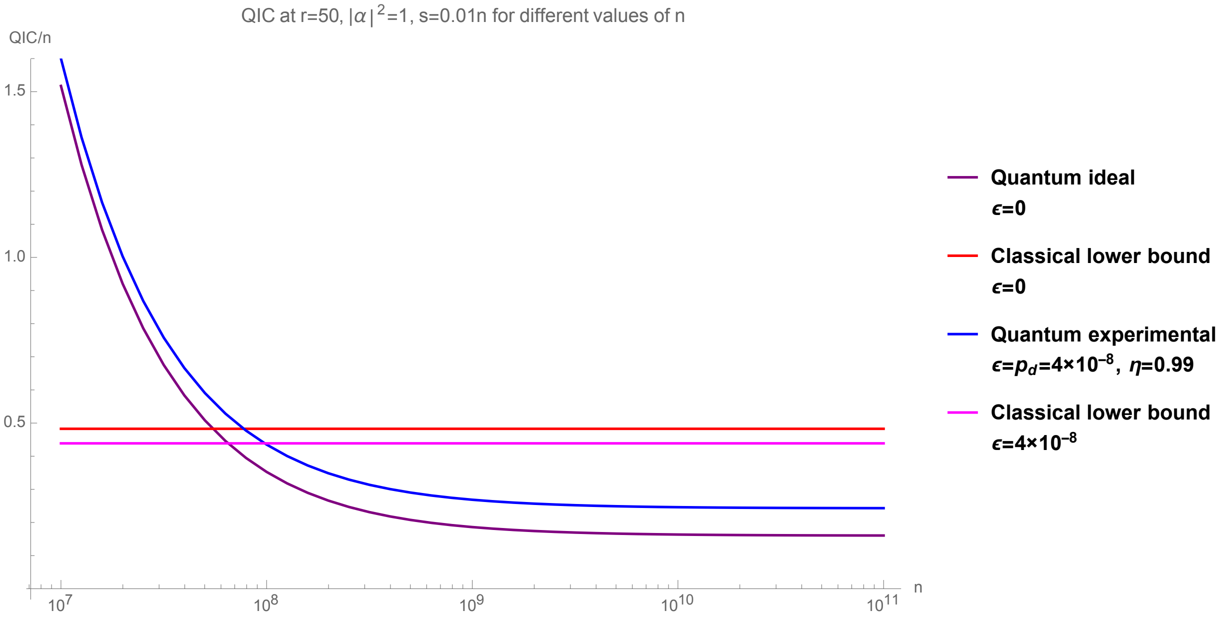

Figure 5.1: QIC$/n$ in the ideal setting for $\varepsilon = 0$, and under experimental imperfections $p_d = 4 \times 10^{-8}$ and $\eta = 0.99$ for $\varepsilon = 4 \times 10^{-8}$, compared with the classical lower bounds of [7], [19]

lower bounds derived in [7] and [19] for the zero-error and non-zero error settings, respectively, and find that our appointment scheduling protocol has lower QIC than the classical lower bound by a factor of nearly 1/2. We have optimized over the mean photon number, the number of rounds, and the number of bits of classical subsampling and settled on $|\alpha|^2 = 1$, $r = 50$, and $s = 0.01n$, respectively.

For the bitwise-AND protocol running $\widetilde{\Pi}_A^1$ as subroutine, we have found a factor of nearly 1/2 improvement for $p_d = \varepsilon = 4 \times 10^{-8}$ and $\eta = 0.999$. As this protocol performs worse than $\widetilde{\Pi}_A^2$ and seems more difficult to implement in practice, we neglect an explicit plot of this result.

5.5 Discussion

In this chapter we developed quantum appointment scheduling protocols using coherent states. We first developed a coherent state version of the original distributed Grover search protocol. This protocol is largely still of theoretical interest due to the experimental infeasibility of the required global linear optics transformation of inversion about the mean. We also developed bitwise AND protocols for appointment scheduling which require coherent states and linear optics on just two modes. Although these protocols seem significantly more realistic for implementation than the distributed Grover search, they still appear to be outside the realm of current technology, primarily due to the low channel loss required.

We have made repeated use of the mapping of [3] (described in Section 2.3) from quantum protocols using pure states in $\mathbb{C}^n$ to optical protocols using a tensor product of n coherent states. The apparent usefulness of this mapping leads us to ask whether similar mappings might exist to, say, protocols using a tensor product of qubits, or some other readily implementable states.

We hope that this work might inspire further developments towards practical appointment scheduling, and that the techniques we have developed in Appendix C to analyze the information cost of our protocols might be useful in other settings.

Chapter 6

Conclusion

In this thesis we have developed quantum communication protocols for quantum fingerprinting and appointment scheduling which are feasible for experimental implementation and outperform their classical counterparts, even under experimental imperfections. Along the way, we have developed tools for analyzing the information content of quantum communication protocols. Looking ahead, we hope that these protocols and our methods to analyze their information content might inspire and inform further work towards practical quantum communication.

References

[1] Scott Aaronson and Andris Ambainis. Quantum search of spatial regions. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '03, pages 200–, Washington, DC, USA, 2003. IEEE Computer Society.

[2] Erika Andersson, Marcos Curty, and Igor Jex. Experimentally realizable quantum comparison of coherent states and its applications. *Phys. Rev. A*, 74:022304, Aug 2006.

[3] Juan Miguel Arrazola and Norbert Lütkenhaus. Quantum communication with coherent states and linear optics. *Phys. Rev. A*, 90:042335, Oct 2014.

[4] Juan Miguel Arrazola and Norbert Lütkenhaus. Quantum fingerprinting with coherent states and a constant mean number of photons. *Phys. Rev. A*, 89:062305, Jun 2014.

[5] Juan Miguel Arrazola and Dave Touchette. Quantum advantage on information leakage for equality. *arXiv:quant-ph/1607.07516*, Jul 2016.

[6] Koenraad M R Audenaert. A sharp continuity estimate for the von neumann entropy. *Journal of Physics A: Mathematical and Theoretical*, 40(28):8127, 2007.

[7] Boaz Barak, Mark Braverman, Xi Chen, and Anup Rao. How to compress interactive communication. *SIAM Journal on Computing*, 42(3):1327–1363, 2013.

[8] Stephen M. Barnett, Anthony Chefles, and Igor Jex. Comparison of two unknown pure quantum states. *Physics Letters A*, 307(4):189 – 195, Feb 2003.

[9] M. Berta, M. Christandl, and D. Touchette. Smooth entropy bounds on one-shot quantum state redistribution. *IEEE Transactions on Information Theory*, 62(3):1425–1439, March 2016.

[10] M. Boyer, G. Brassard, P. Høyer, and A. Tapp. Tight Bounds on Quantum Searching. *Fortschritte der Physik*, 46:493–505, 1998.

[11] Mark Braverman. Coding for interactive computation: progress and challenges. In *Proceedings of the 50th Annual IEEE Allerton Conference on Communication, Control, and Computing*, pages 1914–1921. IEEE, 2012.

[12] Mark Braverman, Ankit Garg, Young Kun Ko, Jieming Mao, and Dave Touchette. Near-optimal bounds on bounded-round quantum communication complexity of disjointness. In *Proceedings of the 56th Annual IEEE Symposium on Foundations of Computer Science*, pages 773–791. IEEE, 2015.

[13] Harry Buhrman, Richard Cleve, Serge Massar, and Ronald de Wolf. Nonlocality and communication complexity. *Rev. Mod. Phys.*, 82:665–698, Mar 2010.

[14] Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum fingerprinting. *Phys. Rev. Lett.*, 87:167902, Sep 2001.

[15] Harry Buhrman, Richard Cleve, and Avi Wigderson. Quantum vs. classical communication and computation. In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC '98, pages 63–68, New York, NY, USA, 1998. ACM.

[16] C Campopiano and B Glazer. A coherent digital amplitude and phase modulation scheme. *IRE Transactions on Communications Systems*, 10(1):90–95, 1962.

[17] Anthony Chefles, Richard Jozsa, and Andreas Winter. On the existence of physical transformations between sets of quantum states. *International Journal of Quantum Information*, 02(01):11–21, 2004.

[18] Goong Chen, Stephen A Fulling, Hwang Lee, and Marlan O Scully. Grover's algorithm for multiobject search in quantum computing. *Lecture Notes in Physics-New York then Berlin-*, pages 165–175, 2001.

[19] Yuval Dagan, Yuval Filmus, Hamed Hatami, and Yaqiao Li. Trading information complexity for error. *arXiv:quant-ph/1611.06650*, 2016.

[20] Hélio MagÃlhaes De Oliveira and Gérard Battail. Sur les constellations en croix généralisées à deux dimensions et le canal secondaire opportunists. *Annales Des Télécommunications*, 47(5):202–213, May 1992.

[21] I. Devetak and P. W. Shor. The capacity of a quantum channel for simultaneous transmission of classical and quantum information. *Communications in Mathematical Physics*, 256(2):287–303, Jun 2005.

[22] Igor Devetak. The private classical capacity and quantum capacity of a quantum channel. *IEEE Transactions on Information Theory*, 51(1):44–55, 2005.

[23] Igor Devetak and Jon Yard. Exact cost of redistributing multipartite quantum states. *Phys. Rev. Lett.*, 100(23):230501, 2008.

[24] Ronald J. Evans and J. Boersma. The entropy of a poisson distribution (c. robert appledorn). *SIAM Review*, 30(2):314–317, 1988.

[25] M. Fannes. A continuity property of the entropy density for spin lattice systems. *Communications in Mathematical Physics*, 31(4):291–294, Dec 1973.

[26] G Folland. Harmonic analysis in phase space. *Annals of mathematical studies*, 122, 1986.

[27] Peter Frankl and Vojtěch Rödl. Forbidden intersections. *Transactions of the American Mathematical Society*, 300(1):259–286, 1987.

[28] C. Gerry and P. Knight. *Introductory Quantum Optics*. Cambridge University Press, 2005.

[29] E. N. Gilbert. A comparison of signalling alphabets. *Bell System Technical Journal*, 31(3):504–522, 1952.

[30] F. Gray. Pulse code communication, March 17 1953. US Patent 2,632,058.

[31] Lov K Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, pages 212–219. ACM, 1996.

[32] Jian-Yu Guan, Feihu Xu, Hua-Lei Yin, Yuan Li, Wei-Jun Zhang, Si-Jing Chen, Xiao-Yan Yang, Li Li, Li-Xing You, Teng-Yun Chen, Zhen Wang, Qiang Zhang, and Jian-Wei Pan. Observation of quantum fingerprinting beating the classical limit. *Phys. Rev. Lett.*, 116:240502, Jun 2016.

[33] Brian Hall. *Lie groups, Lie algebras, and representations: an elementary introduction*, volume 222. Springer, 2015.

[34] Brian C Hall. *Quantum theory for mathematicians*, volume 267. Springer, 2013.

[35] A. S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Transactions on Information Theory*, 44(1):269–273, September 2006.

[36] Alexander Semenovich Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii*, 9(3):3–11, 1973.

