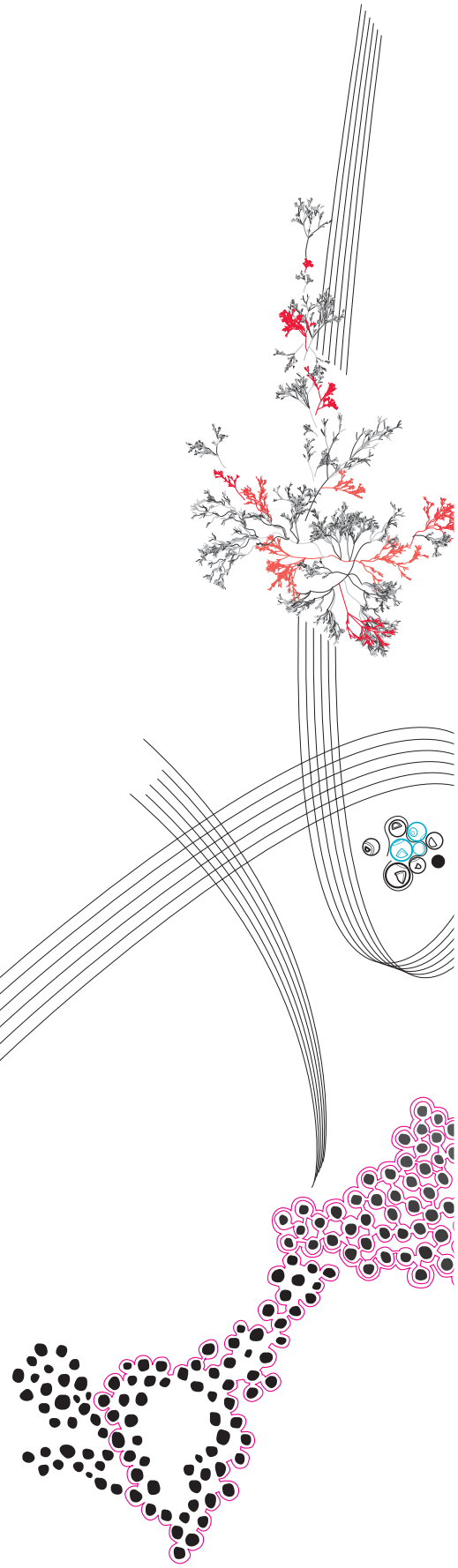


BSc Thesis Applied Mathematics & Applied Physics

Five-Mode Photon Distillation

Boy Kite Sauër

Supervisor Applied Mathematics: M. Schlottbom
Supervisor Applied Physics: J.J. Renema
Daily supervisor: F.H.B. Somhorst

May , 2024

Department of Applied Mathematics
Faculty of Electrical Engineering, Mathematics and Computer Science
Department of Applied Physics
Faculty of Science and Technology

UNIVERSITY OF TWENTE.

Preface

I could not have completed this thesis without the help of my supervisors. I would thank Jelmer, Matthias and Frank dearly for the time they spent helping me gain insight, explain difficult concepts and discussing this subject with me. I would like to thank Frank in special for the daily guidance and all the long discussion that often probably took longer than planned, helped me to understand the material and that have most certainly deepened my interest. I would also like thank all the members of the AQO research group for their support and help during this project.

Contents

Five-Mode Photon Distillation

Boy Kite Sauër

May , 2024

Keywords: photon distillation, indistinguishable photons, partially distinguishable photons, quantum information, quantum computation, linear optical quantum computer

Abstract

Generating highly indistinguishable photons is one of the major challenges in realizing optical quantum computers. Recently the photon distillation protocol was proposed by J. Marshall that uses a standard linear optical network to increase photon indistinguishability to arbitrary accuracy by sacrificing ancillary photons.[11] Ideally, the number of photons that need to be sacrificed to achieve a certain error reduction is kept at a minimum. Via an extensive numerical search, the optimal networks for three-mode and four-mode networks were found. [11] We recognized that the three-mode optimal matrix was not just some random network, but a network that applies the discrete Fourier transform (DFT). We, therefore, research in this report how the five-mode DFT performs in the photon distillation protocol. The results are enlightening; the five-mode DFT could outperform the three-mode DFT, contradicting what was expected in the original paper. The parameters that determine the performance are its ability to reduce the error, success rate, and the number of photons required. In both the three-mode and the five-mode case we see an error reduction by a factor $\frac{1}{n}$ requiring n photons. However, where the three-mode protocol has a success rate of $\frac{1}{3}$, the five-mode has a success rate that is higher than $\frac{1}{4}$, which breaks the expected $\frac{1}{n}$ behavior. Additionally, we found data suggesting that higher-order systems could perform even better.

1 Introduction

Quantum computing and quantum information have been the subject of immense interest from both academia and industry. The compelling nature of this area of study is rooted in the immense capacity of quantum computers to outperform classical computing in solving some specific traditionally difficult problems. The most famous example of this is the prime factorisation of integers that can be solved by Peter Shor's quantum algorithm in polynomial time on an ideal quantum computer (i.e. one that is not affected by errors)[18, 12].

In addition to the factorisation problem, a broad range of effective algorithms and applications have been developed for quantum computers, offering significant increases in efficiency in a variety of industries and research fields. The areas that are likely to be affected the most are those related to shortest-path finding algorithms, machine learning, solving differential equations (which can be used to simulate a variety of complex physical systems), and cryptography. At present, quantum computers exist; yet they have few or no practical applications as a result of their susceptibility to noise [6].

Producing a fault-tolerant quantum computer is therefore of great importance for realising the potential of quantum computers. Currently, there are multiple physical systems suitable for realising quantum computing. Each having its own advantages and challenges. For example, IBM and Google are developing systems based on superconducting circuits [6]. An advantage of using quantum computing with superconducting circuits is that they can be easily created with existing chip technology and combined with classical computers [10]. However, to create the necessary superconducting states, the superconducting circuits must be cooled to cryogenic temperatures. Maintaining these extreme temperatures when scaling up to a large number of qubits presents a major challenge [6, 10].

Another architecture realises qubit states using photons. A major advantage over superconducting circuits is that Qubit states created with photons do not require these extremely low temperatures. Moreover, photons hardly suffer from decoherence, which destroys superposition states, because of their weak interaction with the environment. However, this weak interaction of photons with the environment and with other photons is a major hurdle for the creation of universal two-qubit gates [10, 9]. Before 2001, these interactions were thought to require non-linearities that were much more powerful than those found in existing nonlinear media [10, 9]. This problem was circumvented by the Knill, Laflamme and Milburn (KLM) protocol which shows that "efficient quantum computing is possible using only beam splitters, phase shifters, single-photon sources, and photon-detectors" [8]. However, the KLM protocol requires a large amount of resources for implementation, so the protocol mainly serves as a proof that linear optical quantum computing (LOQC) is achievable [8, 9]. Since then improved protocols that use measurements on large entangled photon states have reduced the resource costs significantly, e.g. measurement-based quantum computers and fusion-based quantum computers [23, 15, 1, 13]. To create these cluster states, we need to entangle a large number of photons. Photon entanglement is a consequence of quantum interference, which is a result of coherent photons, that is, the photons need to be indistinguishable. A prime example of the role photon indistinguishability plays in photon entanglement can best be demonstrated by a two-photon interference experiment known as the Hong-Ou-Mandel (HOM) effect, where the outcome of a two photon experiment depends on whether the photons behave as indistinguishable or distinguishable. This effect can be used to measure the number of events that occur due to nonidentical photons and can therefore be related to the partial indistinguishability of the photons [7, 5, 2]. In reality, the current photon sources are not capable of creating perfectly

indistinguishable photons. Instead the photons partially behave as indistinguishable and partially distinguishable.

The necessity of cluster states and the challenge of generating perfectly indistinguishable photons highlight the need for a protocol that reduces the partial distinguishable behaviour of photons. The recently developed photon distillation protocol aims to do just this [11], see Figure 1.

FIGURE 1: Illustration photon distillation experiment. N "copies" of photons with a state described by $\rho(\epsilon)$ are send into a N-mode interferometer of which $N-1$ ancillary modes are measured. The ϵ gives an indication for the error of the photon. The remaining photon has a state described by $\rho(\epsilon')$, such that $\epsilon' < \epsilon$. SOURCE: *Distillation of Indistinguishable Photons.* [11]

The photon distillation protocol showed that the amount of partial distinguishability can be reduced arbitrarily close to zero by repeated implementation of the protocol. The process of photon distillation involves sending N photons into an N-mode interferometric experiment. The number of photons in N-1 ancillary modes is then measured. By requiring that a valid measurement measures N-1 photons in these ancillary modes, the measurement provides information about the partial indistinguishability of the photon in the remaining mode. By comparing the behaviour of distinguishable and indistinguishable photons in such an experiment, the result can be used to determine whether the state of the remaining photon is more likely to be indistinguishable from the target state of the photon source. The measurement heralds information about the partial distinguishability of the remaining photon. The distribution linked to this outcome is called the heralding pattern. For this protocol, the researchers assume no errors other than errors in the photon source, which are the cause of the partial indistinguishability. This partial distinguishability can be described with the random source model [19]. This model defines an average internal state for the photons assuming that the photons are independently and identically distributed, which for most sources is a reasonable assumption. The error in the system is then the amount of partial distinguishability. The protocol was developed for a three-mode interferometer (tritter) and a four-mode interferometer.

