University of Bristol

Quantum Engineering Centre for Doctoral Training

Advanced Quantum Information

Universal quantum computation by linear optics

Submitted by:
Stasja Stanisic

May 22, 2015

Contents

1 Introduction

With the development of algorithms that can utilize quantum properties for faster computation of certain groups of problems (such as database search [**?**], factorization [**?**] and quantum simulation [**?**]), research into quantum computation has intensified. The possibility of quantum computing was further supported by the discovery that quantum computers can be error-correctable [**?**] [**?**]. Over the years, many different platforms have been researched such as: ion and atom traps, nuclear magnetic resonance, superconducting systems, quantum dots and optical platforms [**?**]. Due to restrictions of bosonic systems, it was believed that it is not possible to build a universal computer using only linear optics, until in 2001, Knill, Laflamme and Milburn (KLM) [**?**] realized that measurement on parts of the circuit can be used to evoke non-linearity and still deliver scalability. This changed the quantum computation landscape making linear optical quantum computing (LOQC) seem possible again.

As the search for the universal quantum computer matured, the need for criteria of what constitutes a quantum computer manifested. In 2000, DiVincenzo [**?**] laid out five criteria for quantum computing. Focus of this essay will be on two of those five criteria, namely qubit definition and implementation of universal set of gates (see more on the criteria in appendix A.1). When it comes to optical architectures, photons lend themselves well due to various degrees of freedom that can represent a qubit, low levels of decoherence and interaction and some types of gates can be easy to implement [**?**]. In the "worst case" scenario where optical quantum computer is not possible, photons will still, most likely, be incorporated into the future of quantum computing as information carriers.

1.1 Preliminary definitions

Fock states will be marked as $|n_1, ..., n_n\rangle$, where n_j gives us the number of photons in mode j and n is the total number of modes. Total number of photons is then $N := \sum_{j=1}^{n} n_j$. Sometimes the notation $|n_j\rangle_j$ will be used, marking the exact number state of the j-th mode.

Definition 1.1. Bosonic creation and annihilation operators are operators acting upon Fock space states in the following fashion

$$\hat{a}_j^\dagger |n\rangle_j = \sqrt{n+1}\,|n+1\rangle_j$$
$$\hat{a}_j |n\rangle_j = \sqrt{n}\,|n-1\rangle_j$$

and $[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}$ is also true $\forall i, j \in \{1, ..., n\}$.

Definition 1.2. Vacuum of all modes is a state $|vac\rangle := |0\rangle^{\otimes n}$. Vacuum of a single mode is just denoted as $|0\rangle_j$.

Lemma 1.1.
a)

$$|n_j\rangle_j = \frac{(\hat{a}_j^\dagger)^{n_j}}{\sqrt{n_j!}}\,|0\rangle\,, \forall j \in \{1, ..., n\}$$

b)

$$|n_1, ..., n_n\rangle = \prod_{j=1}^{n} \frac{(\hat{a}_j^\dagger)^{n_j}}{\sqrt{n_j!}}\,|vac\rangle\,, \forall j \in \{1, ..., n\}$$

Proof is straightforward from above definitions.

Optical elements are defined by the effective interaction Hamiltonian of the medium H [**?**]. The unitary that evolves the state of the system, U, acting on $|n_1, ..., n_j\rangle$ is given as $U = e^{itH}$.

Definition 1.3. Linear optical elements are such that the mode transformation under evolution U can be described by matrices u and v, which transform the modes linearly, that is, $\hat{a}_j^\dagger \to \sum_k u_{kj}\hat{a}_k^\dagger + v_{kj}\hat{a}_k$.

Definition 1.4. Linear optical elements are called passive if the energy of the incoming photons is conserved which implies that the number of photons is conserved thus $v = 0$ (from definition 1.3).

In the case of passive linear optics, the interaction Hamiltonian is bilinear in the creation and annihilation operators and is of the form $H = \sum_{jk} h_{jk}\hat{a}_j^\dagger\hat{a}_k$ where the annihilation operator comes second by convention [**?**].

This gives some nice properties which allow finding u for a given U. Firstly, due to this convention, $U|vac\rangle = |vac\rangle$. Since U is unitary, the transformation of a state defined by an operator $\hat{a}_j^\dagger$ can be written as $U\hat{a}_j^\dagger|vac\rangle = U\hat{a}_j^\dagger U^\dagger U|vac\rangle = U\hat{a}_j^\dagger U^\dagger|vac\rangle$. Finally, u such that $U\hat{a}_j^\dagger U^\dagger = \sum_k u_{kj}\hat{a}_k^\dagger$ can be found by applying the Baker-Campbell-Hausdorff expansion (longer explanation of these claims in appendix A.2)

As a consequence, transformation on modes, u, is unitary. Moreover, for any given mode transformation unitary u, there is a way to construct it using only beam splitters and phase shifters as shown by Reck et al. [**?**].

Definition 1.5. Unitary transformation on optical modes, u, is an isomorphism from the space of input optical operators to output optical operators.

When describing an optical element X, the unitary transformation* $u(\mathrm{X})$ it performs on the optical modes will be given.

Definition 1.6. Phase shifter P_θ is a one mode passive linear optical element with Hamiltonian $H_{P_\theta}(\theta) = -\theta\hat{a}^\dagger\hat{a}$. Its transformation is $u = e^{-i\theta}$.

Definition 1.7. Beam splitter $\mathrm{B}_{\theta,\phi}$ is a two mode passive linear optical element generated by Hamiltonian $H_{B_{\theta,\phi}}(r) = \theta e^{i\phi}\hat{a}^\dagger\hat{b} + \theta e^{-i\phi}\hat{b}^\dagger\hat{a}$. Its transformation is then

$$u(\mathrm{B}_{\theta,\phi}) = \begin{pmatrix} \cos(\theta) & -e^{i\phi}\sin(\theta) \\ e^{-i\phi}\sin(\theta) & \cos(\theta) \end{pmatrix}$$

Let $\mathrm{B}_\theta := \mathrm{B}_{\theta,0}$. The commonly used beam splitter is "50:50" beam splitter, where the transmission and reflection are equal. There are a few different representations of this beam splitter which are equivalent up to a global phase (choosing $\theta \in \{\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}\}$ in the definition

*In the rest of this essay, unless if noted otherwise, when unitary u is discussed, the unitary from definition 1.5 is meant, as opposed to the unitary that is mapping the input states to output states (they act on different spaces, former acting within the operator space while the latter acts within the state Hilbert space).

1.7 would give different types of this beam splitter). The one[†] we will find the most use for will be

$$u(\text{BS}) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

A beam splitter in polarization encoding is also sometimes called "polarization rotator" (commonly used types are quarter wave plate and half wave plate).

Definition 1.8. Polarizing beam splitter (PBS) that separates horizontal and vertical polarization is a four mode passive linear optical element with following transformation

$$u(\text{PBS}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Definition 1.9. A set of quantum gates is universal for quantum computation if any unitary operation can be approximated to arbitrary accuracy by a quantum circuit involving only those gates.

Definition 1.10. CNOT gate is a two qubit "entangling" gate which in matrix representation is

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Definition 1.11. Hadamard gate is a single qubit gate which in matrix representation is

$$\text{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Definition 1.12. A $\frac{\pi}{8}$ gate is a single qubit gate which in matrix representation is

$$\text{T} = \begin{pmatrix} 1 & 0 \\ 0 & e^{\frac{i\pi}{4}} \end{pmatrix}$$

The gate set { CNOT, H, T } is universal for quantum computation. Further, as CNOT can be obtained by applying HCZH, { CZ, H, T } is universal as well.

1.2 Qubit encoding

Different types of qubit encoding can be used in LOQC due to photons having several degrees of freedom. The most commonly used are polarization and spatial. Usually, only one degree of freedom is chosen to encode a qubit, but sometimes two qubits can be encoded on one photon by mixing these degrees of freedom. Types of encoding used are single-rail, dual-rail, mixed polarization and spatial encoding, parity encoding and redundant encoding.

[†]equivalent (up to a global phase) with the "50:50" beam splitters found by applying definition 1.7

In single-rail encoding, a photon being present in the rails is considered to be the logical $|1\rangle$, while vacuum is considered to be $|0\rangle$ (so $|0\rangle_L = |0\rangle_{Fock}$, $|1\rangle_L = |1\rangle_{Fock}$). Notice that using a "50:50" beam splitter, a qubit encoded in the single-rail will get entangled. But, in this encoding single-qubit operations are difficult (see appendix A.3.1) [**?**].

In the dual-rail encoding, two modes are employed to represent a qubit (so $|0\rangle_L = |10\rangle_{12}$ and $|1\rangle_L = |01\rangle_{12}$). These two modes can represent polarization or spatial modes. In the case of polarization, the qubits are also written as $|0\rangle_L = |H\rangle$ and $|1\rangle_L = |V\rangle$. In this encoding, single qubit operations are easy to implement, but entanglement is not straightforward (see appendix A.3.2) [**?**]. Dual-rail encoding of qubits is generally preferred over single-rail encoding because both of dual-rail logical qubits are marked by the presence of a photon as opposed to the absence and presence. From a practical perspective former is less error-prone and can give easier error-detection.

In parity encoding, a qubit is described by an equal superposition of states which can be split into "odd" and "even" [**?**]. For example, with two photons that carry polarization information the encoding could be $|0\rangle_L = \frac{1}{\sqrt{2}}(|HH\rangle + |VV\rangle)$ and $|1\rangle_L = \frac{1}{\sqrt{2}}(|HV\rangle + |VH\rangle)$. This type of encoding is useful for error-correction (see appendix A.3.3).

As an example of mixed encoding polarization and path mode, two qubits can be encoded in a single photon, with the polarization information representing the first qubit and path information representing the second qubit.

Redundant encoding is similar to parity encoding from the error-correction perspective. An example for polarization encodings are logical qubits encoded as $|0\rangle_L = |H\rangle^{\otimes n}$ and $|1\rangle_L = |V\rangle^{\otimes n}$ [**?**].

There are much more complex ways of encoding qubits for the purposes of fault-tolerance, such as surfaces codes [**?**] and the Raussendorf lattice [**?**]. The closer we get to engineering a quantum computer the more important the question becomes of qubit encoding and fault-tolerance.

2 KLM scheme

Knill-Laflamme-Milburn (KLM) scheme was first to show a theoretically scalable LOQC and was based on the idea that non-linearity can be introduced through measurement [**?**]. Before the KLM scheme, another attempt at making LOQC worth mentioning is the Adami-Cerf-Kwiat scheme [**?**]. They use 2^{n-1} paths to represent n qubits, and the qubits are path and polarization encoded. While this scheme can perform the needed gates for universality, it is not scalable and moreover the question is raised of the lack of "non-locality". Further, specific algorithms have been shown to work on an optical system such as Shor's algorithm [**?**] and Grover's search [**?**], but none of these are a demonstration of a universal LOQC.

It was believed that an optics computer would have to use some non-linearities for the purposes of entanglement, such as Zeno gates and Kerr non-linearities [**?**]. In no known element have these non-linearities been strong enough [**?**] thus this solution remains impractical.

2.1 Introduction to KLM

The KLM scheme uses dual-rail qubit encoding (except for the teleportation protocol which is performed in single-rail qubit encoding). The state preparation consists of using the single photon source to prepare a photon in one of the two modes representing a qubit. As mentioned before (section 1.2), in the dual-rail encoding, single qubit gates are easy to implement. Therefore to complete the universal gate set a two-qubit entangling gate, such as CNOT or CZ is missing and is the focus of the paper.

KLM introduces the idea of "non-deterministic quantum computation" which gives a source of non-linearity through measurement. Non-deterministic gates will succeed some of the time and it is known what the probability of them working is. Further, from the measurement result it is known when they succeed.

In the paper, they first demonstrate a way that a CZ gate can be constructed with success probability $\frac{1}{16}$ [**?**]. Then, by employing quantum teleportation, this probability is increased to $\frac{1}{4}$. Next, a near-deterministic gate is constructed at the expense of the size of the ancilla state by combining teleportation with clever ancilla states. Finally, they prove efficiency in the sense of polynomial resources.

