universität wien

MASTERARBEIT

Titel der Masterarbeit

„Quantum Interference and Tomography of Linear Optical Networks“

Verfasser

Christian Schmidt, BSc

angestrebter akademischer Grad

Master of Science (MSc)

Wien, 2014

Studienkennzahl lt. Studienblatt:	A 066 876
Studienrichtung lt. Studienblatt:	Masterstudium Physik
Betreuer:	Assoz. Prof. Dr. Philip Walther

Contents

Abstract

Quantum interference in passive linear optical networks has recently attracted great interest in the field of optical quantum computation. Recent technological progress allowed the fabrication of large integrated interferometric networks that are composed out of many beam splitters and phase shifting elements. A major issue is the reliable tomography of such networks.

Surprisingly, the quantum interference of two photons in the network is sufficient to provide a high quality set of data for tomography purposes. In this thesis, the model for two-photon interference in linear optical networks is improved. This detailed description is based on a precise modeling of the input state, which includes modal mismatch originating from the parametric down conversion process as well as unavoidable frequency correlations.

Then current approaches to tomography of linear optical networks are investigated and benchmarked. It is important to distinguish the approaches into those using a sufficient set of primary data (passive) and those utilizing an over-sufficient set of data (active). It is conjectured that active approaches exhibit better error robustness than the currently known passive schemes, providing analytical and numerical evidence for this conjecture. This is the main reason that active approaches generate better approximations of the actual network in comparison to passive ones.

A major drawback of active approaches has been their inability to reconstruct optical black boxes, i.e. networks of unknown structure. Here a novel hybrid-approach is proposed which combines the advantages of active and passive methods, while eliminating their shortcomings.

Finally, the potential of this hybrid-approach is validated numerically, showing that it outperforms all currently utilized approaches.

Zusammenfassung

Passive linear optische Netzwerke und Ihre Anwendung in Quantenoptik-Experimenten hat jüngst vielfältige Aufmerksamkeit auf sich gezogen. Neue Fabrikationsmethoden ermöglichen die Herstellung großer integrierter interferometrischer Netzwerken, die aus einzelnen Strahlenteiler- und Phasenschieber-Elementen zusammengesetzt sind. Eine große Herausforderung ist allerdings die quantitativ zuverlässige Charakterisierung oder tomographische Rekonstruierung dieser Netzwerke.

Erstaunlicher Weise ist die Detektion von Zwei-Photonen Interferenzereignissen ausreichend für diese Zwecke. In dieser Arbeit wird die Beschreibung von Zwei-Photonen Interferenzereignissen in linear optischen Netzwerken erweitert und verbessert, was eine nötige Voraussetzung für die zuverlässige Charakterisierung ist. Das hierfür herangezogene Modell basiert auf einer detaillierten Beschreibung des Eingangszustands, die Frequenzkorrelationen, sowie die andere Moden-Diskrepanzen der Ein-Photonen Quelle mit einbezieht.

Hiernach werden gegenwärtige Ansätze zur Tomographie von linearen optischen Netzwerken untersucht und bewertet. Es wird unterscheiden zwischen jenen Methoden, die ein über-bestimmte Menge an Primärdaten nutzen (aktiv) und denen, die eine ausreichende Menge verwenden (passiv). Damit kann behauptet werden, dass aktive Ansätze wie eine bessere Fehlerrobustheit besitzen als die gegenwärtig bekannten passiven Herangehensweisen. Weiter werden numerische und analytische Evidenz für diese Behauptung geliefert. Die Fehlerrobustheit ist der entscheidende Grund dafür, dass aktive Ansätze eine bessere Approximation des zugrundeliegenden Netzwerkes liefern.

Bis jetzt war ein großes Nachteil des aktiven Ansatzes allerdings, dass es sich nicht auf Black-Box-artige Netzwerke, also Netzwerke unbekannter Struktur, anwenden ließ. Wir schlagen eine neuartige Herangehensweise vor, die dieses Problem löst und die passiven Ansätze mit dem aktiven Ansatz so kombiniert, dass die jeweiligen

Nachteile beseitigt werden und die Vorteile beider Methoden in einer kombinierten Version zum tragen kommen.
Abschließend wird numerisch gezeigt, dass dieser Ansatz alle gegenwärtig genutzten an Präzision übertrifft.

I. INTRODUCTION

After the pioneering work of Planck [1] and subsequently by Schrödinger, Heisenberg, Bohr, Pauli, Dirac and many others in the beginning of the 19^{th} century, the field of quantum mechanics had spawned many new developments in Physics. Quantum information and quantum computation is prominent example. The major promise, increasing the computational efficiency with respect to classical computers, is believed to be crucial for simulating and solving many fundamental problems in science [2–7].

Quantum computation has been subject to many major developments in the last decades, and many promising approaches towards its physical implementation exist. Experimental approaches to encode quantum information reach from photons to trapped ions, nuclear magnetic resonance, quantum dots, dopants in solids, superconductors, and other systems [8].

Utilizing photons as quantum bits (qubits) has attracted vast attention ever since Knill, Laflamme and Milburn (KLM) proposed their seminal scheme for efficient quantum computation with linear optics in 2001 [9] and the measurement bases quantum computing by Raussendorf et al. [10].

In 2013 Aaronson and Arkhipov introduced a rudimentary quantum computer, often referred to as BosonSampling [11]. This entirely passive scheme solves a sampling problem that is intractable on a classical computer efficiently.[1] An experimental realization of BosonSampling, for large instances of photons ($N \approx 30$), would provide strong evidence against the extended Church-Turing Thesis, which states that all computational problems that are efficiently solvable by a realistic physical device, are efficiently solvable by a probabilistic Turing machine [11, 12]. The KLM scheme requires ancillary photons and adaptive measurements whereas

[1] Mathematically speaking their result is just a conjecture. However, the unproven assumptions are reasonable and widely believed to be true.

BosonSampling implements a classically hard problem by utilizing exclusively resource efficient passive optical elements.
The recent experimental breakthrough in implementing large scale networks for photonic quantum computing is accomplished by utilizing integrated optical circuits [13–16]. These waveguide circuits offer better scalability by shrinking bulk optics into microchip scale devices and are interferometric stable which makes them the current state-of-the-art for linear optical quantum computing.
Generating the necessary computational state – the optical qubits – is still challanging. Today, the most prominent implementation of such photon sources are parametric down converters (PDC) that exploit nonlinear crystals to spontaneously emit pairs of photons [17–20]. Research on single-photon PDC-sources dates back to the work of Louisell [17] and Burnham and Weinberg [18]. This effect has been used in many quantum optics applications and is up to now a nearly inevitable source for optical qubits [8, 21–23].

For decades the common method to manipulate photonic qubits has been bulk optical elements acting on e.g. the spatial or polarization degree of freedom. Usually multiple optical elements are cascaded to form interferometers. Embedding those interferometers in integrated optical chips significantly improves their interferometric stability. However, it is extremely challenging to trace slight fabrication imperfections of each optical element back due to the embedded structure. Therefore a key point is the characterization or tomography of the integrated linear optical circuits. In the last two years three different approaches, 'Bristol' [24], 'Brisbane' [25] and 'Vienna' [26], have been applied for the tomography of such networks. Until now no systematic benchmark has been carried out hence their respective performance remains unknown.

In this thesis two issues will be adressed. First, the quantum interference in passive

linear optical networks which is strongly related to the properties of the single-photon sources. The quantum nature of n-photon interference can be characterized by the n^{th}-order correlation function [27]. In particular a detailed investigation of the 2^{nd} -order correlation function is necessary as an elaborate model is required for the network tomography schemes in [24, 26]. Subsequently, I will investigate the different approaches to network tomography and give an explicit analysis of their respective performances under realistic experimental conditions. Eventually a new hybrid-approach that outperforms the current methods will be proposed and investigated.

Methods

In this and the following section all methods and concepts necessary will be introduced. However, familiarity with the basic methods and concepts of quantum mechanics, quantum computation and quantum information will be assumed (for an introduction see for instance [7]) as well as the standard mathematical techniques. Notations will be introduced were it seems most suitable.
The first section of the methods will cover the theoretical framework of linear quantum optics and other fundamental ideas needed. This will be followed by a brief introduction of the different approaches [24–26] to characterize linear optical networks that will be investigated. In the last section additional statistical and numerical methods will be discussed.

II. THE QUANTUM THEORY OF LINEAR OPTICS

Starting point is the quantum mechanical treatment of optical fields. Namely the quantization of the electromagnetic field which is classically governed by the Maxwell-Equations [27].
In the framework of quantum mechanics the electromagnetic field is described by its quantized states of oscillation. Or more formally the electromagnetic field can be expanded into its (positive and negative) frequency parts

$$\hat{\mathbf{E}}(\mathbf{r},t) = \int_{-\infty}^{\infty} d\omega\, \hat{\boldsymbol{e}}(\omega,\mathbf{r})\, \mathrm{e}^{-i\omega t}. \tag{1}$$

Bold symbols indicate matrices (vectors) throughout this thesis. For now (1) is nothing but the Fourier representation. (1) can be decomposed into positive and

negative frequency parts

$$\hat{\mathbf{E}}(\mathbf{r},t) = \hat{\mathbf{E}}^{(+)}(\mathbf{r},t) + \hat{\mathbf{E}}^{(-)}(\mathbf{r},t) \tag{2}$$

that are given by

$$\hat{\mathbf{E}}^{(+)}(\mathbf{r},t) \;=\; \int_0^\infty d\omega\, \hat{\boldsymbol{e}}(\omega,\mathbf{r})\, \mathrm{e}^{-i\omega t} \tag{3}$$

$$\hat{\mathbf{E}}^{(-)}(\mathbf{r},t) \;=\; \int_0^\infty d\omega\, \hat{\boldsymbol{e}}^\dagger(\omega,\mathbf{r})\, \mathrm{e}^{-i\omega t}. \tag{4}$$

Where the Hermitian property of the electromagnetic field $\hat{\mathbf{E}} = \hat{\mathbf{E}}^\dagger$ (note that hence $\hat{\mathbf{E}}^{(+)\dagger} = \hat{\mathbf{E}}^{(-)}$) has been used. From now on the calculations will always be restrict to $\mathbf{r} \equiv z$.

For a field in a lossless dielectric material the operators take the form [28]

$$\hat{\mathbf{E}}^{(+)}(z,t) = \hat{\mathbf{E}}^{(-)\dagger}(z,t) = \int d\omega\, \underbrace{i \left[\frac{\hbar\omega}{4\pi\epsilon_0 c\, A\, n(\omega)}\right]^{\frac{1}{2}}}_{\coloneqq A(\omega)} \hat{a}(\omega)\, \mathrm{e}^{i\,(k(\omega)z - \omega t)}. \tag{5}$$

Where ϵ_0 is the vacuum permittivity, c the speed of light, A the cross-sectional area in the xy plane and $n(\omega)$ the refractive index.

The annihilation operator $\hat{a}(\omega)$ and its adjoint, the creation operator $\hat{a}^\dagger(\omega)$, are the quantum mechanic harmonic oscillator operators and for now restricted to the frequency modes. Later additional (discrete) modes will be introduced which can encode additional degrees of freedom, such as polarization or spatial modes. As is well known these operators act on the *Fock basis* in the following manner: The annihilation operator $\hat{a}(\omega)$ is annihilating a photon of frequency (in the oscillation mode) ω and the creation operator is creating one with frequency (in the oscillation

mode) ω.

$$\begin{aligned} \hat{a}(\omega)\,|N(\omega)\rangle &= \sqrt{N(\omega)}\,|N(\omega)-1\rangle \\ \hat{a}^{\dagger}(\omega)\,|N(\omega)\rangle &= \sqrt{N(\omega)+1}\,|N(\omega)+1\rangle\,. \end{aligned} \tag{6}$$

Here the occupation number $N(\omega)$ is the total number of photons of frequency ω. From (6) one may define $\hat{N}(\omega) \equiv \hat{a}^{\dagger}(\omega)\hat{a}(\omega)$ as the number operator that acts on a Fock state as follows: $\hat{N}(\omega)\,|N(\omega)\rangle = N(\omega)\,|N(\omega)\rangle$. These basic operators obey the following commutation relations

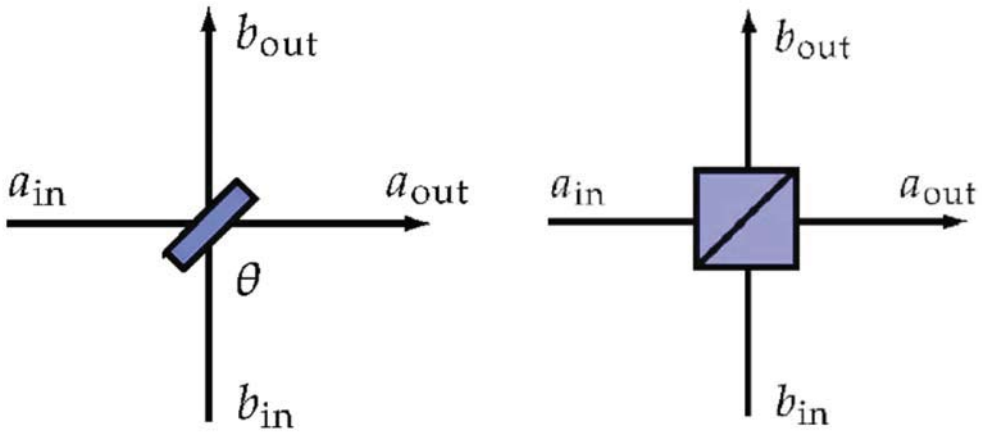
$$\begin{aligned} [\hat{a}(\omega), \hat{a}^{\dagger}(\omega')] &= \delta(\omega-\omega') \text{ and} \\ [\hat{a}(\omega), \hat{a}(\omega')] &= [\hat{a}^{\dagger}(\omega), \hat{a}^{\dagger}(\omega')] = 0\,. \end{aligned} \tag{7}$$

Now that a quantum version of the electromagnetic field has been accomplished, the main building blocks for linear optical quantum computing can be introduced. Most generally speaking linear optics consists of all the elements whose Hamiltonians are of the form

$$\hat{H} = \sum_{kl} M_{lk}\hat{a}_k^{\dagger}\hat{a}_l, \tag{8}$$

or that can be mapped on such a Hamiltonian under a linear transformation of the annihilation and creation operators. However, here it shall be enough to restrict ourselves to the subset of Hamiltonians that is preserving the total number of photons, i.e. $[\hat{H}, \hat{N}] = 0$.