[37] Alfred Horn. Doubly stochastic matrices and the diagonal of a rotation matrix. *American Journal of Mathematics*, 76(3):620–630, 1954.

[38] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A lower bound for the bounded round quantum communication complexity of Set Disjointness. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, pages 220–229, 2003.

[39] Bala Kalyanasundaram and Georg Schintger. The probabilistic communication complexity of set intersection. *SIAM J. Discret. Math.*, 5(4):545–557, November 1992.

[40] Matthias Kleinmann, Hermann Kampermann, and Dagmar Bruss. Generalization of quantum-state comparison. *Phys. Rev. A*, 72(3):032308, 2005.

[41] Niraj Kumar, Eleni Diamanti, and Iordanis Kerenidis. Efficient quantum communications with coherent state fingerprints over multiple channels. *Phys. Rev. A*, 95:032337, Mar 2017.

[42] Mathieu Lauriere and Dave Touchette. The flow of information in interactive quantum protocols: the cost of forgetting. *arXiv:quant-ph/1701.02062*, 2017.

[43] Felix Leditzky, Mark M. Wilde, and Nilanjana Datta. Strong converse theorems using Rényi entropies. *Journal of Mathematical Physics*, 57(8):082202, 2016.

[44] Seth Lloyd. Capacity of the noisy quantum channel. *Phys. Rev. A*, 55:1613–1622, Mar 1997.

[45] Norbert Lütkenhaus. Probabilistic quantum computation and linear optical realizations. In Dagmar Bruß and Gerd Leuchs, editors, *Lectures on quantum information*. John Wiley & Sons, Inc., Hoboken, NJ, USA., 2007.

[46] C. Morgan and A. Winter. Pretty strong converse for the quantum capacity of degradable channels. *IEEE Transactions on Information Theory*, 60(1):317–333, Jan 2014.

[47] Michael A Nielsen. An introduction to majorization and its applications to quantum mechanics. *Lecture Notes, Department of Physics, University of Queensland, Australia*, 2002.

[48] 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.

[49] T. Ogawa and H. Nagaoka. Strong converse to the quantum channel coding theorem. *IEEE Transactions on Information Theory*, 45(7):2486–2489, Nov 1999.

[50] Asher Peres. How to differentiate between non-orthogonal states. *Physics Letters A*, 128(1):19, 1988.

[51] Lin Qing. Linear optical realization of unambiguous quantum state comparison. *Chinese Physics B*, 18(1):51, 2009.

[52] A A Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya: Mathematics*, 67(1):145, 2003.

[53] A.A. Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385 – 390, 1992.

[54] Michael Reck, Anton Zeilinger, Herbert J. Bernstein, and Philip Bertani. Experimental realization of any discrete unitary operator. *Phys. Rev. Lett.*, 73:58–61, Jul 1994.

[55] Stefan Scheel, Kae Nemoto, William J Munro, and Peter L Knight. Measurement-induced nonlinearity in linear optics. *Physical Review A*, 68(3):032310, 2003.

[56] Benjamin Schumacher. Quantum coding. *Physical Review A*, 51(4):2738, 1995.

[57] Benjamin Schumacher and Michael D Westmoreland. Sending classical information via noisy quantum channels. *Physical Review A*, 56(1):131, 1997.

[58] I. Schur. Über eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie. *Sitzungsberichte der Berliner Mathematischen Gesellschaft*, 22:9–20, 1923.

[59] Claude E Shannon. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55, 2001.

[60] Peter Shor. The quantum channel capacity and coherent information. *Lecture Notes, MSRI Workshop on Quantum Computation*, 2002.

[61] Peter W. Shor. Additivity of the classical capacity of entanglement-breaking quantum channels. *Journal of Mathematical Physics*, 43(9):4334–4340, 2002.

[62] ‘The Creative Act’, Duchamp’s lecture in Houston, April 1957; as quoted in *Art News*, 56. no. 4, Summer 1957, p. 28 –29.

[63] Dave Touchette. Quantum information complexity and amortized communication. *arXiv:quant-ph/1404.3733*, 2014.

[64] Dave Touchette. Quantum information complexity. In *Proceedings of the Forty-seventh Annual ACM Symposium on Theory of Computing*, STOC '15, pages 317–326, New York, NY, USA, 2015. ACM.

[65] Jacobus Hendricus Van Lint. *Introduction to coding theory*, volume 86. Springer Science & Business Media, 2012.

[66] R. R. Varshamov. Estimate of the number of signals in error correcting codes. *Dokl. Akad. Nauk SSSR*, 117:739–741, 1957.

[67] John Watrous. *Theory of Quantum Information*. 2017. Manuscript of a book, available at `https://cs.uwaterloo.ca/~watrous/`.

[68] Feihu Xu, Juan Miguel Arrazola, Kejin Wei, Wenyuan Wang, Pablo Palacios-Avila, Chen Feng, Shihan Sajeed, Norbert Lütkenhaus, and Hoi-Kwong Lo. Experimental quantum fingerprinting with weak coherent pulses. *Nature communications*, 6, 2015.

[69] Jon T Yard and Igor Devetak. Optimal quantum source coding with quantum side information at the encoder and decoder. *IEEE Transactions on Information Theory*, 55(11):5339–5351, 2009.

APPENDICES

Appendix A

Quantum fingerprinting error analysis

In Appendix A.0.1 we calculate the worst case error probability of the interpolation for each block-size $k = 1, \ldots, m$, which we use in Section 4.3.1 to show that the interpolation has information leakage $\mathcal{O}(\log n)$.

In Appendix A.0.2 we calculate the worst case error probability of the optical ring protocols described in Section 4.2.2, which we use in Section 4.3.2 and Appendix B to bound the information leakage of these protocols. We also prove optimality of this protocol over several similar protocols which use coherent state signals arranged in a ring in phase space, and state without proof analogous results for the lattice protocols.

In Appendix A.0.3 we determine the optimal one-sided error measurement for any simultaneous message passing model equality protocol, and show that the worst case error probability is lower bounded by the square of the error probability of the beamsplitter measurement.

For most protocols that we consider, each signal encodes several bits of $\mathrm{E}(x) \in \{0,1\}^m$. For different inputs $x \neq y \in \{0,1\}^n$, the error probability depends on the distribution of the bit differences between the codewords across the signals. For a given code, the worst case distribution over the particular 2^n codewords could be difficult to calculate. Instead, for most arguments we make the following simplifying assumption.

Remark 36. *In calculating the worst case error probability of any protocol, we assume that the code is uncharacterized apart from its minimum distance, and take the worst case over all strings* $\mathrm{E}_A \neq \mathrm{E}_B$ *in the output space of the code which differ by at least the minimum distance.*

A.0.1 Interpolation error probability

Here we calculate the worst case error probability of the interpolation protocol with block-size k, and then show that under the choice $|\langle q_0^{(k)}, q_1^{(k)}\rangle| = 1 - k/m$ the error probability is upper bounded by $2^{-\delta r}$, for repetition number r denoting the number of identical copies of $|\psi_x^{(k)}\rangle$ sent. This implies that for a given desired error probability, r can be fixed independent of n. We use this fact in the proof of Proposition 34 that the information leakage of the interpolation for fixed error probability is $\mathcal{O}(\log n)$.

As mentioned in Remark 36, the worst case is taken over all strings $\mathrm{E}_A \neq \mathrm{E}_B \in \{0,1\}^m$ which differ by minimum distance δm bits. To determine the worst case over this set, we first calculate the probability that the referee's measurement on the j-th pair of signals of $|\psi_{xy}^{(k)}\rangle$ returns "dark-port detection" or "anti-symmetric" under the assumption that $\mathrm{E}_{A,i} \neq \mathrm{E}_{B,i}$ for d indices $i \in \mathrm{I}[j,k]$. Recall that a direct product of beamsplitter measurements are performed on the first component of (4.7), and the controlled-swap is performed on the second component. Recall (4.2) that on input $|q_0^{(k)}\rangle\,|q_1^{(k)}\rangle$, the beamsplitter measurement outputs "dark-port detection" with probability $p_k := 1 - |\langle q_0^{(k)}, q_1^{(k)}\rangle|$. Thus, when d bits differ the direct product of beamsplitter measurements on the first component of (4.7) outputs "dark-port detection" ("D") with probability

$$\Pr_k^{\mathrm{I}}(\text{"D"}|d \text{ bits differ}) = \frac{d p_k}{k^2}. \tag{A.1}$$

On input $|\psi\rangle \in \mathcal{A} \otimes \mathcal{B}$, the controlled-swap outputs "anti-symmetric" with probability $\|\frac{1}{2}(\mathbb{1} - W)\,|\psi\rangle\,\|^2$. It follows that the controlled-swap on the second component of (4.7) outputs "anti-symmetric" with probability

$$\begin{aligned}\Pr_k^{\mathrm{I}}(\text{"A-S"}|d) &= \frac{1}{4k^2} \sum_{\substack{i,l \in \mathrm{I}[j,k] \\ i \neq l}} \left\| |q_{\mathrm{E}(x)_i}^{(k)}\rangle\, |q_{\mathrm{E}(y)_l}^{(k)}\rangle - |q_{\mathrm{E}(y)_i}^{(k)}\rangle\, |q_{\mathrm{E}(x)_l}^{(k)}\rangle \right\|^2 \\ &= \frac{1}{2k^2}\Big(k(k-1) - (k-d)(k-d-1) \\ &\qquad - 2d(k-d)(1-p_k) - d(d-1)(1-p_k)^2\Big).\end{aligned}$$

Let "no detection" ("ND") denote the event that neither "dark-port detection" nor "anti-symmetric" occur for a given signal. After simplification, the probability $\Pr_k^{\mathrm{I}}(\text{"ND"}|d)$ of outcome "no detection" when the two k-bit blocks differ by d bits is given by

$$\begin{aligned}\Pr_k^{\mathrm{I}}(\text{"ND"}|d) &= 1 - \Pr_k^{\mathrm{I}}(\text{"D"}|d) - \Pr_k^{\mathrm{I}}(\text{"A-S"}|d) \\ &= 1 - \frac{d}{k} p_k \left(1 - \frac{(d-1)p_k}{2k}\right),\end{aligned} \tag{A.2}$$

so for a given distribution of bit differences $d_1,\ldots,d_{m/k}$ beetween two codewords among the m/k blocks, the total error probability is given by

$$\Pr_k^{\mathrm{I}}(\mathrm{Err}|d_1,\ldots,d_{m/k}) = \prod_{i=1}^{m/k}\Pr_k^{\mathrm{I}}(\text{“ND”}|d_i), \tag{A.3}$$

which reproduces (4.5) for $p_k = 1$ and $k = m$ for the worst case $d = \delta m$ as expected.