2.2 Non-deterministic non-linear sign shift gate

A non-linear sign shift gate is a gate that takes photons in input mode 1 to output mode 1 as follows $\mathrm{NS}_x : \alpha_0 |0\rangle_1 + \alpha_1 |1\rangle_1 + \alpha_2 |2\rangle_1 \xrightarrow{\mathrm{NS}_x} \alpha_0 |0\rangle_1 + \alpha_1 |1\rangle_1 + x\alpha_2 |2\rangle_1$, where x is the phase shift applied. The gate of interest is non-linear sign flip $\mathrm{NS} := \mathrm{NS}_{-1}$ which takes the above input to $\alpha_0 |0\rangle_1 + \alpha_1 |1\rangle_1 - \alpha_2 |2\rangle_1$.

Lemma 2.1. *Let the NS gate*[‡] *be applied to the state* $|\phi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle + \alpha_2 |2\rangle$ *which is found in mode* 1 *with creation operator* $\hat{a}_1^\dagger$. *Let there be two ancilla modes* 2 *and* 3, *such that there is one photon in mode* 2 *and no photons in mode* 3. *Apply the mode transformation marked by unitary*

$$M = \begin{pmatrix} 1-\sqrt{2} & 2^{-\frac{1}{4}} & \sqrt{\frac{3}{\sqrt{2}}-2} \\ 2^{-\frac{1}{4}} & \frac{1}{2} & \frac{1}{2}-\frac{1}{\sqrt{2}} \\ \sqrt{\frac{3}{\sqrt{(2)}}-2)} & \frac{1}{2}-\frac{1}{\sqrt{2}} & \sqrt{2}-\frac{1}{2} \end{pmatrix}$$
$$= \begin{pmatrix} -0.414213562373095 & 0.840896415253715 & 0.348310699749007 \\ 0.840896415253715 & 0.5 & -0.207106781186548 \\ 0.348310699749007 & -0.207106781186548 & 0.914213562373095 \end{pmatrix}$$

such that $\hat{a}_i^\dagger \to \sum_j M_{ji}\hat{a}_j^\dagger$. *If one photon is found in ancilla mode* 2 *and no photons in ancilla mode* 3 *after the transformation, then a sign shift gate has been performed on the state* $|\phi\rangle$ *with probability* $\frac{1}{4}$.

Proof. Let the input state be written as $|\phi_{in}\rangle = |\phi\rangle |10\rangle_{23} = (\alpha_0 |0\rangle_1 + \alpha_1 |1\rangle_1 + \alpha_2 |2\rangle_1) |10\rangle_{23} = (\alpha_0 + \alpha_1\hat{a}_1^\dagger + \frac{1}{\sqrt{2}}\alpha_2(\hat{a}_1^\dagger)^2)\hat{a}_2^\dagger |vac\rangle$. Since $\hat{a}_1^\dagger \to \sum_j M_{j1}\hat{a}_j^\dagger$ and $\hat{a}_2^\dagger \to \sum_j M_{j2}\hat{a}_j^\dagger$ then $|\phi_{in}\rangle \xrightarrow{NS} (\alpha_0 + \alpha_1 \sum_j M_{j1}\hat{a}_j^\dagger + \frac{1}{\sqrt{2}}\alpha_2(\sum_j M_{j1}\hat{a}_j^\dagger)^2)(\sum_j M_{j2}\hat{a}_j^\dagger) |vac\rangle =: |\phi_{out}\rangle$

[‡]In KLM, an implementation of NS_x for any x is given. For the purpose of this essay only NS gate is needed however.

Expanding this state $|\phi_{out}\rangle$ gives a polynomial of the output creation operators of degree 3. The terms of of interest in that polynomial will have $\hat{a}_2^\dagger$ of power 1 and $\hat{a}_3^\dagger$ of power 0, denoting one photon in mode 1 and zero photons in mode 2. Isolate the term containing $\hat{a}_2^\dagger$ (using some symbolic programming language), with $C_{(\hat{a}_2^\dagger)^1}$ being the "coefficient"[§] that corresponds to it:

$$\begin{aligned}
C_{(\hat{a}_2^\dagger)^1}\hat{a}_2^\dagger = \Big(& \frac{\sqrt{2}}{2} M_{11}^2 M_{22}\alpha_2 (\hat{a}_1^\dagger)^2 + \sqrt{2} M_{11}M_{12}M_{21}\alpha_2 (\hat{a}_1^\dagger)^2 \\
& + \sqrt{2} M_{11}M_{21}M_{32}\alpha_2 \hat{a}_1^\dagger \hat{a}_3^\dagger + \sqrt{2} M_{11}M_{22}M_{31}\alpha_2 \hat{a}_1^\dagger \hat{a}_3^\dagger + \sqrt{2} M_{12}M_{21}M_{31}\alpha_2 \hat{a}_1^\dagger \hat{a}_3^\dagger \\
& + \sqrt{2} M_{21}M_{31}M_{32}\alpha_2 (\hat{a}_3^\dagger)^2 + \frac{\sqrt{2}}{2} M_{22}M_{31}^2 \alpha_2 (\hat{a}_3^\dagger)^2 \\
& + M_{22}M_{31}\alpha_1 \hat{a}_3^\dagger + M_{21}M_{32}\alpha_1 \hat{a}_3^\dagger + M_{11}M_{22}\alpha_1 \hat{a}_1^\dagger + M_{12}M_{21}\alpha_1 \hat{a}_1^\dagger \\
& + M_{22}\alpha_0 \Big) \hat{a}_2^\dagger
\end{aligned}$$

This is what the state would be after post-selection on one photon in mode 2. Post-selecting is also needed of 0 photons in mode 3, so terms that only depend on $(\hat{a}_3^\dagger)^0 = 1$ and have no $(\hat{a}_3^\dagger)^1$ or $(\hat{a}_3^\dagger)^2$ should be included, which reduces above to

$$\begin{aligned}
C_{(\hat{a}_2^\dagger)^1,(\hat{a}_3^\dagger)^0}\hat{a}_2^\dagger =& \big(\frac{\sqrt{2}}{2} M_{11}^2 M_{22}\alpha_2 (\hat{a}_1^\dagger)^2 + \sqrt{2} M_{11}M_{12}M_{21}\alpha_2 (\hat{a}_1^\dagger)^2 \\
& + M_{11}M_{22}\alpha_1 \hat{a}_1^\dagger + M_{12}M_{21}\alpha_1 \hat{a}_1^\dagger \\
& + M_{22}\alpha_0 \big) \hat{a}_2^\dagger \\
=& \big(\frac{\sqrt{2}}{2} \alpha_2 M_{11}(M_{11}M_{22} + 2M_{12}M_{21})(\hat{a}_1^\dagger)^2 \\
& + \alpha_1 (M_{11}M_{22} + M_{12}M_{21}) \hat{a}_1^\dagger + \alpha_0 M_{22} \big) \hat{a}_2^\dagger
\end{aligned}$$

where $C_{(\hat{a}_2^\dagger)^1,(\hat{a}_3^\dagger)^0}$ denotes the coefficient corresponding to the term $(\hat{a}_2^\dagger)^1(\hat{a}_3^\dagger)^0$.

When the pattern $|10\rangle_{23}$ is detected, the state after measurement is

$$|\phi_{out}\rangle_1 = \big(\frac{1}{\sqrt{2}} \alpha_2 \lambda_2 (\hat{a}_1^\dagger)^2 + \alpha_1 \lambda_1 \hat{a}_1^\dagger + \alpha_0 \lambda_0 \big) |vac\rangle$$

where

$$\begin{aligned}
\lambda_2 &= M_{11}(M_{11}M_{22} + 2M_{12}M_{21}) = -\frac{1}{2} \\
\lambda_1 &= M_{11}M_{22} + M_{12}M_{21} = \frac{1}{2} \\
\lambda_0 &= M_{22} = \frac{1}{2}
\end{aligned}$$

[§]The coefficient here is actually a function of the remaining operators, grouped into a "coefficient" of the given term.

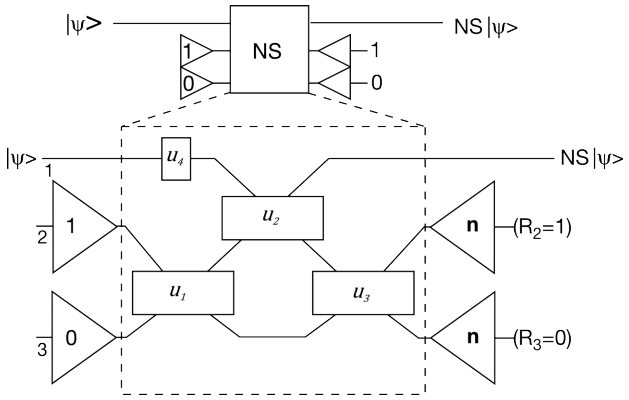

Figure 1: Image taken from KLM[**?**]. Optical implementation of a single qubit NS gate in mode 1. Ancilla modes 2 and 3 used as mentioned in Lemma 2.1. Unitaries u_1, u_2 and u_3 are beam splitters, B_θ with θ taking values 22.5°, 65.5302° and −22.5° respectively. Unitary u_4 is a phase shifter, P_θ with $\theta = 180°$.

which gives

$$\begin{aligned}|\phi_{out}\rangle_1 &= \frac{1}{2}(-\frac{1}{\sqrt{2}}\alpha_2(\hat{a}_1^\dagger)^2 + \alpha_1\hat{a}_1^\dagger + \alpha_0)\hat{a}_2^\dagger\,|vac\rangle \\ &= \frac{1}{2}(\alpha_0\,|0\rangle_1 + \alpha_1\,|1\rangle_1 - \alpha_2\,|2\rangle_1)\,|10\rangle_{23}\end{aligned}$$

An NS gate has indeed been applied to this state and from the normalization of $|\phi_{out}\rangle_1$ it follows that the detection probability of $|10\rangle_{23}$ and the success rate of the gate is $\frac{1}{4}$. □

To find out how to implement this unitary, they used the results of Reck et al. [**?**].

Lemma 2.2. *The unitary mode transformation matrix M found in Lemma 2.1 can be constructed using elements $B_{22.5°}$, $B_{-22.5°}$, $B_{65.5302°}$ and $P_{180°}$ in a setup as seen in figure 1.*

Proof. The unitaries of each component are

$$\begin{aligned}u_1 := u(B_{22.5°}) &= \begin{pmatrix}1 & 0 & 0\\ 0 & 0.923879532511287 & -0.38268343236509\\ 0 & 0.38268343236509 & 0.923879532511287\end{pmatrix}\\
u_2 := (B_{65.5302°}) &= \begin{pmatrix}0.414229709262441 & -0.910172372665944 & 0\\ 0.910172372665944 & 0.414229709262441 & 0\\ 0 & 0 & 1\end{pmatrix}\\
u_3 := u(B_{-22.5°}) &= \begin{pmatrix}1 & 0 & 0\\ 0 & 0.923879532511287 & 0.38268343236509\\ 0 & -0.38268343236509 & 0.923879532511287\end{pmatrix}\\
u_4 := u(P_{180°}) &= \begin{pmatrix}-1.0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\end{aligned}$$

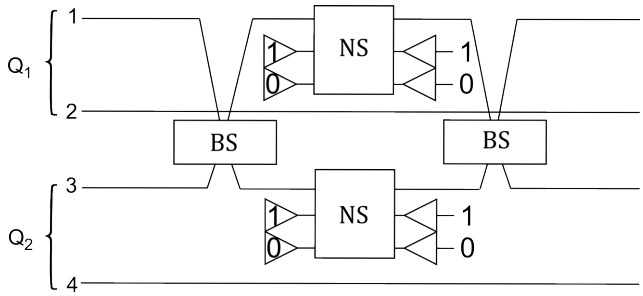

Figure 2: Image taken from KLM[?]. CZ is a two qubit gate. Qubits are dual-rail encoded. A NS gate is used on one mode from each qubit, so two are needed. Each of the NS gates need two ancilla modes, thus a total of 8 modes are used for this CZ gate. Qubit one ($|Q_1\rangle$) is found in modes 1 and 2, qubit two ($|Q_2\rangle$) is found in modes 3 and 4, while ancilla modes are 5 to 8. A "50:50" beam splitter is used on mode 1 which belongs to the first qubit and on mode 3 which belongs to the second qubit. An NS gate is then applied to each of these two modes. The first beam splitter is then undone using another "50:50" beam splitter.