Within the scope of this thesis mainly the phase shifter and beam splitter are of interest. However, for completeness also the polarizing beam splitter and polarization rotations will be introduced. *Note* that in the following $\hbar$ is usually set to one.



(a) (b)

Figure 1: Figurative illustration of (a) a lossless beam splitter with transmission amplitude $\cos\theta$ and (b) a polarizing beam splitter – taken from [21].

A. Elements of Linear Optics

The Single Mode Phase Shifter (PS) varies the phase of the electromagnetic field in a particular mode a (Fig. 1a)

$$\hat{H}_{\text{PS}} = \varphi\, \hat{a}_{\text{in}}^{\dagger} \hat{a}_{\text{in}}. \tag{9}$$

Then the evolution of an operator $\hat{o}$ can be obtained in the interaction picture by $e^{i\hat{H}_{\text{PS}}t}\, \hat{o}\, e^{-i\hat{H}_{\text{PS}}t}$ and if the time dependence is absorbed into φ the evolution of a given mode can eventually be written as

$$\hat{a}_{\text{out}}^{\dagger} = e^{i\varphi} \hat{a}_{\text{in}}^{\dagger}. \tag{10}$$

The Lossless Beam Splitter (BS) is converging input state $\hat{a}_{\text{in}}$ ($\hat{b}_{\text{in}}$) into a linear

superposition of output states

$$\begin{aligned}\hat{a}^\dagger_{\text{out}} &= \cos\theta\,\hat{a}^\dagger_{\text{in}} + ie^{-i\varphi}\sin\theta\,\hat{b}^\dagger_{\text{in}} \\ \hat{b}^\dagger_{\text{out}} &= ie^{i\varphi}\sin\theta\,\hat{a}^\dagger_{\text{in}} + \cos\theta\,\hat{b}^\dagger_{\text{in}}.\end{aligned} \tag{11}$$

And consequently the Hamiltonian can be written as

$$\hat{H}_{\text{BS}} = \theta\, e^{i\varphi}\hat{a}^\dagger_{\text{in}}\hat{b}_{\text{in}} + \theta\, e^{-i\varphi}\hat{a}_{\text{in}}\hat{b}^\dagger_{\text{in}}. \tag{12}$$

Physically speaking, when a photon arrives at the BS in mode a_{in}, it is transmitted with probability $|T|^2 = \cos^2\theta$ and reflected with probability $|R|^2 = 1-|T|^2 = \sin^2\theta$ and vice versa if it arrives in mode b_{in}. Additionally the two possible events can pick up the relative phase shift $ie^{\pm i\varphi}$.

Mathematically speaking, (11) is one possible way to parametrize the special unitary group $SU(2)$ and θ and φ are the angles of rotation on the Poincaré sphere. In that way the unitary constraints of (lossless) evolutions in quantum mechanics are satisfied. However, φ is just a phase in the input modes and the constraints are still satisfied when we choose it to be $\varphi = \frac{\pi}{2}$. Thus we end up with the very handy and intuitive representation of the BS

$$B = \begin{pmatrix} T & -R^* \\ R & T^* \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{13}$$

which is just an arbitrary two dimensional rotation.

The Polarization Rotation can be treated in the same manner, with the only difference that the modes being rotated are not spatial (a, b), but polarization

modes, i.e. some set of orthogonal coordinates x and y

$$\begin{aligned} \hat{a}^\dagger_{x'} &= \cos\theta\, \hat{a}^\dagger_x + ie^{-i\varphi} \sin\theta\, \hat{a}^\dagger_y \\ \hat{a}^\dagger_{y'} &= ie^{i\varphi} \sin\theta\, \hat{a}^\dagger_x + \cos\theta\, \hat{a}^\dagger_y. \end{aligned} \tag{14}$$

Where $\hat{a}^\dagger_\mu$ creates a photon with polarization μ, with $\mu \in \{x, y\}$.

The Polarizing Beam Splitter (PBS) is a polarization sensitive beam splitter that transmits a photon of one particular mode of polarization, for example horizontally polarized, while it reflects photons with orthogonal polarization, for example vertically polarized, independently of the spatial modes a, b the photons are in (Fig. 1b). This can be formalized in the following notation for the in- and output transformations

$$\begin{aligned} \hat{a}_{\text{in},H} &\to \hat{a}_{\text{out},H} \\ \hat{b}_{\text{in},H} &\to \hat{b}_{\text{out},H} \end{aligned} \quad \text{and} \quad \begin{aligned} \hat{a}_{\text{in},V} &\to \hat{b}_{\text{out},V} \\ \hat{b}_{\text{in},V} &\to \hat{a}_{\text{out},V}\,. \end{aligned} \tag{15}$$

B. Linear Optical Quantum Networks

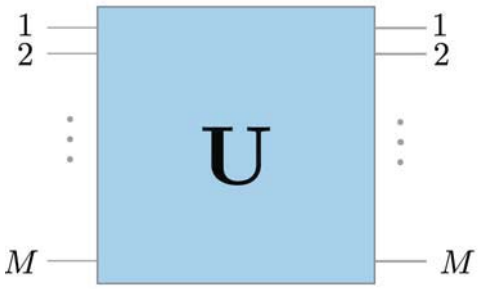


Figure 2: Black box like abstraction of a passive linear optical circuits with M in- and output modes. Any passive linear optical circuit can be seen as a unitary transformation of a input state into an output state.

With the elements introduced above one can set up any passive linear optical circuit (PLOC). Note that passive linear optical circuits are not known to be universal

for the bounded error quantum polynomial time complexity class (BQP) unless we allow *adaptive measurements* [9, 11]. However, as pointed out in the introduction, such passive circuits are of great interest for various quantum computational applications and in this work we will merely be concerned about these passive linear optical circuits.

Such a general passive circuit can be seen as a unitary transformation $\mathbf{U}$ of the set of incoming modes (Fig. 2).

Let the t^{th} input mode be described by $\hat{a}_t^\dagger$. Then we might define a circuit of M modes to apply the following transformation to a photon in mode t

$$\hat{a}_t^\dagger \rightarrow \sum_{k=1}^{M} u_{kt}\, \hat{a}_k^\dagger. \tag{16}$$

For an input of more than one photon we need to apply (16) to each of the creation operators $\hat{a}_t^\dagger$ appearing in the input.

Note that technically the two matrix representations

$$\mathbf{U} = \mathbf{D}(\boldsymbol{\mu})\, \mathbf{U}'\, \mathbf{D}(\boldsymbol{\nu}), \tag{17}$$

where $\mathbf{D}(\boldsymbol{\mu}) = \mathrm{diag}(e^{i\mu_1}, e^{i\mu_2}, \ldots, e^{i\mu_m})$, yield the same observable results and thus represent the same physical network. Or in other words: global in- and output phases can not be determined and are therefore regarded as trivial in quantum mechanics.

At this point it is suitable to insert an interlude and introduce what is commonly known as *BosonSampling* [11].

1. BosonSampling

The problem can be stated as follows:

Alice inputs a Fock state $|in\rangle$ into a network $\mathbf{U}$ which is drawn from the Haar

measure [11, 29]. Note that in the standard version Alice inputs the standard initial state $|1_n\rangle \coloneqq \hat{a}_1^\dagger \hat{a}_2^\dagger \cdots \hat{a}_n^\dagger |0\rangle$. $\mathbf{U}$ operates on $|in\rangle$, according to (16) and thus the different modes interfere and the probability amplitudes mix according to the unitary transformation that is imposed. Subsequently Bob takes a measurement on the output side of the network, i.e. Bob samples from the probability distribution over all possible outputs weighted by their corresponding coefficients.

As we shall see, the probability distribution is connected to the permanent of an $N \times N$ matrix $\mathbf{A}$

$$\mathrm{Per}(\mathbf{A}) = \sum_{\sigma \in S_N} \prod_{i=1}^{N} a_{i,\sigma(i)} \,, \tag{18}$$

where S_n is the set of all permutations of the symmetric group S_N.

A general Fock input state of N photons can be written as

$$|in\rangle = |T\rangle \equiv \sum_{T=(t_1,\ldots,t_M)} \alpha_T \left(\hat{a}_1^\dagger\right)^{t_1} \cdots \left(\hat{a}_M^\dagger\right)^{t_M} |0\rangle \,. \tag{19}$$

Here T goes over all possible tupels such that $t_1 + \cdots + t_M = N$, $|0\rangle$ is the vacuum state: $\langle 0| \hat{a}_t^\dagger = 0$, and α_T are the corresponding amplitudes occurring in the superposition. Similarly the output can be written as

$$|out\rangle = |S\rangle \equiv \sum_{S=(s_1,\ldots,s_M)} \beta_S \left(\hat{a}_1^\dagger\right)^{s_1} \cdots \left(\hat{a}_M^\dagger\right)^{s_M} |0\rangle \,. \tag{20}$$

The probability that Bob measures the N-photon output $|S\rangle$, given the N photon input $|T\rangle$ and $M \times M$ unitary transformation $\mathbf{U}$ is then given by $P_{|S\rangle,|T\rangle}$

$$P_{|S\rangle,|T\rangle} = \frac{|\mathrm{Per}(\mathbf{U}_{T,S})|^2}{s_1! \cdots s_m! t_1! \cdots t_m!}. \tag{21}$$

The $N \times N$ matrix $\mathbf{U}_{T,S}$ is defined by first taking t_i copies of the i^{th} column which leads to am matrix $\mathbf{U}_T$. Then take s_j copies of the j^{th} row of $\mathbf{U}_T$. The result is

$\mathbf{U}_{T,S}$.

Remark: Aaronson and Arkhipov [11] gave strong evidence that if there exists a polynomial-time classical algorithm for approximating (21) for a matrix of complex entries $\mathbf{U}_{T,S} \in \mathbb{C}^{N\times N}$ the polynomial hierarchy collapses. This is a striking result since successfully sampling from (21) for large instances of photons ($N \approx 30$) would give strong evidence against the extended Church-Turing Thesis that says that all computational problems that are efficiently solvable by a realistic physical device, are efficiently solvable by a probabilistic Turing machine [11, 12]. Let us now get back to the circuit itself.

2. N-Port Interferometer

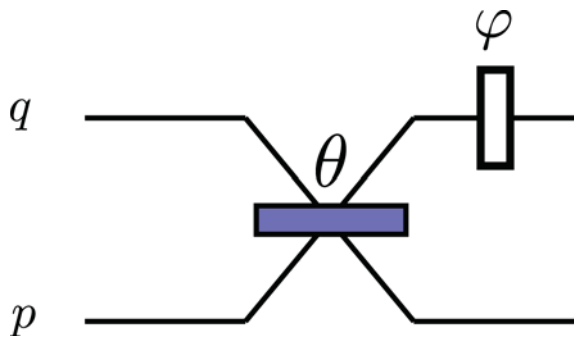


Figure 3: The beam splitter device with transmission amplitude $\cos\theta$ and phase φ in mode q as the main building block for any unitary transformation. It is essentially a beam splitter with an additional phase shift in one of the output modes, but can also be implemented in an analogous manner by a Mach-Zehnder interferometer.

As introduced above, a passive linear optical circuit with M modes is formally described by an $M \times M$ unitary transformation $\mathbf{U}(M)$. However, this is leaving open the question of how to implement such a network in a laboratory. Fortunately, Reck et al. [30] could show that it is possible to decompose any $M \times M$ unitary into a product of unitary beam splitter transformations $T_{p,q}$ acting on at most two

modes $p,\ q$ $(p > q)$. The maximum number of beam splitter devices needed is $\binom{M}{2} = \frac{M(M-1)}{2}$, i.e. only quadratic in M.

$$\mathbf{T}_{p,q} \coloneqq T_{p,q} \otimes \mathbf{I} = \begin{array}{c} \\ q \\ \\ p \\ \\ \end{array} \begin{array}{c} \begin{array}{cccccc} & & q & & p & \end{array} \\ \begin{pmatrix} \mathbf{I} & \vdots & & \vdots & & \\ \cdots & e^{i\varphi}\sin\theta & \cdots & e^{i\varphi}\cos\theta & \cdots & \\ & \vdots & \mathbf{I} & \vdots & & \\ \cdots & \cos\theta & \cdots & -\sin\theta & \cdots & \\ & \vdots & & \vdots & \mathbf{I} & \end{pmatrix} \end{array} \tag{22}$$

Where $\mathbf{I}$ is the identity matrix, θ parametrizes the beam splitting ratio and φ the phase in the output mode q (Fig.3). Here we have adapted the definition made by Reck et al. in the original paper which is slightly different, though equivalent, to (11). One only has to rename the output modes.

The algorithmic proof of the decomposition into such building blocks goes as follows. Given an arbitrary unitary $\mathbf{U}(M)$ we right multiply it with $T_{M,M-1}\cdots T_{M,1}$ to eliminate the last row (and column, by unitarity) except for the diagonal element. Subsequently we right multiply it by appropriate $T_{M-1,M-2}\cdots T_{M-1,1}$ to zero out all not diagonal elements of the $(M-1)^{th}$ row and column. This can be repeated until only diagonal elements are left, which can be turned into ones by right multiplying the appropriate diagonal matrix $\mathbf{D}$. Hence we have implemented the inverse operation $\mathbf{U}(M)^{-1} = \mathbf{U}(M)^{\dagger}$. In general the transformation just established is rotating one of the unit vectors $(0,\ldots,0,1,0,\ldots,0)$ into an arbitrary M-dimensional vector, parametrized by generalized spherical coordinates $\theta_1,\ldots\theta_{M-1}$ and the phases $\varphi_1,\ldots,\varphi_{M-1}$.

A generalized M-dimensional unitary transformation can be written as a stack of M such vectors. Therefore we can find a mapping for any unitary transformation

into physical devices building up an M-port interferometer parametrized by $\binom{M}{2}$ beam splitter ratios and $\binom{M}{2}$ phases. *Note* that this is not a unique mapping, i.e. a certain unitary can also have a physical realization differing from the one just introduced.