Now we prove that the worst case error probability for each block-size k occurs when all bit differences between the codewords are consolidated into the fewest number of signals. Indeed, by straightforward calculation it can be shown that for all $0 \le p_k \le 1$, for every pair of integers $1 \le d, c \le k-1$ such that $d \ge c$,

$$\Pr_k^{\mathrm{I}}(d)\Pr_k^{\mathrm{I}}(c) \le \Pr_k^{\mathrm{I}}(d+1)\Pr_k^{\mathrm{I}}(c-1),$$

which proves the claim. Thus, the worst case error probability is given by

$$\Pr_k^{\mathrm{I}}(\mathrm{Err}) = \left[1 - p_k\left(1-\frac{(k-1)p_k}{2k}\right)\right]^{\lfloor\frac{\delta m}{k}\rfloor r}\left[1-p_k\frac{t}{k}\left(1-\frac{(t-1)p_k}{2k}\right)\right]^r \tag{A.4}$$

for t the remainder of $\frac{\delta m}{k}$ given by $\frac{\delta m}{k} = \left\lfloor\frac{\delta m}{k}\right\rfloor + \frac{t}{k}$.

In the remainder of this section, we show that under the choice $p_k = k/m$ the worst case error proability is upper bounded by a constant to the power r. This implies that any fixed error $\varepsilon > 0$ can be attained with fixed r. We use this fact in Proposition 34 to show that the information leakage of the interpolation under this choice is $\mathcal{O}(\log n)$.

Under the choice $p_k = k/m$, the worst case error probability is upper bounded by $2^{-\delta r}$:

$$\begin{aligned}
\Pr_k^{\mathrm{I}}(\mathrm{Err}|p_k = k/m) &= \left[1-\frac{k}{m}\left(1-\frac{(k-1)}{2m}\right)\right]^{\lfloor\frac{\delta m}{k}\rfloor r}\left[1-\frac{t}{m}\left(1-\frac{(t-1)}{2m}\right)\right]^r \\
&\le \left[1-\frac{k}{m}\left(1-\frac{k}{2m}\right)\right]^{\lfloor\frac{\delta m}{k}\rfloor r}\left[1-\frac{t}{m}\left(1-\frac{t}{2m}\right)\right]^r \\
&\le \left[1-\frac{k}{m}\left(1-\frac{k}{2m}\right)\right]^{\frac{\delta m}{k} r} \\
&\le 2^{-\delta r}
\end{aligned} \tag{A.5}$$

for all $1 \le k \le m$. The equality follows from simplification of (A.4). The first and third inequalities are straightforward, and the second follows from the fact that the function $[1-x(1-\frac{x}{2})]^{1/x}$ is strictly increasing with x for all $0 < x < 1$.

A.0.2 Error analysis of optical protocols

Here we derive the worst case error probabilities of the optical ring protocols described in Section 4.2.2. We also prove that for all $k = 1, \ldots, 4$ the ring Gray encoding is an optimal encoding of k-bit blocks of binary codewords into coherent states arranged in a ring in phase space. We then show that the ring Gray encoding outperforms an alternative which uses $q-$ary codewords. We state without proof analogous statements for the optical lattice protocols.

Lemma 37. *For any positive integer k, the ideal ring protocol described in Section 4.2.2 using the Gray encoding, states of total mean photon number μ_k, and the beamsplitter measurement described in Section 4.1.1, satisfies the following.*

The worst case error probability $\mathrm{Pr}_k^{\mathrm{O}}(\mathrm{Err})$ *satisfies*

$$\begin{aligned}\mathrm{Pr}_k^{\mathrm{O}}(\mathrm{Err}) \leq \exp\Bigg[-\mu_k\Bigg[&(1-(\delta k - \lfloor \delta k \rfloor))\left(1-\cos\left(\frac{2\pi\lfloor \delta k \rfloor}{2^k}\right)\right) \\ &+(\delta k - \lfloor \delta k \rfloor)\left(1-\cos\left(\frac{2\pi(\lfloor \delta k \rfloor+1)}{2^k}\right)\right)\Bigg]\Bigg]. \qquad \text{(A.6)}\end{aligned}$$

Furthermore, if the worst case error probability is taken over all codewords $\mathrm{E}_A \neq \mathrm{E}_B \in \{0,1\}^m$ which differ by at least the minimum distance δm bits (as described in Remark 36), this bound is attained with equality for all $0 \leq \delta \leq 3/k$.

As a corollary, since every binary code with more than two codewords has minimum distance $\delta \leq 1/2$, then for all $k = 1, \ldots, 6$, the worst case error probability $\mathrm{Pr}_k^{\mathrm{O}}(\mathrm{Err})$ is given by (A.6) with equality in the ideal setting, and can be used to derive equation (4.13) for the approximate error probability in the experimental setting.

Proof. The error probability of the beamsplitter measurement is given by the probability that "no dark-port detection" ("ND") occurs for every pair of signals. The worst case error probability depends on the worst case distribution of the bit differences between the codewords across the signals. To upper bound the worst case error probability, we first bound the probability $\mathrm{Pr}_k^{\mathrm{O}}(\text{“ND”}|d)$ of "ND" when a pair of k-bit blocks differ by d bits. As a property of the Gray encoding, nearest-neighbour coherent state signals in phase space correspond to blocks which differ by one bit. Thus, for any pair of blocks which differ by d bits, the corresponding pair of coherent state signals must be spaced at least d steps apart on the ring. It follows that $\mathrm{Pr}_k^{\mathrm{O}}(\text{“ND”}|d)$ is upper-bounded by the probability of "ND" when the pair of coherent state signals are spaced d steps apart on the ring, which is given by

$$\mathrm{Pr}_k^{\mathrm{O}}(\text{“ND”}|d) \leq \exp\left[\frac{-\mu_k}{m/k}\left(1-\cos\left(\frac{2\pi d}{2^k}\right)\right)\right], \qquad \text{(A.7)}$$

which follows from straightforward calculation using equation 4.2, along with the fact that the ring has amplitude $\beta_k = \mu_k/(m/k)$ (a necessary condition for total mean photon number μ_k). For $d = 1,2,3$ under the Gray code, (A.7) is satisfied with equality (see Figure 4.1 or [30]).

Now we upper bound the worst case error probability using the bound (A.7) for each signal. For a given distribution of bit differences $d_1,\ldots,d_{m/k}$ among the blocks, the error probability is given by

$$\begin{aligned}\Pr_k^{\mathrm{O}}(\mathrm{Err}|d_1,\ldots,d_{m/k}) &= \prod_{i=1}^{m/k}\Pr_k^{\mathrm{O}}(\text{“ND”}|d_i)\\ &\leq \exp\left[\frac{-\mu}{m/k}\left(m/k-\sum_{i=1}^{m/k}\cos\left(\frac{2\pi d_i}{2^k}\right)\right)\right]. \qquad \text{(A.8)}\end{aligned}$$

Note that

$$\cos\left(\frac{2\pi d}{2^k}\right)+\cos\left(\frac{2\pi c}{2^k}\right)\leq\cos\left(\frac{2\pi}{2^k}\left\lfloor\frac{d+c}{2}\right\rfloor\right)+\cos\left(\frac{2\pi}{2^k}\left\lceil\frac{d+c}{2}\right\rceil\right)$$

for all $1 \leq d,c \leq k$. This can be proven by direct calculation for $k = 3$; for $k \geq 4$ it follows from the fact that the function $\cos(2\pi\nu)+\cos(2\pi(a-\nu))$ is a strictly decreasing function of ν whenever $0 \leq \nu < a-\nu < \frac{1}{4}$, along with the fact that $\frac{k}{2^k} \leq \frac{1}{4}$.

It follows that the righthand side of (A.8) is maximized for strings $\mathrm{E}_A \neq \mathrm{E}_B \in \{0,1\}^m$ such that the bit differences $d_1,\ldots,d_{m/k}$ are evenly distributed among the m/k blocks and differ by minimum distance δm bits (this is the bound (A.6)).

If $\delta \leq 3/k$, then $\delta m \leq 3m/k$, so when the δm bit differences are evenly distributed among the m/k blocks, $d_i \leq 3$ for all $i = 1,\ldots,m/k$. As mentioned above, under the Gray code such inputs satisfy (A.7) with equality for every signal, and thus also satisfy (A.6) with equality. □

Now we prove statements of optimality for the Gray encoding in the optical ring protocols. Analogous results also hold for the lattice protocols.

Proposition 38. *In all protocols considered below, assume the referee uses the beamsplitter measurement described in Section 4.1.1, and that the worst case error probability is taken over all codewords $\mathrm{E}_A \neq \mathrm{E}_B \in \{0,1\}^m$ which differ by at least the minimum distance δm bits, as described in Remark 36.*

For any positive integer k, the ideal ring protocol described in Section 4.2.2 using total mean photon number μ_k satisfies the following: For all $0 \leq \delta \leq 2/k$, the Gray encoding minimizes the worst case error probability over all encodings of k-bit blocks of binary codewords into a ring of equally spaced coherent state signals in phase space.

As a corollary, since every binary code with more than two codewords has minimum distance $\delta \leq 1/2$, the Gray encoding is optimal for all $k = 1, \ldots, 4$.

Proof. We prove optimality of any encoding in which nearest neighbour coherent state signals differ by one bit (this set of encodings includes the Gray encoding). First we show that for any such encoding, the worst case error probability is given by (A.6) with equality. For any such encoding, any pair of coherent state signals which are second-nearest neighbours correspond to a pair of blocks which differ by two bits, and thus satisfy (A.7) with equality. Thus, by the same arguments used to prove Lemma 37, the worst case error probability for all $\delta \leq 2/k$ under any such encoding is given by (A.6) with equality.

Now we show that any encoding for which there exist nearest-neighbour coherent state signals which differ by $d' \geq 2$ bits has worst case error probability greater than (A.6). The inequalities $\delta \leq 2/k$ and $d' \geq 2$ imply $\delta m \leq d'm/k$. Thus, there exist codewords which differ in δm bits and for which the bit differences $d_1, \ldots, d_{m/k}$ are distributed among the m/k blocks such that $d_i = d'$ for $\delta m/d'$ indices i, and all other bit differences are zero. Furthermore, the codewords can be chosen so that for every index i satisfying $d_i = d'$, the corresponding pair of coherent state signals are nearest neighbours in phase space. Such codewords clearly give rise to error probability greater than (A.6). Here we have assumed d' divides δm, but similar techniques can be used to prove the same statement in the case when d' does not divide δm. □

We have shown that under certain conditions, the Gray encoding is an optimal encoding of binary codewords into a ring. We now consider another family of optical ring protocols which map q-ary codewords into q equally-spaced nodes on a ring. Under the assumption that all codes saturate the Gilbert -Varshamov bound, we show that the Gray encoding of binary codewords outperforms this family for all q powers of two. An analogous result holds for the lattice protocols.

Proposition 39. *In all protocols considered below, assume all codes saturate the Gilbert-Varshamov bound. Furthermore, assume that the referee uses the beamsplitter measurement described in Section 4.1.1, and that the worst case error probability is taken over all codewords $\mathrm{E}_A \neq \mathrm{E}_B$ in the output space of the code which differ by at least the minimum distance of the code, as described in Remark 36.*

Let q be any power of two. For all $\varepsilon > 0$, for any q-ary ring protocol described above which attains error probability ε with total mean photon number μ_q and m_q signals, the $k = \log q$-bit ring protocol described in Section 4.2.2 with Gray encoding can attain the same error probability ε with the same total mean photon number μ_q and fewer than m_q signals.