which gives us

$$u_4u_1u_2u_3 = \begin{bmatrix} -0.414229709262441 & 0.840889626163301 & 0.348307887615681 \\ 0.840889626163301 & 0.500013782232149 & -0.207101072399072 \\ 0.348307887615681 & -0.207101072399072 & 0.914215927030292 \end{bmatrix} = M$$

□

2.3 Non-deterministic CZ gate

A non-deterministic CZ gate can be implemented using two NS gates as described in figure 2. Since the inputs are logical dual-rail encoded qubits, the possible combinations are $|00\rangle_L = |0101\rangle_{1234}$, $|01\rangle_L = |0110\rangle_{1234}$, $|10\rangle_L = |1001\rangle_{1234}$ and $|11\rangle_L = |1010\rangle_{1234}$. The beam splitter acts on modes 1 and 3, so that $\hat{a}_1^\dagger \to \frac{1}{\sqrt{2}}(\hat{a}_1^\dagger + \hat{a}_3^\dagger)$ and $\hat{a}_3^\dagger \to \frac{1}{\sqrt{2}}(\hat{a}_1^\dagger - \hat{a}_3^\dagger)$. There are four qubit combinations

$$\begin{aligned}
|00\rangle_L = |0101\rangle_{1234} = \hat{a}_2^\dagger\hat{a}_4^\dagger\,|vac\rangle &\xrightarrow{\mathrm{BS}^{(1,3)}} \hat{a}_2^\dagger\hat{a}_4^\dagger\,|vac\rangle = |0101\rangle_{1234} \\
|01\rangle_L = |0110\rangle_{1234} = \hat{a}_2^\dagger\hat{a}_3^\dagger\,|vac\rangle &\xrightarrow{\mathrm{BS}^{(1,3)}} \frac{1}{\sqrt{2}}\hat{a}_2^\dagger(\hat{a}_1^\dagger - \hat{a}_3^\dagger)\,|vac\rangle \\
&= \frac{1}{\sqrt{2}}(|1100\rangle_{1234} - |0110\rangle_{1234}) \\
|10\rangle_L = |1001\rangle_{1234} = \hat{a}_1^\dagger\hat{a}_4^\dagger\,|vac\rangle &\xrightarrow{\mathrm{BS}^{(1,3)}} \frac{1}{\sqrt{2}}(\hat{a}_1^\dagger + \hat{a}_3^\dagger)\hat{a}_4^\dagger\,|vac\rangle \\
&= \frac{1}{\sqrt{2}}(|1001\rangle_{1234} + |0011\rangle_{1234}) \\
|11\rangle_L = |1010\rangle_{1234} = \hat{a}_1^\dagger\hat{a}_3^\dagger\,|vac\rangle &\xrightarrow{\mathrm{BS}^{(1,3)}} \frac{1}{2}(\hat{a}_1^\dagger + \hat{a}_3^\dagger)(\hat{a}_1^\dagger - \hat{a}_3^\dagger)\,|vac\rangle
\end{aligned}$$

$$= \frac{1}{\sqrt{2}}(\frac{1}{\sqrt{2}}(\hat{a}_1^\dagger)^2 - \frac{1}{\sqrt{2}}(\hat{a}_3^\dagger)^2)\,|vac\rangle$$
$$= \frac{1}{\sqrt{2}}(|2000\rangle_{1234} + |0020\rangle_{1234})$$
$$= \frac{1}{\sqrt{2}}(|20\rangle_{13} + |02\rangle_{13})\,|vac\rangle$$

As the NS gate only flips the phase of the state $|2\rangle$, in the last case only it will produce an interesting change. The other 3 cases will not be changed and with application of the BS again, they will return to the original state (the matrix used for BS (Def. 1.1) is essentially a Hadamard, which is its own inverse). Thus

$$|00\rangle_L \xrightarrow{\text{BS}^{(1,3)},\text{NS}^{(1)},\text{NS}^{(3)},\text{BS}^{(1,3)}} |00\rangle_L$$
$$|01\rangle_L \xrightarrow{\text{BS}^{(1,3)},\text{NS}^{(1)},\text{NS}^{(3)},\text{BS}^{(1,3)}} |01\rangle_L$$
$$|10\rangle_L \xrightarrow{\text{BS}^{(1,3)},\text{NS}^{(1)},\text{NS}^{(3)},\text{BS}^{(1,3)}} |10\rangle_L$$
$$|11\rangle_L \xrightarrow{\text{BS}^{(1,3)}} \frac{1}{2}(\hat{a}_1^\dagger + \hat{a}_3^\dagger)(\hat{a}_1^\dagger - \hat{a}_3^\dagger)\,|vac\rangle = \frac{1}{\sqrt{2}}(\frac{1}{\sqrt{2}}(\hat{a}_1^\dagger)^2 - \frac{1}{\sqrt{2}}(\hat{a}_3^\dagger)^2)\,|vac\rangle$$
$$\xrightarrow{\text{NS}^{(1)},\text{NS}^{(3)}} = \frac{1}{\sqrt{2}}(\frac{1}{\sqrt{2}}(\hat{a}_3^\dagger)^2 - \frac{1}{\sqrt{2}}(\hat{a}_1^\dagger)^2)\,|vac\rangle = \frac{1}{2}(\hat{a}_3^\dagger + \hat{a}_1^\dagger)(\hat{a}_3^\dagger - \hat{a}_1^\dagger)\,|vac\rangle$$
$$\xrightarrow{\text{BS}^{(1,3)}} \frac{1}{2}\left(\frac{1}{\sqrt{2}}(\hat{a}_3^\dagger - \hat{a}_1^\dagger) + \frac{1}{\sqrt{2}}(\hat{a}_3^\dagger + \hat{a}_1^\dagger)\right)\left(\frac{1}{\sqrt{2}}(\hat{a}_3^\dagger - \hat{a}_1^\dagger) - \frac{1}{\sqrt{2}}(\hat{a}_3^\dagger + \hat{a}_1^\dagger)\right)|vac\rangle$$
$$= \frac{1}{2}\frac{2}{\sqrt{2}}\hat{a}_3^\dagger\frac{2}{\sqrt{2}}(-\hat{a}_1^\dagger)\,|vac\rangle = -\hat{a}_3^\dagger\hat{a}_1^\dagger\,|vac\rangle = -\,|11\rangle_L$$

From here it follows that for any $|Q_1\rangle = \alpha_0\,|0\rangle_L + \alpha_1\,|1\rangle_L$ and $|Q_2\rangle = \beta_0\,|0\rangle_L + \beta\,|1\rangle_L$, applying CZ as defined above gives

$$\begin{aligned} CZ\,|Q_1\rangle\,|Q_2\rangle &= CZ(\alpha_0\beta_0\,|00\rangle_L + \alpha_0\beta_1\,|01\rangle_L + \alpha_1\beta_0\,|10\rangle_L + \alpha_1\beta_1\,|11\rangle_L) \\ &= \alpha_0\beta_0 CZ\,|00\rangle_L + \alpha_0\beta_1 CZ\,|01\rangle_L + \alpha_1\beta_0 CZ\,|10\rangle_L + \alpha_1\beta_1 CZ\,|11\rangle_L \\ &= \alpha_0\beta_0\,|00\rangle_L + \alpha_0\beta_1\,|01\rangle_L + \alpha_1\beta_0\,|10\rangle_L - \alpha_1\beta_1\,|11\rangle_L \end{aligned}$$

which is exactly the expected behaviour of a CZ gate.

Since NS is used twice, the probability of success for this gate will be $P(\text{CZ}) = P(\text{NS}) \cdot P(\text{NS}) = \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}$.

2.4 Teleportation and near-deterministic CZ gate

In the next section of KLM, they increase the success rate of CZ to $\frac{1}{4}$ using teleportation. The idea stems from the "teleportation trick" introduced by Gottesman and Chuang [**?**]. In their paper, they use the properties of the Clifford group to move the entanglement gate from entangling input qubits to entangling ancilla states instead. By applying CNOT or CZ on the ancillas, the entanglement can now be teleported through to the input qubits that were supposed to get entangled (see A.4).

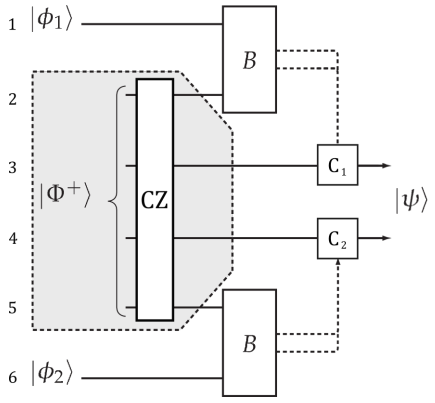

Figure 3: Based on an image taken from Pieter Kok lectures [**?**]. Qubits $|\phi_1\rangle$ (rail 1) and $|\phi_2\rangle$ (rail 6) are single-rail encoded. In the dual-rail KLM scheme, only one of the two physical qubits per logical qubit would be teleported. The physical qubits in rails 2 and 3 need to be entangled, same is true for 4 and 5. If this was just a teleportation of qubits from rail 1 and 6 to 3 and 4, there would be no CZ gate in the shaded region. Since a CZ gate is being teleported in, the dual-rail logical ancilla qubits are entangled using a dual-rail CZ gate, forming a Bell state $|\Phi^+\rangle$. Bell measurement is taken in the box marked as B. This is a single-rail Bell measurement (otherwise four modes would be coming into the box from the left). Depending on the outcome of the measurement, corrections C_1 and C_2 are applied to the qubits which were not measured. In the case of the KLM CZ gate, the correction is a phase shifter P_{180° [**?**].

Unfortunately, this doesn't mean that we can apply a CZ gate deterministically. Teleportation protocol requires a Bell measurement to be performed, and in linear optics we can only be certain about which Bell state we have in 2 out of 4 Bell states, thus giving the probability of successful measurement to be $\frac{1}{2}$. This teleportation has to be done twice, once per qubit, thus the total probability of CZ gate being teleported is $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.

This is still not high enough probability and KLM paper goes a step further, expanding the idea of Bell measurement on two qubits, to a Bell measurement on $n+1$ qubits.

Lemma 2.3. *Let* $|t_n\rangle = \frac{1}{\sqrt{n+1}} \sum_{j=0}^{n} |1\rangle^j |0\rangle^{n-j} |0\rangle^j |1\rangle^{n-j}$ *and* $\alpha|0\rangle + \beta|1\rangle$ *be the state that is being teleported. Let* BM_n *be a Bell measurement involving the mode of the state to be teleported and the first* n *modes of* $|t_n\rangle$*, implemented using an* $(n+1)$*-point discrete quantum Fourier transform. Let modes* 0 *to* n *be measured and* k *be the number of photons that has been detected. Then*

- *If* $0 < k < n+1$ *then the teleported state appears in mode* $n+k$ *and only needs to be corrected by applying a phase shift. Modes* $2n-l$ *are in state* 1 *for* $0 \leq l \leq (n-k)$ *and for* $n-k < l < n$ *are in state 0.*
- *If* $k = 0$*, the input state has been measured and projected to* $|0\rangle_0$
- *If* $k = n+1$*, the input state has been measured and projected to* $|1\rangle_0$

Proof. Let $|t_n^{(j)}\rangle = |1\rangle^j |0\rangle^{n-j} |0\rangle^j |1\rangle^{n-j}$. Then, this state contains j photons in its first n modes. Also $|t_n\rangle = \frac{1}{\sqrt{n+1}} \sum_{j=0}^{n} |t_n^{(j)}\rangle$. The discrete quantum Fourier transform $\hat{F}_{n+1}$ in matrix notation

is

$$(F_{n+1})_{jk} = \frac{1}{n+1}\omega^{kl} \text{ , where } \omega = \exp(\frac{2\pi i}{n+1}) \text{ and } k,l \in \{1,...,n\} \tag{1}$$

QFT $\hat{F}_{n+1}$ can be considered to be n mode generalization of the Hadamard matrix or in quantum optics terms, the beam splitter.