Marshall shows that the protocol using the tritter reduces the error by a factor 1/3, with a success rate of 1/3. The success rate is determined by the probability of measuring the heralding pattern that reduces the error and can be directly related to the resources (the number of photons) required to achieve successful distillation. He demonstrates that the 4-mode interferometer protocol reduces the error even further by a factor of 1/4. This would be very useful if it were not for the fact that the success rate decreases to 1/4. The average number of photons required to successfully distil the photon would be 3*3 for the tritter and 4*4 for the 4-mode interferometer. This suggests that an adequate N-mode system that could reduce the error by a factor of $1/N$ would have a success probability of $1/N$ and would therefore need N^2 photons to perform successful distillation. The fact that the amount of resources required to realise N-mode photon distillation scales with N^2

suggests that it is not useful to employ interferometers with a greater number of modes for photon distillation.

This paper investigates whether a five-mode interferometer actually is capable of reducing the error by a factor 1/5 and also whether the probability of success is 1/5 to investigate whether the suggested $1/N$ behaviour holds.

We recognised that in J.Marshall's paper the found to be "best" tritter performs the quantum Fourier transform (QFT) and the "best" 4-mode interferometer implements a Hadamard transform. Here, "best" was defined to be the operation that achieved a certain error reduction with the lowest number of photon sacrificed. The best matrices where found via intense numerical search. To spare the intense numerical search for the best 5-mode network, we decided that it would be best to start with an operation that had already proved it works. Consequently, we are now exploring the behaviour of the 5-mode QFT in the 5-mode interferometer. In subsection 3.1, we estimate the reduction capabilities of an N-mode linear optical network in the photon distillation protocol, assuming that the linear optical network in use can be considered as an balanced beamsplitter. That implies that the photon has equal probability of being in one of the output modes, regardless of the input mode it started in. In section 3, we investigate the error-reducing capabilities of the 5-mode Quantum Fourier Transform (QFT) when heralding with different patterns, as well as the success rate of the 5-mode QFT. Interestingly, we can use multiple patterns to indicate error reduction. If we could act quickly enough to make decisions based on the observed pattern, we could filter using all patterns that reduce the error simultaneously. This naturally leads to a higher success rate as it is much more likely to detect any error-reducing pattern. If we could achieve this, then the success rate seems to be as high as 1/4. In subsubsection 3.3.1, we simulate the success rate for the 3 to 8-mode QFT inspired by this interesting result from the 5-mode QFT. Even more interesting is that systems with more modes also achieve a success rate of about 1/4.

2 Theory

The photon distillation protocol is based on the effect that partial distinguishability has on the probability of measuring certain outcomes in a boson sampling experiment [11, 20, 21, 17]. In the protocol one partial distinguishable photon is injected into each one of the N modes of an interferometer. After the interferometer, the number of photons in N-1 ancillary output modes is measured. Given that we have measured the number of photons in these N-1 ancillary modes, we know the number of photons in the remaining mode. The measurement of these modes is therefore heralding information about the state in the remaining mode. The outcome is therefore the distribution of the photons over the output modes. The effect that partial distinguishability has on the outcome of the experiment can be used to filter the remaining mode. The protocol aims to distil the photons to a single indistinguishable photon, and therefore it is required that there is only one photon in the remaining mode. In order to simulate the photon distillation protocol we have to simulate the outcome probabilities of this boson sampling experiment. The theory of boson sampling with partial distinguishable photons is well described in Refs. [17, 20].

2.1 The Photon State

The state of a photon depends on its spatial components, its polarisation and its frequency [17]. The variables can be classified into two categories, the spatial components and the spectral components (polarisation and frequency). The photons occupying the same inter-

ferometer mode are spatially coherent and therefore the spatial component of the photon state is completely described by the interferometer mode that the photon occupies, which is enforced by the single-mode character of the optical wave-guides [22]. The spectral components will not be directly controlled in this experiment. The partial distinguishability arises from the differences in the spectral components as result from our imperfect internal state realisation by the nonideal photon source. We refer to the spatial components of the photon state as the external state and to the spectral component as the internal state, due to the control we have over the components.

2.1.1 The Internal State

The internal state of an arbitrary photon can be described by the summation over its polarisation and frequency modes [11, 4, 14, 17].

$$|\psi\rangle = \sum_{s\in\{h,v\}} \int_{-\infty}^{\infty} d\omega c_{s,\omega} |s,\omega\rangle = \sum_{i=0}^{\infty} c_i \hat{a}_i^\dagger |\mathbf{0}\rangle = \sum_{i=0}^{\infty} c_i |\psi_i\rangle \tag{1}$$

Here, we can see that the modes of the internal state can be indexed with a single index i.

Photons are indistinguishable if their states completely overlap. That is, two photons with wave function $|\psi_A\rangle$ and $|\psi_B\rangle$ are indistinguishable if $|\langle\psi_A|\psi_B\rangle| = 1$. If there is partial overlap between the photon states, then the absolute value of the inner product takes some value between zero and one, that is, $0 < |\langle\psi_A|\psi_B\rangle| < 1$ and distinguishable photons have zero overlap.

2.1.2 Modelling the Partial Indistinguishability

Our nonideal photon source creates a photon with internal state according to Equation 1, where the coefficients are realisation dependent [11]. This introduces uncertainty in the preparation of the state and is the cause of the partial distinguishability between the photons. We can thus model the partial distinguishability by considering the internal state as an prepared statistical ensemble of states and move to the density matrix formalism. The density matrix describing the internal state for the photon α is given by the following equation according to Ref. [17]

$$\hat{\rho}_\alpha = \int_{-\infty}^{\infty} p_\alpha(x) |\psi_\alpha(x)\rangle\langle\psi_\alpha(x)| dx. \tag{2}$$

Here, $p_\alpha(x)$ is the density distribution of the photon being prepared in state $|\psi_\alpha(x)\rangle\langle\psi_\alpha(x)|$.

We make the assumption that the photons are independently and identically distributed, that is, $p_\alpha(x) = p(x)$ so that we can define an average internal state for all the photons.

$$\hat{\rho} = \int_{-\infty}^{\infty} p(x) |\psi(x)\rangle\langle\psi(x)| dx = \sum_{i=0}^{\infty} p_i |\psi_i\rangle\langle\psi_i|. \tag{3}$$

In this equation, $|\psi(x)\rangle$ denotes a possible distribution of the coefficients c_i from Equation 1. This assumption and the resulting density matrix is known in the literature as the random source model (RSM) [19, 11, 17].

We should interpret the coefficients p_i in Equation 3 as the probability of finding the photon in the i^{th} internal mode. Without loss of generality, we can assume that our nonideal

photon source creates photons in the internal mode $|\psi_0\rangle$ with probability $p_0 = 1 - \epsilon$ and with probability p_i in some other (error) mode, with $\sum_{i>0}^{\infty} p_i = \epsilon$ [11]. This allows us to write the state as

$$\rho(\epsilon) = (1-\epsilon)|\psi_0\rangle\langle\psi_0| + \sum_{i>0}^{\infty} p_i|\psi_i\rangle\langle\psi_i|. \tag{4}$$

Since it does not matter in which error mode the photon is, we can say that the photon is an error state orthogonal to $|\psi_0\rangle$ with probability ϵ, and we can describe the photon state with the following equation. This model is known in literature as the orthogonal bad bit model (OBBM) [19, p. 151];

$$\rho(\epsilon) = (1-\epsilon)|\psi_0\rangle\langle\psi_0| + \epsilon|\psi_\perp\rangle\langle\psi_\perp|. \tag{5}$$

In the limit where ϵ is very small, the contribution of photon entanglement due to photons occupying similar error modes to the probability of finding a specific measurement outcomes is negligible. Therefore, we assume that the i^{th} photon occupies the i^{th} error mode.

$$\rho_i(\epsilon) = (1-\epsilon)|\psi_0\rangle\langle\psi_0| + \epsilon|\psi_i\rangle\langle\psi_i| \tag{6}$$

Here, we let ρ_α denote the density operator of the α^{th} photon.

Describing the state of the photons in the high indistinguishability regime with equation Equation 6 is experimentally well justified [16].