So far we have mainly focused on the network properties itself. As already introduced in the BosonSampling interlude, for physical applications these networks are operating on states. For that reason let us now draw our attention to the input to these networks (and the resulting output).

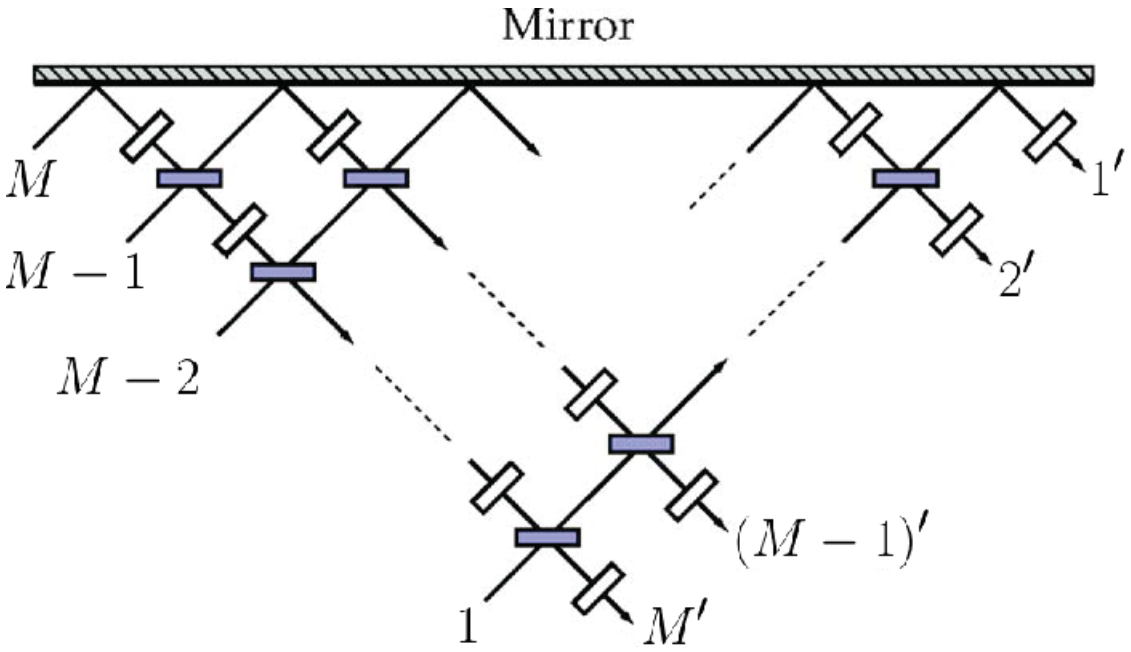


Figure 4: (4) Representation of a general M-port interferometer that is applying the unitary transformation $\mathbf{U}(M)$. Readapted from [21].

III. THE PARAMETRIC DOWN CONVERSION PHOTON SOURCE

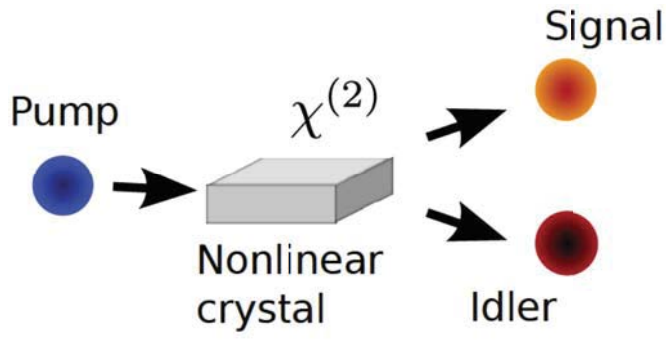


Figure 5: In a parametric down conversion photon source, photons are produced by pumping a nonlinear χ^2 crystal. Inside the crystal a pump photon can spontaneously decay into two photons: the so called "signal" and "idler" photon. Figure taken from [31].

Different from the ideal, frequency independent input states in the BosonSampling example, the physical reality of what can be achieved to date looks different. This section shall be devoted to a description of these more realistic conditions by introducing the predominant parametric down conversion source and its properties. We first investigate the properties of the parametric down conversion (PDC) photon source that is utilized to carry out many quantum optics experiments. The PDC process is a scattering process inside a χ^2 crystal - i.e. a *nonlinear* process - in which one photon of an incident electromagnetic field decays spontaneously (that's why it is often abbreviated as SPDC) into two photons obeying energy and momentum conservation (Fig. 5). It has been intensively studied ever since the first experiments by Burnham and Weinberg in 1970 [18]. Up to now it is the most utilized method to create single-photon Fock states in quantum optics laboratories.

The PDC interaction Hamiltonian of the pump field with the signal and idler fields

reads [32, 33]

$$\hat{H}(t) = \frac{\epsilon_0}{3} \sum_{ijk} \chi^{(2)}_{ijk} \int_{-\frac{L}{2}}^{\frac{L}{2}} dz\, \hat{E}_i^{(p)}(z,t) \hat{E}_j(z,t) \hat{E}_k(z,t). \tag{23}$$

For a crystal of length L (where the origin of the coordinate system has been chosen in the center of the crystal), pump field $\hat{E}^{(p)}$ and outgoing fields $\hat{E}_i$ and $\hat{E}_j$ (recall that $\hat{E}^\dagger = \hat{E}$), i, j, k are describing the different polarizations and $\chi^{(2)}_{ijk}$ is a tensor depending on the crystal properties. We will now restrict ourselves to uniaxial crystals and only consider the quadratic term. This so called type-II PDC can formally be put in the following interaction Hamiltonian for the three interacting modes

$$\hat{H}(t) = \frac{\epsilon_0}{3} \chi^{(2)} \int_{-\frac{L}{2}}^{\frac{L}{2}} dz\, \hat{E}_p^{(p)}(z,t) \hat{E}_s(z,t) \hat{E}_i(z,t). \tag{24}$$

With (2) we may write

$$\begin{aligned} \hat{H}(t) \;=\; & \frac{\epsilon_0}{3} \chi^{(2)} \int_{-\frac{L}{2}}^{\frac{L}{2}} dz\, \left(\hat{E}_p^{(+)}(z,t) + \hat{E}_p^{(-)}(z,t) \right) \\ & \left(\hat{E}_s^{(-)}(z,t) + \hat{E}_s^{(+)}(z,t) \right) \left(\hat{E}_i^{(-)}(z,t) + \hat{E}_i^{(+)}(z,t) \right). \end{aligned} \tag{25}$$

In the last step we apply the rotating wave approximation [34] in which we can neglect all processes different from the pump photon decay into one signal and one idler photon and the reverse process were the signal and idler photon decay into a pump photon (the Hermitian conjugate of the previous one that is guaranteeing

the Hermitian property of the Hamiltonian)

$$\hat{H}(t) = \frac{\epsilon_0}{3}\chi^{(2)} \int_{-\frac{L}{2}}^{\frac{L}{2}} dz\, \hat{E}_p^{(+)}(z,t)\hat{E}_s^{(-)}(z,t)\hat{E}_i^{(-)}(z,t) + \text{H.c.} \tag{26}$$

The positive frequency parts for $\hat{E}_\mu^{(+)}(z,t)$ are given by (5). Let us introduce two further approximations. First, $A(\omega)$ is a slowly varying function of frequency [23] and can thus be taken outside the integral in (26) . Second, PDC is a very inefficient process and to obtain suitable photon numbers the pump field must be large. Hence the pump field $\hat{E}_p^{(+)}(z,t)$ might be treated as classical field. In the Fourier representation we have

$$E_p^{(+)}(z,t) = A_P \int d\omega_p\, \alpha(\omega_p)\, e^{i(k_p(\omega_p)z - \omega_p t)}. \tag{27}$$

Now, if we plug (27) and (5) into (26) and make use of the $A(\omega) \approx$ const. approximation, we may write

$$\hat{H}_{\text{PDC}}(t) \;=\; A \int_{-\frac{L}{2}}^{\frac{L}{2}} dz \int_{-\infty}^{\infty} d\omega_p \int_{-\infty}^{\infty} d\omega_s \int_{-\infty}^{\infty} d\omega_i \alpha(\omega_p)\, e^{i[(k_p(\omega_p) - k_i(\omega_i) - k_s(\omega_s))\cdot z - (\omega_p - \omega_i - \omega_s)\cdot t]} \hat{a}_s^\dagger(\omega_s)\hat{a}_i^\dagger(\omega_i) + \text{H.c.} \tag{28}$$

Where we have combined all constants in A.
In the interaction picture the state evolution is given by

$$|\psi\rangle_{\text{PDC}} = \mathcal{T}\, \exp\left[-\frac{i}{\hbar}\int_{t_0}^{t} dt'\, \hat{H}_{PDC}(t')\right], \tag{29}$$

where $\mathcal{T}$ is the time ordering operator that is taking account for the fact that the PDC interaction Hamiltonian does not commute in time [35]. In the extend of this

thesis we shall only work with the *first order perturbation* of (29). That is

$$|\psi\rangle_{\mathrm{PDC}} \approx |0\rangle + \frac{1}{i\hbar}\int_{t_0}^{t} dt' \hat{H}_{PDC}(t')\,|0\rangle \tag{30}$$

Furthermore we extend the range of integration to minus- and plus infinity. This assumes that before t_0 the pump field is zero and that we consider times long after the interaction, i.e. the interaction Hamiltonian vanishes before t_0 and after t. This enables us to carry out the time integral in (30) yielding $2\pi\delta(\omega_p - \omega_s - \omega_i)$ which makes it trivial to carry out the ω_p integral as well. Introducing

$$\Delta k(\omega_s, \omega_i) = k_p(\omega_s + \omega_i) - k_s(\omega_s) - k_i(\omega_i) \tag{31}$$

we obtain

$$|\psi\rangle_{\mathrm{PDC}} = \frac{2\pi A}{i\hbar}\int_{-\frac{L}{2}}^{\frac{L}{2}} dz \int d\omega_s \int d\omega_i\, \alpha(\omega_s+\omega_i)\, e^{i\Delta k(\omega_s,\omega_i)\cdot z}\, \hat{a}_s^\dagger(\omega_s)\hat{a}_i^\dagger(\omega_i)\,|0\rangle\,. \tag{32}$$

The part corresponding to the Hermitian conjugate vanishes due to (6) ($\hat{a}(\omega)\,|0\rangle = 0$). The remaining z integration gives

$$|\psi\rangle_{\mathrm{PDC}} = \underbrace{\frac{2\pi A}{i\hbar}\cdot L}_{=B}\int d\omega_s \int d\omega_i \alpha(\omega_s+\omega_i)\underbrace{\frac{\sin\left(\Delta k(\omega_s,\omega_i)\cdot\frac{L}{2}\right)}{\Delta k(\omega_s,\omega_i)\cdot\frac{L}{2}}}_{=:\phi(\omega_s,\omega_i)}\cdot \tag{33}$$
$$\hat{a}_s^\dagger(\omega_s)\hat{a}_i^\dagger(\omega_i)\,|0\rangle\,.$$

Which can be written more compact by introducing the *joint spectral amplitude* (JSA) $f(\omega_s,\omega_i) = \alpha(\omega_s+\omega_i)\cdot\phi(\omega_s,\omega_i)$.

$$|\psi\rangle_{\mathrm{PDC}} = B\int d\omega_s \int d\omega_i\, f(\omega_s,\omega_i)\,\hat{a}_s^\dagger(\omega_s)\hat{a}_i^\dagger(\omega_i)\,|0\rangle \tag{34}$$

The joint spectral amplitude, also known as *spectral correlation function* and it is the product of the *pump envelope* $\alpha(\omega_s + \omega_i)$ with the *phase matching function* $\phi(\omega_s, \omega_i)$. The joint spectral amplitude $f(\omega_s, \omega_i)$ contains all the properties of the PDC source and shall be subject to further investigation in the following.

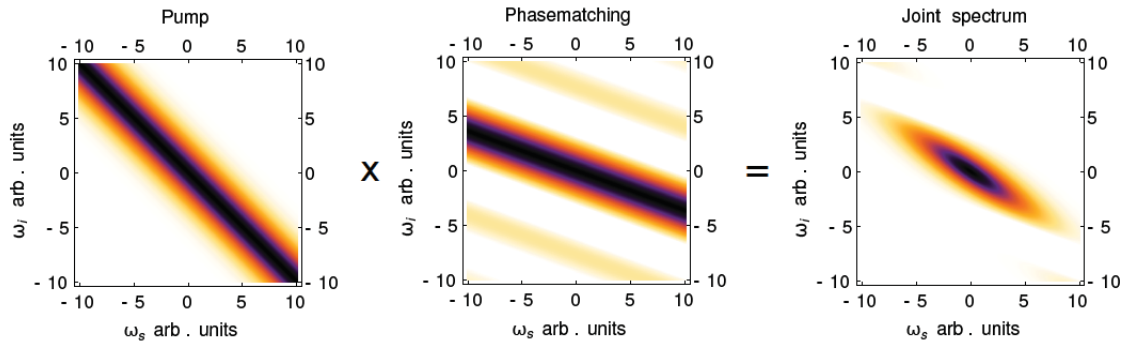


Figure 6: Exemplary spectral correlation function, composed of the product of the pump envelope $\alpha(\omega_s + \omega_i)$ and the phase matching function $\phi(\omega_s, \omega_i)$. Here without the Gaussian approximation for the phase matching function. Taken from [31].

A. The Spectral Correlation Function

To study the spectral correlation function analytically, some further approximations and assumptions have to be made.