As a corollary, because the $\log q$-bit ring protocol uses the same total mean photon number and fewer signals than the q-ary ring protocol, then it has lower information leakage under the bound derived in Appendix B. It can also be shown that this statement holds under the stronger bound derived in Section 4.3.2 for $n \gg \mu_q$ using standard approximation techniques.

Proof. Let m_q denote the length of the $q-$ary code (i.e. the number of "qits" in the code), and let $\delta_q m_q$ denote the minimum distance (i.e. the minimum number of differing qits between codewords). It is easy to see that the worst case error probability occurs when the codewords $\mathrm{E}_A \neq \mathrm{E}_B \in [q]^{m_q}$ differ by $\delta_q m_q$ qits, and every pair of differing qits correspond to nearest-neighbour coherent state signals; and is given by

$$\Pr_q(\mathrm{Err}) = \exp\left[-\mu_q \delta_q \left(1 - \cos\left(\frac{2\pi}{q}\right)\right)\right]. \tag{A.9}$$

As the $q-$ary code saturates the Gilbert-Varshamov bound, the quantity δ_q satisfies

$$\frac{n}{m_q} = \log q - \delta_q \log(q-1) - \mathrm{h}(\delta_q) \tag{A.10}$$

and $0 \leq \delta_q < 1 - 1/q$ [65]. Note that (A.10) is the ratio of n to the number of signals contained in each state.

Now consider the $k = \log q$-bit ring protocol using the Gray encoding, the same total mean photon number, the beamsplitter measurement, and minimum distance

$$\delta = \frac{\delta_q}{\log q} \leq \frac{1 - \frac{1}{q}}{\log q} \leq \frac{2}{\log q}. \tag{A.11}$$

The first inequality follows from $\delta_q < 1 - 1/q$ and the second is straightforward. By Lemma 37, the inequality (A.11) implies that the worst case error probability is given by (A.9) with equality. As the code saturates the Gilbert-Varshamov bound, the ratio of n to the number signals is given by

$$\frac{n}{m/\log q} = (1 - \mathrm{h}(\delta_q/\log q)) \log q. \tag{A.12}$$

For all $0 \leq \delta_q < 1 - 1/q$, we now show that the righthand side of (A.10) is no greater than the righthand side of (A.12), which implies that this protocol sends fewer signals than the q-ary ring protocol, completing the proof.

After substituting $\log q = k$ and $\delta_q = \delta k$, the desired inequality becomes

$$\frac{\mathrm{h}(\delta)}{\delta} - \frac{\mathrm{h}(\delta k)}{\delta k} \leq \log(2^k - 1) \tag{A.13}$$

for all $0 \leq \delta \leq (1 - 2^{-k})/k$. Using standard calculus techniques, it can be shown that the lefthand side of (A.13) is a strictly increasing function of δ. Thus, the inequality need only be shown for $\delta = (1 - 2^{-k})/k$, which is proven using standard techniques. □

A.0.3 Optimal measurement

Here we derive the optimal one-sided error (i.e. zero error for equal inputs) measurement for any simultaneous message passing model equality protocol, and show that the worst case error probability is lower bounded by the square of the error probability of the beamsplitter measurement.

Consider a general setting in which the referee receives a state ρ_z^{AB} for $z \in \mathrm{Z}$ chosen according to some probability distribution $P \in \Pr(\mathrm{Z})$. For some partition $\mathrm{Z} = \mathrm{EQ} \sqcup \mathrm{NEQ}$, the referee wishes to determine whether z is contained EQ or NEQ under the constraint that if $z \in \mathrm{EQ}$ she never errs. Assuming $p(z) > 0$ for all $z \in \mathrm{EQ}$, it is straightforward to show that the measurement $\{\Pi_{\mathrm{EQ}}, \mathbb{1} - \Pi_{\mathrm{EQ}}\}$ (the projection onto the space spanned by the image of the states $\rho_{z\in\mathrm{EQ}}^{AB}$ and its orthogonal complement) minimizes the worst case error probability (and the average error probability) of this task.

Proposition 40. *In the setting described above, under the definitions*

$$\begin{aligned} \mathrm{Z} &= \{(x,y) : x,y \in \{0,1\}^n\} \\ \mathrm{EQ} &= \{(x,x) : x \in \{0,1\}^n\} \\ \mathrm{NEQ} &= \{(x,y) : x \neq y \in \{0,1\}^n\}, \end{aligned} \tag{A.14}$$

if

$$\rho_{(x,y)}^{AB} = |\psi_x\rangle\langle\psi_x| \otimes |\psi_y\rangle\langle\psi_y| \tag{A.15}$$

for all $x,y \in \{0,1\}^n$, then the error probability on input $x \neq y \in \{0,1\}^n$ is lower bounded by $\left|\left\langle \psi_x, \psi_y \right\rangle\right|^2$.

As a corollary, since the error probability of the beamsplitter measurement in the optical protocols is given by $|\left\langle \psi_x, \psi_y \right\rangle|$ (see (4.2) and the subsequent discussion), then the error probability of the optimal measurement is lower bounded by the square of the error probability of the beamsplitter measurement.

Proof. The error probability is lower bounded by

$$\left\langle \psi_x\psi_y \middle| \Pi_{\mathrm{EQ}} \middle| \psi_x\psi_y \right\rangle \geq \frac{2\left|\left\langle \psi_x, \psi_y \right\rangle\right|^2}{1 + \left|\left\langle \psi_x, \psi_y \right\rangle\right|^2} \geq \left|\left\langle \psi_x, \psi_y \right\rangle\right|^2.$$

The second inequality is straightforward and the first is derived by considering only the first two terms of the decomposition

$$\Pi_{\mathrm{EQ}} = |\psi_x\psi_x\rangle\langle\psi_x\psi_x| + |\phi\rangle\langle\phi| + \ldots, \tag{A.16}$$

where $|\phi\rangle$ is the normalized component of $\left|\psi_y\psi_y\right\rangle$ orthogonal to $|\psi_x\psi_x\rangle$. □

Appendix B

Information leakage of optical quantum fingerprinting protocols

Here we bound the information leakage of any pure state simultaneous message passing model protocol in which every state $\left|\psi_{xy}\right\rangle$ is a tensor product of m_k coherent states with total mean photon number lying in a fixed range $[\mu_{\min}, \mu_{\max}]$ for all $x \in \mathrm{X}, y \in \mathrm{Y}$.

In Appendix B.0.1 we give a practical bound on the information leakage using a continuity bound on entropy. Due to the dimension dependence of continuity bounds, this bound does not give the desired $\mathcal{O}(\log m_k)$ limiting behaviour, but has the advantage of being straightforward to calculate in practice. In Appendix B.0.2 we bound the asymptotic behaviour as $\mathcal{O}(\log m_k)$.

B.0.1 Practical information leakage bound

Here we use an extension of Theorem 1 of [4] and a continuity bound on entropy to bound the information leakage of any simultaneous message passing model protocol satisfying the conditions outlined above.

Recall that for pure state protocols the information leakage is equal to the entropy of ρ_P^{AB}, maximized over prior distributions $P \in \Pr(\mathrm{X} \times \mathrm{Y})$. We use the Fannes-Audenart continuity bound, which bounds $\mathrm{H}(\rho_P^{AB})$ in terms of $\mathrm{H}(\sigma)$, $\left\|\rho_P^{AB} - \sigma\right\|_1$, and $|\mathrm{A} \times \mathrm{B}|$ for any state σ [25][6]. We choose $\sigma = \Pi_0 \rho_P^{AB} \Pi_0 / \mathrm{Tr}(\Pi_0 \rho_P^{AB})$, where Π_0 projects onto a "typical subspace" of ρ_P^{AB}, given by the set of Fock states of total photon number lying within some radius $\Delta \in \mathbb{N}$ of the interval $[\mu_{\min}, \mu_{\max}]$, as in [4]. By straightforward extension of Theorem 1 of [4],

$$\left\langle \psi_{xy} \,\middle|\, \Pi_0 \,\middle|\, \psi_{xy} \right\rangle \geq 1 - \max\left\{0, e^{-\mu_{\min}} \left(\frac{e\mu_{\min}}{\mu_{\min}-\Delta}\right)^{\mu_{\min}-\Delta}\right\} \\ - e^{-\mu_{\max}} \left(\frac{e\mu_{\max}}{\mu_{\max}+\Delta}\right)^{\mu_{\max}+\Delta}, \tag{B.1}$$

and

$$\log\dim(\Pi_0) \leq (\mu_{\max}+\Delta)\log(\mu_{\max}+\Delta+m_k-1) \\ +\log(\mu_{\max}-\mu_{\min}+2\Delta+1). \tag{B.2}$$

We choose any $\varepsilon' > 0$ (which can be optimized over), and fix Δ such that $\left\langle \psi_{xy} \,\middle|\, \Pi_0 \,\middle|\, \psi_{xy} \right\rangle \geq 1-\varepsilon'$. We now bound the quantities $\mathrm{H}(\sigma)$, $\left\|\rho_P^{AB}-\sigma\right\|_1$, and $|\mathrm{X}\times\mathrm{Y}|$ for a given choice of ε'. First, by the dimension bound on entropy, $\mathrm{H}(\sigma)$ is upper bounded by (B.2). Second,

$$\begin{aligned} &\left\|\rho_P^{AB}-\sigma\right\|_1 \\ &\leq \sum_{x\in\mathrm{X},y\in\mathrm{Y}} P(x,y)\left\| |\psi_{xy}\rangle\langle\psi_{xy}| - \frac{\Pi_0|\psi_{xy}\rangle\langle\psi_{xy}|\Pi_0}{\mathrm{Tr}(\Pi_0\rho^{AB})}\right\|_1 \\ &\leq 2\sum_{x\in\mathrm{X},y\in\mathrm{Y}} P(x,y)\sqrt{1-\left|\left\langle \psi_{xy} \,\middle|\, \Pi_0 \,\middle|\, \psi_{xy} \right\rangle\right|^2} \\ &\leq 2\sqrt{2\varepsilon'} \end{aligned} \tag{B.3}$$

where the first inequality is the triangle inequality, the second is the Fuchs-van de Graaf inequality along with $\mathrm{Tr}(\Pi_0\rho_P^{AB}) \leq 1$, and the third is (B.1). Third, as there are $|\mathrm{X}\times\mathrm{Y}|$ states $|\psi_{xy}\rangle$, they span at most a $|\mathrm{A}\times\mathrm{B}|$-dimensional space. Combining these bounds, for $2\sqrt{2\varepsilon'} \leq 1/2$ the Fannes-Audenart continuity bound gives

$$\mathrm{H}(\rho_P^{AB}) \leq \log\dim(\Pi_0) + \log(|\mathrm{X}\times\mathrm{Y}|)\sqrt{2\varepsilon'} + \mathrm{h}(\sqrt{2\varepsilon'}). \tag{B.4}$$

The quantity (B.4) is independent of the distribution P, and thus upper bounds the information leakage. Although this bound is easily calculable in practice, it is linear in $\log(|\mathrm{X}\times\mathrm{Y}|)$. For the optical quantum fingerprinting protocols, $\log(|\mathrm{X}\times\mathrm{Y}|) = 2n = \mathcal{O}(m_k)$, so this bound does not give the desired $\mathcal{O}(\log m_k)$ asymptotic behaviour in this case.