Instead of a single vector, a statistical mixture of states is used to describe the state. In this case, previous definition of the photon indistinguishability does not work quite well. Naturally, as we make use of average states, the average indistinguishability would seem to be the right extension. Thus we will use the mean of the overlap as our measure of indistinguishability. Let $|\psi\rangle$ and $|\phi\rangle$ denote the state of the photons, then the average indistinguishability is $\mathcal{I} = \text{mean}\left(|\langle\psi|\phi\rangle|\right)$ [11]. For photons with an average internal state ρ, this is equivalent to the trace of the product of the density matrices; $\mathcal{I} = \text{tr}(\rho^2)$ [11]. This value is closely related to the fidelity of the system. The fidelity of quantum states $\mathcal{F}(\sigma, \rho)$ is one of the most common measures to indicate the closeness of two quantum states with density matrices σ and ρ [12]. If one of the two states is pure, say $\sigma = |\psi_0\rangle\langle\psi_0|$, then the fidelity reduces to $\mathcal{F}(\sigma, \rho) = \langle\psi_0|\rho|\psi_0\rangle$. This can be interpreted as the probability that the photon in the state ρ will behave like a photon in the state $|\psi_0\rangle\langle\psi_0|$ [12].

2.1.3 Internal State of the Composite System of N Photons

The photon distillation protocol uses N photons. The density operator describing the internal state of the composite photon state is obtained by taking the tensor product of the density operators describing the internal states of the single photons; see Equation 7 [17].

$$\rho_{\text{composite}} = \bigotimes_{i=0}^{N} \hat{\rho}_i. \tag{7}$$

Following, Marshall we have found the following expression for the composite density operator [11],

$$\rho(\epsilon) = (1-\epsilon)^n |\Psi_0\rangle\langle\Psi_0| + \epsilon(1-\epsilon)^{n-1} \sum_{k=1}^{n} |\Psi_k\rangle\langle\Psi_k| + O(\epsilon^2), \tag{8}$$

where we employ notation, $|\Psi_0\rangle = |\psi_0\rangle^{\otimes n}$ and

$$|\Psi_k\rangle = |\Psi_k\rangle = |\psi_0\rangle^{\otimes(k-1)} |\psi_k\rangle |\psi_0\rangle^{\otimes(n-k)},$$

which is slightly different from the notation originally used by Marshall due to the assumption that the photons do not occupy the same error mode.

An important observation from Equation 8 is that the system can behave in different ways. With probability $(1-\epsilon)^n$ the system behaves as if there were only completely indistinguishable photons, with probability $n(1-\epsilon)^{n-1}\epsilon$ the system behaves like a system with only one distinguishable photon, etc.

2.1.4 The Mode Occupation Representation

The previous section described the internal state of the photon and shows that we can construct an orthonormal basis that describes the internal state. Furthermore, we described a partial distinguishability model for the photons which shows that we consider the problem as different probabilities happening. For the external state we made the assumption that the photons are spatially coherent in each of the interferometer modes. The state of the photon is then described by the combination of its internal mode and its spatial mode and we can construct an orthonormal basis for the photon state by employing the creation operators $\hat{a}^\dagger_{k,i}$ to create a photon in the k^{th} spatial mode and in the i^{th} internal mode [17]. With these creation operators we can define the Fock state that keeps track of the number of photons in each photon mode; see Equation 9.

$$|n_{1,0} n_{1,1} \dots n_{2,0} \dots n_{N,N}\rangle = \prod_{i=0}^{N} \prod_{j=0}^{N} \frac{(\hat{a}^\dagger_{i,j})^{n_{i,j}}}{n_{i,j}!} |\mathbf{0}\rangle \tag{9}$$

Here, $n_{i,j}$ denotes the number of photons with internal mode i and spatial mode k.

Although Equation 9 describes the total photon state, in practice we cannot access all the information that is needed to describe this state. The only information that we can access is the number of photons in each spatial mode, regardless of the internal mode they have. That is, we can measure $n_k = \sum_i n_{k,i}$ and therefore can only describe the input and output with a vector of the form $|n\rangle = |n_1 n_2 \dots n_N\rangle$, named the mode occupation representation. This tracks the number of photons occupying each internal mode.

The density matrix of Equation 8 can be rewritten in mode occupation form by using a tilde to indicate the presence of an error mode. If we let $|\mathbf{1}\rangle = |1_1 1_2 \dots 1_N\rangle$ denote the state with one photon in each mode, and we let $|\tilde{1}_k\rangle = |1_1 1_2 \dots \tilde{1}_k \dots 1_N\rangle$ denote the state with the photon in the k^{th} spatial mode in an error mode, then Equation 10 describes the input of the photon distillation experiment.

$$\hat{\rho}(\epsilon) = (1-\epsilon)^n |\mathbf{1}\rangle\langle\mathbf{1}| + \epsilon(1-\epsilon)^{n-1} \sum_{k=0}^{N} |\tilde{1}_k\rangle\langle\tilde{1}_k| + O(\epsilon^2) \tag{10}$$

2.1.5 The Interferometer

The input of the interferometer is the mode occupation vector $|n\rangle = |n_1 n_2 \dots n_N\rangle$ and the output is described by the mode vector $|m\rangle = |m_1 m_2 \dots m_N\rangle$. We are interested in the probability of measuring the output $|m\rangle$, given that the input is $|n\rangle$.

The behaviour of the interferometer is completely described by a unitary operator $\hat{U}$ that acts on the creation operators of Equation 9 [17, 20, 9],

$$\hat{U}(\hat{a}^\dagger_{k,i}) = \sum_{s=0}^{N} U_{kl} \hat{a}^\dagger_{l,i}. \tag{11}$$

The interferometer does not interact with the internal states. However, the outcome of the experiment does not only depend on the interferometer input and output, but also on the partial distinguishability of the photons. This is the result of the ability of indistinguishable photons to interfere, so that some waves might extinguish completely.

2.1.6 The Hong Ou Mandel Effect

The most famous and simplest example of the influence partial distinguishability has on the outcome of an interferometric experiment is the Hong Ou Mandel effect [7]. The Hong Ou Mandel effect is a phenomenon of two-photon interference that occurs when two photons are sent into each mode of a 50-50 beamsplitter; see Figure 2. The original experiment was proposed to measure the partial distinguishability between photons. The experiment provides a method for calculating the partial distinguishability of one photon if the partial distinguishability of the other photon is already known.

FIGURE 2: An illustration of the HOM interferometric experiment. Two photons are inserted into each of the input modes of the HOM-interferometer. The HOM black box consists of a 50:50 beamsplitter and its operation on the creation operators can be described by the matrix U_{HOM}. The number of photons at each of the output modes is then measured. The possible outcomes are measuring both photons in either one of the output modes or one photon in both output modes. The probability of measuring such outcome depends on the partial distinguishability of the photons.

The unitary matrix describing the Hong Ou Mandel interferometer is U_{Hom} [5, 7, 3].

$$U_{\text{HOM}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{12}$$

Now consider two photons with internal states ρ_i and ρ_j, respectively, that have different degrees of distinguishability from the ideal state $|\psi_0\rangle\langle\psi_0|$. One photon will be injected in either modes, so that the total states of the photon can be described as;

$$\hat{\rho}_A = (1-\epsilon)\hat{a}^\dagger_{1,0}|\mathbf{0}\rangle\langle\mathbf{0}|\hat{a}_{1,0} + \epsilon\hat{a}^\dagger_{1,i}|\mathbf{0}\rangle\langle\mathbf{0}|\hat{a}_{1,i}$$

and

$$\hat{\rho}_B = (1-\epsilon')\hat{a}^\dagger_{2,0}|\mathbf{0}\rangle\langle\mathbf{0}|\hat{a}_{2,0} + \epsilon'\hat{a}^\dagger_{2,j}|\mathbf{0}\rangle\langle\mathbf{0}|\hat{a}_{2,j}.$$

The composite state of this system $\hat{\rho}_{AB}$ is given by the tensor product of these states.

$$\begin{aligned}\hat{\rho}_{AB} = {} & (1-\epsilon)(1-\epsilon')\hat{a}^\dagger_{1,0}\hat{a}^\dagger_{2,0}|\mathbf{0}\rangle\langle\mathbf{0}|\hat{a}_{1,0}\hat{a}_{2,0} + \epsilon(1-\epsilon')\hat{a}^\dagger_{1,i}\hat{a}^\dagger_{2,0}|\mathbf{0}\rangle\langle\mathbf{0}|\hat{a}_{1,i}\hat{a}_{2,0} \\ & + (1-\epsilon)\epsilon'(\hat{a}^\dagger_{1,0}\hat{a}^\dagger_{2,j}|\mathbf{0}\rangle\langle\mathbf{0}|\hat{a}_{1,0}\hat{a}_{2,j}) + \epsilon\epsilon'(\hat{a}^\dagger_{1,i}\hat{a}^\dagger_{2,j}|\mathbf{0}\rangle\langle\mathbf{0}|\hat{a}_{1,i}\hat{a}_{2,j}).\end{aligned}$$