The first assumption is a Gaussian pump envelope:

$$\alpha(\omega_s + \omega_i) \propto e^{-\frac{(\omega p - \omega_p^0)^2}{2\sigma_p^2}} . \tag{35}$$

Other approximations made are the Gaussian approximation for $\phi(\omega_s, \omega_i)$ and the linear expansion of $\Delta k(\omega_s, \omega_i)$. The Gaussian approximation reads

$$\mathrm{sinc}(x) \approx e^{-\gamma x^2}, \quad \gamma = 0.193 . \tag{36}$$

And for the expansion of (31) around the central frequencies ω_μ^0 we find

$$\Delta k(\omega_s, \omega_i) \approx \tag{37}$$
$$\Delta k(\omega_s^0, \omega_i^0) + \left.\frac{\partial \Delta k(\omega_s, \omega_i)}{\partial \omega_s}\right|^{\omega_s^0, \omega_i^0} (\omega_s - \omega_s^0) + \left.\frac{\partial \Delta k(\omega_s, \omega_i)}{\partial \omega_i}\right|^{\omega_s^0, \omega_i^0} (\omega_i - \omega_i^0)\,.$$

Introducing the group velocity mismatch $\kappa_\mu = \frac{dk_\mu(\omega_\mu^0)}{d\omega_\mu} - \frac{dk_p(\omega_p^0)}{d\omega_p}$ and the mismatch – as κ_μ is always expressed with respect to the group velocity of the pump – from the central frequencies $\nu_\mu = \omega_\mu - \omega_\mu^0$ and considering momentum conservation $\Delta k(\omega_s^0, \omega_i^0) = 0$, the phase matching function (omitting the constant) then becomes

$$\phi(\nu_s, \nu_i) \approx e^{-\gamma \frac{L^2}{4}(\kappa_s \nu_s + \kappa_i \nu_i)^2} \tag{38}$$

Thus, the spectral correlation function takes the form

$$f(\nu_s, \nu_i) \;=\; \alpha(\nu_s, \nu_i)\phi(\nu_s, \nu_i) \approx \frac{1}{\sqrt{N}}\, e^{-\gamma \frac{L^2}{4}(\kappa_s \nu_s + \kappa_i \nu_i)^2} \cdot e^{-\frac{(\nu_s + \nu_i)^2}{2\sigma_p^2}} \tag{39}$$

Bare in mind that we have introduced ν_μ as new variables, i.e. we have changed the reference frame, which will be important whenever one wants to calculate (34) explicitly. Additional we made a slight change in notation and introduced the $1/\sqrt{N}$ instead of B that is taking care of normalization.

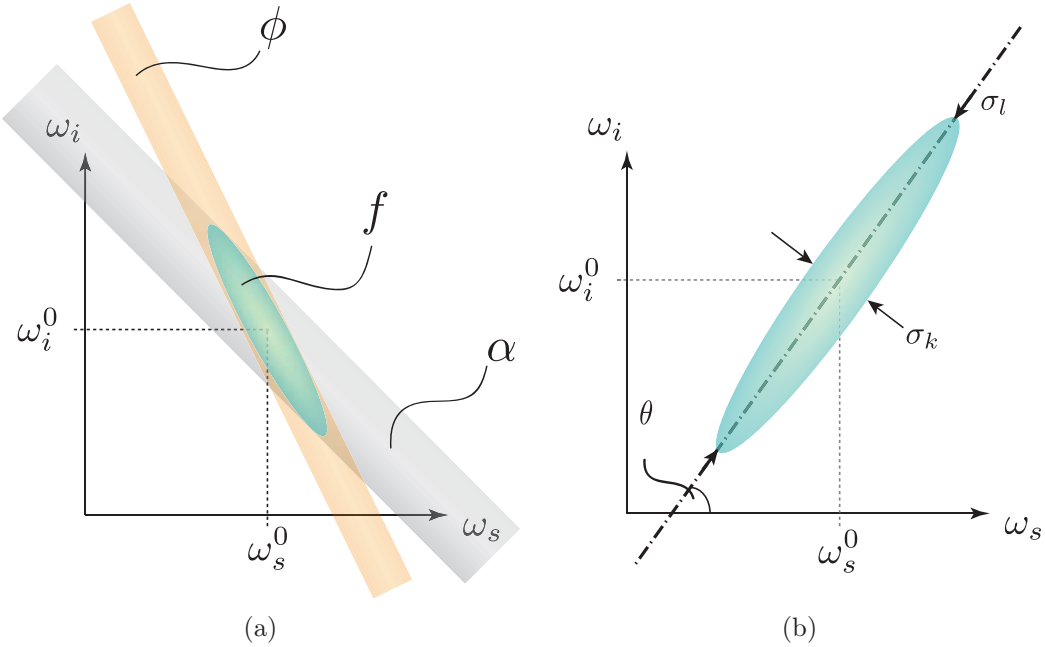


Figure 7: Composition and parameterization of the joint spectral amplitude (JSA). (a) Sketch of an examplary spectral correlation function f, formed by the phase matching function ϕ and the pump envelope α. (b) Parametrization of the ellipse by its major and minor axis σ_l and σ_k, and the angle of the major axis θ to the ω_s axis.

To enable a more intuitive view on the spectral correlation function and since it will be of great use in general, it might be useful to rewrite (39) in the following manner

$$f(\nu_s, \nu_i) = B \exp\left[-\frac{1}{2}\boldsymbol{\nu}^T \mathbf{M} \boldsymbol{\nu}\right]. \tag{40}$$

Which is just the matrix representation of the argument of the exponential function with $\boldsymbol{\nu}^T = (\nu_s, \nu_i) = (\omega_s - \omega_s^0, \omega_i - \omega_i^0)$ and

$$\mathbf{M} = \begin{pmatrix} \frac{1}{\sigma_p^2} + \gamma\frac{L^2}{2}\kappa_s^2 & \frac{1}{\sigma_p^2} + \gamma\frac{L^2}{2}\kappa_s\kappa_i \\ \frac{1}{\sigma_p^2} + \gamma\frac{L^2}{2}\kappa_s\kappa_i & \frac{1}{\sigma_p^2} + \gamma\frac{L^2}{2}\kappa_i^2 \end{pmatrix}. \tag{41}$$

Here (40) is the bivariate normal distribution, which is well known and has been subject to extensive studies in mathematics as well as physics.

The matrix $\mathbf{M}$ depends on the crystal and pump properties. Varying these parameters allows us to manufacture different JSAs.

Note that $\boldsymbol{\nu}^T\mathbf{M}\boldsymbol{\nu} = \text{const.}$ represents an *ellipse* in the $\omega_s - \omega_i$-plane centered at (ω_s^0, ω_i^0). The direction of the two axis of the ellipse are given by the eigenvectors of $\mathbf{M}^{-1}$ and the extend of the great and small radii is given by $r\sqrt{\lambda_{1/2}}$, where λ_μ is the respective eigenvalue [36]. Hence restricting oneself to the eigenvectors and eigenvalues of $\mathbf{M}$ already provides full information about shape of the JSA. The eigenvalues are reading

$$\lambda_{1,2} = \frac{1}{\sigma_p^2} + \gamma(\kappa_s^2 + \kappa_i^2)\frac{L^2}{4} \pm \sqrt{\frac{1}{\sigma_p^4} + \gamma^2(\kappa_s^2 + \kappa_i^2)^2\frac{L^4}{16} + \gamma\frac{\kappa_s\kappa_i}{\sigma_p^2}L}\,. \tag{42}$$

But since $\mathbf{M}$ and therefore $f(\nu_s, \nu_i)$ can be parametrized as an ellipse, we might as well choose the most convenient way to do so. Instead of restricting ourselves to the parametrization in (41) we now introduce a much handier one in which the ellipse $\mathbf{M}$ is parametrized by a long and a short axis, σ_l and σ_k respectively, and an additional angle θ that defines the angle between the long axis of the ellipse and the ω_s-axis

$$\mathbf{M}(\sigma_k, \sigma_k, \theta) = \begin{pmatrix} \frac{\cos^2\theta}{\sigma_l^2} + \frac{\sin^2\theta}{\sigma_k^2} & \sin\theta\cos\theta\left(\frac{1}{\sigma_l^2} - \frac{1}{\sigma_s^2}\right) \\ \sin\theta\cos\theta\left(\frac{1}{\sigma_l^2} - \frac{1}{\sigma_s^2}\right) & \frac{\cos^2\theta}{\sigma_k^2} + \frac{\sin^2\theta}{\sigma_l^2} \end{pmatrix}. \tag{43}$$

At this point a very intuitive access to frequency correlations can be given [37–39]. A state is correlated in two variables, whenever it cannot be written as a product state of two functions that depend on just one of the variables respectively. Now (41) can be factorized into its two frequency parts whenever either the axis of the ellipse are equal or the angle θ is an integer multiple of $\frac{\pi}{2}$. Therefore (43) contains all necessary information about the frequency correlations and (40) is factorizable in the $\omega_s\omega_i$ system whenever the off-diagonal elements of (43) vanish. As we shall see in the next section it is possible to define a quantitative measure by use of

the *Schmidt modes* of the spectral correlation function. Let us also introduce the following notation: The range $0 < \theta < \frac{\pi}{2}$ shall be called the correlated case, whereas $-\frac{\pi}{2} < \theta < 0$ shall be called the anti-correlated case. Since $\mathbf{M}$ is π-periodic in θ it is enough to restrict θ to $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$ (or likewise $0 \leq \theta \leq \pi$).
Prior to this, let us introduce some modifications of the spectral correlation function that will be necessary to describe general quantum interference events. We have derived (40) as the spectral weights in the PDC state in (34). However, we have implicitly ignored the time delay between the signal and idler mode as well as filtering processes that might be – and in most quantum optics applications will be – performed on the PDC state.

1. The Temporal Delay

A temporal delay in mode μ introduces a phase $e^{i\omega_\mu \tau_\mu}$ to the JSA $f(\nu_s, \nu_i)$ and therefore if we introduce temporal delays to signal and idler modes

$$f(\nu_s, \nu_i) \to \beta(\nu_s,\nu_i; \tau_s, \tau_i) = f(\nu_s, \nu_i) \cdot e^{i(\nu_s+\omega_s^0)\tau_s} \cdot e^{i(\nu_i+\omega_i^0)\tau_i}. \tag{44}$$

Introducing $\boldsymbol{\tau}^T = (\tau_s, \tau_i)$ and $\boldsymbol{\omega}_0^T = (\omega_s^0, \omega_i^0)$, equation (44) can also be written as

$$\beta(\omega_s,\omega_i; \tau_s, \tau_i) = \frac{1}{\sqrt{N}} \exp\left[-\frac{1}{2}(\boldsymbol{\omega} - \boldsymbol{\omega_0})^T \mathbf{M} (\boldsymbol{\omega} - \boldsymbol{\omega_0}) + i\boldsymbol{\omega}^T \boldsymbol{\tau}\right]. \tag{45}$$

For the explicit form of $\frac{1}{\sqrt{N}}$ see XII.

2. Introducing Filtering

The process of filtering in the frequency domain can be modeled by a frequency dependent beam splitter [37–39]. Similar to (11) we obtain

$$\hat{a}_\mu^\dagger(\omega) \to T_\mu(\omega)\hat{a}_\mu^\dagger(\omega) + R(\omega)\hat{v}^\dagger(\omega). \tag{46}$$

Where $\hat{v}^\dagger(\omega)$ is introducing the vacuum state to the system and $T_\mu(\omega)$ is the filter function that describes the shape of the filter in mode μ and $|T(\omega)|^2 + |R(\omega)|^2 = 1$. Inserting (46) and (45) into (34) it is easy to find that the PDC state eventually takes the form

$$\begin{aligned} |\psi\rangle_{\mathrm{PDC}} &= \int\int d\omega_s d\omega_i \underbrace{\beta(\omega_s,\omega_i;\tau_s,\tau_i)T_s(\omega_s)T_i(\omega_i)}_{:=\beta^{(\mathrm{F})}(\omega_s,\omega_i)} \hat{a}_s^\dagger(\omega_s)\hat{a}_i^\dagger(\omega_i)\,|0\rangle \\ &= \int\int d\omega_s d\omega_i\, \beta^{(\mathrm{F})}(\omega_s,\omega_i)\,\hat{a}_s^\dagger(\omega_s)\hat{a}_i^\dagger(\omega_i)\,|0\rangle\,. \end{aligned} \tag{47}$$

The explicit form of $\beta_{\mathrm{F}}(\omega_s,\omega_i)$ depends on the applied filters and in the following it shall always take the form of a *Gaussian filter* that explicitly reads

$$T_\mu(\omega_\mu) = \sqrt{N_\mu^{(\mathrm{G})}}\exp\left[-\frac{(\omega_\mu-\omega_\mu^F)^2}{2\sigma_\mu^2}\right]. \tag{48}$$

Where ω_μ^{F} is the centroid frequency and $\sqrt{N_\mu^{(\mathrm{G})}}$ a normalizing factor (derived in XII) of the respective filter and σ_μ is related to the full width half maximum (FWHM). Introducing $\mathbf{M}' = \mathrm{diag}(\frac{1}{\sigma_s^2},\frac{1}{\sigma_i^2})$ and writing the centroid frequencies of the filters as $\boldsymbol{\omega_F}^T = (\omega_s^{\mathrm{F}},\omega_i^{\mathrm{F}})$ allows to write $\beta^{(\mathrm{F})}(\omega_s,\omega_i)$ as

$$\beta^{(F)}(\omega_s,\omega_i) \;=\; \frac{\sqrt{N_s^{(G)}N_i^{(G)}}}{\sqrt{N}}\cdot \tag{49}$$
$$\exp\left[-\frac{1}{2}(\boldsymbol{\omega}-\boldsymbol{\omega_0})^T\mathbf{M}(\boldsymbol{\omega}-\boldsymbol{\omega_0})-\frac{1}{2}(\boldsymbol{\omega}-\boldsymbol{\omega_F})^T\mathbf{M}'(\boldsymbol{\omega}-\boldsymbol{\omega_F})+i\boldsymbol{\omega}^T\boldsymbol{\tau}\right].$$

And reordering yields

$$\begin{aligned}\beta^{(\mathrm{F})}(\omega_s,\omega_i) \;=\; & \frac{\sqrt{N_s^{(\mathrm{G})}N_i^{(\mathrm{G})}}}{\sqrt{N}} \\ & \cdot\exp\left[-\frac{1}{2}\boldsymbol{\omega}^T(\mathbf{M}+\mathbf{M}')\boldsymbol{\omega}+\boldsymbol{\omega}^T\left(\mathbf{M}\boldsymbol{\omega_0}+\mathbf{M}'\boldsymbol{\omega_F}\right)\right] \\ & \cdot\exp\left[-\frac{1}{2}\boldsymbol{\omega_0}^T\mathbf{M}\boldsymbol{\omega_0}-\frac{1}{2}\boldsymbol{\omega_F}^T\mathbf{M}'\boldsymbol{\omega_F}\right] \\ & \cdot\exp\left[i\boldsymbol{\omega}^T\boldsymbol{\tau}\right].\end{aligned} \tag{50}$$

Equation (50) has still a Gaussian form and contains many information on the effect of filtering that we will briefly discuss in the following.