Full input state is $|\phi\rangle = (\alpha|0\rangle + \beta|1\rangle)|t_n\rangle$. Let the number of photons detected in mode j be n_j so that $\sum_j n_j = k$. QFT can be implemented using passive optical elements [**?**]. From the definition of passive optical elements it follows that if the photon number after QFT is k then it was also k beforehand. Further, it is only applied to the input state and first n photons of $|t_n\rangle$, thus

$$\begin{aligned}\hat{F}_{n+1}|\phi\rangle &= \hat{F}_{n+1}(\alpha|0\rangle + \beta|1\rangle)\frac{1}{\sqrt{n+1}}\sum_{j=0}^{n}|t_n^{(j)}\rangle \\ &= \frac{1}{\sqrt{n+1}}\sum_{j=0}^{n}F_{n+1}(\alpha|0\rangle + \beta|1\rangle)|t_n^{(j)}\rangle\end{aligned}$$

Knowing that the state after the measurement had k photons, the collapsed state must either have had one photon from the input state and $(k-1)$ photons from $|t_n\rangle$ (which only corresponds to the term $|t_n^{(k-1)}\rangle$ in the superposition), or there were no photons from the input state and k photons from $|t_n\rangle$ (which only corresponds to the term $|t_n^{(k)}\rangle$ in the superposition).

• If k photons have been measured where $k \neq 0$ and $k \neq n+1$, then the state that had k photons before the measurement will be of interest

$$\begin{aligned}|\phi_k\rangle = \hat{F}_{n+1}\beta|1\rangle|t_n^{(k-1)}\rangle + \hat{F}_{n+1}\alpha|0\rangle|t_n^{(k)}\rangle = \hat{F}_{n+1}\beta|1\rangle|1\rangle^{k-1}|0\rangle^{n-k+1}|0\rangle^{k-1}|1\rangle^{n-k+1} \\ + \hat{F}_{n+1}\alpha|0\rangle|1\rangle^{k}|0\rangle^{n-k}|0\rangle^{k}|1\rangle^{n-k}\end{aligned}$$

Denote the two states in the superposition as follows

$$|\phi_k^{(\alpha)}\rangle = \hat{F}_{n+1}\alpha|0\rangle|1\rangle^{k}|0\rangle^{n-k}|0\rangle^{k}|1\rangle^{n-k} \tag{2}$$

$$|\phi_k^{(\beta)}\rangle = \hat{F}_{n+1}\beta|1\rangle|1\rangle^{k-1}|0\rangle^{n-k+1}|0\rangle^{k-1}|1\rangle^{n-k+1} \tag{3}$$

Isolate the first $n+1$ modes of the state $|\phi_k^{(\beta)}\rangle$ and apply $\hat{F}_{n+1}$

$$\begin{aligned}|\phi_k^{(\beta)}\rangle_{0\ldots n} = \hat{F}_{n+1}\beta|1\rangle|1\rangle^{k-1}|0\rangle^{n-k+1} = \hat{F}_{n+1}\beta\prod_{j=0}^{k}\hat{a}_j^{\dagger}|vac\rangle \\ = \beta\prod_{j=0}^{k}\left(\sum_{l=0}^{n+1}\frac{1}{\sqrt{n+1}}w^{jl}\hat{a}_l^{\dagger}\right)\end{aligned}$$

Let a phase shift $\mathrm{P}_{2\pi l/(n+1)} = \exp(i2\pi l/(n+1)) = w^l$ be applied to all modes l such that $0 \leq l \leq n$. Define the tensor product of all the phase shifts applied to be $\hat{\mathbf{P}}$. Then $\hat{a}_l^{\dagger} \to \exp(i2\pi l/(n+1))\hat{a}_l^{\dagger}$ and

$$
\begin{aligned}
\hat{\mathbf{P}}\left|\phi_k^{(\beta)}\right\rangle_{0\ldots n} &= \hat{\mathbf{P}}\beta\prod_{j=0}^{k}\left(\sum_{l=0}^{n+1}\frac{1}{\sqrt{n+1}}w^{jl}\hat{a}_l^\dagger\right)|vac\rangle && \text{(use definition of } \hat{\mathbf{P}}\text{)}\\
&= \beta\prod_{j=0}^{k}\left(\sum_{l=0}^{n+1}\frac{1}{\sqrt{n+1}}w^{jl}e^{i2\pi l/(n+1)}\hat{a}_l^\dagger\right)|vac\rangle && \text{(use definition of } w\text{)}\\
&= \beta\prod_{j=0}^{k}\left(\sum_{l=0}^{n+1}\frac{1}{\sqrt{n+1}}e^{i2\pi jl/(n+1)}e^{i2\pi l/(n+1)}\hat{a}_l^\dagger\right)|vac\rangle\\
&= \beta\prod_{j=0}^{k}\left(\sum_{l=0}^{n+1}\frac{1}{\sqrt{n+1}}e^{i2\pi (j+1)l/(n+1)}\hat{a}_l^\dagger\right)|vac\rangle && \text{(change product index)}\\
&= \beta\prod_{j=1}^{k+1}\left(\sum_{l=0}^{n+1}\frac{1}{\sqrt{n+1}}e^{i2\pi jl/(n+1)}\hat{a}_l^\dagger\right)|vac\rangle && \text{(use definition of } w\text{)}\\
&= \beta\prod_{j=1}^{k+1}\left(\sum_{l=0}^{n+1}\frac{1}{\sqrt{n+1}}w^{jl}\hat{a}_l^\dagger\right)|vac\rangle && \text{(use definition of QFT)}\\
&= \hat{F}_{n+1}\beta\,|0\rangle\,|1\rangle^k\,|0\rangle^{n-k} && (4)
\end{aligned}
$$

Notice that the end state in equation (4) matches the first $n+1$ modes of $|\phi_k^{(\alpha)}\rangle$ from equation (2) (upto constants α and β which would give a global phase difference). The measurement is made in the number basis on each of the $n+1$ modes individually. So if two states have the same number of photons in a mode, but differ by a phase, then they will not be distinguishable. Since $\hat{\mathbf{P}}$ introduced only a phase shift to each of the $n+1$ modes to make a state $\hat{\mathbf{P}}\left|\phi_k^{(\beta)}\right\rangle_{0\ldots n}$ and the states are being measured in the number basis, the measurement will not be able to distinguish it from $\hat{\mathbf{P}}\left|\phi_k^{(\alpha)}\right\rangle_{0\ldots n}$.

Let the state after QFT be described in the most general possible superposition of states in $n+1$ modes, $|\phi_k\rangle = \left(\sum_{m=0}^{M} C_{n_0^m,\ldots,n_n^m}\,|n_0^m,\ldots,n_n^m\rangle\right)|\phi_k\rangle_{n+1\ldots 2n}$, where n_j gives the number of detected photons in j-th mode and M is the number of different combinations of $(n_0,\ldots,n_n)$ for which $\sum_{j=0}^{n} n_j = k$. Then what the previous argument shows is that $C_{n_0^m,\ldots,n_n^m} = (\alpha\,|\phi_k^\alpha\rangle_{n+1\ldots 2n} + G_{n_0^m,\ldots,n_n^m}\beta\,|\phi_k^\beta\rangle_{n+1\ldots 2n})$, where $G_{n_0^m,\ldots,n_n^m}$ is the phase between the appropriate states in $\left|\phi_k^{(\alpha)}\right\rangle_{0\ldots n}$ and $\left|\phi_k^{(\beta)}\right\rangle_{0\ldots n}$ (the appropriate states are $\langle n_0,\ldots,n_n|\phi_k^{(\alpha)}\rangle_{0\ldots n}\,|n_0,\ldots,n_n\rangle_{0\ldots n}$ and $\langle n_0,\ldots,n_n|\phi_k^{(\beta)}\rangle_{0\ldots n}\,|n_0,\ldots,n_n\rangle_{0\ldots n}$). The phase shift can be calculated easily due to definition of $\hat{\mathbf{P}}$. For n_j photons in j-th mode, the phase shift would accrue to w^{ln_j}. Combining this for all modes gives a total phase shift of $G_{n_0^m,\ldots,n_n^m} = \prod_{l=0}^{n} w^{ln_l}$.

Let the result of the measurement of the first $n+1$ modes be $(n_0,\ldots,n_n)$. Let $G := G_{n_0,\ldots,n_n}$. The state after measurement is then

$$
\begin{aligned}
\langle n_0,\ldots,n_n|_{0\ldots n}\,|\phi_k\rangle &= \alpha\,|0\rangle^k\,|1\rangle^{n-k} + G\beta\,|0\rangle^{k-1}\,|1\rangle^{n-k+1}\\
&= \alpha\,|0\rangle^{k-1}\,|0\rangle\,|1\rangle^{n-k} + G\beta\,|0\rangle^{k-1}\,|1\rangle\,|1\rangle^{n-k}\\
&= |0\rangle^{k-1}\,(\alpha\,|0\rangle + G\beta\,|1\rangle)\,|1\rangle^{n-k}
\end{aligned}
$$

$$= \bigotimes_{i=n+1}^{n+k-1} |0\rangle (\alpha |0\rangle + G\beta |1\rangle) \bigotimes_{i=n+k+1}^{2n} |1\rangle$$

Applying a phase shift operation to mode $n+k$ retrieves the teleported input state in mode $n+k$ and this proves first part of the claim.

• If $k = 0$ then the state from the superposition must have been

$$|\phi_0\rangle = \frac{1}{\sqrt{n+1}} \hat{F}_{n+1} \alpha |0\rangle |t_n^0\rangle$$

and the resulting state is just $|0\rangle^n$. As $\hat{F}_{n+1}$ is unitary, then $P(\frac{1}{\sqrt{n+1}} \hat{F}_{n+1} \alpha |0\rangle |t_n^0\rangle) = P(\frac{1}{\sqrt{n+1}} \alpha |0\rangle |t_n^0\rangle) = \| \frac{1}{\sqrt{n+1}} \alpha |0\rangle |t_n^0\rangle \| = \frac{|\alpha|^2}{n+1}$.

• If $k = n+1$ then the state from the superposition must have been

$$|\phi_{n+1}\rangle = \frac{1}{\sqrt{n+1}} \hat{F}_{n+1} \beta |1\rangle |t_n^n\rangle$$

and the resulting state is just $|1\rangle^n$. Similarly to above $P(\frac{1}{\sqrt{n+1}} \hat{F}_{n+1} \beta |1\rangle |t_n^n\rangle) = \frac{|\beta|^2}{n+1}$.