The creation operators change according to

$$\begin{aligned}
\hat{a}^\dagger_{1,0}\hat{a}^\dagger_{2,0}|\mathbf{0}\rangle &\xrightarrow{U_{Hom}} \frac{1}{2}\left(\hat{a}^\dagger_{1,0}+\hat{a}^\dagger_{2,0}\right)\left(\hat{a}^\dagger_{1,0}-\hat{a}^\dagger_{2,0}\right)|\mathbf{0}\rangle = \frac{1}{2}\left(\hat{a}^\dagger_{1,0}\hat{a}^\dagger_{1,0}-\hat{a}^\dagger_{2,0}\hat{a}^\dagger_{2,0}\right)|\mathbf{0}\rangle \quad (13)\\
\hat{a}^\dagger_{1,i}\hat{a}^\dagger_{2,0}|\mathbf{0}\rangle &\xrightarrow{U_{Hom}} \frac{1}{2}\left(\hat{a}^\dagger_{1,i}+\hat{a}^\dagger_{2,i}\right)\left(\hat{a}^\dagger_{1,0}-\hat{a}^\dagger_{2,0}\right)|\mathbf{0}\rangle = \frac{1}{2}\left(\hat{a}^\dagger_{1,i}\hat{a}^\dagger_{1,0}-\hat{a}^\dagger_{1,i}\hat{a}^\dagger_{2,0}+\hat{a}^\dagger_{2,i}\hat{a}^\dagger_{1,0}-\hat{a}^\dagger_{2,i}\hat{a}^\dagger_{2,0}\right)|\mathbf{0}\rangle\\
\hat{a}^\dagger_{1,0}\hat{a}^\dagger_{2,j}|\mathbf{0}\rangle &\xrightarrow{U_{Hom}} \frac{1}{2}\left(\hat{a}^\dagger_{1,0}+\hat{a}^\dagger_{2,0}\right)\left(\hat{a}^\dagger_{1,j}-\hat{a}^\dagger_{2,j}\right)|\mathbf{0}\rangle = \frac{1}{2}\left(\hat{a}^\dagger_{1,0}\hat{a}^\dagger_{1,j}-\hat{a}^\dagger_{1,0}\hat{a}^\dagger_{2,j}+\hat{a}^\dagger_{2,0}\hat{a}^\dagger_{1,j}-\hat{a}^\dagger_{2,0}\hat{a}^\dagger_{2,j}\right)|\mathbf{0}\rangle\\
\hat{a}^\dagger_{1,i}\hat{a}^\dagger_{2,j}|\mathbf{0}\rangle &\xrightarrow{U_{Hom}} \frac{1}{2}\left(\hat{a}^\dagger_{1,i}+\hat{a}^\dagger_{2,i}\right)\left(\hat{a}^\dagger_{1,j}-\hat{a}^\dagger_{2,j}\right)|\mathbf{0}\rangle = \frac{1}{2}\left(\hat{a}^\dagger_{1,i}\hat{a}^\dagger_{1,j}-\hat{a}^\dagger_{1,i}\hat{a}^\dagger_{2,j}+\hat{a}^\dagger_{2,i}\hat{a}^\dagger_{1,j}-\hat{a}^\dagger_{2,i}\hat{a}^\dagger_{2,j}\right)|\mathbf{0}\rangle.
\end{aligned}$$

In Equation 13 it immediately becomes clear how the indistinguishability of the photons suppresses the coincidental outcome.

To understand what the measurements outcome will be we would like to write this down in the mode occupation representation. However, we mark the mode which is occupied by an error mode with an tilde for error mode i and an asterisk for error mode j to not lose sight of the different distinguishable photons.

$$\begin{aligned}
|11\rangle &\xrightarrow{U_{Hom}} \frac{\sqrt{2}}{2}\left(|20\rangle - |02\rangle\right) \quad (14)\\
|\tilde{1}1\rangle &\xrightarrow{U_{Hom}} \frac{1}{2}\left(\sqrt{2}|\tilde{2}0\rangle - |\tilde{1}1\rangle + |1\tilde{1}\rangle - \sqrt{2}|0\tilde{2}\rangle\right)\\
|11^*\rangle &\xrightarrow{U_{Hom}} \frac{1}{2}\left(\sqrt{2}|2^*0\rangle - |11^*\rangle + |1^*1\rangle - \sqrt{2}|02^*\rangle\right)\\
|\tilde{1}1^*\rangle &\xrightarrow{U_{Hom}} \frac{1}{2}\left(\sqrt{2}|\tilde{2}^*0\rangle - |\tilde{1}1^*\rangle + |1^*\tilde{1}\rangle - \sqrt{2}|0\tilde{2}^*\rangle\right).
\end{aligned}$$

Here we used that $\hat{a}^\dagger_{1,0}|00\rangle = |10\rangle$ and $\hat{a}^\dagger_{1,0}|10\rangle = \sqrt{2}|20\rangle$. These equation can be used to describe the state after the interferometer. The total density matrix transforms from $\hat{\rho}_{AB}$ into $\hat{\rho}'_{AB}$.

After the interferometer, the number of photons in the output is measured. This measurement performs an operation on the state that depends on the outcome. Suppose that one photon is measured in either modes. The measurement operator projects the total state onto every state where one photon is measured in either mode, that is, for the outcome $|11\rangle$ the measurement operator is

$$\hat{M}_{|11\rangle} = |11\rangle\langle 11| + |\tilde{1}1\rangle\langle\tilde{1}1| + |1\tilde{1}\rangle\langle 1\tilde{1}| + |1^*1\rangle\langle 1^*1| + |11^*\rangle\langle 11^*| + |1^*\tilde{1}\rangle\langle 1^*\tilde{1}| + |\tilde{1}1^*\rangle\langle\tilde{1}1^*|. \quad (15)$$

The probability of obtaining the measurement outcome $|11\rangle$ is then given by

$$\begin{aligned}
P_{|11\rangle} &= \text{trace}(\rho' \hat{M}_{|11\rangle}) = \text{trace}(\rho'_{AB}\big(|11\rangle\langle 11| + |\tilde{1}1\rangle\langle\tilde{1}1| + |1\tilde{1}\rangle|1\tilde{1}\rangle \\
&+ |1^*1\rangle\langle 1^*1| + |11^*\rangle\langle 11^*| + |1^*\tilde{1}\rangle\langle 1^*\tilde{1}| + |\tilde{1}1^*\rangle\langle\tilde{1}1^*|\big) \\
&= \text{trace}(\rho'_{AB}|11\rangle\langle 11|) + \text{trace}(\rho'_{AB}|\tilde{1}1\rangle\langle\tilde{1}1|) + \text{trace}(\rho'_{AB}|1\tilde{1}\rangle|1\tilde{1}\rangle) \\
&+ \text{trace}(\rho'_{AB}|1^*1\rangle\langle 1^*1|) + \text{trace}(\rho'_{AB}|11^*\rangle\langle 11^*|) + \text{trace}(\rho'_{AB}|1^*\tilde{1}\rangle\langle 1^*\tilde{1}|) \\
&+ \text{trace}(\rho'_{AB}|\tilde{1}1^*\rangle\langle\tilde{1}1^*|) \\
&= \langle 11|\hat{\rho}'_{AB}|11\rangle + \langle\tilde{1}1|\hat{\rho}'_{AB}|\tilde{1}1\rangle + \langle 1\tilde{1}|\hat{\rho}'_{AB}|1\tilde{1}\rangle \\
&+ \langle 1^*1|\hat{\rho}'_{AB}|1^*1\rangle + \langle 11^*|\hat{\rho}'_{AB}|11^*\rangle + \langle\tilde{1}1^*|\rho'_{AB}|\tilde{1}1^*\rangle + \langle 1^*\tilde{1}|\hat{\rho}'_{AB}|1^*\tilde{1}\rangle. \qquad (16)
\end{aligned}$$

This illustrates that only the diagonal elements of the density matrix $\hat{\rho}'_{AB}$ are interesting in calculating the probabilities. The density matrix can then be written as

$$\begin{aligned}
\hat{\rho}'_{AB} = \frac{1}{4}\Big[&2(1-\epsilon)(1-\epsilon')\left(|20\rangle\langle 20| + |02\rangle\langle 02|\right) \\
&+\epsilon(1-\epsilon')\left(2|\tilde{2}0\rangle\langle\tilde{2}0| + |\tilde{1}1\rangle\langle\tilde{1}1| + |1\tilde{1}\rangle\langle 1\tilde{1}| + 2|0\tilde{2}\rangle\langle 0\tilde{2}|\right) \\
&+(1-\epsilon)\epsilon'\left(2|\tilde{2}0\rangle\langle\tilde{2}0| + |\tilde{1}1\rangle\langle\tilde{1}1| + |1\tilde{1}\rangle\langle 1\tilde{1}| + 2|0\tilde{2}\rangle\langle 0\tilde{2}|\right) \\
&+\epsilon\epsilon'\left(2|\tilde{2}^*0\rangle\langle\tilde{2}^*0| + |\tilde{1}1^*\rangle\langle\tilde{1}1^*| + |1^*\tilde{1}\rangle\langle 1^*\tilde{1}| + 2|0\tilde{2}^*\rangle\langle 0\tilde{2}^*|\right) \\
&+...\Big]. \qquad (17)
\end{aligned}$$