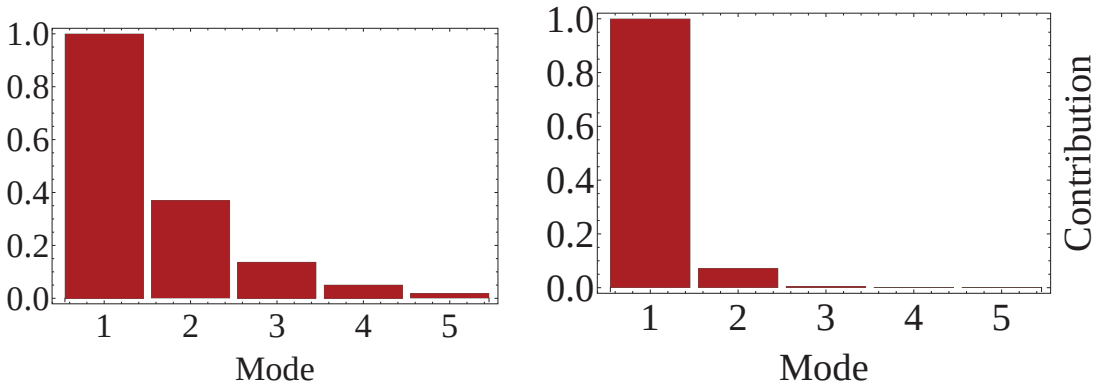
First, we note that considering the signal and idler modes only, they will in general not be normalized to one after filtering.

Second, the effective centroid frequencies of $\beta^{(\mathrm{F})}$ do in general differ from that of β, unless $\boldsymbol{\omega_F}=\boldsymbol{\omega_0}$.

Last, taking a close look at $\mathbf{M}+\mathbf{M}'$ reveals that narrow band filtering, i.e. $\frac{1}{\sigma_\mu^2}\gg\frac{1}{\sigma_k^2}$, effectively suppresses the off-diagonal elements of $\mathbf{M}$ in the sense that they become very small w.r.t. the new diagonal elements. But recall the result of III A, claiming that these off-diagonal elements are carrying the information about the spectral correlations. Therefore it is now required to introduce a new quantitative measure for the spectral correlations between signal and idler mode. Such a measure can be introduced by means of the so called Schmidt modes. These are obtained by performing a Schmidt decomposition on (50) that shall be briefly introduced in

the following.

3. *Schmidt or Singular Value Decomposition*



(a) (b)

Figure 8: The Schmidt Modes of an exemplary spectral correlation function. Here with $\boldsymbol{\omega_F} = \boldsymbol{\omega_0}$ and a FWHM for the filter that is half of the ellipsis minor axis. (a) depicts the contribution of the first ten Schmidt modes before filtering and (b) after filtering. Clearly, the filtering suppresses higher Schmidt modes and thus frequency correlations. Nevertheless, even after filtering the remaining contribution of the second mode is about 10%. Note that we have carried out a renormalization with respect to the first mode in order to obtain a better comparison for the contribution of the modes. Without this renormalization, an overall decrease in contribution due to filtering would be observed.

To reveal the quantum correlations of the spectral correlation function $\beta^{(\mathrm{F})}(\omega_s, \omega_i)$ a Schmidt or singular value decomposition (SVD) can be performed [7, 40, 41]. That is accomplished by finding the eigenbasis of $\beta^{(\mathrm{F})}(\omega_s, \omega_i)$ such that it can be written as

$$\beta^{(\mathrm{F})}(\omega_s, \omega_i) = \sum_k \sqrt{\lambda_k}\psi_n(\omega_s)\phi_k(\omega_i). \tag{51}$$

Where λ_k, ψ_k and ϕ_k are solutions of the integral eigenvalue equations

$$\int d\omega' K_1(\omega,\omega')\psi_k(\omega') \;=\; \lambda_k\psi_k(\omega), \tag{52}$$

$$\int d\omega' K_2(\omega,\omega')\phi_k(\omega') \;=\; \lambda_k\phi_k(\omega), \tag{53}$$

with the kernels (one photon spectral correlations)

$$\begin{aligned} K_1(\omega,\omega') \;&=\; \int d\omega_i \beta^{(\mathrm{F})}(\omega,\omega_i)\beta^{(\mathrm{F})*}(\omega',\omega_i) \\ K_2(\omega,\omega') \;&=\; \int d\omega_i \beta^{(\mathrm{F})}(\omega_s,\omega)\beta^{(\mathrm{F})*}(\omega_s,\omega)\,, \end{aligned}$$

here ψ_k and ϕ_k are the corresponding eigenfunctions. For (49) an analytic expression of (51) can be given by Hermite polynomials [22].

Carrying out the SVD numerically is of great interest, especially for more general filter functions, and shall be introduced briefly.

The discretization of $\beta^{(\mathrm{F})}(\omega,\omega')$ may be written as $\beta^{(\mathrm{F})}_{\omega,\omega'}$ and yields the following matrix decomposition:

$$\beta^{(\mathrm{F})}_{\omega,\omega'} = \sum_k U_{\omega.k} D_{k,k} V_{k,\omega'}. \tag{54}$$

Where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices and $\mathbf{D}$ is a diagonal matrix. And for a step size of $\Delta\omega$ we have the following relations

$$\begin{aligned} U_{\omega,k}/\Delta\omega \;&\leftrightarrow\; \phi_k(\omega), \\ V_{k,\omega'}/\Delta\omega \;&\leftrightarrow\; \psi_k(\omega'), \\ D_{k,k}\Delta\omega \;&\leftrightarrow\; \sqrt{\lambda_k}. \end{aligned} \tag{55}$$

Fig. 8 depicts an exemplary Schmidt decomposition of an unfiltered and filtered state. There exists many different measures to extract the amount of correlations from the Schmidt modes [41, 42], but for now it shall be enough to bare in mind

that we can find a quantitative measure harnessing Schmidt modes.

IV. QUANTUM INTERFERENCE IN LINEAR OPTICAL NETWORKS

Combining sections II and III provides everything required for a detailed description of a general quantum interference scheme in linear optical networks. The next step is to apply the previously derived PDC state to a linear optical quantum experiment, like the BosonSampling from section II B 1.
Main interest of this thesis is the tomography of the underlying linear optical network which is the basic building block for many optical quantum experiments. As we will see in the next section, two-photon interference is sufficient for that purpose. Therefore we start by introducing the concept of two-photon correlation measurements. To the best of our knowledge the first experiments addressing this topic were carried out Hanbury Brown and Twist in 1956 [43] and Hong, Ou and Mandel in 1987 [44]. A first complete formal analysis for general n^{th}-order correlation functions of a quantum optical field was given 1963 by Glauber [27].

A. Two-Photon Correlation Measurements

Two-photon interference is characterized by the second-order correlation function $G^{(2)}$ that shall now be introduced. Following [27] a $G^{(n)}$-measurement of the electromagnetic field expresses the correlations of the field at $2n$ different points of space time and values of these functions are measured by n-fold delayed coincidence detection of photons. Therefore a $G^{(2)}$-measurement is a fourth-order interference effect between two photons. Measuring a photon is associated with the annihilation operator (3). Now, a transition of the field from the initial state $|i\rangle$ to the final state $|f\rangle$ in which one photon had been absorbed at (z,t) in space time

is associated with the matrix element

$$\langle f|\, E^{(+)}(z,t)\,|i\rangle\,. \tag{56}$$

And correspondingly for the absorption of two photons at (z,t) and (z',t') we have the matrix element

$$\langle f|\, E^{(+)}(z',t')E^{(+)}(z,t)\,|i\rangle\,. \tag{57}$$

For a frequency independent detector the probability of detection is given by the sum (integral) over the modulo square of all these frequency states

$$\begin{aligned}
\sum_f \left|\langle f|\, E^{(+)}(z',t')E^{(+)}(z,t)\,|i\rangle\right|^2 \;&=\\
=\sum_f \langle i|\, E^{(-)}(z,t)E^{(-)}(z',t')\,|f\rangle\,\langle f|\, E^{(+)}(z',t')E^{(+)}(z,t)\,|i\rangle \;&=\\
= \langle i|\, E^{(-)}(z,t)E^{(-)}(z',t')E^{(+)}(z',t')E^{(+)}(z,t)\,|i\rangle\,.&
\end{aligned} \tag{58}$$

This result is independent of whether the $|f\rangle$ is discrete or continuous. In the second case we need to take the integral over all final states though $\sum \to \int$.

1. *An Example: The Hong-Ou-Mandel-Dip (HOM-Dip)*

Let us give an instance for a two-photon correlation measurement by assuming a beam splitter of transmissivity $|T|^2$ and reflectivity $|R|^2$. In the case of a perfectly monochrome input into the two beam splitter modes, i.e. $|in\rangle = |1,1\rangle$ we are back at the BosonSampling II B 1 and with (16), (13) we find that our initial state undergoes the following transition

$$|1,1\rangle \to (T^2 - R^2)\,|1,1\rangle - TR^*\,|2,0\rangle + T^*R\,|0,2\rangle\,. \tag{59}$$

Note that this is consistent with the probabilities given by (21). Therefore, in case of a 50 : 50 beam splitter the transformation yields $\frac{1}{\sqrt{2}}(|0,2\rangle - |2,0\rangle)$. However, in practice the photons are never monochrome as we have discussed extensively in III.

Assuming an input of (34) with temporal delay (44) the probability to detecting a coincidence of detecting a photon in mode 1 and 2 after the beam splitter $|out\rangle = \hat{a}_1^\dagger(\omega)\hat{a}_2^\dagger(\omega')\,|0\rangle$ is given by (50). Again, with (13) this results in

$$P_{1,2} = \int d\omega \int d\omega' \, \left| |T|^2 f(\omega,\omega') - |R|^2 f(\omega',\omega) \right|^2 . \tag{60}$$

We have omitted the explicit calculations to arrive at (60) and we have not carried out the integrals. We have done so to prevent redundancy with the rest of this section in which these calculations will be given explicitly for the more general case of an arbitrary unitary. Though, it might be a nice instance for preview to see how finding the $G^{(2)}$ function works and why the shape of the JSA matters for that purpose. Therefore, let us now proceed with the more general case.

B. Two-Photon Correlation Measurement for a Passive Linear Optical Circuit

In this section a derivation of a general two-photon interference will be given. The input shall be of the form as (47) with a JSA containing correlation and filtering effects as in (49) and is feed into a general $M \times M$-unitary $\mathbf{U}$. The interest lays in measuring a two-fold coincidence of two output modes, i.e. the corresponding $G^{(2)}$-function.

Feeding one photon in mode t_1 and one in t_2 yields an input of the form

$$|in\rangle = |\psi\rangle_{\mathrm{PDC}} = \int\int d\omega_s d\omega_i \, \beta^{(\mathrm{F})}(\omega_s,\omega_i)\, \hat{a}_{t_1}^\dagger(\omega_s)\hat{a}_{t_2}^\dagger(\omega_i)\,|0\rangle\,, \tag{61}$$

with

$$\beta^{(\mathrm{F})}(\omega_s,\omega_i) \;=\; \frac{\sqrt{N_s^{(\mathrm{G})}N_i^{(\mathrm{G})}}}{\sqrt{N}} \qquad (62)$$
$$\exp\left[-\frac{1}{2}(\boldsymbol{\omega}-\boldsymbol{\omega_0})^T\mathbf{M}(\boldsymbol{\omega}-\boldsymbol{\omega_0})-\frac{1}{2}(\boldsymbol{\omega}-\boldsymbol{\omega_F})^T\mathbf{M}'(\boldsymbol{\omega}-\boldsymbol{\omega_F})+i\boldsymbol{\omega}^T\boldsymbol{\tau}\right].$$

In the sense of (16) the linear optical circuit is imposing the following transition to the input $|in\rangle$

$$\hat{\mathbf{U}}\,|in\rangle \;=\; \int\int d\omega_s d\omega_i\,\beta^{(\mathrm{F})}(\omega_s,\omega_i) \qquad (63)$$
$$\left[u_{1,t_1}\hat{a}_1^\dagger(\omega_s)+\cdots+u_{M,t_1}\hat{a}_M^\dagger(\omega_s)\right]\left[u_{1,t_2}\hat{a}_1^\dagger(\omega_i)+\cdots+u_{M,t_2}\hat{a}_M^\dagger(\omega_i)\right]|0\rangle\,.$$

And - as in the example of the HOM-Dip - the output shall be a coincidence in two modes, say s_1 and s_2

$$|out\rangle=\hat{a}_{s_1}^\dagger(\omega)\hat{a}_{s_2}^\dagger(\omega')\,|0\rangle\,. \qquad (64)$$

And thus the matrix element are

$$\langle out|\,\hat{\mathbf{U}}\,|in\rangle \;=\; \int\int d\omega_s d\omega_i\,\beta^{(\mathrm{F})}(\omega_s,\omega_i) \qquad (65)$$
$$\langle 0|\,\hat{a}_{s_2}(\omega')\hat{a}_{s_1}(\omega)\left[u_{1,t_1}\hat{a}_1^\dagger(\omega_s)+\cdots+u_{M,t_1}\hat{a}_M^\dagger(\omega_s)\right]$$
$$\left[u_{1,t_2}\hat{a}_1^\dagger(\omega_i)+\cdots+u_{M,t_2}\hat{a}_M^\dagger(\omega_i)\right]|0\rangle\,.$$