As the last two cases are failure cases, the probability of success for teleportation is

$$P(\text{success}) = 1 - P(\text{failure}) = 1 - \frac{|\alpha|^2}{n+1} - \frac{|\beta|^2}{n+1} = 1 - \frac{|\alpha|^2 + |\beta|^2}{n+1} = 1 - \frac{1}{n+1} = \frac{n}{n+1}$$

□

This Lemma proves that near-deterministic Bell measurement can be done. For the same reasons as before, CZ gate can be commuted and applied to two of these new ancilla states $|t_n\rangle$. The full ancilla is then $|t_n^{(1)}\rangle |t_n^{(2)}\rangle = \sum_{i,j=0}^{n} |1\rangle^j |0\rangle^{n-j} |0\rangle^j |1\rangle^{n-j} \times |1\rangle^i |0\rangle^{n-i} |0\rangle^i |1\rangle^{n-i}$. CZ is now applied to pairs of modes $(n+i, 3n+j)$ where $i, j \in \{1, ..., n\}$. Notice the effect of CZ on a state for specific i and j, $|1\rangle^j |0\rangle^{n-j} |0\rangle^j |1\rangle^{n-j} \times |1\rangle^i |0\rangle^{n-i} |0\rangle^i |1\rangle^{n-i}$. It will only incur a phase flip when both modes have one photon in them, which will be true for CZ applications on $(n+j, ..., 2n)$ and $(n+i, ..., 2n)$, which gives $(n-j)(n-i)$ different combinations where the flip occurs. Thus the resulting state is

$$|t_n'\rangle = \sum_{i,j=0}^{n} (-1)^{(n-j)(n-i)} |1\rangle^j |0\rangle^{n-j} |0\rangle^j |1\rangle^{n-j} \times |1\rangle^i |0\rangle^{n-i} |0\rangle^i |1\rangle^{n-i} \tag{5}$$

2.5 Scalability discussion

The idea behind the KLM paper is summarized well by Scott Aaronson [**?**].

Theorem 2.4. *BosonP$_{adap}$*[¶] *= BQP*

[¶]As defined by Scott Aaronson[**?**]: "Define BosonP$_{\text{adap}}$ to be the class of languages that are decidable in BPP, augmented with the ability to prepare k-photon state (for any $k = \text{poly}(n)$) in any of $m = \text{poly}(n)$ modes; apply arbitrary optical elements to pairs of modes measure the photon number of any mode at any time; and condition future optical elements and classical computations on the outcomes of the measurements."

Proof. (Informal) The ancilla states $|t_n\rangle$ can be produced using $O(n)$ gates. The state given in equation 5 can be generated using $O(n^2)$ CZ gates and two ancilla states $|t_n\rangle$ which gives a total of $O(n^2)$ gates and this is polynomial in resources. QFT can be implemented using fast Fourier transform in $O(n \log(n))$ gates [**?**]. This will give a scalable CZ gate which succeeds with probability $\frac{n^2}{(n+1)^2}$. Single qubit gates were already mentioned to be implementable in $O(1)$ resources (section 1.2). This gives a full set of universal gates that can be performed in polynomial resources on a dual-rail LOQC. Then all the algorithms which can be computed in polynomial resources using the universal set of gates, can also be computed in polynomial time on this architecture. Thus any algorithm that belongs to BQP also belongs to $\text{BosonP}_{\text{adap}}$. □

There are various ways in which the failure of a gate or Bell measurement can be used to error correct. There are also codes that can protect against losing some of the information when a gate is applied, thus the gates can actually be tried more times without the need to make them near deterministic. Some of these techniques are mentioned in KLM but this is not the focus of this essay.

After KLM scheme, various simplifications of the NS gate were proposed, either reducing the number of ancillas or beam splitters or with higher success probability. For example an implementation of a CZ gate with success rate $\frac{2}{27}$ [**?**] exists. Another result worth mentioning is from 2006, by Spedalieri et al. [**?**] in which the teleported qubits are dual-rail polarization qubits (advantages of dual-rail in section 1.2).

3 Cluster state computing

The KLM scheme proved it was scalable from the perspective of information theory, but from an implementation perspective, the resource overhead is large and the level of control required to keep this very large interferometer stable is high. Further, the depth of the quantum circuit is a problem considering the exponential loss of photons, ideally it should not be more than $O(1)$ and in KLM it is poly(n).

The search for new, improved schemes continued, most prominently Raussendorf and Briegel introduced one-way or measurement based quantum computing (MBQC) using cluster states in 2001 [**?**].

3.1 Cluster states

Definition 3.1. Let $G(V, E)$ be some graph, where E is the set of edges of the graph and V is the set of vertices. Denote a Pauli-X performed on some single qubit a as $\sigma_x^{(a)}$ and similarly, a Pauli-Z as $\sigma_z^{(a)}$ and Pauli-Y as $\sigma_y^{(a)}$. A graph state is then a pure quantum state of qubits represented by the vertices of G that obeys the following eigenvalue equation for every node a in the graph

$$K^{(a)} |\phi_\kappa\rangle_V = (-1)^{\kappa_a} |\phi_\kappa\rangle_V \tag{6}$$

where $\kappa := \{\kappa_a \in \{0, 1\} | a \in V\}$ and $K^{(a)} := \sigma_x^{(a)} \bigotimes_{b \in n(a)} \sigma_z^{(b)}$, a is a node in the cluster and $n(a)$ is the set of all nodes in the graph connected to a as defined by this graph's adjacency matrix.

Note that while $K^{(a)}$ is an operator defined for node a, it acts on the qubit a as well as its neighbouring qubits.

Lemma 3.1. *Let a graph $G'(V', E')$ be such that all of its vertices are disconnected and initialized to $|+\rangle$ and $\kappa_a = 0, \forall a \in V$. Then this graph represents a graph state and applying CZ gates to pairs of vertices of this graph creates a collection of graph states $\{G(V, E)|V = V'\}$ and $(a, b) \in E$ if CZ has been applied to qubits $a \in V$ and $b \in V$.*[‖]

Proof. Let $G'(V', E')$ such that all of its vertices are disconnected, thus $E' = \emptyset$. Notice that for all $a \in V'$ we get $K^{(a)} = \sigma_x^{(a)} \bigotimes_{b \in n(a)} \sigma_z^{(b)} = \sigma_x^{(a)}$ since there is no connection between the qubits ($n(a) = \emptyset, \forall a \in V'$). Let $|+\rangle_{V'} := \bigotimes_{a \in V'} |+\rangle_a$, where $|+\rangle_a$ denotes that qubit a is in state $|+\rangle$. This is indeed a graph state since $K^{(a)} |+\rangle_{V'} = \sigma_x^{(a)} |+\rangle_{V'} = |+\rangle_{V'}$, and we see that $\kappa_a = 0$ for all $a \in V'$.

Applying a CZ gate to a pair of vertices a and b can be described as $S^{(ab)} = |0\rangle_a \langle 0| \otimes I^{(b)} + |1\rangle_a \langle 1| \otimes \sigma_z^{(b)}$ for $a \neq b$. Let $G(V, E)$ be a graph constructed from G' by application of CZ gates such that $V = V'$ (but now E might not be empty). Let $S := S^G = \prod_{(b,c) \in E} S^{(bc)}$. Applying S to state $|+\rangle_V$, gives a state $|\phi\rangle_V = S |+\rangle_V$. Further $S^\dagger |\phi\rangle_V = S^\dagger S |+\rangle_V = |+\rangle_V$ (since S is unitary as it is a product of unitaries). Since $\sigma_x^{(a)} |+\rangle_{V'} = |+\rangle_{V'}$ and $V' = V$ then $S\sigma_x^{(a)} S^\dagger |\phi\rangle_V = S\sigma_x^{(a)} |+\rangle_V = S |+\rangle_V = |\phi\rangle_V$. Observe 3 distinct vertices a, b and c, and how $S^{(ab)}$ affects $\sigma_x^{(a)}$, $\sigma_x^{(b)}$ and $\sigma_x^{(c)}$. In the case of $S^{(ab)} \sigma_x^{(c)} (S^{(ab)})^\dagger$, since the vertices are distinct, $S^{(ab)}$ and $\sigma_x^{(c)}$ will commute, thus

$$S^{(ab)} \sigma_x^{(c)} (S^{(ab)})^\dagger = \sigma_x^{(c)} S^{(ab)} (S^{(ab)})^\dagger = \sigma_x^{(c)} \tag{7}$$

In the case of $S^{(ab)} \sigma_x^{(a)} (S^{(ab)})^\dagger$

$$\begin{aligned}
S^{(ab)} \sigma_x^{(a)} (S^{(ab)})^\dagger =& \Big(|0\rangle_a \langle 0| \otimes I^{(b)} + |1\rangle_a \langle 1| \otimes \sigma_z^{(b)} \Big) \\
& \Big(|0\rangle_a \langle 1| + |1\rangle_a \langle 0| \Big) \Big(|0\rangle_a \langle 0| \otimes I^{(b)} + |1\rangle_a \langle 1| \otimes \sigma_z^{(b)} \Big) \\
=& \Big(|0\rangle_a \langle 1| \otimes I^{(b)} + |1\rangle_a \langle 0| \otimes \sigma_z^{(b)} \Big) \Big(|0\rangle_a \langle 0| \otimes I^{(b)} + |1\rangle_a \langle 1| \otimes \sigma_z^{(b)} \Big) \\
=& (|0\rangle_a \langle 1| \otimes I^{(b)})(|0\rangle_a \langle 0| \otimes I^{(b)}) + (|0\rangle_a \langle 1| \otimes I^{(b)})(|1\rangle_a \langle 1| \otimes \sigma_z^{(b)}) + \\
& (|1\rangle_a \langle 0| \otimes \sigma_z^{(b)})(|0\rangle_a \langle 0| \otimes I^{(b)}) + (|1\rangle_a \langle 0| \otimes \sigma_z^{(b)})(|1\rangle_a \langle 1| \otimes \sigma_z^{(b)}) \\
=& (|0\rangle_a \langle 1| \otimes \sigma_z^{(b)}) + (|1\rangle_a \langle 0| \otimes \sigma_z^{(b)}) = (|0\rangle_a \langle 1| + |1\rangle_a \langle 0|) \otimes \sigma_z^{(b)} \\
=& \sigma_x^{(a)} \otimes \sigma_z^{(b)}
\end{aligned} \tag{8}$$

In the case of $S^{(ab)} \sigma_x^{(b)} (S^{(ab)})^\dagger$

$$\begin{aligned}
S^{(ab)} \sigma_x^{(b)} (S^{(ab)})^\dagger =& (|0\rangle_a \langle 0| \otimes \sigma_x^{(b)} + |1\rangle_a \langle 1| \otimes (\sigma_z^{(b)} \sigma_x^{(b)}))(|0\rangle_a \langle 0| \otimes I^{(b)} + |1\rangle_a \langle 1| \otimes \sigma_z^{(b)}) \\
=& |0\rangle_a \langle 0| \otimes \sigma_x^{(b)} + |1\rangle_a \langle 1| \otimes (\sigma_z^{(b)} \sigma_x^{(b)} \sigma_z^{(b)}) \\
=& |0\rangle_a \langle 0| \otimes \sigma_x^{(b)} + |1\rangle_a \langle 1| \otimes \sigma_x^{(b)} \\
=& \sigma_z^{(a)} \otimes \sigma_x^{(b)}
\end{aligned}$$

It is similarly easy to check that $S^{(ab)} \sigma_z^{(c)} (S^{(ab)})^\dagger = \sigma_z^{(c)}$ for any $c \in V$.

[‖] Adapted from [**?**]

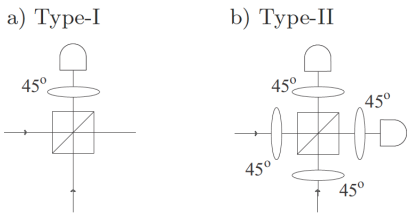

Figure 4: Taken from Browne and Rudolph [**?**]. (a) Type-I fusion gate. It consists of a PBS and a half wave plate. One output arm of the PBS is detected with a polarization discriminating detector. (b) Type-II fusion gate. It consists of 4 half-wave plates (two in the input modes, two in the output modes) and a PBS. Both output arms of the PBS are detected with polarization discriminating detectors.

For a given a it can now be shown using the above and the fact that all $S^{(bc)}$ in S commute that $S\sigma_x^{(a)}S^\dagger = \bigotimes_{b\in n(a)} \sigma_z^{(b)} \otimes \sigma_x^{(a)} = K^{(a)}$ (for more detailed calculation see appendix A.5.1).

Thus from the definition 3.1 it follows that this state is also a graph state with $\kappa_a = 0, a \in V$. □

This collection of states is used when cluster state computing is being discussed. Notice the consequences of Pauli measurements and how the states are usually depicted in figure 5.