B.0.2 Information leakage asymptotic analysis

Here we prove the $\mathcal{O}(\log m_k)$ asymptotic information leakage of any simultaneous message passing model protocol satisfying the conditions outlined above.

Recall that the information leakage of any pure state protocol is equal to the entropy of ρ_P^{AB}, maximized over prior distributions $P \in \mathrm{Pr}(\mathrm{X} \times \mathrm{Y})$. We bound $\mathrm{H}(\rho_P^{AB})$ as $\mathcal{O}(\log m_k)$ by projecting ρ_P^{AB} onto Fock states lying within telescoping neighbourhoods $\Delta_0, \Delta_1, \ldots$ of $[\mu_{\min}, \mu_{\max}]$.

In general, consider any projective measurement $\{\Pi_0, \Pi_1 \ldots, \}$ (with possibly infinitely many measurement operators), and define an isometry

$$V = \sum_{i=0}^{\infty} \Pi_i \otimes |i\rangle \otimes |i\rangle \in \mathcal{U}(AB, ABD_1D_2). \tag{B.5}$$

Then,

$$\begin{aligned} \mathrm{H}(AB)_{\rho_P} &= \mathrm{H}(ABD_1D_2)_{V\rho_P V^\dagger} \\ &= \mathrm{H}(D_1) + \mathrm{H}(AB|D_1D_2) \\ &\leq \mathrm{H}(D_1) + \mathrm{H}(AB|D_1) \\ &\leq \mathrm{H}(D_1) + \sum_{i=0}^{\infty} \mathrm{Pr}(D_1 = i) \log \dim(\Pi_i), \end{aligned} \tag{B.6}$$

where the first equality follows from the fact that isometries preserve entropy and the second equality follows from the chain rule and $\mathrm{H}(D_1D_2) = \mathrm{H}(D_1)$. The first inequality follows from strong subadditivity, and the second inequality follows from the dimension bound on quantum entropy.

For a fixed positive integer Δ, let

$$\begin{aligned} \Delta_0 &= \{N \geq 0 : \mu_{\min} - \Delta \leq N \leq \mu_{\max} + \Delta\} \\ \Delta_i &= \{N \geq 0 : i\Delta + 1 \leq \mu_{\min} - N \leq (i+1)\Delta \\ &\qquad \text{or} \quad i\Delta + 1 \leq N - \mu_{\max} \leq (i+1)\Delta\} \end{aligned} \tag{B.7}$$

for each $i = 0, 1, \ldots.$ Let Π_i be the projection onto the space of Fock states with total photon number lying in the set Δ_i. Then the set $\{\Pi_0, \Pi_1, \ldots\}$ forms a measurement.

We now show that $\mathrm{Pr}(D_1 = i)$ decreases exponentially with i and $\log \dim(\Pi_i)$ is $\mathcal{O}(\log m_k)$ for each i (with prefactors not growing too quickly with i) to bound (B.6) as $\mathcal{O}(\log m_k)$. Using similar techniques to those used to prove Theorem 1 of [4], it can be shown that

$$\mathrm{Pr}(D_1 = i) \leq e^{-\mu_{\max}} \left(\frac{e\mu_{\max}}{\mu_{\max} + i\Delta} \right)^{\mu_{\max} + i\Delta} \tag{B.8}$$

for all $i = 0, 1, \ldots$ under the simplifying assumption $\Delta > \mu_{\max}$, and

$$\begin{aligned}\log\dim(\Pi_i) \leq (\mu_{\max} + (i+1)\Delta)\log(\mu_{\max} + (i+1)\Delta + m_k - 1) \\ + \log(\Delta) \qquad \text{for all } i = 1, 2, \ldots\end{aligned} \tag{B.9}$$

It is straightforward to show that under these bounds the infinite sum appearing in the second term of (B.6) converges and is $\mathcal{O}(\log m_k)$. It is also straightforward to show that $\mathrm{H}(C_1)$ is no greater than the entropy of the $\mu_{\max}$ Poisson distribution, which is finite and is well-approximated by $\frac{1}{2}\log(2\pi e\mu_{\max})$ when $\mu_{\max} \gg 1$ [24]. Thus, the asymptotic information leakage is $\mathcal{O}(\log m_k)$.

Appendix C

Information cost for appointment scheduling protocols

C.1 Properties of information cost and entropic quantities

Before analyzing the information cost of the appointment scheduling protocols, we state some properties of the quantum information cost and other entropic quantities that we will use.

C.1.1 Properties of entropic quantities

We first introduce useful properties of entropic quantities that we will use. Proofs of many of these statements can be found in, e.g. [67].

Lemma 41 (Conditioning on a classical register is taking the average)**.** *If*

$$\rho = \rho^{ABCD} = \sum_c p(c)|c\rangle\langle c| \otimes \rho_c^{ABD}$$

is a classical-quantum state with classical register C, then

$$\mathrm{H}(A|CD)_\rho = \mathbb{E}_c\left[\mathrm{H}(A|D)_{\rho_c}\right],$$
$$\mathrm{I}(A:B|CD)_\rho = \mathbb{E}_c\left[\mathrm{I}(A:B|D)_{\rho_c}\right],$$

Lemma 42 (CQ dimension bound)**.** *If $\rho = \rho^{ABCD} = \sum_c p(c)|c\rangle\langle c| \otimes \rho_c^{ABD}$ is a classical-quantum state with classical register C, then*

$$\mathrm{I}(A:C|D)_\rho \leq \log\dim(\mathcal{C}).$$

Lemma 43 (Pure states have no entropy)**.** *If $\rho = \rho^{AB} = |\psi\rangle\langle\psi|^A \otimes \rho^B$ is pure on register A, then*

$$\mathrm{H}(A|B)_\rho = 0.$$

Lemma 44 (Strong subadditivity)**.** *For any quantum state ρ^{ABC},*

$$\mathrm{H}(A|BC)_\rho \leq \mathrm{H}(A|B)_\rho. \tag{C.1}$$

Lemma 45 (Isometric invariance)**.** *For any quantum state ρ^A and any isometry $V \in \mathcal{U}(A,B)$,*

$$\mathrm{H}(A)_\rho = \mathrm{H}(B)_{V\rho V^\dagger}. \tag{C.2}$$

Lemma 46 (Dimension bound)**.** *For any quantum state $\rho^A \in \mathcal{D}(\mathcal{A})$,*

$$\mathrm{H}(A) \leq \log \dim(\mathcal{A}). \tag{C.3}$$

We also use the following bound on the entropy of any rank-two state:

Lemma 47. *Consider any state $\rho \in \mathcal{D}(A)$ which can be written as a convex combination*

$$\rho = p|\psi\rangle\langle\psi| + (1-p)|\phi\rangle\langle\phi|$$

of two pure states $|\psi\rangle\langle\psi|$ and $|\phi\rangle\langle\phi|$. Let $F = |\langle\psi,\phi\rangle|$. Then,

$$\mathrm{H}(A)_\rho = \mathrm{H}\left(\frac{1}{2} - \frac{1}{2}\sqrt{1-4p(1-p)(1-F^2)}\right) \tag{C.4}$$

$$\leq \mathrm{H}\left(\frac{1}{2}(1-F)\right). \tag{C.5}$$

C.1.2 Properties of the information cost of safe interactive protocols on classical inputs

We also make use of the following properties of the information cost of safe interactive protocols on classical inputs, the proofs of which can be found in [12, 64, 63, 42]. For brevity, we implicitly assume Π is a safe interactive protocol on classical inputs for every statement in this section, and avoid writing down this assumption for each statement.

Lemma 48 (QIC: Concavity in input distribution)**.** *Let ν be a distribution over a set of input distributions on $\mathrm{X} \times \mathrm{Y}$ and denote $\bar{P} = \mathbb{E}_{P\sim\nu}[P]$. Then*

$$\mathbb{E}_{P\sim\nu}[\mathrm{QIC}(\Pi,P)] \leq \mathrm{QIC}(\Pi,\bar{P}). \tag{C.6}$$

Lemma 49 (QIC: Quasi-convexity in input distribution)**.** *For any $p \in [0,1]$ and any two input distributions P_1 and P_2 on $\mathrm{X} \times \mathrm{Y}$, let $P = pP_1 + (1-p)P_2$. The following then holds for any M-message protocol Π:*

$$\mathrm{QIC}(\Pi, P) \leq p\mathrm{QIC}(\Pi, P_1) + (1-p)\mathrm{QIC}(\Pi, P_2) + 2M\mathrm{H}(p). \tag{C.7}$$

Note that Lemma 49 gives the following bound for any distribution P with mass $w \leq 1/2$ on a particular element $x' \times y' \in \mathrm{X} \times \mathrm{Y}$: Let $P_1(x,y) = \frac{1}{1-w}P(x,y)$ for all $(x,y) \neq (x',y')$ and $P_1(x',y') = 0$; $P_2(x',y') = 1$, and $p = 1-w$. Then

$$\mathrm{QIC}(\Pi, P) \leq (1-w)\mathrm{QIC}(\Pi, P_1) + 2M\mathrm{H}(w), \tag{C.8}$$

which follows from $\mathrm{QIC}(\Pi, P_2) = 0$.

Lemma 50 (QIC: subadditivity)**.** *Let Π_1 be a protocol acting on input distributions on $\mathrm{X}_1 \times \mathrm{Y}_1$ and Π_2 be a protocol acting on input distributions on $\mathrm{X}_2 \times \mathrm{Y}_2$. For any input distribution $P_{12} \in \Pr(\mathrm{X}_1 \times \mathrm{Y}_1 \times \mathrm{X}_2 \times \mathrm{Y}_2)$,*

$$\mathrm{QIC}(\Pi_1 \otimes \Pi_2, P_{12}) \leq \mathrm{QIC}(\Pi_1, P_1) + \mathrm{QIC}(\Pi_2, P_2), \tag{C.9}$$

where P_1 and P_2 denote the marginals of P_{12} restricted to $\mathrm{X}_1 \times \mathrm{Y}_1$ and $\mathrm{X}_2 \times \mathrm{Y}_2$, respectively.

Information cost of protocols with side-information

Here we present known results on the information cost of safe interactive protocols on classical inputs with side-information. Side-information to a protocol is information contained in an additional register held by Alice and/or Bob which could affect their knowledge of eachother's inputs (i.e. it could be correlated with registers X and Y), and thus it could affect the information cost of the protocol.

These results will be useful, for example, in bounding the information cost of the two-bit AND protocol Π_A^j, $j = 1,2$, which runs protocol $\widetilde{\Pi}_A^j$ as a subroutine with both Alice and Bob storing side information about the outcome of each run, along with potentially any leftover quantum information from each run.