To carry out (65) we need to recall the commutator relations (7) and extend them for the case of additional discrete spatial modes, say i and j

$$[\hat{a}_i(\omega),\hat{a}_j^\dagger(\omega')]=\delta(\omega-\omega')\delta_{i,j}. \qquad (66)$$

We can exploit (66) to rewrite the four-operator terms $\hat{a}_i(\omega_i)\hat{a}_j(\omega_j)\hat{a}_k^\dagger(\omega_k)\hat{a}_l^\dagger(\omega_l)$

appearing in (65). This is done by reordering the appearing two-operator terms $\hat{a}_j(\omega_j)\hat{a}_k^\dagger(\omega_k) = \delta(\omega_j - \omega_k)\delta_{j,k} + \hat{a}_k^\dagger(\omega_k)\hat{a}_j(\omega_j)$ successively and using $\hat{a}_i(\omega_i)\left|0\right\rangle = 0$. This leaves us with

$$\hat{a}_i(\omega_i)\hat{a}_j(\omega_j)\hat{a}_k^\dagger(\omega_k)\hat{a}_l^\dagger(\omega_l) = \delta_{i,j}\delta_{j,k}\delta(\omega_j - \omega_k)\delta(\omega_i - \omega_l) + \delta_{i,k}\delta_{j,l}\delta(\omega_i - \omega_k)\delta(\omega_j - \omega_l). \tag{67}$$

Using this result, (65) becomes

$$\left\langle out\right|\hat{\mathbf{U}}\left|in\right\rangle = \left[\underbrace{u_{s_2,t_2}u_{s_1,t_1}}_{:=b_e}\beta^{(\mathrm{F})}(\omega,\omega') + \underbrace{u_{s_2,t_1}u_{s_1,t_2}}_{:=b_o}\beta^{(\mathrm{F})}(\omega',\omega)\right]. \tag{68}$$

The resulting $G^{(2)}$-function (58) reads

$$\begin{aligned} G^{(2)} &= \int\int d\omega d\omega' \left|\left\langle out\right|\hat{\mathbf{U}}\left|in\right\rangle\right|^2 \\ &= \int\int d\omega d\omega' \left[|b_e|^2\left|\beta^{(\mathrm{F})}(\omega,\omega')\right|^2 + |b_o|^2\left|\beta^{(\mathrm{F})}(\omega',\omega)\right|^2\right. \\ &\quad \left. + \left(b_e^* b_o\left(\beta^{(\mathrm{F})}\right)^*(\omega,\omega')\beta^{(\mathrm{F})}(\omega',\omega) + \text{c.c.}\right)\right] \\ &= \left(|b_e|^2 + |b_o|^2\right)\int\int d\omega d\omega' \left|\beta^{(\mathrm{F})}(\omega,\omega')\right|^2 \\ &\quad + \left(b_e^* b_o\int\int d\omega d\omega' \beta^{(\mathrm{F})*}(\omega,\omega')\beta^{(\mathrm{F})}(\omega',\omega) + \text{c.c.}\right). \end{aligned} \tag{69}$$

In (69) we get two cases for the integrand which we can write down explicitly by (50). Let us rename $\omega \to \omega_s$ and $\omega' \to \omega_i$ to allow ourselves to keep the notation of III A 2. Then the first case gives the explicit integrand

$$\begin{aligned} \left|\beta^{(\mathrm{F})}(\omega_s,\omega_i)\right|^2 &= \frac{N_s^{(\mathrm{G})}N_i^{(\mathrm{G})}}{N}\exp\left[-\frac{1}{2}\boldsymbol{\omega}^T 2\cdot(\mathbf{M}+\mathbf{M}')\boldsymbol{\omega} + \boldsymbol{\omega}^T 2\cdot(\mathbf{M}\boldsymbol{\omega_0} + \mathbf{M}'\boldsymbol{\omega_F})\right]\cdot \\ &= \exp\left[-\boldsymbol{\omega_0}^T\mathbf{M}\boldsymbol{\omega_0} - \boldsymbol{\omega_F}^T\mathbf{M}'\boldsymbol{\omega_F}\right]. \end{aligned} \tag{70}$$

And the second integrand reads

$$\begin{aligned}
&\beta^{(\mathrm{F})*}(\omega_s,\omega_i)\beta^{(\mathrm{F})}(\omega_i,\omega_s) = \qquad (71)\\
&= \frac{N_s^{(\mathrm{G})}N_i^{(\mathrm{G})}}{N}\cdot\exp\left[-\boldsymbol{\omega_0}^T\mathbf{M}\boldsymbol{\omega_0}-\boldsymbol{\omega_F}^T\mathbf{M}'\boldsymbol{\omega_F}\right]\\
&\cdot\exp\left[-\frac{1}{2}\boldsymbol{\omega}^T(\mathbf{M}+\mathbf{M}')\boldsymbol{\omega}+\boldsymbol{\omega}^T\left(\mathbf{M}\boldsymbol{\omega_0}+\mathbf{M}'\boldsymbol{\omega_F}\right)-i\boldsymbol{\omega}^T\boldsymbol{\tau}\right]\\
&\cdot\exp\left[-\frac{1}{2}(\boldsymbol{\sigma_x}\boldsymbol{\omega})^T(\mathbf{M}+\mathbf{M}')\boldsymbol{\sigma_x}\boldsymbol{\omega}+(\boldsymbol{\sigma_x}\boldsymbol{\omega})^T\left(\mathbf{M}\boldsymbol{\omega_0}+\mathbf{M}'\boldsymbol{\omega_F}\right)+i(\boldsymbol{\sigma_x}\boldsymbol{\omega})^T\boldsymbol{\tau}\right].
\end{aligned}$$

In (71) we have introduced the Pauli matrix $\boldsymbol{\sigma_x}=\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$ to perform the switch of arguments. Lets define the following quantities for a clearer notation:

$$\begin{aligned}
\mathbf{A_0} &\equiv 2(\mathbf{M}+\mathbf{M}'), && (72)\\
\mathbf{A} &\equiv \frac{1}{2}\mathbf{A_0}+\frac{1}{2}\boldsymbol{\sigma_x}\mathbf{A_0}\boldsymbol{\sigma_x}, && (73)\\
\mathbf{F_0} &\equiv 2\left(\mathbf{M}\boldsymbol{\omega_0}+\mathbf{M}'\boldsymbol{\omega_F}\right) && (74)\\
\mathbf{F} &\equiv \frac{1}{2}\mathbf{F_0}+\frac{1}{2}\boldsymbol{\sigma_x}\mathbf{F_0}, && (75)\\
\boldsymbol{\Delta\tau} &\equiv \boldsymbol{\sigma_x}\boldsymbol{\tau}-\boldsymbol{\tau}=(\tau_i-\tau_s)\cdot\begin{pmatrix}1\\-1\end{pmatrix}, && (76)\\
C &\equiv \exp\left[-\boldsymbol{\omega_0}^T\mathbf{M}\boldsymbol{\omega_0}-\boldsymbol{\omega_F}^T\mathbf{M}'\boldsymbol{\omega_F}\right]. && (77)
\end{aligned}$$

Thus we obtain

$$\left|\beta^{(\mathrm{F})}(\omega_s,\omega_i)\right|^2 = \frac{N_s^{(\mathrm{G})}N_i^{(\mathrm{G})}}{N}\cdot C\cdot\exp\left[-\frac{1}{2}\boldsymbol{\omega}^T\mathbf{A_0}\boldsymbol{\omega}+\boldsymbol{\omega}^T\mathbf{F_0}\right] \qquad (78)$$

and

$$\beta^{(\mathrm{F})*}(\omega_s,\omega_i)\beta^{(\mathrm{F})}(\omega_i,\omega_s) = \frac{N_s^{(\mathrm{G})}N_i^{(\mathrm{G})}}{N}\cdot C\cdot\exp\left[-\frac{1}{2}\boldsymbol{\omega}^T\mathbf{A}\boldsymbol{\omega}+\boldsymbol{\omega}^T\mathbf{F}+i\boldsymbol{\omega}^T\boldsymbol{\Delta\tau}\right]. \qquad (79)$$

The normalization constants appearing in $\frac{N_s^{(\mathrm{G})}N_i^{(\mathrm{G})}}{N}$ is given in XII.

We can now carry out the integrals in (69). Both (78) and (79) are two instances of the following kind of N dimensional integral

$$\int d\mathbf{v} \exp\left[-\frac{1}{2}\mathbf{v}^T\mathbf{X}\mathbf{v} + \mathbf{j}^T\mathbf{v}\right] = \sqrt{\frac{(2\pi)^N}{\det \mathbf{X}}} \cdot \exp\left[\frac{1}{2}\mathbf{j}^T\mathbf{X}^{-1}\mathbf{j}\right]. \tag{80}$$

For $\mathbf{X}$ being symmetric (and $\mathbf{j}$ being a general N dimensional vector), which is here the case, this can be derived by substituting $\mathbf{v} \to \mathbf{v} + \mathbf{X}^{-1}\mathbf{j}$ and then by analogy with (144) in XII. Thus (69) can be calculated explicitly and with (72)-(77) we obtain

$$\begin{aligned} G^{(2)} \;=\; & \frac{N_s^{(G)} N_i^{(G)}}{N} \cdot C \cdot \left\{ \left(|b_e|^2 + |b_o|^2\right) \frac{2\pi}{\sqrt{\det \mathbf{A_0}}} e^{\frac{1}{2}\mathbf{F_0}^T \mathbf{A_0}^{-1} \mathbf{F_0}} + \right. \\ & \left. (b_e^* b_o + \text{c.c.}) \frac{2\pi}{\sqrt{\det \mathbf{A}}} e^{\frac{1}{2}\mathbf{F}^T \mathbf{A}^{-1} \mathbf{F}} \cdot \exp\left[-\frac{1}{2}\boldsymbol{\Delta\tau}^T \mathbf{A}^{-1} \boldsymbol{\Delta\tau}\right] \right\} \end{aligned} \tag{81}$$

Factoring out all the constants, combining them in $K = \frac{N_s^{(G)} N_i^{(G)}}{N} C \frac{2\pi}{\sqrt{\det \mathbf{A_0}}} e^{\frac{1}{2}\mathbf{F_0}^T \mathbf{A_0}^{-1} \mathbf{F_0}}$ and renaming $G^{(2)} \equiv P_c(\Delta\tau)$ yields

$$\begin{aligned} P_c(\Delta\tau) \;=\; & K \cdot \left(|b_e|^2 + |b_o|^2 + \right. \\ & \left. + 2\text{Re}\left(b_e^* b_o\right) \underbrace{\sqrt{\frac{\det \mathbf{A_0}}{\det \mathbf{A}}} e^{\frac{1}{2}\mathbf{F}^T \mathbf{A}^{-1} \mathbf{F} - \frac{1}{2}\mathbf{F_0}^T \mathbf{A_0}^{-1} \mathbf{F_0}}}_{\equiv \Lambda} \cdot \exp\left[-\frac{1}{2}\boldsymbol{\Delta\tau}^T \mathbf{A}^{-1} \boldsymbol{\Delta\tau}\right] \right). \end{aligned} \tag{82}$$

It will be vital to include all the discussed influences (frequency mismatch, correlations and filtering) in the network tomography since the extraction of the visibilities is such that we are fitting a theoretical model (the above) to our set of data. To improve the procedure we need to have a theoretical description that is as complete as possible.

As we shall see in the next section an ubiquitous measure for distinguishability in quantum optics experiments - and in particular for the network tomography - is the so called *visibility*.

The visibility is defined [38] as the fractional reduction of the probability of coincidence from its *temporal* uncorrelated case, i.e.

$$\mathcal{V} = 1 - \frac{P_c(\Delta\tau = 0)}{P_c(\Delta\tau \to \infty)}\,. \tag{83}$$

Which, for (82), reads

$$\begin{aligned} \mathcal{V}_{out,\,in} &= -\frac{b_e^* b_o + b_e b_o^*}{|b_e|^2 + |b_o|^2}\,\Lambda. \\ \mathcal{V}_{\{s_1,s_2\},\{t_1,t_2\}} &= -\Lambda\,\frac{u^*_{s_2,t_2}u^*_{s_1,t_1}u_{s_2,t_1}u_{s_1,t_2} + u_{s_2,t_2}u_{s_1,t_1}u^*_{s_2,t_1}u^*_{s_1,t_2}}{|u_{s_2,t_2}u_{s_1,t_1}|^2 + |u_{s_2,t_1}u_{s_1,t_2}|^2} \end{aligned} \tag{84}$$

If we had perfectly indistinguishable photons, the visibility $\mathcal{V}_0$ would only depend on the elements of the network $\mathbf{U}$

$$\mathcal{V}_0 = -\frac{u^*_{s_2,t_2}u^*_{s_1,t_1}u_{s_2,t_1}u_{s_1,t_2} + u_{s_2,t_2}u_{s_1,t_1}u^*_{s_2,t_1}u^*_{s_1,t_2}}{|u_{s_2,t_2}u_{s_1,t_1}|^2 + |u_{s_2,t_1}u_{s_1,t_2}|^2}. \tag{85}$$

Which deviates by the factor Λ.

1. Extraction of Visibilities from Experimental Data

The procedure to extract visibilities from experimental data shall now be briefly introduced. The given experimental data is a collection of sample points of the coincident count rate that corresponds to (82) for different time-delays $\Delta\tau$. However, as individual delays are set by translating a fiber coupler with a motorized screw there can be a small drift in coupling efficiency over the whole delay-range. Without this drift, the background of the visibility would be a horizontal straight line. For drifts smaller than 5% of the two-photon flux the drift is in good approximation linear and can be modeled with an additional parameter, T [15]. Therefore

our fit model reads

$$N_c(t) = (1 + T \cdot t)\left(\mathcal{O} + \mathcal{A} \cdot 2 \cdot \Lambda \cdot \exp\left[-\frac{1}{2}\left[(1,-1)\,\mathbf{A}^{-1}(1,-1)^T\right] \cdot (t - t_c)^2\right]\right). \tag{86}$$

Here$T$, $\mathcal{O}$, $\mathcal{A}$ and t_c are parameters to be fitted. The *adjusted* visibilities (85) can be extracted from the fitting-parameters via

$$\tilde{\mathcal{V}} = 1 - \frac{\mathcal{O} + \mathcal{A}}{\mathcal{O}} = -\frac{\mathcal{A}}{\mathcal{O}}. \tag{87}$$

V. TOMOGRAPHY OF PASSIVE LINEAR OPTICAL NETWORKS

This section shall cover a brief introduction to the three different approaches to tomography of passive linear optical networks as well as brief systematic discussions of the respective method. Thereafter we propose some new methods in order to benchmark these approaches. At last we propose a novel approach to network tomography.