"Cluster state" is sometimes used interchangeably with "graph state", but by the definitions of Raussendorf [**?**], graph state is a generalization of cluster state. Cluster states are restricted to connected subsets of a simple d-dimensional lattice.

In 2003, Yoran and Reznik had some ideas similar to cluster state model [**?**]. They used the polarization and path qubit encoding. To add a link between two qubits in their chain model, they did not need a CZ gate to be near-deterministic, but just succeed with a probability higher than $\frac{1}{2}$ so the expected value of the change in number of qubits in a chain would be positive. This was followed by a more cluster-like proposal by Nielsen in 2004. Taking the KLM scheme implementation of a CZ gate, the paper presents a cluster state equivalent and thus shows that MBQC is universal. He introduces the idea of using failure cases (of gates) to save resources. He also describes microclusters which are located at the end of our larger cluster state. These microclusters are small and have a lot of links, thus further protecting from failure affecting the rest of the cluster. This implementation removes the need for any physical depth for photons to go through, present in KLM [**?**].

The scheme presented here is the Browne-Rudolph scheme in which they introduce two fusion gates, so called because they can fuse clusters together [**?**]. Essentially, they are parity gates, but Browne and Rudolph have an innovative way of applying them and give insight on how they can be used for cluster computing.

3.2 Type-I fusion gate

Type-I fusion gate consists of a PBS and a half wave plate (polarization beam splitter) and a detector (see figure 4). The most general state impinging on a Type-I fusion gate is $|\phi_a\rangle |\phi_b\rangle = (\alpha_H \hat{a}_H^\dagger + \alpha_V \hat{a}_V^\dagger)(\beta_H \hat{b}_H^\dagger + \beta_V \hat{b}_V^\dagger)$. This transforms as follows

$$
\begin{aligned}
|\phi_a\rangle |\phi_b\rangle &\xrightarrow{\text{PBS}} (\alpha_H \hat{a}_H^\dagger + \alpha_V \hat{b}_V^\dagger)(\beta_H \hat{b}_H^\dagger + \beta_V \hat{a}_V^\dagger) \\
&\xrightarrow{BS} (\alpha_H \hat{a}_H^\dagger + \alpha_V \frac{1}{\sqrt{2}}(\hat{b}_H^\dagger - \hat{b}_V^\dagger))(\beta_H \frac{1}{\sqrt{2}}(\hat{b}_H^\dagger + \hat{b}_V^\dagger) + \beta_V \hat{a}_V^\dagger) \\
&= \frac{\alpha_H \beta_H}{2}\sqrt{2}\hat{a}_H^\dagger \hat{b}_H^\dagger + \frac{\alpha_H \beta_H}{2}\sqrt{2}\hat{a}_H^\dagger \hat{b}_V^\dagger + \alpha_H \beta_V \hat{a}_H^\dagger \hat{a}_V^\dagger \\
&+ \frac{\alpha_V \beta_H}{2}(\hat{b}_H^\dagger)^2 - \frac{\alpha_V \beta_H}{2}(\hat{b}_V^\dagger)^2 + \frac{\alpha_V \beta_V}{2}\sqrt{2}\hat{a}_V^\dagger \hat{b}_H^\dagger - \frac{\alpha_V \beta_V}{2}\sqrt{2}\hat{a}_V^\dagger \hat{b}_V^\dagger \\
&= \frac{1}{\sqrt{2}}(\alpha_H \beta_H \hat{a}_H^\dagger + \alpha_V \beta_V \hat{a}_V^\dagger)\hat{b}_H^\dagger + \frac{1}{\sqrt{2}}(\alpha_H \beta_H \hat{a}_H^\dagger - \alpha_V \beta_V \hat{a}_V^\dagger)\hat{b}_V^\dagger \\
&+ \alpha_H \beta_V \hat{a}_H^\dagger \hat{a}_V^\dagger + \frac{\alpha_V \beta_H}{2}\left((\hat{b}_H^\dagger)^2 - (\hat{b}_V^\dagger)^2\right)
\end{aligned}
$$

The following cases can happen on the detector:

- one H polarized photon was detected then state is collapsed to $\frac{1}{\sqrt{2}}(\alpha_H \beta_H \hat{a}_H^\dagger + \alpha_V \beta_V \hat{a}_V^\dagger)$
- one V polarized photon was detected then state is collapsed to $\frac{1}{\sqrt{2}}(\alpha_H \beta_H \hat{a}_H^\dagger - \alpha_V \beta_V \hat{a}_V^\dagger)$
- no photons were detected then state is collapsed $\alpha_H \beta_V \hat{a}_H^\dagger \hat{a}_V^\dagger$
- two photons of H or V polarization where detected then state is collapsed to $\frac{\alpha_V \beta_H}{2}$

The first two cases are the interesting ones, since the result is a combination of the input states from both arms of the beam splitter with the output arm. If this gate was applied to couple of qubits which are both attached to a cluster, then in the first two outcomes the clusters would combine with the new qubit in mode $\hat{a}^\dagger$ inheriting all of the connections. This also means that a photon would be lost due to detection, thus the two combined clusters would have a qubit less.

Kraus operators that describe the success cases are

$$
\begin{aligned}
U_{typeI}^{(H)} &= \frac{1}{\sqrt{2}}(|H\rangle \langle HH| + |V\rangle \langle VV|) \\
U_{typeI}^{(V)} &= \frac{1}{\sqrt{2}}(|H\rangle \langle HH| - |V\rangle \langle VV|)
\end{aligned}
$$

Lemma 3.2. *Let $|\phi_A\rangle$ and $|\phi_B\rangle$ be two clusters of size n_A and n_B respectively. Let qubit a from cluster A and qubit b from cluster B interact on Type-I fusion gate. In the case of gate success, marked by detection of a single H or V polarized photon at the detector, the two clusters are joined into a single cluster state $|\phi\rangle$ of size $n_A + n_B - 1$. In the case of gate failure, both clusters stay separate and lose a qubit.*[**]

Proof can be found in appendix A.5.2.

3.3 Type-II fusion gate

One dimensional cluster states are not enough for universal quantum computation [**?**]. To create vertical links between 1D clusters, another type of fusion gate is introduced.

[**] Adapted from [**?**]

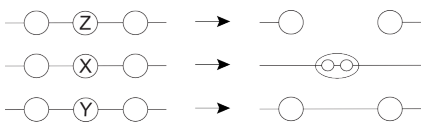

Figure 5: Taken from Browne and Rudolph [**?**]. Effects of the Pauli measurements on a linear cluster [**?**] [**?**]. Circles denote qubits in $|+\rangle$ state. Edges denote that a CZ gate was applied to the two qubits it connects. Taking a Z measurement on a qubit in a cluster removes the qubit from the cluster as well as its edges (1 qubit and 2 edges are lost from the graph). Taking an X measurement on a qubit in a cluster removes the measured qubit and redundantly encodes (see 1.2) the two neighbouring qubits (1 qubit and 0 edges are lost). Taking a Y measurement removes the measured qubit and links its neighbours (1 qubit and 1 edge are lost).

In the failure case of a Type-I fusion gate, a Z-measurement is performed which is not desirable as it would destroy big clusters (figure 5). On the other hand, the result of an X measurement is a redundantly encoded qubit (figure 5). This can be used to the advantage of the scheme since, as mentioned in section 1.2, when measuring a redundantly encoded qubit, information is not lost (see figure 6). Type-II fusion gates for polarization encoded qubits are built out of a PBS which is rotated at 45° which is also equivalent to 4 polarization rotators (two on input modes, two on output modes) and a PBS. In the case of general input modes $|\phi_a^{(0)}\rangle = \left(\alpha_H \hat{a}_H^\dagger + \alpha_V \hat{a}_V^\dagger\right)|vac\rangle$ and $|\phi_b^{(0)}\rangle = \left(\beta_H \hat{b}_H^\dagger + \beta_V \hat{b}_V^\dagger\right)|vac\rangle$

$$
\begin{aligned}
|\phi_a^{(0)}\rangle\,|\phi_b^{(0)}\rangle \xrightarrow{BS_{a_H,a_V},BS_{b_H,b_V}} & \left(\alpha_H \frac{1}{\sqrt{2}}\left(\hat{a}_H^\dagger + \hat{a}_V^\dagger\right) + \alpha_V \frac{1}{\sqrt{2}}\left(\hat{a}_H^\dagger - \hat{a}_V^\dagger\right)\right) \\
& \left(\beta_H \frac{1}{\sqrt{2}}\left(\hat{b}_H^\dagger + \hat{b}_V^\dagger\right) + \beta_V \frac{1}{\sqrt{2}}\left(\hat{b}_H^\dagger + \hat{b}_V^\dagger\right)\right)|vac\rangle \\
\xrightarrow{PBS_{a_H,a_V,b_H,b_V}} & \frac{1}{2}\left(\alpha_H\left(\hat{a}_H^\dagger + \hat{b}_V^\dagger\right) + \alpha_V\left(\hat{a}_H^\dagger - \hat{b}_V^\dagger\right)\right) \\
& \left(\beta_H\left(\hat{b}_H^\dagger + \hat{a}_V^\dagger\right) + \beta_V\left(\hat{b}_H^\dagger - \hat{a}_V^\dagger\right)\right)|vac\rangle \\
\xrightarrow{BS_{a_H,a_V},BS_{b_H,b_V}} & \frac{1}{4}\left(\alpha_H\left(\hat{a}_H^\dagger + \hat{a}_V^\dagger + \hat{b}_H^\dagger - \hat{b}_V^\dagger\right) + \alpha_V\left(\hat{a}_H^\dagger + \hat{a}_V^\dagger - \hat{b}_H^\dagger + \hat{b}_V^\dagger\right)\right) \\
& \left(\beta_H\left(\hat{b}_H^\dagger + \hat{b}_V^\dagger + \hat{a}_H^\dagger - \hat{a}_V^\dagger\right) + \beta_V\left(\hat{b}_H^\dagger + \hat{b}_V^\dagger - \hat{a}_H^\dagger + \hat{a}_V^\dagger\right)\right)|vac\rangle
\end{aligned}
$$

Two photons are always detected, thus there are 8 detection patterns. Group the output states before the detection using symbolic programming

- two photons in mode a_H is possible for $\frac{1}{4}(\alpha_H + \alpha_V)(\beta_H - \beta_V)(\hat{a}_H^\dagger)^2\,|vac\rangle$
- two photons in mode a_V is possible for $\frac{1}{4} - (\alpha_H + \alpha_V)(\beta_H - \beta_V)(\hat{a}_V^\dagger)^2\,|vac\rangle$
- two photons in mode b_H is possible for $\frac{1}{4}(\alpha_H - \alpha_V)(\beta_H + \beta_V)(\hat{b}_H^\dagger)^2\,|vac\rangle$
- two photons in mode b_V is possible for $\frac{1}{4} - (\alpha_H - \alpha_V)(\beta_H + \beta_V)(\hat{b}_V^\dagger)^2\,|vac\rangle$
- one photon in mode a_H and one in b_H is possible for $\frac{1}{2}(\alpha_H\beta_H + \alpha_V\beta_V)\hat{a}_H^\dagger\hat{b}_H^\dagger\,|vac\rangle$
- one photon in mode a_V and one in b_V is possible for $\frac{1}{2}(\alpha_H\beta_H + \alpha_V\beta_V)\hat{a}_V^\dagger\hat{b}_V^\dagger\,|vac\rangle$

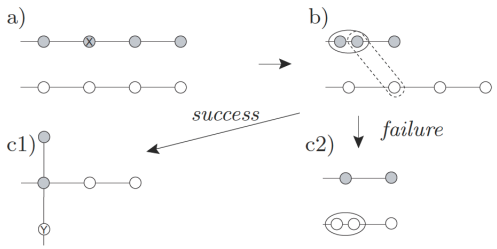

Figure 6: Image and caption taken from Browne and Rudolph [**?**]: "Here we illustrate the construction of the "L-shape": a) A σ_x-measurement causes the neighbouring qubits to be joined into a single logical qubit in the redundant encoding. b) Type II-fusion is now attempted between this logical qubit and a qubit in the lower cluster. The fusion succeeds with probability 1/2. c1) If the fusion succeeds, a single further σ_y measurement creates the desired L-shape (see Fig. 4c). c2) If it fails, a redundantly encoded qubit is created on the lower cluster. The qubits are now in a pattern similar to step b, so with the addition of two further qubits another Type-II fusion can be attempted. These steps are repeated until a successful fusion is accomplished. On average, creating the L-shape uses up 8 bonds from the linear clusters involved."