In general, say a protocol Π acts on the state of register XY. Additionally, assume Alice has side-information stored in register $\tilde{A}$, and Bob has side-information stored in register $\tilde{B}$. Then the quantum information cost of protocol Π in this case is equal to that defined in (C.1.2) with registers $\tilde{A}$ and $\tilde{B}$ added to Alice and Bob's memories, respectively:

$$\mathrm{QIC}(\Pi, P|\sigma) = \sum_i \mathrm{I}(C_i : X|YB_i\tilde{B})_{\rho_i} + \mathrm{I}(C_i : Y|XA_i\tilde{A})_{\rho_i}$$

in which $\sigma \in \mathcal{D}(XY\tilde{A}\tilde{B})$ is the combined state of Alice and Bob's inputs along with their side information, and satisfies $\sigma^{XY} = P$. Note that we have written $\mathrm{QIC}(\Pi, P|\sigma)$ to denote the information cost of protocol Π on input distribution $\sigma^{XY\tilde{A}\tilde{B}}$. Of course, the protocol Π still only acts on the reduced state P.

We will make use of the following inequality:

Lemma 51 (QIC: Increasing under discarding of side-information)**.**

$$\mathrm{QIC}(\Pi, P|\sigma) \leq \mathrm{QIC}(\Pi, P).$$

Now we consider the case in which $\sigma^{XY\tilde{A}\tilde{B}}$ takes the form

$$\sigma^{XY\tilde{A}\tilde{B}} = \sum_{o\in\mathrm{O}} Q(o) \sum_{x\in\mathrm{X}, y\in\mathrm{Y}} P_o(x,y)|xy\rangle\langle xy| \otimes |o\rangle\langle o|^{\tilde{A}} \otimes |o\rangle\langle o|^{\tilde{B}} \tag{C.10}$$

for some classical state set O and some set of probability distributions $P_o \in \Pr(\mathrm{X}\times\mathrm{Y})$. For us, this analysis will be helpful in evaluating the information cost of Π_A^j: for each iteration of the subroutine $\widetilde{\Pi}_A^j$, the classical state set O will represent the classical outcomes of the previous runs of the protocol (which both Alice and Bob posess at the end of each run).

For any such input σ, the information cost takes the following form, which is an easy corollary to the above expression for information cost with side information and Lemma 41 (Conditioning on classical register is taking the average).

Lemma 52 (QIC: Conditioning on classical register is taking the average)**.** *For any input distribution $\sigma^{XY\tilde{A}\tilde{B}}$ taking the form* (C.10) *for some classical state set* O*,*

$$\mathrm{QIC}(\Pi, P|\sigma) = \sum_{o\in\mathrm{O}} Q(o)\mathrm{QIC}(\Pi, P_o).$$

Note that, using this result, one can prove the special case of Lemma 51 for when σ takes the form (C.10) using Lemma 48 (concavity in input distribution) along with $\sum_{o\in\mathrm{O}} Q(o)P_o = P$.

Information cost of pure state protocols

Here we simplify the expression for the information cost of safe interactive protocols on classical input distributions P for those in which, conditioned on fixed inputs x, y, pure states $\left|\phi_i^{x,y}\right\rangle^C$ are exchanged. All appointment scheduling protocols we consider fall into this category.

In this case, $\mathrm{QIC}_i(\Pi, P)$ simplifies to

$$\begin{aligned}
\mathrm{QIC}_i(\Pi, P) &= \mathrm{I}(X : C_i|YB_i) + \mathrm{I}(Y : C_i|XA_i) \\
&= \mathrm{H}(C_i|YB_i) - \mathrm{H}(C_i|XYB_i) + \mathrm{H}(C_i|XA_i) - \mathrm{H}(C_i|XYA_i) \\
&= \mathrm{H}(C_i|YB_i) + \mathrm{H}(C_i|XA_i). \qquad \text{(C.11)}
\end{aligned}$$

The first and second equalities follow from the definitions. The third equality follows from $H(C_i|XYB_i) = H(C_i|XYA_i) = 0$, which follows from purity of the state of register C_i conditioned on fixed inputs (x, y).

C.2 Information cost of bitwise AND protocols

In this section we analyze the quantum information cost of the modified appointment scheduling protocol Π_D^l under experimental imperfections. We have used the index $l = 1, 2$ to indicate the choice of subroutine $\widetilde{\Pi}_A^l$. As the modified protocol converges to the ideal protocol for $p_d = 0, \eta = 1$, the following bound also applies in the ideal setting. First we bound the information cost of Π_D^l in terms of the information cost of protocol $\widetilde{\Pi}_A^l$ for $l = 1, 2$ and protocol parameters. In Section C.2.1 we bound the information cost of $\widetilde{\Pi}_A^l$ in terms of protocol parameters.

If $p_d = 0$, define $K = 1$, and if $p_d > 0$ (and $\varepsilon > 0$), define $K = \lceil \log(1/\varepsilon)/\log(1/p_s) \rceil$.

Lemma 53. *For all values of n and s satisfying $n \geq 4$ and $8\ln(n) \leq s \leq n$, the protocol $\Pi_D^l(\eta, p_d, \varepsilon)$ described in Section 5.3 satisfies*

$$\begin{aligned}
&\mathrm{QIC}(\Pi_D^l) \leq s + \log s + 1 \\
&\qquad + \frac{n}{1-p} \frac{1 - \left(\frac{p_z}{1-p}\right)^K}{1 - \frac{p_z}{1-p}} \left[\frac{2(2r+2)}{n} + \mathrm{QIC}_0(\widetilde{\Pi}_A^l) + 2(2r+2)\mathrm{H}\left(\frac{2\ln n}{s} + \frac{1}{n}\right) \right]
\end{aligned}$$

for $l = 1,2$, where

$$\mathrm{QIC}_0(\widetilde{\Pi}^l_A) := \sup_{P_0 \in \Pr(\{0,1\}^2):P_0(1,1)=0} \mathrm{QIC}(\widetilde{\Pi}^l_A, P_0),$$

$0/0 := 1$, *p is the probability of an inconclusive outcome, p_s is the probability that the second mode clicks given non-intersecting inputs, and p_z is the probability of outcome "Zero?" given an input distribution with zero mass on (1,1). These probabilities are given by*

$$\begin{aligned} p &= e^{-|\alpha|^2}(1-p_d)^2 + (1-e^{-|\alpha|^2} + e^{-|\alpha|^2}p_d)p_d \\ p_s &= e^{-|\alpha|^2}(1-p_d)p_d \\ p_z &= (1-e^{-|\alpha|^2} + e^{-|\alpha|^2}p_d)(1-p_d). \end{aligned}$$

To prove the statement, we first prove the following bound on the information cost of the subroutine bitwise-AND protocols.

Lemma 54. *Let P be an input distribution with weight $w \leq 1/2$ on $(1,1)$. The protocol $\Pi^l_A\left(\widetilde{\Pi}^l_A(\eta), p_d, \varepsilon\right)$ described in Section 5.3 satisfies*

$$\mathrm{QIC}(\Pi^l_A, P) \leq \frac{1}{1-p} \frac{1-\left(\frac{p_z}{1-p}\right)^K}{1-\frac{p_z}{1-p}} [\mathrm{QIC}_0(\widetilde{\Pi}^l_A) + 2(2r+2)\mathrm{H}(w)] \tag{C.12}$$

and

$$\mathrm{QIC}(\Pi^l_A, P) \leq \frac{1}{1-p} \frac{1-\left(\frac{p_z}{1-p}\right)^K}{1-\frac{p_z}{1-p}} 2(2r+2) \tag{C.13}$$

for $l = 1,2$.

Proof. We prove only the first inequality. The second can be proven similarly by replacing Lemma 49 (quasi-convexity in input distribution) with Lemma 42 (CQ dimension bound) below, and using the fact that $2r+2$ messages are communicated in $\widetilde{\Pi}^l_A$, each based solely on the two bits $x_i, y_i \in \{0,1\}$. We write $\Pi^l_{A,j}$ to denote the protocol Π^l_A described in Section 5.3 with K' initialized to j. Let $\mathrm{X}_i = \mathrm{Y}_i = \{0,1\}$, and let ν be any input distribution on $\mathrm{X}_i \times \mathrm{Y}_i$ with mass $w' \leq 1/2$ on $(1,1)$.

Protocol $\Pi^l_{A,j}$ on input distribution ν first runs protocol $\widetilde{\Pi}^l_A$. With probability p, $\widetilde{\Pi}^l_A$ is inconclusive and is run again. As p is independent of the state of X_iY_i, then ν conditioned on outcome "Inconclusive" is equal to ν. With some probability $\Pr(\text{“Zero”}|\nu)$, protocol $\Pi^l_{A,j+1}$ is run on input distribution ν_z, (ν conditioned on outcome "Zero"). Therefore, the information cost of $\Pi^l_{A,j}$ on input distribution ν is given by

$$\begin{aligned}\mathrm{QIC}(\Pi^l_{A,j},\nu) &= \mathrm{QIC}(\widetilde{\Pi}^l_A,\nu) + p\mathrm{QIC}(\Pi^l_{A,j},\nu|\sigma_I) + \Pr(\text{“Zero”}|\nu)\mathrm{QIC}(\Pi^l_{A,j+1},\nu_z|\sigma_z)\\ &\leq \mathrm{QIC}_0(\widetilde{\Pi}^l_A) + 2(2r+2)\mathrm{H}(w') + p\mathrm{QIC}(\Pi^l_A,\nu) + p_z\mathrm{QIC}(\Pi^l_{A,j+1},\nu_z),\end{aligned}$$

where σ_I and σ_Z are the states of $XY\tilde{A}\tilde{B}$ conditioned on outcomes "Inconclusive" and "Zero", respectively (the registers $\tilde{A}\tilde{B}$ contain all of Alice and Bob's leftover quantum information from this and previous steps of the protocol Π^l_D). The inequality follows from Lemma 51 (QIC: increasing under discarding of side information), the discussion subsequent to Lemma 49 (quasi-convexity in input distribution), and $\Pr(\text{“Zero”}|\nu) \leq p_z$ (recall p_z is the probability of "Zero?" given any input distribution with zero weight on (1,1), and thus upper bounds the probability of "Zero?" given any input distribution).

Thus,

$$\mathrm{QIC}(\Pi^l_{A,j},\nu) \leq \frac{1}{1-p}(\mathrm{QIC}_0(\widetilde{\Pi}^l_A) + 2(2r+2)\mathrm{H}(w') + p_z\mathrm{QIC}(\Pi^l_{A,j+1},\nu_z)) \tag{C.14}$$

for any $j = 0,1,\ldots,K-2$. For the case $j = K-1$, the protocol ends on any outcome other than "Inconclusive", so we have

$$\mathrm{QIC}(\Pi^l_{A,K-1},\nu) \leq \frac{1}{1-p}(\mathrm{QIC}_0(\widetilde{\Pi}^l_A) + 2(2r+2)\mathrm{H}(w')). \tag{C.15}$$

Note that outcome "Zero" cannot increase the mass w' of ν on $(1,1)$. Thus, the mass of ν on $(1,1)$ for each iteration of the above recursion is upper bounded by the original mass w. Under this bound, equations (C.14) and (C.15) define a geometric series with ratio $\frac{p_z}{1-p}$ and prefactor given by (C.15). The result follows. □

Note that the above bound on $\mathrm{QIC}(\Pi^l_A,\nu)$ strongly depends on the weight w on $(1,1)$. As mentioned previously, the classical subsampling component of Π^l_D serves to keep w small for each subroutine AND protocol.