A. Vienna

This approach is based on minimizing a certain matrix function $D(\mathbf{U})$, where $\mathbf{U}$ describes the underlying linear optical circuit. And $\min_{\mathbf{U}} D(\mathbf{U})$ corresponds to the best possible guess for $\mathbf{U}$, given a certain set of reference data. The reference data shall just be the set of visibilities extracted from experimental measurements by the method derived in IV B 1, i.e. $\tilde{\mathcal{V}}$ in (87). Therefore we define

$$D(\mathbf{U}) \coloneqq \sum_{i=1}^{N} \frac{\left|\tilde{\mathcal{V}}_i - \mathcal{V}_i(\mathbf{U})\right|^2}{\sigma_i^2 \Gamma}. \tag{88}$$

The explicit form of $\mathcal{V}_i(\mathbf{U})$ is given in (85), σ_i is the standard deviation derived from the extraction of $\tilde{\mathcal{V}}_i$. Γ is a constant taking care of the right normalization

w.r.t the degrees of freedom [26] and N is the set of all possible visibilities and therefore $N = \binom{M}{2}^2$. At this point we might already discuss some properties that follow from the definition of $D(\mathbf{U})$.

1. Note that $2 \cdot \binom{M}{2}$ parameters are sufficient to describe any unitary (however, it can be more or fewer as we have mentioned in II B 2). As a consequence we are optimizing an over-complete set of equations. This is done in the hope that the noise – that is in general imposed on $\tilde{\mathcal{V}}_i$ – is not biased towards any direction. If that is true, the set of visibilities $\tilde{\boldsymbol{\mathcal{V}}}$ scatters randomly around the expectation value and hence, calculating the expectation value, they average out in the limit $N \to \infty$. Of course N is limited by $N = \binom{M}{2}^2$ but that should still be better than using for instance $2 \cdot \binom{M}{2}$ visibilities and provide some robustness against errors. However, this applies only if every physical effect is taken into account. If it was not, the data $\tilde{\boldsymbol{\mathcal{V}}}$ could be biased (see for instance IX C). Therefore it is crucial to consider a description of the physical effects that is as detailed as possible to perform the extraction of the primary data $\tilde{\boldsymbol{\mathcal{V}}}$. Which can be considered a weak point of this method.

2. Let us anticipate another drawback at this point. The main problem of this approach is to minimize $D(\mathbf{U})$ which can in general only be done numerically. And, as we will see, the numerical success (cost) depends on the initial points from which we start the optimization. I.e. we have the additional requirement that we need a well educated guess for where the parameters (the unitary) are located. And furthermore the minimization might lead to an overhead in computational cost. However, only using a minimization approach, some information on the quality of the reconstruction is gained.

3. Another important and in contrast very convenient property is that visibility measurements are independent of loss in the in- and output modes [24].

4. Again anticipating what will be discussed in more detail in the following

we would like to mention another advantage. If one knows the underlying structure of the network, one can apply unitary constraints throughout the optimization process by expressing $\mathbf{U}$ as a function of the physical building blocks, described by the set of beam splitting and phase shifting parameters $\mathbf{p}$, i.e. $\mathbf{U}(\mathbf{p})$. By choosing to implement the unitary constraints in such a way, we implicitly impose additional conditions that are restricting the predicted set of visibilities and therefore $\mathbf{U}(\mathbf{p})$ to possess the right structure. In fact this property is a very strong argument for choosing an algorithm that is intrinsically taking into account the unitary constraints in such a natural manner as it does not allow to deviate to much from the actual structure, as it automatically restricts the transformation to the unitary manifold on which it has to be located.

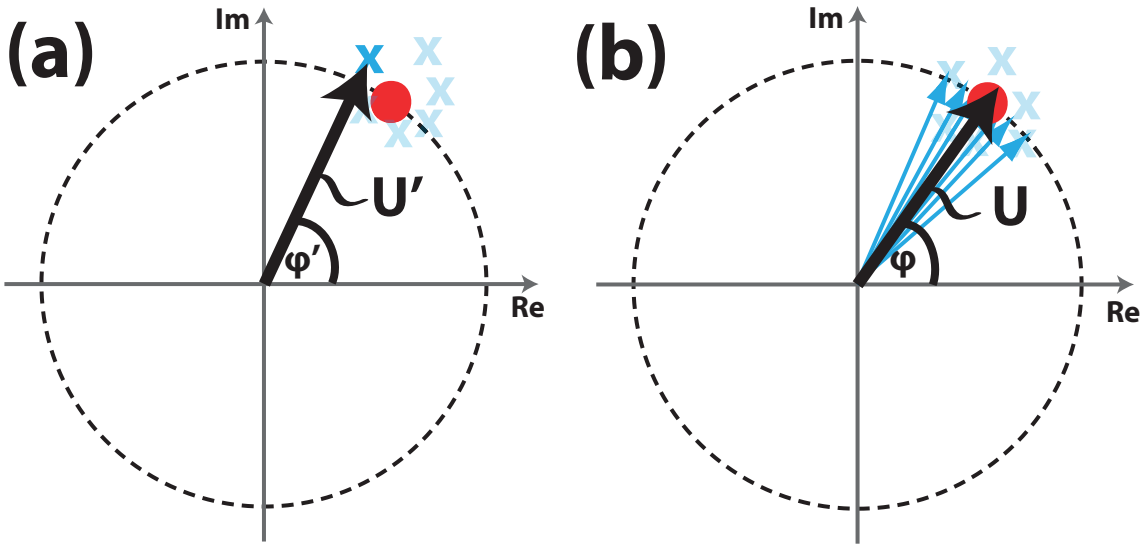


Figure 9: Two dimensional exemplary depiction of the difference between (a) an approach that forces the unitary property by applying a polar decomposition and (b) an approach that samples from an over-complete set and applies unitary constraints in the natural manner (i.e. it optimizes φ, given the structure $e^{i\varphi}$) as in 'Vienna'. The red dot is the actual value that describes the unitary transformation $\mathbf{U}$ and the crosses correspond to measurements. (a) applies for the passive approaches 'Bristol' and 'Brisbane'. Note that here just one measurement is realized and enters the reconstruction, different from (b) where several measurements, originating from just one measurement run, enter into the reconstruction.

For solving the minimization problem (88) we have employed an interior-point method (IPM) [45–47] to solve $\min_{\mathbf{p}} D\left(\mathbf{U}(\mathbf{p})\right)$.[2] However, the convergence of IPM does depend on the initial values $\mathbf{p}_0$ and only recovers local minima in general [48]. Therefore the success of the minimization crucially depends on the educated guess for the initial values.[3] But if we guess a $\mathbf{p}_0$ that is close enough to the true value of $\mathbf{p}_{\min}$ we will converge into the global minimum. The remaining question is: "How

[2] Implemented via the nonlinear optimization toolbox in matlab.

[3] We have also practically investigated the possibility of utilizing pattern-search algorithms that are often doing better in finding global minima and are less depending on the starting point. These algorithm do not depend on gradients and are utilized for highly nonlinear or ill-defined problems and are usually doing better in finding global minima [49]. However, we gave up following this approach as the convergence took much longer and the results were much worse.

close is close enough?". To guarantee convergence towards the right minimum we have implemented the following numerical trick: We start with some parameter configuration $\mathbf{p}_0$ of that we know that it is within a range of $\boldsymbol{\epsilon}$ away from the true value $\mathbf{p}_{\min}$ that minimizes (88). I.e. we can trust the initial guess within the same region $\mathbf{p}_0 \pm \boldsymbol{\epsilon}$. In the next step we choose r random initial points $\mathbf{p}_r \in \mathbf{p}_0 \pm \boldsymbol{\epsilon}$ from that region and run the optimization. Eventually, we pick the resulting $\mathbf{p}_{\min}$ that minimizes (88) over the r optimization runs.

B. Bristol

For this approach single photon transition amplitudes $\boldsymbol{\tau}$ and visibilities are utilized in order to infer the network. The aim is to find an approach to linear optical network tomography that is independent of constant loss in the in- and output modes. Let us briefly introducethe procedure, described in detail in [24].
Consider the 2×2 device, depicted in Fig. 10: The respective visibility for a coincidence measurement at output modes j and g, if one inputs two indistinguishable photons into modes k and h, is given as

$$\mathcal{V}_{g,h,j,k} = -2\cos(\alpha_{j,k} - \alpha_{j,h} - \alpha_{g,k} + \alpha_{g,h}) \frac{\tau_{j,k}\tau_{j,h}\tau_{g,k}\tau_{g,h}}{\tau_{j,k}^2\tau_{g,h}^2 + \tau_{g,k}^2\tau_{j,h,}^2}. \tag{89}$$

This follows immediately from (85) by rewriting $u_{\mu,\nu} = \tau_{\mu,\nu} \cdot e^{i\alpha_{\mu,\nu}}$. The visibility (89) is independent of loss in in- and output modes because the loss factors cancel out as they appear in the same manner in the nominator and denominator [24].
One now defines another loss insensitive quantity $x_{g,h,j,k}$ that reads

$$x_{g,h,j,k} = \frac{\tau_{j,k}\tau_{g,h}}{\tau_{j,h}\tau_{g,k}}. \tag{90}$$

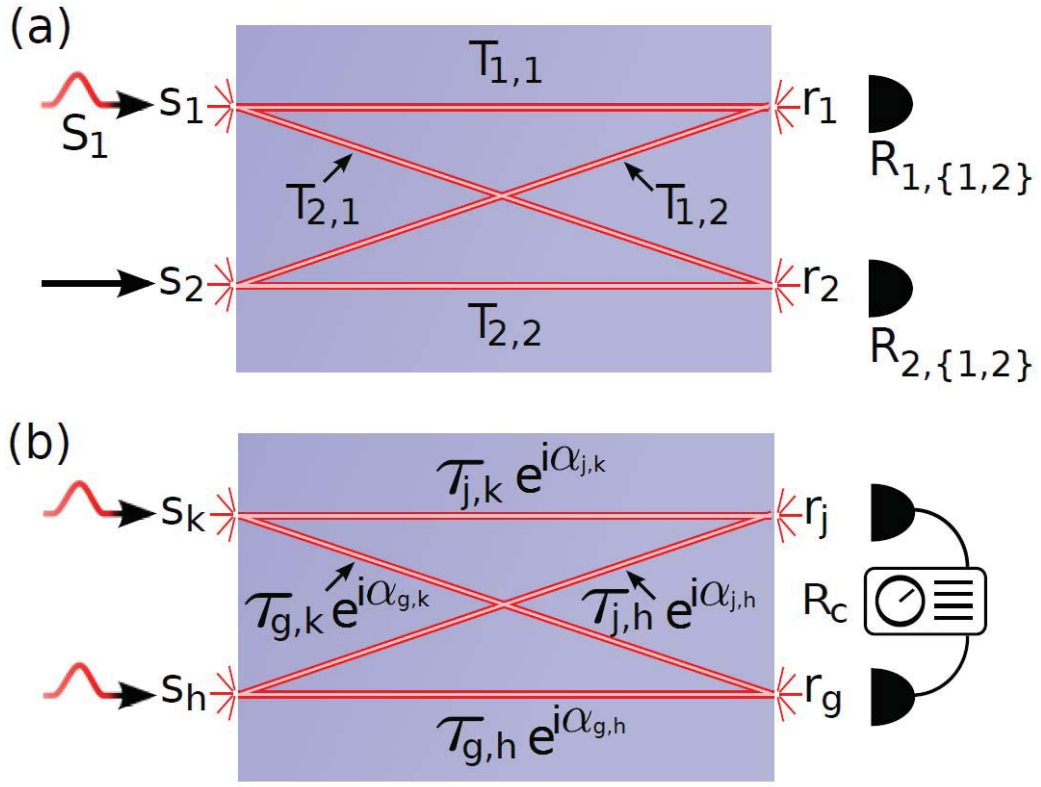


Figure 10: Underlying 2×2-building blocks. (a) For one-photon transition amplitudes that are obtained from a one-photon state that is fed into one mode while the other(s) are in the vacuum mode. (b) Visibilities are obtained from two-photon probing state and are phase sensitive. They allow the reconstruction of the phases. Taken from [24].

And for practical purposes

$$y_{g,h,j,k} \coloneqq x_{g,h,j,k} + x_{g,h,j,k}^{-1}. \tag{91}$$

Note that $x_{g,h,j,k}$ and $\mathcal{V}_{g,h,j,k}$ are both fully determined by experimental measurements of the one photon transition amplitudes ($\Rightarrow x_{g,h,j,k}$) and the HOM-Dip ($\Rightarrow \mathcal{V}_{g,h,j,k}$).