Here, the off-diagonal elements are collected in "...". Equation 16 can now be expressed in terms of ϵ and ϵ';

$$P_{|11\rangle} = \frac{1}{4}\Big[0 + \epsilon(1-\epsilon') + \epsilon(1-\epsilon') + (1-\epsilon)\epsilon' + (1-\epsilon)\epsilon' + \epsilon\epsilon' + \epsilon\epsilon'\Big]. \qquad (18)$$

The same can be done for the other probabilities to obtain the following

$$\begin{aligned}
P_{|20\rangle} &= \frac{1}{4}\left(2(1-\epsilon)(1-\epsilon') + \epsilon(1-\epsilon') + \epsilon'(1-\epsilon) + \epsilon\epsilon'\right) = \frac{(1-\epsilon)(1-\epsilon')+1}{4} \\
P_{|11\rangle} &= \frac{1}{4}\left(2\epsilon(1-\epsilon') + 2\epsilon'(1-\epsilon) + 2\epsilon\epsilon'\right) \\
P_{|02\rangle} &= \frac{1}{4}\left(2(1-\epsilon)(1-\epsilon') + \epsilon(1-\epsilon') + \epsilon'(1-\epsilon) + \epsilon\epsilon'\right) = \frac{(1-\epsilon)(1-\epsilon')+1}{4}.
\end{aligned}$$

The probabilities are a function of the distinguishability of the photons ϵ and ϵ'. Now suppose that we "know" the indistinguishability of the first photon. That is, we know ϵ, then we can use this to calculate ϵ'.

$$\epsilon' = 1 - \frac{1 - 2P(|11\rangle)}{(1-\epsilon)}. \qquad (19)$$

The HOM effect that we have described can be considered as a HOM experiment between a "distilled" photon and an "original" photon. Therefore this experiment can be employed to calculate the average reduction in partial distinguishability of the photons.

2.1.7 The Outcome Probabilities for Scattering with N photons

In the previous section we illustrated the role partial distinguishability plays in the outcome of the experiment by individually manipulating the creation operators and calculating the outcome probabilities by hand. We have seen that the probabilities of obtaining certain photon number measurement outcomes does not only depend on the input and the output, but also on the partial distinguishability of the photons.

Experiments involving more photons get very complex to calculate. Given that the input is the coincidental state, we can calculate the probability of measuring the output $|\vec{m}\rangle$ using the formalism described by M.Tichy [20].

This formalism considers the following. The probability to measure the outcome $|m\rangle$ given that the input state is $|n\rangle$ does not depend on the unoccupied interferometer modes and the corresponding elements of the matrix U are therefore irrelevant. The relevant elements can be collected in a sub-matrix of U called the *effective scattering matrix* M. Let $\vec{d}(\vec{s}) = d(d_1, d_2, \ldots, d_N)$ denote the mode assignment list that tracks the mode of the j^{th} photon. The effective scattering matrix is then defined as;

$$M = U_{\vec{d}(\vec{r}),\vec{d}(\vec{s})}. \tag{20}$$

The partial distinguishability between the photons i and j, with internal state $|\psi_i\rangle$ and $|\psi_j\rangle$ can be described by the partial distinguishability matrix $\mathcal{S}$

$$\mathcal{S}_{ij} = \langle\psi_i|\psi_j\rangle. \tag{21}$$

If the photons are completely indistinguishable then all elements of the partial distinguishability matrix are one, $\mathcal{S}_{ij} = 1\ \forall i, j$, if the photons are completely distinguishable then the partial distinguishability matrix is just the identity matrix.

Assuming that the input state is the coincidental state, the probability of finding the outcome $|m\rangle$ can be described by calculating the multi-dimensional permanent of the three dimensional tensor W

$$\mathcal{P}_{\mathcal{S}}(|m\rangle) = \mathrm{perm}(W) = \sum_{\sigma,\rho\in S_n}\prod_{j=1}^{n} W_{\sigma_j,\rho_j,j}. \tag{22}$$

Where the elements of W are given by

$$W_{i,j,k} = M_{k,j}M *_{l,j} \mathcal{S}_{l,k}. \tag{23}$$

2.1.8 The Output State

The density operator of the output state $\hat{\rho}'$ is given by the transformation of the density operator of the input state $\hat{\rho}$ by the operator T that depends on the interferometer $\hat{U}$ and the partial distinguishability of the photons

$$\hat{\rho}' = \hat{T}\hat{\rho}\hat{T}^{\dagger}. \tag{24}$$

2.1.9 Measuring The Ancillary Modes

The measurement of the ancillary modes changes the state. The operation that is performed by the measurement of the ancillary modes with the outcome $|m\rangle = |m_1 m_2 \ldots m_{N-1}\rangle$ is the tensor product between the 1-mode identity operator and the projection onto $|m\rangle$

and all the states $|\tilde{m}\rangle$ containing distinguishable photons; see Equation 25 [12, 11]. The identity operator conserves the remaining mode

$$\hat{M} = I \otimes \left(|\vec{m}\rangle\langle\vec{m}| + \sum_k |\tilde{m}_k\rangle\langle\tilde{m}_k| + \dots \right). \tag{25}$$

Here, the dots involve modes containing two or more distinguishable photons.

The density matrix of the remaining photon can then be computed by

$$\hat{\rho}_{p.s.} = \frac{\hat{M}\hat{\rho}'\hat{M}^\dagger}{\mathcal{P}_{\mathcal{S}(\epsilon)}(|m\rangle)}. \tag{26}$$

Consider the special case where we measure the coincidental state and suppose that the interferometer can be considered to be a balanced beamsplitter. That is, the photon has equal probability to end up in each one of the output modes, regardless of its input mode. We conjecture that the probability of measuring the coincidental state does not depend on the input mode of the distinguishable photon does not depend on the input mode of the distinguishable photon $\mathcal{P}_{\tilde{\mathcal{S}}_k} = \mathcal{P}_{\tilde{\mathcal{S}}_l}$ in the case that there is only one distinguishable photon and we conjecture that we have equal probability of finding the photon in each of the output modes

$$\hat{\rho}_{p.s.} = \frac{\mathcal{P}_{\mathbb{E}}(|\mathbf{1}\rangle)(1-\epsilon)^n|1\rangle\langle 1| + \mathcal{P}_{\tilde{\mathcal{S}}}(|1\rangle)n\epsilon(1-\epsilon)^{n-1}(\frac{n-1}{n}|1\rangle\langle 1| + \frac{1}{n}|\tilde{1}\rangle\langle\tilde{1}|) + O(\epsilon^2)}{\mathcal{P}_{\mathcal{S}(\epsilon)}(|\mathbf{1}\rangle)}. \tag{27}$$

2.2 Numerical Model

This section describes the model that we employed to simulate the reduced error of the distilled photon when using the five-mode QFT for the photon distillation protocol, given that the measurement outcome of the ancillary modes is described by the measurement $|m\rangle = |m_1 m_2 m_3 m_4\rangle$. We simulate a six-mode interferometric experiment that carries out a five-mode QFT on the initial five photons and a HOM experiment on the last two photons; as illustrated in Figure 3. Subsequently, we use the HOM effect to compute the reduced error by employing the relation of Equation 19.

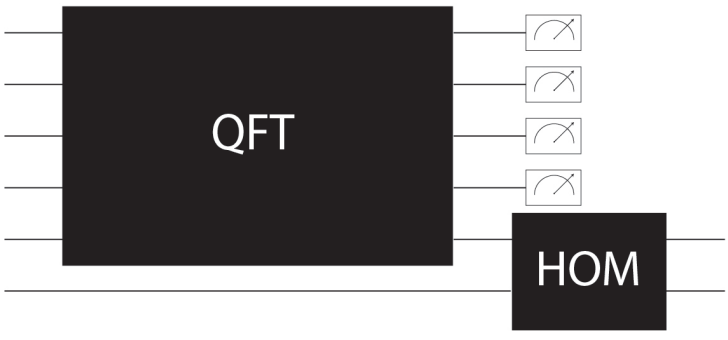


FIGURE 3: This figure shows a schematic overview of the interferometer which is used to simulate the reduced error.