- one photon in mode a_H and one in b_V is possible for $\frac{1}{2}(\alpha_H\beta_V + \alpha_V\beta_H)\hat{a}_H^\dagger\hat{b}_V^\dagger\,|vac\rangle$
- one photon in mode a_V and one in b_H is possible for $\frac{1}{2}(\alpha_H\beta_V + \alpha_V\beta_H)\hat{a}_V^\dagger\hat{b}_H^\dagger\,|vac\rangle$

Notice that in the latter 4 cases the coefficients would correspond to an entangled state. Also, for this gate, if there are only have two photons, then there is actually no photon output. The Kraus operators for the four success cases are

$$U_{typeII}^{(HH)} = U_{typeII}^{(VV)} = \frac{1}{\sqrt{2}}(\langle HH| + \langle VV|)$$
$$U_{typeII}^{(HV)} = U_{typeII}^{(VH)} = \frac{1}{\sqrt{2}}(\langle HV| + \langle VH|)$$

Lemma 3.3. *Let $|\phi_A\rangle$ and $|\phi_B\rangle$ be two clusters of size n_A and n_B respectively. Let qubit a from cluster A and qubit b from cluster B interact on a Type-II fusion gate. In the case of gate success, marked by detection of a single H or V polarized photon at each of the detectors, the two clusters are joined into a single cluster state $|\phi\rangle$ of size $n_A + n_B - 2$. In the case of gate failure, both clusters stay separate and lose a qubit.*

Proof is similar to that of lemma 3.2 that can be found in appendix A.5.2.

3.4 Resource requirements

Let N be the size of a main cluster chain and n be the size of chain to be added. Let p be the probability of adding the chain to the main cluster chain and d_s be the number of lost qubits from the main cluster when the gate has succeeded and d_f be the number of lost qubits from the main cluster when the gate has failed [**?**]. Then the change in the size of the main chain $E[\Delta N] = p(n - d_s) - d_f(1 - p)$. This quantity needs to always be positive for the main cluster

chain to grow.

In Type-I fusion gate, the gate is successful $\frac{1}{2}$ of the time, and the main cluster loses a qubit no matter the outcome. This means that $0.5(n-1)-0.5 = 0.5n-1 > 0$ thus $n > 2$. So the size of the smaller chain being added has to be bigger than 2. For cluster of size 3 expected growth of the main chain is $E[\Delta N] = \frac{1}{2}$.

To get a 3 cluster state, two Bell states need to be connected. As a Type-I fusion gate succeeds with $\frac{1}{2}$, on average the gate needs to be attempted twice, and thus, on average 4 Bell states are used to get a 3 cluster state. To grow the main cluster by one qubit, a cluster of size 3 will need to be attached twice, which gives 8 Bell states needed to grow the cluster.

Further, in a Type-II fusion gate, success probability is $p = \frac{1}{2}$, succeeding removes 2 qubits, failing removes 1. Thus $0.5(n-2) - 0.5 = 0.5n - 1.5 > 0$ that is $n > 3$. A 4 qubit cluster state will be required for the chain to grow. This 4 qubit cluster state can be constructed using Type-I gate. A cluster that is a 3 cluster state, needed 4 Bell states to be constructed. Further, using Type-I fusion gate and an additional 8 Bell state, this 3 cluster state can grow to a 4 qubit cluster. So a 4 qubit state requires 12 Bell pairs.

When the Type-II gate is successful with a 4 cluster state, the growth is $E[\Delta N] = 0.5$. Two of 4 qubit clusters are needed for successful addition of one qubit to the main cluster, which gives a total of 24 Bell pairs.

More interesting thing to do with Type-II fusion is create vertical edges between linear clusters (see figure 6). Further, there are various strategies of how to approach growing clusters and the idea of Browne and Rudolph was to use a Type-I fusion gate to create some microclusters and then a Type-II fusion gate would be used to add them to the main cluster.

3.4.1 Future of cluster computing

Some apparent limitations of cluster computing, such as creation of large cluster lattices, can be overcome by making the lattice “on-the-go” - while it is being measured from one side, it is being built up from the other. Exploiting the properties of the Clifford group could help in cluster preparation. Finally, being clever about what computation needs to be done and not thinking in the circuit model would make quite a difference. Both KLM and MBQC presented so far rely heavily on active switching. New schemes are passive, so called ”ballistic” schemes which use percolation theory to build clusters. This new idea started with Kieling’s paper [**?**] and recently a paper by Gimeno-Segovia et al. [**?**] significantly improves the number of Bell pairs needed. While both of these schemes actually need more Bell pairs than architectures based on Browne and Rudolph, they are still exploring the limits of their scalability.

4 Conclusion

These protocols opened doors to new ways of thinking about optical quantum computing. While there are still some supporters of the quantum circuit model, it is getting more likely that a cluster state optical computer will be a reliable architecture of choice. This essay has touched upon just some of the many problems and proposed solutions to them when it comes to linear

optical quantum computing. The quest for a truly scalable and engineerable quantum computer is still very active.

A Appendix

A.1 DiVincenzo criteria

DiVincenzo [**?**] gives the following five requirements for quantum computing:

- Qubits need to be well defined
- Qubit specific measurements can be carried out
- Initialization to a simple pure state such as $|00...0\rangle_L$.
- Universal set of quantum gates has to be implementable
- Long coherence times

When it comes to optical architectures, photons lend themselves well due to various degrees of freedom that can represent a qubit. Further they have low levels of decoherence and interaction, thus long coherence times are achievable. Some types of qubit encoding are discussed in 1.2. A universal set of gates (as defined in 1.9) is the greatest hurdle for photonics due to photon lack of interaction. The biggest differences in LOQC architectures rest on the scalability of these gates.

Initialization to a simple, pure state is usually done using single photon sources and measurements are carried out using single or number resolving photodetectors. The initialization and detection problems are shared between optical architectures.

Single photon sources are difficult, currently the biggest contenders are quantum dots, NV centres and multiplexed probabilistic sources [**?**]. One of the problems is that processes that create photons are often probabilistic thus not all of, for example, a pump laser beam is converted into single photons. Further, in certain types of sources more than one photon could be created some of the time. This can be fixed by heralding in sources that produce a pair of photons or by multiplexing. Sources that are deterministic have a different problem. Photons cannot always be extracted from the system, thus they suffer from inefficiency of another kind.

Another challenge are detectors [**?**]. There are currently a variety of detectors used, with different detection efficiency that sometimes depend on the wavelength. Superconducting detectors have reached very high efficiency, $> 99\%$, but they need to be in a cryostat. Some "more standard" detectors such as avalanche photodiode are still widely used, but have much lower efficiency and are highly dependant on wavelength. Further problems occur when the detectors used need to be number resolving, which is sometimes circumvented by using lots of beam splitters and single photon detectors, increasing the chance that the photons will be split up onto different paths.

A.2 Unitary evolution of states and modes

First, an informal proof of the claim that for passive linear optics $U|vac\rangle = |vac\rangle$ is true. Apply the Taylor expansion on U: $U|vac\rangle = e^{itH}|vac\rangle = (1 + itH + \frac{1}{2!}(itH)^2 + ...)|vac\rangle$. Since $H = \sum_{jk} h_{jk}\hat{a}_j^\dagger \hat{a}_k$, each term in the Taylor expansion except for the first will be a polynomial of creation and annihilation operators. The i-th polynomial contains H^i - notice that raising H to the i-th power will leave an annihilation operator as the right-most operator for each term of the i-th polynomial. But $\hat{a}_j|0\rangle^j = 0$ for any j. Thus all of the terms in the i-th polynomial are 0 and the polynomial is actually 0. That is, all of the terms except for the first one in the

Taylor expansion will be 0 and thus $U|vac\rangle = |vac\rangle$

Second, an informal proof of the that the modes transform as $\hat{a}_j^\dagger \to U\hat{a}_j^\dagger U^\dagger$. Let some state have a single photon in the mode a_i. Write this state as $\hat{a}_i^\dagger|vac\rangle$. Then its evolution under some unitary operator U for a passive linear optical element is given as $U\hat{a}_i^\dagger|vac\rangle = U\hat{a}_i^\dagger U^\dagger U|vac\rangle = U\hat{a}_i^\dagger U^\dagger|vac\rangle$ as U is unitary and $U|vac\rangle = |vac\rangle$. This shows that the state modes transform as $\hat{a}_j^\dagger \to U\hat{a}_j^\dagger U^\dagger = \sum_k u_{kj}\hat{a}_k^\dagger$ for a given U.

To find out what this unitary transformation u is, Baker–Campbell–Hausdorff expansion of $U\hat{a}_i^\dagger U^\dagger$ can be used. This expansion states that $e^{\hat{A}}\hat{B}e^{-\hat{A}} = B + [A,B] + \frac{1}{2!}[A,[A,B]] + ...$ where $\hat{A}$ and $\hat{B}$ are some unitary operators. Combining this formula with $U = e^{itH}$, an expression can be derived for u which, in most common optical elements, tidies away to a nice equation (see definitions of phase shifters and beam splitters in section 1.1).

A.3 Qubit encoding

A.3.1 Single-rail encoding

As mentioned, single-rail encoding can do entanglement in a straightforward way using a 50:50 beam splitter, BS. Take two qubits such that $|10\rangle_L = |10\rangle_{Fock} = \hat{a}_1^\dagger|vac\rangle \xrightarrow{BS} \frac{1}{\sqrt{2}}(\hat{a}_1^\dagger + \hat{a}_2^\dagger)|vac\rangle = \frac{1}{\sqrt{2}}(|10\rangle_{Fock} + |01\rangle_{Fock}) = \frac{1}{\sqrt{2}}(|10\rangle_L + |01\rangle_L)$.

Single-qubit operations, on the other hand, are tricky. For example, take X rotation, so $|0\rangle_1 \to |1\rangle_1$ and $|1\rangle_1 \to |0\rangle_1$. In single-rail encoding, this can not be done by a number-preserving (passive) linear optical element, since in the first case a photon is created and in the second case a photon is lost. Ancilla states would be needed and in case they need to be measured, the operation becomes non-deterministic. More specifically, there is a way to implement an arbitrary phase rotation deterministically, but it is not as straight forward as it would be in dual-rail encoding and Hadamard gate is, so-far, only implementable non-deterministically.

A.3.2 Dual-rail encoding

Single qubit operations are easy to implement using phase shifters and beam splitters in dual-rail encoding. For example, Hadamard is achieved using a beam splitter and a phase shifter can be used to implement arbitrary phase rotation. On the other hand entanglement is not straightforward. Consider two qubits entangled in dual-rail encoding, so that first qubit is described by modes 1 and 2 and the second qubit is described by modes 3 and 4. Without a lack of generality, assume the photons are in mode 1 and 3. Then all the possible states reachable from these two qubits are described as $|00\rangle_L = |1010\rangle_{1234} = \hat{a}_1^\dagger\hat{a}_3^\dagger|vac\rangle \to (\sum_{j=1}^4 u_{j1}\hat{a}_j^\dagger)(\sum_{j=1}^4 u_{j3}\hat{a}_j^\dagger)$. We would like our output state to be a Bell state, for example $|00\rangle + |11\rangle = |1010\rangle + |0101\rangle = \hat{a}_1^\dagger\hat{a}_3^\dagger + \hat{a}_2^\dagger\hat{a}_4^\dagger$. Any output state reachable though is still separable and no matrix can make it not be separable, yet the Bell state expression is very much the opposite. Similar to single-rail and Hadamard, in dual-rail encoding, entanglement is so far only implementable non-deterministically.