To bound $\mathrm{QIC}(\Pi^l_D) = \max_P \mathrm{QIC}(\Pi^l_D,P)$, we first bound the information cost of the classical subsampling component of Π^l_D. Alice chooses s indices uniformly at random from the set $[n]$

and sends them to Bob. As these indices are independent of the state of XY, this component of the protocol has zero information cost. Next, Alice sends x_i to Bob for each $i \in S$. By Lemma 42 (CQ dimension bound), the information cost of this component is upper bounded by s. Then, if Bob finds that $x_i = y_i = 1$ for any index $i \in S$, he sends the minimum such i to Alice, and if not he tells Alice they should continue with the quantum bitwise AND portion of Π_D^l. Thus, by Lemma 42 (CQ dimension bound) the information cost of this stage is upper bounded by $\log(s)+1$. In sum, the information cost of the classical subsampling component of Π_D^l is upper bounded by $s+\log(s)+1$.

Now we bound $\mathrm{QIC}(\Pi_D^l, P)$ directly. Let S_A be a binary random variable taking value 1 with the probability that Alice and Bob successfuly find an intersecting coordinate during the classical subsampling component of Π_D^l. Then

$$\begin{aligned}
\mathrm{QIC}(\Pi_D^l, P) &\leq s+\log s+1+\Pr[S_A=0]\mathrm{QIC}(\Pi_A^{l\otimes n}, \nu|\sigma_0) \\
&\leq s+\log s+1+\Pr[S_A=0]\mathrm{QIC}(\Pi_A^{l\otimes n}, \nu) \\
&\leq s+\log s+1+\Pr[S_A=0]\sum_{i\in[n]}\mathrm{QIC}(\Pi_A^l, \nu_i) \\
&\leq s+\log s+1+\Pr[S_A=0]n\mathrm{QIC}(\Pi_A^l, \frac{1}{n}\sum_{i\in[n]}\nu_i), \qquad \text{(C.16)}
\end{aligned}$$

where ν is the probability distribution P conditioned on $S_A = 0$, σ_0 is the leftover (quantum) information from the classical subsampling component of Π_D, and ν_i is the marginal of ν on X_iY_i. The first inequality follows from the above classical subsampling bound along with Lemma 52 (QIC: conditioning on classical register is taking the average) and the fact that the protocol terminates if $S_A = 1$. For the first inequality we have also upper bounded the information cost of the portion of the protocol which runs the bitwise-AND protocol on the remaining $n-s$ coordinates by the information cost of the bitwise-AND protocol run on all n coordinates. This choice gives a negligibly worse bound at the advantage of being easier to present. The second inequality follows from Lemma 51 (QIC: increasing under discarding of side-information). The third inequality follows from Lemma 50 (QIC: subadditivity). The fourth inequality follows from Lemma 48 (QIC: concavity in input distribution).

Note that the bound (C.16) strongly depends on the expected number of intersections $\frac{1}{n}\sum_i \nu_i$. We now formalize our intuition that the classical subsampling component should keep this expected value low.

If $\Pr[S_A = 0] \leq 1/n$ we don't even need to use the effect of classical subsampling to obtain a good bound. Simply by the dimension bound (C.13) of Lemma 54,

$$\Pr[S_A=0]n\mathrm{QIC}(\Pi_A,\frac{1}{n}\sum_{i\in[n]}\nu_i)\le\frac{1}{1-p}\frac{1-\left(\frac{p_z}{1-p}\right)^K}{1-\frac{p_z}{1-p}}2(2r+2). \tag{C.17}$$

If $\Pr[S_A=0]>1/n$ and $\frac{1}{n}\sum_{i\in[n]}\nu_i(1,1)\le 1/2$, then by the bound (C.12) of Lemma 54 and $\Pr[S_A=0]\le 1$,

$$\begin{aligned}&\Pr[S_A=0]n\mathrm{QIC}(\Pi_A^l,\frac{1}{n}\sum_{i\in[n]}\nu_i)\\&\qquad\le n\frac{1}{1-p}\frac{1-\left(\frac{p_z}{1-p}\right)^K}{1-\frac{p_z}{1-p}}\left[\mathrm{QIC}_0(\widetilde{\Pi}_A^l)+2(2r+2)H\left(\frac{1}{n}\sum_{i\in[n]}\nu_i(1,1)\right)\right],\end{aligned}$$

which completes the proof of Lemma 53.

To finish it off, we need only show that the classical subsampling stage ensures the inequality

$$\frac{1}{n}\sum_{i\in[n]}\nu_i(1,1)\le\frac{2\ln(n)}{s}+\frac{1}{n}\le 1/2 \tag{C.18}$$

for all values of n and s satisfying $n\ge 4$ and $8\ln(n)\le s\le n$. The second inequality in (C.18) is straightforward. For the first, let $\mathrm{N}(X,Y)$ be a random variable outputting the number of intersecting coordinates of (x,y). Note that

$$\begin{aligned}\sum_{i\in[n]}\nu_i(1,1)&=\mathbb{E}_\nu\mathrm{N}(X,Y)=\mathbb{E}_{P|S_A=0}\mathrm{N}(X,Y)\\&=\sum_{1\le d\le n}\Pr[\mathrm{N}(X,Y)=d|S_A=0]d\end{aligned} \tag{C.19}$$

and

$$\begin{aligned}\Pr[\mathrm{N}(X,Y)=d|S_A=0]&=\frac{\Pr[\mathrm{N}(X,Y)=d]}{\Pr[S_A=0]}\cdot\Pr[S_A=0|\mathrm{N}(X,Y)=d]\\&\le n\Pr[\mathrm{N}(X,Y)=d](1-d/n)^s\\&\le n\Pr[\mathrm{N}(X,Y)=d]\exp(-ds/n),\end{aligned} \tag{C.20}$$

where the first inequality follows from $\Pr[S_A = 0] \geq 1/n$ and $\Pr[S_A = 0|\mathrm{N}(X,Y) = d] \leq (1 - d/n)^s$, and the second inequality follows from the Taylor series expansion of the exponential function. Thus,

$$\begin{aligned}
\sum_{i\in[n]} \nu_i(1,1) = & \sum_{1\leq d\leq \left\lfloor \frac{2n\ln(n)}{s} \right\rfloor} \Pr[\mathrm{N}(X,Y) = d|S_A = 0]d \\
& + \sum_{\left\lfloor \frac{2n\ln(n)}{s} \right\rfloor < d \leq n} Pr[\mathrm{N}(X,Y) = d|S_A = 0]d \\
& \leq \sum_{1\leq d\leq \left\lfloor \frac{2n\ln(n)}{s} \right\rfloor} \Pr[\mathrm{N}(X,Y) = d|S_A = 0]d \\
& + (n\exp(-ds/n))n \\
& \leq \sum_{1\leq d\leq \left\lfloor \frac{2n\ln(n)}{s} \right\rfloor} \Pr[\mathrm{N}(X,Y) = d|S_A = 0]d \\
& + 1 \\
& \leq \frac{2n\ln(n)}{s} + 1.
\end{aligned}$$

The first inequality follows from (C.20), the upper bound $d \leq n$ for all d in the range of the second sum, and the bound n on the number of terms in the sum. The second inequality follows from $\exp(-ds/n) \leq 1/n^2$ (which results from $2\ln(n) \leq s \leq n$), and the third inequality follows from $d \leq \frac{2n\ln(n)}{s}$ for all terms in the sum, along with the fact that the sum is upper bounded by a convex combination of $1 \leq d \leq \left\lfloor \frac{2n\ln(n)}{s} \right\rfloor$, which is upper bounded by the largest term $\left\lfloor \frac{2n\ln(n)}{s} \right\rfloor$. This completes the proof of the inequality (C.18), and thus the proof of Lemma 53.

C.2.1 Information cost of subroutine AND protocol

Here we detail the framework we will use to bound $\mathrm{QIC}_0(\widetilde{\Pi}^l_A(\eta))$ for $l = 1,2$. A lossy channel acting on coherent states can be modeled by a beamsplitter with transmittivity η. We assume that channel loss resides in the communication register, and after it has been communicated it resides in Bob's memory. This corresponds to the case in which Bob is "honest but curious", i.e. he honestly performs the protocol while attempting to gain as much information as possible about Alice's input using his memory and the environment.

By the simplification (C.11) of the information cost of pure state protocols, the information cost of protocol $\widetilde{\Pi}_A^l$ simplifies to

$$\mathrm{QIC}(\widetilde{\Pi}_A^l, P_0) = \sum_{i=1}^{2r} \mathrm{H}(C_i|YB_i) + \mathrm{H}(C_i|XA_i) + \mathrm{QIC}_{2r+1}(\widetilde{\Pi}_A^l, P_0) + \mathrm{QIC}_{2r+2}(\widetilde{\Pi}_A^l, P_0). \tag{C.21}$$

Note that the last two messages are classical, and are mixed because they depend on Alice's measurement outcome.

Now we simplify the above expression for $P_0 \in \Pr(\mathrm{X}_i \times \mathrm{Y}_i)$ satisfying $P_0(1,1) = 0$. For both protocols, $\mathrm{H}(C_i|X=0) = 0$ for all $i \leq 2r$, and $\mathrm{H}(C_i|Y=0) = 0$ for even $i \leq 2r$. In the final two messages C_{2r+1}, C_{2r+2}, if the second mode clicks, Alice sends x to Bob and Bob sends y to Alice. For an input distribution with zero mass on $(1,1)$, this occurs with probability p_s. Otherwise, in C_{2r+1}, with probability p Alice sends "Inconclusive" and with probability p_z Alice sends "Zero?". In this case, Bob sends nothing in the final message, so C_{2r+2} is trivial. For an input distribution with zero mass on $(1,1)$, these probabilities and the content of these messages is independent of (x,y) and the content of register B_{2r+1}, so these messages do not contribute to the information cost. Thus, the only contribution to the information cost for the final two messages is when the second mode clicks. By Lemma 42 (CQ dimension bound),

$$\mathrm{QIC}_{2r+1}(\widetilde{\Pi}_A^l(\eta), P_0) + \mathrm{QIC}_{2r+2}(\widetilde{\Pi}_A^l(\eta), P_0) \leq 2p_s. \tag{C.22}$$

By (C.11), Lemma 41 (conditioning on classical register is taking the average), and the above analysis,

$$\mathrm{QIC}(\widetilde{\Pi}_A^l, P_0) \leq 2p_s + \sum_{i=1,\mathrm{odd}}^{2r-1} \mathrm{H}(C_i B_i|Y=0) - \mathrm{H}(B_i|Y=0). \tag{C.23}$$

In the following sections we apply the above bound to each protocol $\widetilde{\Pi}_A^l$.

C.2.2 Information cost for zero-state injection protocol

Here we prove the following bound on the information cost $\mathrm{QIC}_0(\widetilde{\Pi}_A^2(\eta))$ of the subroutine $\widetilde{\Pi}_A^2(\eta)$ for the zero-state injection protocol $\Pi_A^2(\eta, p_d, \varepsilon)$ for AND.