The measurements that are interesting for the photon distillation protocol are required to have one photon in the fifth mode so that $|m\rangle|1\rangle = |m_1 m_2 m_3 m_4\rangle|1\rangle$. The possible outcomes after the HOM experiment are then $P_{|m\rangle|20\rangle}, P_{|m\rangle|11\rangle}, P_{|m\rangle|02\rangle}$.

The calculation of these probabilities can be done by applying Tichy's formalism; see Equation 22. The interferometer is then described by the following matrix U, which is the product of the matrices that describe the five-mode QFT and the HOM interferometer

$$U = U_{\text{QFT}} U_{\text{HOM}} = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & e^{\frac{2i\pi}{5}} & e^{\frac{4i\pi}{5}} & e^{-\frac{4i\pi}{5}} & e^{-\frac{2i\pi}{5}} & e^{-\frac{2i\pi}{5}} \\ 1 & e^{\frac{4i\pi}{5}} & e^{-\frac{2i\pi}{5}} & e^{\frac{2i\pi}{5}} & e^{-\frac{4i\pi}{5}} & e^{-\frac{4i\pi}{5}} \\ 1 & e^{-\frac{4i\pi}{5}} & e^{\frac{2i\pi}{5}} & e^{-\frac{2i\pi}{5}} & e^{\frac{4i\pi}{5}} & e^{\frac{4i\pi}{5}} \\ 1 & e^{-\frac{2i\pi}{5}} & e^{-\frac{4i\pi}{5}} & e^{\frac{4i\pi}{5}} & e^{\frac{2i\pi}{5}} & e^{\frac{2i\pi}{5}} \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}. \tag{28}$$

Where we have the following for the individual matrices.

$$U_{\text{QFT}} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & e^{\frac{2i\pi}{5}} & e^{\frac{4i\pi}{5}} & e^{-\frac{4i\pi}{5}} & e^{-\frac{2i\pi}{5}} & 0 \\ 1 & e^{\frac{4i\pi}{5}} & e^{-\frac{2i\pi}{5}} & e^{\frac{2i\pi}{5}} & e^{-\frac{4i\pi}{5}} & 0 \\ 1 & e^{-\frac{4i\pi}{5}} & e^{\frac{2i\pi}{5}} & e^{-\frac{2i\pi}{5}} & e^{\frac{4i\pi}{5}} & 0 \\ 1 & e^{-\frac{2i\pi}{5}} & e^{-\frac{4i\pi}{5}} & e^{\frac{4i\pi}{5}} & e^{\frac{2i\pi}{5}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{29}$$

$$U_{\text{HOM}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \tag{30}$$

The partial distinguishability matrix $\mathcal{S}$ in the simulation is given by

$$\mathcal{S}_{ij} = \begin{cases} 1 - \epsilon & \text{if } i = j \\ \epsilon & \text{if } i \neq j \end{cases}$$

To calculate the reduced error, we have to consider the probability of obtaining these outcomes, given that the measurement outcome has happened. That is, we have to normalise these probabilities.

$$\begin{aligned} P(|20\rangle) &= \frac{P(|m\rangle|02\rangle)}{P(|m\rangle|20\rangle) + P(|m\rangle|11\rangle) + P(|m\rangle|02\rangle)} \\ P(|11\rangle) &= \frac{P(|m\rangle|02\rangle)}{P(|m\rangle|20\rangle) + P(|m\rangle|11\rangle) + P(|m\rangle|02\rangle)} \\ P(|02\rangle) &= \frac{P(|m\rangle|02\rangle)}{P(|m\rangle|20\rangle) + P(|m\rangle|11\rangle) + P(|m\rangle|02\rangle)} \end{aligned}$$

The reduced error ϵ' is then calculated for $0 < \epsilon < 1$ with MATLAB.

3 Performance of the 5-mode QFT in photon distillation

3.1 Estimation of the Error Reduction for Balanced Beamsplitters

The density matrix after measurement $\rho_{p.s.}$ can be represented in a form similar to Equation 6. In this form we can relate the fidelity of the remaining photon to the reduced error.

$$\mathcal{F} = \langle\psi_0|\rho_{p.s.}|\psi_0\rangle = 1 - \epsilon'. \tag{31}$$

By using Equation 31 Equation 27, we can calculate an upper limit for the diminished error in the low error situation. Although Equation 27 hides terms in $O(\epsilon^2)$, it is evident that it follows the general form of Equation 32.

$$\begin{aligned}
1 - \epsilon' = \mathcal{F} &\geq \frac{(1-\epsilon)^n \binom{n}{0} p_0\alpha_0 + \epsilon(1-\epsilon)^{n-1}\binom{n}{1}p_1\alpha_1 + ... + \epsilon^n\binom{n}{n}p_n\alpha_n}{(1-\epsilon)^n\binom{n}{0}\alpha_0 + \epsilon(1-\epsilon)^{n-1}\binom{n}{1}\alpha_1 + ... + \epsilon^n\binom{n}{n}\alpha_n} && (32)\\
&= \frac{\sum_{k=0}^n p_k\binom{n}{k}\alpha_k\epsilon^k(1-\epsilon)^{n-k}}{\sum_{k=0}^n \binom{n}{k}\alpha_k\epsilon^k(1-\epsilon)^{n-k}} && (33)\\
&= \frac{\sum_{k=0}^n \frac{n-k}{n}\binom{n}{k}\alpha_k\epsilon^k(1-\epsilon)^{n-k}}{\sum_{k=0}^n \binom{n}{k}\alpha_k\epsilon^k(1-\epsilon)^{n-k}} && (34)\\
&= \frac{\sum_{k=0}^n \frac{n}{n}\beta_k\epsilon^k(1-\epsilon)^{n-k}}{\sum_{k=0}^n \beta_k\epsilon^k(1-\epsilon)^{n-k}} - \frac{\sum_{k=0}^n \frac{k}{n}\beta_k\epsilon^k(1-\epsilon)^{n-k}}{\sum_{k=0}^n \beta_k\epsilon^k(1-\epsilon)^{n-k}} && (35)\\
&= 1 - \frac{\sum_{k=0}^n \frac{k}{n}\beta_k\epsilon^k(1-\epsilon)^{n-k}}{\sum_{k=0}^n \beta_k\epsilon^k(1-\epsilon)^{n-k}} && (36)\\
& && (37)
\end{aligned}$$

The binomial coefficient is the result of the number of ways the distinguishable photons can be distributed over the output modes. Given that there are i indistinguishable photons present, we let α_i denote the transition probability from the state in mode occupation representation $|\mathbf{1}\rangle$ to $|\mathbf{1}\rangle$ and we let $p_k = \frac{n-k}{n}$ denote the probability that the photon is indistinguishable, and $\beta_i = \binom{n}{i}\alpha_i$. This gives an upper-bound for the reduced error ϵ'.

$$\begin{aligned}
\epsilon' &\leq \frac{\sum_{k=0}^n \frac{k}{n}\beta_k\epsilon^k(1-\epsilon)^{n-k}}{\sum_{k=0}^n \beta_k\epsilon^k(1-\epsilon)^{n-k}} && (38)\\
&= \frac{(1-\epsilon)^n}{(1-\epsilon)^n}\frac{\sum_{k=0}^n \frac{k}{n}\beta_k(\frac{\epsilon}{1-\epsilon})^k}{\sum_{k=0}^n \beta_k(\frac{\epsilon}{1-\epsilon})^k} && (39)\\
&= \frac{\sum_{k=0}^n \frac{k}{n}\beta_k(\frac{\epsilon}{1-\epsilon})^k}{\sum_{k=0}^n \beta_k(\frac{\epsilon}{1-\epsilon})^k} && (40)\\
&= \frac{1}{n}\frac{\sum_{k=0}^n k\beta_k(\frac{\epsilon}{1-\epsilon})^k}{\sum_{k=0}^n \beta_k(\frac{\epsilon}{1-\epsilon})^k} && (41)\\
&\approx \frac{1}{n}\frac{\beta_1(\frac{\epsilon}{1-\epsilon})}{\beta_0 + \beta_1(\frac{\epsilon}{1-\epsilon})} = \frac{1}{n}\frac{\beta_1\epsilon}{\beta_0 + \epsilon(\beta_1 - \beta_0)}. && (42)
\end{aligned}$$

By Taylor expanding Equation 42 we find the first order approximation of the reduced error ϵ' as a function of the specific probabilities.

$$\epsilon' \approx \frac{1}{n}\frac{\beta_1}{\beta_0}\epsilon + O(\epsilon^2) \approx \frac{1}{n}\frac{\binom{n}{1}\alpha_1}{\binom{n}{0}\alpha_0}\epsilon = \frac{1}{n}\frac{\frac{n!}{(n-1)!1!}\alpha_1}{\frac{n!}{n!0!}\alpha_0}\epsilon = \frac{\alpha_1}{\alpha_0}\epsilon = \frac{\mathcal{P}_{\tilde{\mathcal{S}}}(|1\rangle)}{\mathcal{P}_{\mathbb{E}}(|1\rangle)}\epsilon \tag{43}$$

The success of the 5-mode QFT in the photon distillation protocol is dependent on two elements. Firstly, the amount of reduction, if any, that is achieved. Secondly, the probability of detecting a heralding pattern which indicates a reduced error. This probability is a key factor in determining the resources needed to distil the photon. If the probability of detecting such a pattern is very low, then a large number of photons will be required to achieve the desired outcome.