A.3.3 Parity encoding error-correction

In the text an example is given with two polarization photons so that $|0\rangle_L = \frac{1}{\sqrt{2}}(|HH\rangle + |VV\rangle)$ and $|1\rangle_L = \frac{1}{\sqrt{2}}(|HV\rangle + |VH\rangle)$. Consider a state $\alpha|0\rangle_L + \beta|1\rangle_L = (\alpha|HH\rangle + \beta|HV\rangle) + (\alpha|VV\rangle +$

$\beta |VH\rangle$) and a probabilistic gate was tried on this state and it has failed, and in the failure case the first qubit gets measured. Measurement of one qubit with result $|H\rangle$ would leave the other qubit in state $\alpha |H\rangle + \beta |V\rangle$ and similarly result $|V\rangle$ would leave the other qubit in the state $\alpha |V\rangle + \beta |H\rangle$. The second state can be easily corrected since it is knows that it has occurred.

The parity encoding can be further expanded onto n physical qubits, so that $|0\rangle_L = \frac{1}{\sqrt{2}}((|H\rangle + |V\rangle)^{\otimes n} + (|H\rangle - |V\rangle)^{\otimes n})$ and $|1\rangle_L = \frac{1}{\sqrt{2}}((|H\rangle + |V\rangle)^{\otimes n} - (|H\rangle - |V\rangle)^{\otimes n})$. We can now take this another step further and parity encode the new logical qubits into a single logical qubit, so that $|0\rangle_L^{(2)} = \frac{1}{\sqrt{2}}(|00\rangle_L + |11\rangle_L)$ and $|1\rangle_L^{(2)} = \frac{1}{\sqrt{2}}(|01\rangle_+ |10\rangle_L)$, and so on, thus making the loss of a logical qubit also error correctable.

A.4 Teleportation trick

The "teleportation trick" used in KLM is based on a paper by Gottesman and Chuang[**?**]. In this paper they show that two qubits can be entangled by using teleportation on two entangled ancillas. For this to work, the qubits are now in single-rail encoding. In the teleportation protocol, performing a Bell basis measurement on the first and second qubit when the full state is $(\alpha |0\rangle_1 + \beta |1\rangle_1)\frac{1}{\sqrt{2}}(|00\rangle_{23} + |11\rangle_{23})$ teleports the first qubit to the third one up to a phase difference. Based on the measurement outcomes, phase difference can be corrected using Pauli operators and the remaining state in mode 3 will the be that of the original qubit - $\alpha |0\rangle_3 + \beta |1\rangle_3$.

Take another qubit, $|\psi_2\rangle = \gamma |0\rangle_4 + \delta |1\rangle_4$, that needs to be entangled with qubit $|\psi_1\rangle$ in the first mode. Then both of these qubits could be teleported to modes 3 and 6 and then entangled using a CZ gate.

Here is where they use the definition of the Clifford group, of which CZ is a member, they transform Pauli gates into Pauli gates. Since, based on the measurement result, qubits in modes 3 and 6 only need to be corrected by using Pauli gates and CZ is applied afterwards to modes 3 and 6, the CZ gate can actually be commuted through these Pauli gates and applied to 3 and 6 before the Bell measurement even takes place.

Let $P_3 \otimes P_6$ be the two Pauli group corrections that need to be applied to modes 3 and 6. Then based on the definition of the Clifford group, $CZ_{36}^\dagger (P_3 \otimes P_6) CZ_{36} = (P_3' \otimes P_6')$. Since $CZ_{(3,6)}^\dagger = CZ_{(3,6)}$ and it is unitary, then $(P_3 \otimes P_6) CZ_{(3,6)} = CZ_{(3,6)} (P_3' \otimes P_6')$, which will gives which correction needs to be applied if CZ gate is done before measurement. When CZ gate has been successful on the ancilla states (which will be true $\frac{1}{16}$ of the time), those ancilla state can be used to teleport the entanglement onto the qubits we are interested in entangling. If we were just to entangle the two qubits directly, we would only succeed in doing CZ $\frac{1}{16}$ of the time, and in the rest of the cases we ruin our state.

A.5 Cluster states

A.5.1 Cluster states lemma

For a given a, $S\sigma_x^{(a)} S^\dagger = \prod_{(b,c)\in E} S^{(bc)} \sigma_x^{(a)} \prod_{(d,e)\in E} (S^{(de)})^\dagger$ can be calculated. Since all $S^{(bc)}$ in S commute, the product can be reordered so that the right product mirrors the left one, that is $S\sigma_x^{(a)} S^\dagger = S^{(b_1 c_1)} ... S^{(b_n c_n)} \sigma_x^{(a)} (S^{(b_n c_n)})^\dagger ... (S^{(b_1 c_1)})^\dagger$ where $n = |E|$. From the equations (7) it follows that whenever a is not one of the vertices in the edge, nothing happens. Thus move the operators in the above expression again so that the "inside" ones don't have a as a vertex and

the most "outside" ones have a as a target, that is
$S^{(ab_1)}...S^{(ab_m)}S^{(c_1a)}...S^{(c_la)}S^{(d_1e_1)}...S^{(d_ke_k)}\sigma_x^{(a)}(S^{(d_ke_k)})^\dagger...(S^{(d_1e_1)})^\dagger(S^{(c_la)})^\dagger...(S^{(c_1a)})^\dagger(S^{(ab_m)})^\dagger...(S^{(ab_1)})^\dagger$
where m is the number of edges with a as the first vertex and l is the number of edges with a as the second vertex and k is the number of edges without a.

From above, the inside product with no vertex a mentioned will leave $\sigma_x^{(a)}$ as $\sigma_x^{(a)}$. The above is then reduced to

$S^{(ab_1)}...S^{(ab_m)}S^{(c_1a)}...S^{(c_la)}\sigma_x^{(a)}(S^{(c_la)})^\dagger...(S^{(c_1a)})^\dagger(S^{(ab_m)})^\dagger...(S^{ab_1)})^\dagger$

using eq. (3.1):

$= S^{(ab_1)}...S^{(ab_m)}S^{(c_1a)}...S^{(c_{l-1}a)}\sigma_z^{(c_l)} \otimes \sigma_x^{(a)}(S^{(c_{l-1}a)})^\dagger...(S^{(c_1a)})^\dagger(S^{(ab_m)})^\dagger...(S^{(ab_1)})^\dagger$

$= S^{(ab_1)}...S^{(ab_m)}S^{(c_1a)}...S^{(c_{l-2}a)}\sigma_z^{(c_{l-1})} \otimes \sigma_z^{(c_l)} \otimes \sigma_x^{(a)}(S^{(c_{l-2}a)})^\dagger...(S^{(c_1a)})^\dagger(S^{(ab_m)})^\dagger...(S^{(ab_1)})^\dagger$

use eq. (3.1) $l-2$ more times

$$= S^{(ab_1)}...S^{(ab_m)}\bigotimes_{j=1}^{l}\sigma_z^{(c_j)} \otimes \sigma_x^{(a)}(S^{(ab_m)})^\dagger...(S^{(ab_1)})^\dagger$$

use eq. (8):

$$= S^{(ab_1)}...S^{(ab_{m-1})}\bigotimes_{j=1}^{l}\sigma_z^{(c_j)} \otimes \sigma_x^{(a)} \otimes \sigma_z^{(b)}(S^{(ab_{m-1})})^\dagger...(S^{(ab_1)})^\dagger$$

use eq. (8) $m-1$ more times

$$= \bigotimes_{j=1}^{l}\sigma_z^{(c_j)} \otimes \sigma_x^{(a)}\bigotimes_{i=1}^{l}\sigma_z^{(b_i)}$$

$$= \bigotimes_{b\in n(a)}\sigma_z^{(b)} \otimes \sigma_x^{(a)}$$

A.5.2 Type-I fusion gate lemma

Proof. Remember that an edge in a cluster state represents that a CZ gate was applied to two end point qubits. Qubits all start in the state $|+\rangle$ and the CZ gate is then applied to them based on the adjacency matrix. First, write the cluster state $|\phi_A\rangle$ in a more convenient way, isolating the qubit a. Let $\mathrm{CZ}^{(ac)}$ denote that CZ gate has been applied to qubits a and c with a as control (symmetry of the gate allows this without a lack of generality).

$$\prod_{c\in n(a)}\mathrm{CZ}^{(ac)}\,|+\rangle_a\,|\phi_A\rangle_{V_a\setminus\{a\}} = |0\rangle_a\,|\phi_A\rangle_{V_a\setminus\{a\}} + |1\rangle_a\prod_{c\in n(a)}\mathrm{Z}^{(c)}\,|\phi_A\rangle_{V_a\setminus\{a\}} \tag{6}$$

The Type-I fusion gate is now applied to qubits a and b from cluster states $|\phi_A\rangle$ and $|\phi_B\rangle$. If the measurement result was H then the Kraus operator that needs to be applied is $U_{typeI}^{(H)} = \frac{1}{\sqrt{2}}(|H\rangle\langle HH| + |V\rangle\langle VV|)$, which using 0 and 1 instead gives, $U_{typeI}^{(0)} = \frac{1}{\sqrt{2}}(|0\rangle\langle 00| + |1\rangle\langle 11|)$.

Then

$$(U_{typeI}^{(0)})^{(ab)}\,|\phi_A\rangle\,|\phi_B\rangle = \frac{1}{\sqrt{2}}(|0\rangle_e\langle 00|_{ab} + |1\rangle_e\langle 11|_{ab})\left(|0\rangle_a\,|\phi_A\rangle_{V_a\setminus\{a\}} + |1\rangle_a\prod_{c\in n(a)}\mathrm{Z}^{(c)}\,|\phi_A\rangle_{V_a\setminus\{a\}}\right)$$

$$\begin{aligned}
&\left(|0\rangle_b |\phi_B\rangle_{V_b\setminus\{b\}} + |1\rangle_b \prod_{d\in n(b)} \mathrm{Z}^{(d)} |\phi_B\rangle_{V_b\setminus\{b\}} \right) \\
=&\frac{1}{\sqrt{2}} |0\rangle_e |\phi_A\rangle_{V_a\setminus\{a\}} |\phi_B\rangle_{V_b\setminus\{b\}} \\
&+ \frac{1}{\sqrt{2}} |1\rangle_e \prod_{c\in n(a)} \mathrm{Z}^{(c)} |\phi_A\rangle_{V_a\setminus\{a\}} \prod_{d\in n(b)} \mathrm{Z}^{(d)} |\phi_B\rangle_{V_b\setminus\{b\}}
\end{aligned}$$

Notice that when qubit e is $|1\rangle_e$, the Z gate gets applied to qubits from $n(a)$ and from $n(b)$. Then, they can be collected into one set, call it $n(e) = n(a) \cup n(b)$.

Define $V' := (V_a \setminus \{a\}) \cup (V_b \setminus \{b\})$ and $|\phi\rangle_{V'} = |\phi_A\rangle_{V_a\setminus\{a\}} |\phi_B\rangle_{V_b\setminus\{b\}}$.

Then

$$\begin{aligned}
(U^{(0)}_{typeI})^{(ab)} |\phi_A\rangle |\phi_B\rangle =& \frac{1}{\sqrt{2}} \left(|0\rangle_e |\phi\rangle_{V'} + |1\rangle_e \prod_{f\in n(e)} \mathrm{Z}^{(f)} |\phi\rangle_{V'} \right) \\
=& \prod_{f\in n(e)} \mathrm{CZ}^{(ef)} |+\rangle_e |\phi\rangle_{V'} =: |\phi\rangle
\end{aligned}$$

This $|\phi\rangle$ is then a cluster state, and $n(e)$ represents the neighbourhood of qubit e since all the qubits from this set are connected to e by CZ gate. The graph describing it is $G(V, E)$ where $V = V' \cup \{e\}$ and $E = E_a \cup E_b$.

The proof when V is detected is identical. The proof for the failure cases causing the cluster to stay separate and lose a qubit is given in the text.

□