Lemma 55.

$$\mathrm{QIC}_0(\widetilde{\Pi}_A^2(\eta)) \leq 2p_s + \mathrm{H}(\frac{1}{2}(1 - F_1 F_3 \dots F_{2r-1})),$$

in which

$$p_s = e^{-|\alpha|^2}(1-p_d)p_d \tag{C.24}$$

is the probability that only the second mode clicks for non-intersecting inputs, and

$$F_i = \exp\left[-\frac{|\alpha|^2}{\eta^{2r-(i-1)}}\left[1-\cos\left(\frac{\pi}{2r}\right)\right]\right] \tag{C.25}$$

for all $i = 1, 3, \ldots, 2r-1$.

Corollary 56. *The protocol $\Pi_D^2(\eta, p_d, \varepsilon)$ described in Section 5.3 satisfies*

$$\mathrm{QIC}(\Pi_D^2(\eta, p_d, \varepsilon)) \leq s + \log s + 1 + \frac{n}{1-p}\frac{1-\left(\frac{p_z}{1-p}\right)^K}{1-\frac{p_z}{1-p}}\left[\frac{2(2r+2)}{n} + 2p_s\right.$$
$$\left. + \mathrm{H}\left(\frac{1}{2}(1-F_1F_3\ldots F_{2r-1})\right) + 2(2r+2)\mathrm{H}\left(\frac{2\ln n}{s} + \frac{1}{n}\right)\right],$$

From expression (C.23) it is clear that any content of B_i which produces an uncorrellated pure state when conditioned on $Y = 0$ can be safely discarded without changing the information cost. Therefore, we assume B_i contains only elements which do not produce an uncorellated pure state when conditioned on $Y = 0$. Under this assumption, the state of the registers C_iB_i for odd i in the $(0,0), (0,1), (1,0)$ cases are as follows:

State of registers C_iB_i for odd i on different inputs for protocol $\widetilde{\Pi}_A^2$:

On (0,0):

i odd $(A \to B)$:

$$\left(\left|\frac{1}{\eta^{r-i/2}}\alpha,0\right\rangle\left|\frac{\sqrt{1-\eta}}{\eta^{r-(i-1)/2}}\alpha,0\right\rangle\right)\bigotimes_{l=1,\text{odd}}^{i-2}\left(\left|\frac{1}{\eta^{r-l/2}}\alpha,0\right\rangle\left|\frac{\sqrt{1-\eta}}{\eta^{r-(l-1)/2}}\alpha,0\right\rangle\right)$$

On (0,1): Identical to (0,0).

On (1,0):

i odd $(A \to B)$:

$$\left(\left|\frac{1}{\eta^{r-i/2}}\alpha\cos 2\theta,\frac{1}{\eta^{r-i/2}}\alpha\sin 2\theta\right\rangle\left|\frac{\sqrt{1-\eta}}{\eta^{r-(i-1)/2}}\alpha\cos 2\theta,\frac{\sqrt{1-\eta}}{\eta^{r-(i-1)/2}}\alpha\sin 2\theta\right\rangle\right)$$
$$\bigotimes_{k=1,\text{odd}}^{i-2}\left(\left|\frac{1}{\eta^{r-k/2}}\alpha\cos 2\theta,\frac{1}{\eta^{r-k/2}}\alpha\sin 2\theta\right\rangle\left|\frac{\sqrt{1-\eta}}{\eta^{r-(k-1)/2}}\alpha\cos 2\theta,\frac{\sqrt{1-\eta}}{\eta^{r-(k-1)/2}}\alpha\sin 2\theta\right\rangle\right)$$

where the first two modes are contained in register C_i and the rest are contained in register B_i. Note that the content of registers $C_{i-2}B_{i-2}$ is identical to that of register B_i. Thus, (C.23) simplifies to

$$\begin{aligned}\mathrm{QIC}(\widetilde{\Pi}_A^l,P_0) &= 2p_s+\mathrm{H}(C_{2r-1}B_{2r-1}|Y=0)-\mathrm{H}(B_1|Y=0)\\ &= 2p_s+\mathrm{H}(C_{2r-1}B_{2r-1}|Y=0)\\ &\leq 2p_s+\mathrm{H}\left(\frac{1}{2}(1-F_1F_3\dots F_{2r-1})\right),\end{aligned}\tag{C.26}$$

where the second equality follows from $\mathrm{H}(B_1|Y=0)=0$, which results from the fact that the state of B_1 is pure conditioned on $Y=0$. The product $F_1F_3\dots F_{2r-1}$ is the overlap of the two possible states of $C_{2r-1}B_{2r-1}$ when $Y=0$ and is given by (C.25). The inequality follows from Lemma 47. As the above bound holds for all $P_0\in\Pr(\{0,1\}^2)$ such that $P_0(1,1)=0$, then it also bounds $\mathrm{QIC}_0(\widetilde{\Pi}_A^2)$.

C.2.3 Information cost for coherent state version of qubit protocol

We prove the following bound on the information cost $\mathrm{QIC}_0(\widetilde{\Pi}_A^1(\eta))$ of the subroutine $\widetilde{\Pi}_A^1(\eta)$ for the coherent-state protocol $\Pi_A^1(\eta, p_d, \varepsilon)$ for AND.

Lemma 57.

$$\mathrm{QIC}_0(\widetilde{\Pi}_A^1(\eta)) \leq 2p_s + \sum_{i=1,\mathrm{odd}}^{2r-1} \left[\mathrm{H}\left(\frac{1}{2}(1-F_i) \right) \right],$$

in which

$$p_s = e^{-|\alpha|^2}(1-p_d)p_d \tag{C.27}$$

is the probability that only the second mode clicks for non-intersecting inputs, and

$$F_i = \exp\left[-\frac{|\alpha|^2}{\eta^{2r-(i-1)}} \left[1 - \cos\left(\frac{\pi}{2r} \right) \right] \right] \tag{C.28}$$

for all $i = 1, 3, \ldots, 2r-1$.

Corollary 58. *The protocol* $\Pi_D^1(\eta, p_d, \varepsilon)$ *described in Section 5.3 satisfies*

$$\mathrm{QIC}\left(\Pi_D^1(\eta, p_d, \varepsilon) \right) \leq s + \log s + 1 + \frac{n}{1-p} \frac{1 - \left(\frac{p_z}{1-p} \right)^K}{1 - \frac{p_z}{1-p}} \left[\frac{2(2r+2)}{n} + 2p_s \right.$$
$$\left. + \sum_{i=1,\mathrm{odd}}^{2r-1} \left[\mathrm{H}\left(\frac{1}{2}(1-F_i) \right) \right] + 2(2r+2)\mathrm{H}\left(\frac{2\ln n}{s} + \frac{1}{n} \right) \right]$$

Now we prove the lemma. The state of the register C_i for odd i in the $(0,0), (0,1), (1,0)$ cases are as follows:

State of register C_i for odd i on different inputs for protocol $\widetilde{\Pi}_A^1$:

On (0,0):

i odd $(A \to B)$:

$$\left|\frac{1}{\eta^{r-i/2}}\alpha, 0\right\rangle \left|\frac{\sqrt{1-\eta}}{\eta^{r-(i-1)/2}}, 0\right\rangle \tag{C.29}$$

On (1,0): Identical to (0,0).

On (0,1):

i odd $(A \to B)$:

$$\left|\frac{1}{\eta^{r-i/2}}\alpha\cos 2\theta, \frac{1}{\eta^{r-i/2}}\alpha\sin 2\theta\right\rangle \left|\frac{\sqrt{1-\eta}}{\eta^{r-(i-1)/2}}\alpha\cos 2\theta, \frac{\sqrt{1-\eta}}{\eta^{r-(i-1)/2}}\alpha\sin 2\theta\right\rangle \tag{C.30}$$

By (C.23), Lemma 44 (subadditivity), and Lemma 47, we have

$$\begin{aligned}\mathrm{QIC}(\widetilde{\Pi}_A^1(\eta), P_0) &\leq 2p_s + \sum_{i=1,\mathrm{odd}}^{2r-1} \mathrm{H}(C_i|Y=0)\\ &\leq 2p_s + \sum_{i=1,\mathrm{odd}}^{2r-1} \mathrm{H}\left(\frac{1}{2}(1-F_i)\right)\end{aligned}$$

where F_i is the fidelity between the two possible states of register C_i when $Y=0$, and is given by (C.28). This completes the proof.

C.3 Information cost of coherent state Grover search protocol

Here we bound the information cost of any r'-round safe interactive communication protocol in which, conditioned on fixed inputs $x \in \mathrm{X}, y \in \mathrm{Y}$, pure states $\left|\phi_i^{x,y}\right\rangle^{C_i}$ are exchanged, each of which is a tensor product of m coherent states with total mean photon number lying in a fixed range $[\mu_{\min}, \mu_{\max}]$ for all $x \in \mathrm{X}, y \in \mathrm{Y}$, as $\mathcal{O}(r' \log m)$. We then apply this bound to the coherent state Grover search protocol.

By (C.11) and Lemma 44 (subadditivity), for any such protocol Π and any input distribution Π, $\mathrm{QIC}_i(\Pi, P) \leq 2\mathrm{H}(C_i)$, which is $\mathcal{O}(\log n)$ by the analysis of Appendix B.0.2. Thus, after r' rounds, the total information cost is $\mathcal{O}(r' \log m)$.

Now we apply this bound to the coherent state version of the Grover search protocol. This is a pure state protocol, and every state has total mean photon number $|\alpha|^2$. This follows from $V_S\,|\psi_\alpha\rangle = f_\alpha(U_S\,|\psi\rangle)$ and $V_A\,|\psi_\alpha\rangle = f_\alpha(U_A\,|\psi\rangle)$ for every state $|\psi\rangle$ used in the original protocol, and that Alice and Bob's manipulations of the state to jointly perform V_A do not change the total mean photon number. Each state communicated between Alice and Bob is a tensor product of n coherent states. For K repetitions, by straightforward application of Lemma 51 (QIC: increasing under discarding of side information), the fact that this protocol uses $Kr = \mathcal{O}(\sqrt{n/k})$ rounds of quantum communication, and the above information cost bound, the information cost of this stage is $\mathcal{O}(\sqrt{n/k}\log n)$.

For each repetition of the protocol, Alice sends Bob her measurement outcome i (which is $\log n$ bits) along with x_i (which is one bit), and Bob sends Alice y_i (which is one bit). Or, if Alice received no clicks she uses one bit to tell Bob. Thus, the amount of communication in these stages is upper bounded by $K(\log n + 2 + 1) = \mathcal{O}(\log n)$ bits, which also upper bounds the information cost of these stages by Lemma 42 (dimension bound). Thus, in total, this protocol has information cost $\mathcal{O}(\sqrt{n/k}\log(n))$. The coherent state Grover search protocol thus has a nearly quadratic improvement over the classical information cost lower bound of $\Omega(n)$ proven in [7] and [19] for the zero-error and nonzero-error cases, respectively.