3.2 Error reduction capabilities of the heralding patterns

We simulate the reduced error of the photon state for the remaining photon given some initial error. Figure 4 shows the error-reduction capabilities of each of the heralding patterns. There are clearly three lines that reduce the error given that the error is smaller than some cutoff error $\epsilon \sim 0.48$. Hence, multiple patterns can be used to reduce the error. In contrast to the initial idea for the 3- and 4- mode photon distillation protocol, which relied solely on the coincidental state.

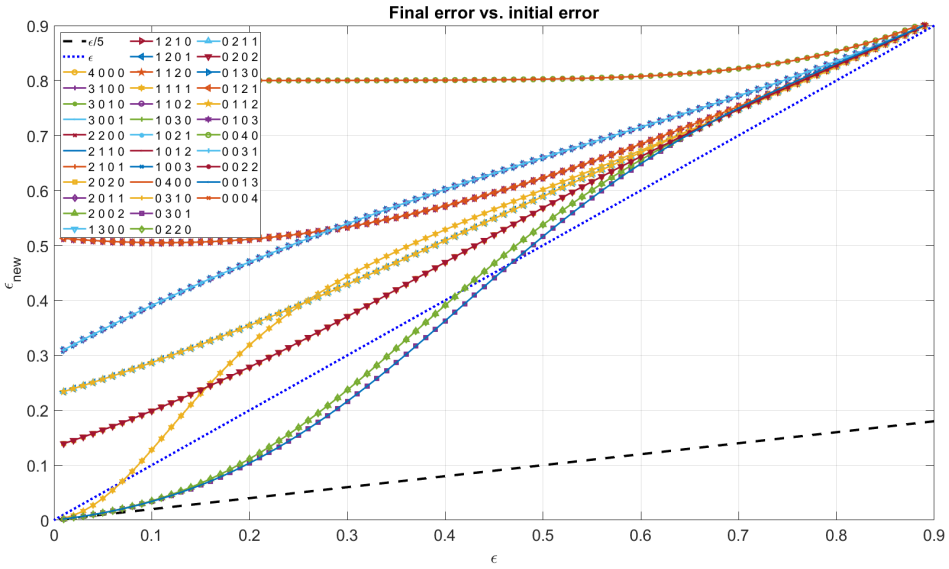


FIGURE 4: Error Reducing Capabilities: Each of the curves describes the new error as function of the initial error for each of the heralding patterns, described by the mode occupancy representation. The blue dotted line represents no change in error. The black striped line represents an error reduction by a factor 1/5.

Figure 5 shows that seven heralding patterns improve the error. The factor of error reduction for these valid patterns seems to converge to 1/5, which supports our initial guess. Notable is that the coincidental pattern (1111) seems to achieve a worse error reduction than the other patterns.

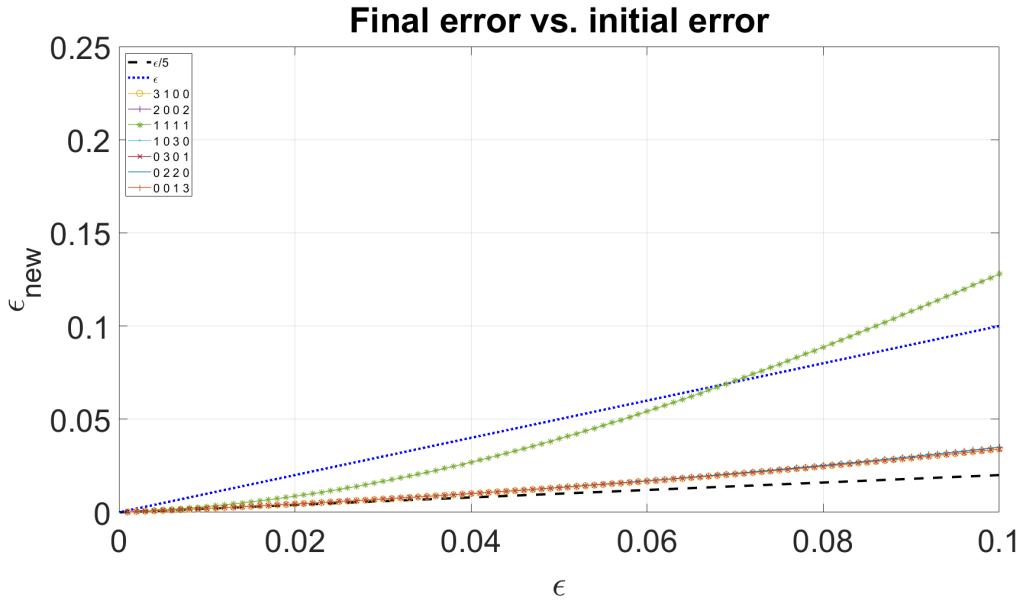


FIGURE 5: Convergence of the New Error: Each of the curves describes the new error as function of the initial error for each of the heralding patterns. The blue dotted line represents no change in error. The black striped line represents an error reduction by a factor 1/5.

3.3 Success Probability of the 5-mode QFT

The other important aspect of the performance of the 5-mode QFT photon distillation is the success probability. Figure 6 shows the probability of measuring each valid pattern, focused on the low error regime. Additionally, it shows the sums of these probabilities of measuring any pattern, that is, the probability of measuring a pattern that leads to error reduction. Considering that multiple patterns lead to error reduction, the natural question is whether we could combine the error reduction efforts of the patterns. If decisions can be made based on the observed pattern, then any valid pattern can be employed to reduce the mistake in the photon state. Consequently, the probability of successful heralding is the total of the probabilities of measuring a valid pattern, significantly increasing the success rate of the system.

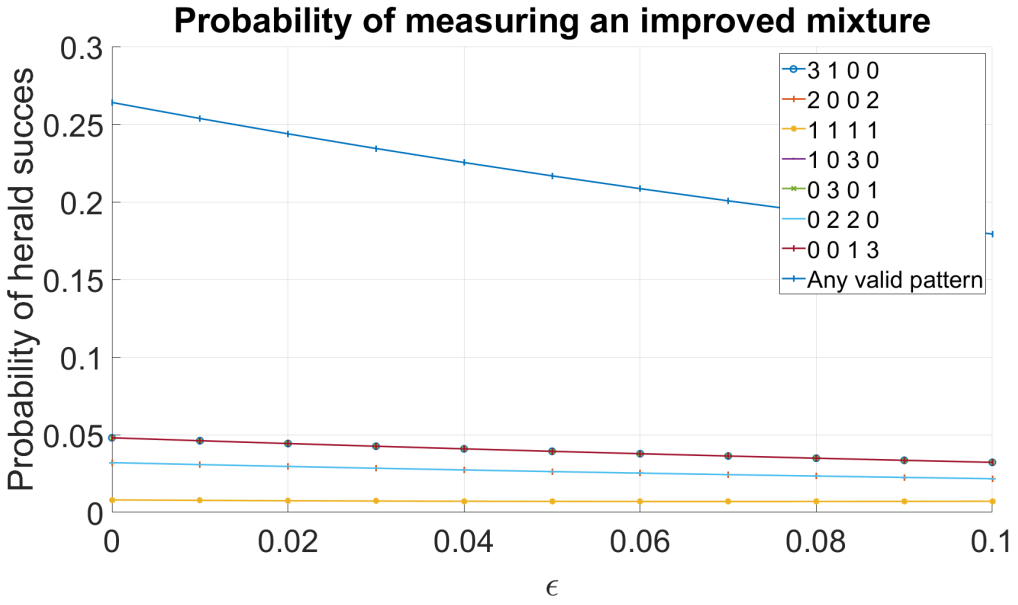
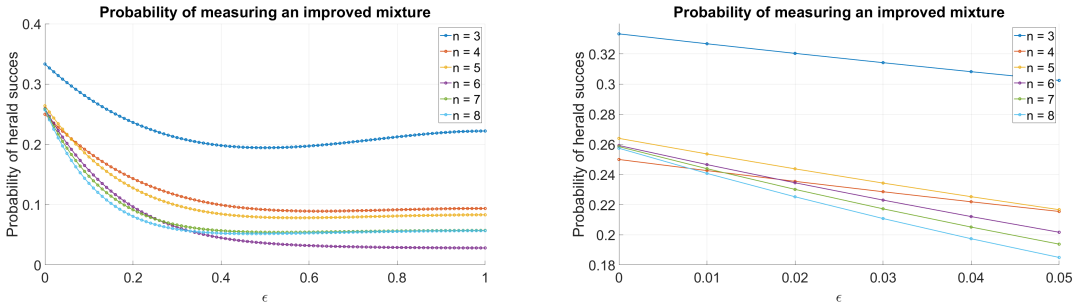


FIGURE 6: The probability of measuring an heralding pattern that heralds an reduced error. The probability of measuring a specific pattern is rather low, for all valid heralding patterns. The probability of measuring any pattern that indicates a reduced error is roughly 0.25.

3.3.1 Total success probability for QFT with different number of modes

The 5-mode QFT system's success rate appears to be larger than 1/5, so we conducted a simulation to compare the success rate of higher-order systems. The results of this simulation are presented in figure 7, which shows the success rate for the three- to eight-mode systems.



(A) Full error range

(B) Low error regime

FIGURE 7: Probability of heralding an improved state for 3- to 8-mode QFT photon distillation

Remarkably, the 5-8 modes have a success probability that is much higher than the expected $1/N$ behaviour.

4 Discussion

The 5-mode quantum Fourier transform can be used in a linear optical network to distil the photon state. In the low-error regime, it reduces the error by a factor 1/5 as expected. In contrast to the 3- and 4-mode photon distillation protocol, there are different patterns which can herald an improved photon state. The probability of measuring the coincidental state is much lower than in the 3-mode photon distillation protocol, and so is the probability of measuring a specific valid heralding pattern. This implies that the system would require a lot of resources when looking at a single heralding pattern. However, the probability of success increases a lot when decisions can be made based on the measured heralding pattern. The probability of success is then also higher than expected 1/5, which raised the question how higher-order systems react. A quick simulation of photon distillation protocols relying on QFT with more modes showed that the five-to-eighth-mode systems all have a higher probability of success than the expected $1/n$ behaviour and the probability of success for these modes seems to lie around 0.25. This suggests that scaling this system could improve its resource efficiency (number of photons required).

Consider, for example, a nested 3-mode QFT photon distillation scheme as in figure 7. The probability of successful distillation, assuming that all 3-mode QFT schemes must be successful, is then $1/3^4 = 1/81$. Therefore, we expect this scheme to succeed every 81 attempts, costing 9 photons each attempt, yielding a total of 729 photons to reduce the error by a factor 1/9. Consider a 9 mode QFT with success probability ~ 0.25, having the same error reduction factor. This would only require 36 photons. We could of course consider other nested schemes; this example is just to highlight the possible efficiency of a system using QFT's with more modes.

The 5-mode quantum Fourier transform seems to have the capabilities to perform well in a photon distillation scheme if decisions can be based on the measurement of a heralding pattern.

Furthermore, we have found an estimation for the error reduction that is linear with the ratio between the behaviour of the indistinguishable system and the system containing a single fully distinguishable photon by making use of some assumptions that are reasonable given our current understanding of the subject. The assumption that the probability to obtain the outcome $|\mathbf{1}\rangle$ with one fully indistinguishable photon does not depend on the input mode of the indistinguishable photon could be investigated by investigating the influence of row echelon operations on the partial distinguishability matrix on the outcome of Tichy's formalism. An exercise which could provide further insight into the theory of sampling with partial distinguishable photons. The estimation could be employed to estimate the functionality of a certain interferometer in the photon distillation protocol.

Lastly, in the 5-mode QFT scheme the valid heralding patterns seem to achieve an error reduction of $1/n$. The estimation of Equation 42 suggest this is can be traced back to a ratio between two distinct cases. Is there a possible explanation for why the ratio of these heralding patterns appears to be $1/n$, such as the symmetry of the linear optical network?

References

[1] Sara Bartolucci, Patrick Birchall, Hector Bombín, Hugo Cable, Chris Dawson, Mercedes Gimeno-Segovia, Eric Johnston, Konrad Kieling, Naomi Nickerson, Mihir Pant, Fernando Pastawski, Terry Rudolph, and Chris Sparrow. Fusion-based quantum com-

putation. *Nature Communications*, 14(1):912, February 2023. Number: 1 Publisher: Nature Publishing Group.

[2] R. Barzel and C. Lämmerzahl. Role of indistinguishability and entanglement in Hong-Ou-Mandel interference and finite-bandwidth effects of frequency-entangled photons. *Physical Review A*, 107(3), 2023.

[3] Agata M. Brańczyk. Hong-Ou-Mandel Interference, October 2017. arXiv:1711.00080 [quant-ph].

[4] B. Brecht, Dileep V. Reddy, C. Silberhorn, and M.G. Raymer. Photon Temporal Modes: A Complete Framework for Quantum Information Science. *Physical Review X*, 5(4):041017, October 2015. Publisher: American Physical Society.

[5] Nicolas Fabre, Maria Amanti, Florent Baboux, Arne Keller, Sara Ducci, and Pérola Milman. The Hong-Ou-Mandel experiment: from photon indistinguishability to continuous variables quantum computing. *The European Physical Journal D*, 76(10):196, October 2022. arXiv:2206.01518 [quant-ph].

[6] Vikas Hassija, Vinay Chamola, Vikas Saxena, Vaibhav Chanana, Prakhar Parashari, Shahid Mumtaz, and Mohsen Guizani. Present landscape of quantum computing. *IET Quantum Communication*, 1(2):42–48, 2020. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1049/iet-qtc.2020.0027.

[7] C. K. Hong, Z. Y. Ou, and L. Mandel. Measurement of subpicosecond time intervals between two photons by interference. *Physical Review Letters*, 59(18):2044–2046, November 1987. Publisher: American Physical Society.

[8] E. Knill, R. Laflamme, and G. J. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409(6816):46–52, January 2001. Number: 6816 Publisher: Nature Publishing Group.

[9] Pieter Kok, W. J. Munro, Kae Nemoto, T. C. Ralph, Jonathan P. Dowling, and G. J. Milburn. Linear optical quantum computing with photonic qubits. *Reviews of Modern Physics*, 79(1):135–174, January 2007. Publisher: American Physical Society.

[10] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien. Quantum computers. *Nature*, 464(7285):45–53, March 2010. Number: 7285 Publisher: Nature Publishing Group.

[11] J. Marshall. Distillation of Indistinguishable Photons. *Physical Review Letters*, 129(21), 2022.

[12] M.A. Nielsen and I.L. Chuang. *Quantum computation and quantum information*. Quantum computation and quantum information. 2000.

[13] Michael A. Nielsen. Optical Quantum Computation Using Cluster States. *Physical Review Letters*, 93(4):040503, July 2004. Publisher: American Physical Society.

[14] T. Opatrný, N. Korolkova, and G. Leuchs. Mode structure and photon number correlations in squeezed quantum pulses. *Physical Review A*, 66(5):053813, November 2002. Publisher: American Physical Society.

[15] R. Raussendorf, J. Harrington, and K. Goyal. A fault-tolerant one-way quantum computer. *Annals of Physics*, 321(9):2242–2270, September 2006.

[16] Jelmer J. Renema, Hui Wang, Jian Qin, Xiang You, Chaoyang Lu, and Jianwei Pan. Sample-efficient benchmarking of multiphoton interference on a boson sampler in the sparse regime. *Physical Review A*, 103(2):023722, February 2021. Publisher: American Physical Society.

[17] V. S. Shchesnovich. Partial indistinguishability theory for multiphoton experiments in multiport devices. *Physical Review A*, 91(1):013844, January 2015. Publisher: American Physical Society.

[18] Peter W. Shor. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer, August 1995.

[19] Christopher Sparrow. Quantum interference in universal linear optical devices for quantum computation and simulation. September 2017. Accepted: 2019-03-14T13:36:05Z Publisher: Imperial College London.

[20] Malte C. Tichy. Sampling of partially distinguishable bosons and the relation to the multidimensional permanent. *Physical Review A*, 91(2):022316, February 2015. arXiv:1410.7687 [physics, physics:quant-ph].

[21] M.C. Tichy, M. Tiersch, F. De Melo, F. Mintert, and A. Buchleitner. Zero-transmission law for multiport beam splitters. *Physical Review Letters*, 104(22), 2010.

[22] C. Yeh, K. Ha, S. B. Dong, and W. P. Brown. Single-mode optical waveguides. *Applied Optics*, 18(10):1490–1504, May 1979. Publisher: Optica Publishing Group.

[23] N. Yoran and B. Reznik. Deterministic Linear Optics Quantum Computation with Single Photon Qubits. *Physical Review Letters*, 91(3):037903, July 2003. Publisher: American Physical Society.