The matrix of a $m \times m$ network that is to be recovered can be parameterized as

$$\mathbf{M}_1 = \begin{pmatrix} \tau_{1,1} & \tau_{1,2} & \cdots & \tau_{1,m} \\ \tau_{2,1} & \tau_{2,2}e^{i\alpha_{2,2}} & \cdots & \tau_{2,m}e^{i\alpha_{2,m}} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{m,1} & \tau_{m,2}e^{i\alpha_{m,2}} & \cdots & \tau_{m,m}e^{i\alpha_{m,m}} \end{pmatrix}. \tag{92}$$

In this real bordered representation it has already been encapsulated that the global input and output phases can be regarded as trivial as they can not be determined (see (17)). Therefore the phase angles in the first row and first column can be regarded as vanishing elements

$$\alpha_{\mu,1} = \alpha_{1,\mu} = 0. \tag{93}$$

Another condition that applies to quantum mechanical measurements is that the measured photon statistic is invariant under the complex conjugation transformation $\mathbf{M} \to \mathbf{M}^*$. Therefore the sign of one of the remaining angles in (92) can be regarded as arbitrary. We choose

$$\text{sign}(\alpha_{2,2}) = +1. \tag{94}$$

With these last two assumptions it is possible to recover all remaining phases in (92) by means of (89). Rewriting (89) with (91) results in

$$\cos(\alpha_{j,k} - \alpha_{j,h} - \alpha_{g,k} + \alpha_{g,h}) = -\frac{\mathcal{V}_{g,h,j,k}y_{g,h,j,k}}{2}. \tag{95}$$

Applying the $\cos^{-1}$ and defining $\rfloor x \lfloor \coloneqq \cos^{-1}(\cos(x))$ yields

$$\beta_{g,h,j,k} \coloneqq \rfloor \alpha_{j,k} - \alpha_{j,h} - \alpha_{g,k} + \alpha_{g,h} \lfloor = \cos^{-1}\left(-\frac{\mathcal{V}_{g,h,j,k}y_{g,h,j,k}}{2}\right). \tag{96}$$

Note that $\rfloor x \lfloor$ is not just the absolute value of x. It does also restrict $0 \leq \rfloor x \lfloor < \pi$ and hence we loose the information on the sign of the argument. However, if the absolute value of the desired phase angle, say $\alpha_{g,h}$, as well as the other three reference angles is known, the sign can be recovered:

$$\begin{aligned} \text{sign}(\alpha_{g,h}) \;=\; \text{sign}\,(&\rfloor \beta_{g,h,j,k} - \rfloor \alpha_{j,k} - \alpha_{j,h} - \alpha_{g,k} - \rfloor \alpha_{g,h} \lfloor \lfloor - \\ &\rfloor \beta_{g,h,j,k} - \rfloor \alpha_{j,k} - \alpha_{j,h} - \alpha_{g,k} + \rfloor \alpha_{g,h} \lfloor \lfloor)\,. \end{aligned} \tag{97}$$

Now we can recover all phases in (92) from (96) in two steps. First we find $\rfloor \alpha_{g,h} \lfloor$ by setting $j = k = 1$ in (96). Thereafter we recover the signs of the phase angles in the second column by fixing $\alpha_{j,k} = \alpha_{2,1}$ while $\alpha_{g,h} = \alpha_{3\ldots m,2}$. And for the second row by fixing $\alpha_{j,k} = \alpha_{1,2}$ and $\alpha_{g,h} = \alpha_{2,3\ldots m}$. The remaining angles are found by fixing $\alpha_{j,k} = \alpha_{2,2}$. Hence, with the introduced method it is possible to infer the remaining angles from the conditions (93) and (94) we are able to recover the angles from loss insensitive measurement data

$$\tilde{\alpha}_{g,h} = \text{sign}(\alpha_{g,h})\beta_{g,h,j,k}. \tag{98}$$

However, the matrix to be recovered does still contain loss sensitive single photon transition amplitudes $\tau_{\mu,\nu}$. We can circumvent this by applying the unitary constraints to $\mathbf{M}_1$. Note that this is usually not the case as the system undergoes interactions with the environment. In that case one needs to force unitarity. This is usually done by applying the polar decomposition as it gives a best estimate of the true unitary [50].

Assuming the unitary property, we first rewrite $\mathbf{M}_1$as

$$\mathbf{M}_2 = \begin{pmatrix} \tau_{1,1} & \tau_{1,2} & \cdots & \tau_{1,m} \\ \tau_{2,1} & \tilde{x}_{2,2,1,1}e^{i\tilde{\alpha}_{2,2}} & \cdots & \tilde{x}_{2,m,1,1}e^{i\tilde{\alpha}_{2,m}} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{m,1} & \tilde{x}_{m,2,1,1}e^{i\tilde{\alpha}_{m,2}} & \cdots & \tilde{x}_{m,m,1,1}e^{i\tilde{\alpha}_{m,m}} \end{pmatrix}, \tag{99}$$

where $\tilde{x}_{g,h,j,k} = x_{g,h,j,k}\frac{\tau_{j,h}\tau_{g,k}}{\tau_{j,k}}$. When $\mathbf{M}$ can be assumed to be unitary we can apply the unitary conditions to the rows and columns of $\mathbf{M}$ (e.g. $M^*_{\mu,\nu}M_{\mu,\kappa} = \delta_{\nu,\kappa}$ in Einstein notation). This results in systems of equations for which the coefficient matrix is given as

$$\mathbf{M}_s = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & x_{2,2,1,1}e^{i\tilde{\alpha}_{2,2}} & \cdots & x_{2,m,1,1}e^{i\tilde{\alpha}_{2,m}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{m,2,1,1}e^{i\tilde{\alpha}_{m,2}} & \cdots & x_{m,m,1,1}e^{i\tilde{\alpha}_{m,m}} \end{pmatrix}. \tag{100}$$

To recover $\mathbf{M}_2$ we need to solve the following systems of equations to obtain $\tau_{1,\mu}$ and $\tau_{\mu,1}$:

$$\mathbf{M}_s^\dagger \begin{pmatrix} \tau_{1,1}^2 \\ \tau_{2,1}^2 \\ \vdots \\ \tau_{m,1}^2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{M}_s \begin{pmatrix} \tau_{1,1}^2 \\ \tau_{1,2}^2 \\ \vdots \\ \tau_{1,m}^2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{101}$$

Note that $\mathbf{M}_s$ only contains quantities that are loss insensitive and purely determined by experimental measurements. As a consequence the single photon transition amplitudes obtained from solving (101) are also loss insensitive. And with the labeling introduced above (99) we can recover the underlying network from the loss insensitive $x_{g,h,j,k}$, $\alpha_{g,h}$ and the single photon transition amplitudes obtained via (101). Some remarks: In order to apply the procedure, the first

row and column should not contain any vanishing elements. If they do, one only needs to relabel in- and/or output modes accordingly to move the vanishing elements into another column or row. Another assumption is constant loss in in- and output modes which might not apply in reality.
Now that the method has been introduced, let us discuss some systematical pros and cons:

1. The advantage of this method is that it is loss insensitive (for the assumed kind of constant loss).

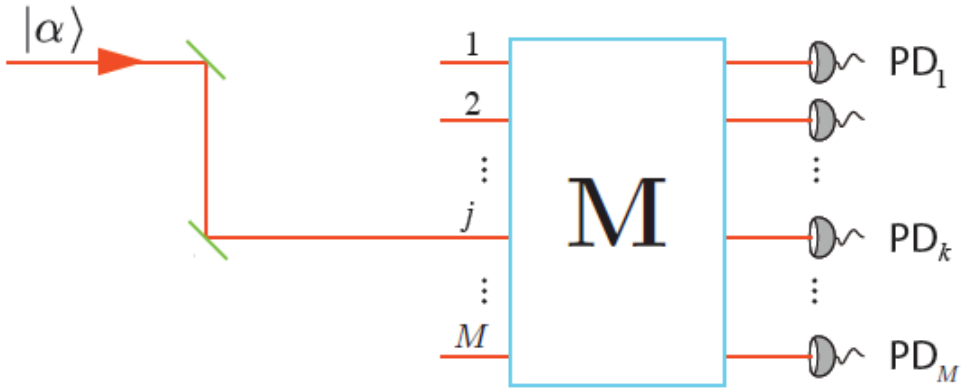
2. Another advantage is that it is computationally very cheap and feasible for large networks. One only needs to perform a polar decomposition and a matrix inversion which can be done for any practical size of the underlying network as they both scale like $\mathcal{O}(n^3)$ in the worst case [51].

3. However, a major drawback is that this approach is completely passive: It contains no feedback like the approach from the previous section did. It is very strongly conditioned by the measurement data and the reliability of these. If they contain slight errors they propagate through the whole reconstruction process. Consider also Fig. 9 to see this. One measures a certain set of experimental data $\boldsymbol{\mathcal{V}}$ and $\boldsymbol{\tau}$ which reliability depends on the experimental properties of the observable measured. (For instance: Collecting a set of one photon transmission amplitudes always carries errors, such as higher order noise etc.) Then the taken experimental values enter without any additional modification. In the worst case scenario one could not even tell if it was completely off, whereas in the previous method in such a case $D(\mathbf{U})$ would attain an untypically high value.

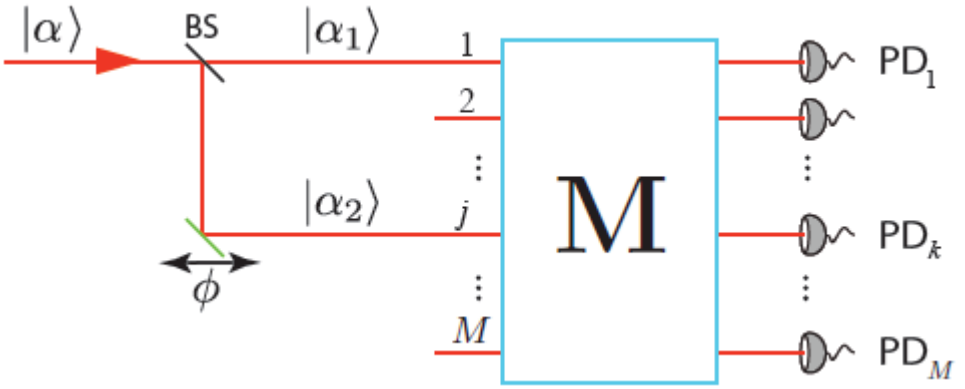
4. Additionally, another step must be performed in order to obtain the missing one photon transition amplitudes. Namely the inversion of (101). This can lead to further amplification of the imposed errors.

We will investigate the last two points in quantitative detail in section X B.

C. Brisbane



(a)



(b)

Figure 11: Schematic sketch of the steps (probing states and measurements) necessary. (11a) To obtain the amplitudes. And (11b) to obtain the phases. Readopted from [25].

In this section we introduce the last approach that shall be investigated. Further details might be found in [25]. This approach deploys coherent probe states to characterize the linear optical network. As previously the implemented network

shall be described by the matrix $\mathbf{M}$ that only is unitary in the ideal case of no loss. As in (16) the in- and output relations are characterized by this matrix via $\hat{a}_t^\dagger \to \sum_{k=1}^{M} M_{kt}\,\hat{a}_k^\dagger$. And therefore the input-output relations for any multimode coherent state reads

$$\beta_t = \sum_{k}^{M} M_{kt}\alpha_k. \tag{102}$$

Where the multimode coherent input state is noted as $|\alpha_1, \alpha_2, \ldots, \alpha_M\rangle$ and the output as $|\beta_1, \beta_2, \ldots, \beta_M\rangle$. Inferring $\mathbf{M}$ from *intensity measurements* can be accomplished by executing the following steps to obtain the amplitudes and phases of $M_{jk} = r_{jk}e^{i\theta_{jk}}$:

1. Amplitudes

1. Feed coherent state with intensity I into mode j
2. Measure all the intensities I_k at output mode k simultaneously to obtain the amplitudes
$$r_{jk} = \sqrt{\frac{I_k}{I}}, \quad k = 1, 2, \ldots, M. \tag{103}$$
3. Do this for $j = 1, 2, \ldots, M$

2. Phases

1. Send a coherent state $|\alpha\rangle$ to a $50:50$ BS and control the phase ϕ between them. Then send $|\alpha_1\rangle$ to input 1 and $|\alpha_2\rangle = \left|\alpha_1 e^{i\phi}\right\rangle$ to input mode j.
2. The intensity at output mode k is given as
$$I_k = I\left|M_{1k} + M_{jk}e^{i\phi}\right|^2 = \begin{cases} I\left(r_{11}^2 + r_{j1}^2 + 2r_{11}r_{j1}\cos\phi\right) & k = 1 \\ I\left(r_{1k}^2 + r_{jk}^2 + 2r_{1k}r_{jk}\cos(\phi + \theta_{jk})\right) & k \neq 1 \end{cases}. \tag{104}$$

3. Choose $\phi = 0$ for I_1 at its maximum and subsequently sweep ϕ until I_k attains its maximum: At this point we have

$$\theta_{jk} = 2\pi - \phi. \tag{105}$$

4. Repeat this for $j = 2, 3, \ldots, M$ and thereby obtain all non-trivial phases θ_{jk}.

In order to obtain the unitary that approximates best the underlying network the polar decomposition is applied.
In the same manner as before we shall now discuss some systematic properties of this method:

1. An experimental advantage is that this method deploys coherent states for probing. These are in good approximation given by standard laser sources. They are known to be much brighter and thus provide a very good signal to noise ratio than current single photon source. Therefore we expect the imposed errors to be smaller. However, considering XII the errors should neither be underestimated.

2. The experimental and computational effort is well under control. The only post-processing step required is that of performing a polar decomposition. Consult X B for an explicit discussion. Another consequence is that no further amplifications of the imposed errors are caused (consider e.g. the previous method where this was the case because additional analytical treatment was required in order to obtain the network).

3. The drawback that we would like to highlight is the same as point 3 of the previous method: The scheme is entirely passive and contains no additional feedback (see Fig. 9).

We will now introduce some methods to analyze the three previously introduced methods.

VI. WITNESS OF QUALITY

Let us define the following witness as a quality measure for how good the method works. One instance would be to compare the trace distance of the reconstructed matrix $\tilde{\mathbf{U}}$ and the actual matrix $\mathbf{U}$, i.e. $T(\tilde{\mathbf{U}}, \mathbf{U}) \coloneqq \frac{1}{2} \left\| \tilde{\mathbf{U}} - \mathbf{U} \right\|_1$. However, in practice we do usually not know the actual matrix $\mathbf{U}$ (even in the best case we only know a close approximation of it) which is precisely why we were interested in network tomography in the first place. In other words: $\tilde{\mathbf{U}}$ is somehow a function of a set of primary data, i.e. $\tilde{\mathbf{U}}(f(\mathbf{U}))$.[4] As discussed in the previous section and as we will see in section X B, performing the actual tomography can, in general, lead to additional errors. Consequently it seems more reasonable to consult $f(\mathbf{U})$, i.e. comparing $f(\mathbf{U})$ and $f\left(\tilde{\mathbf{U}}(f(\mathbf{U}))\right)$ in some measure (Let f for instance be an unspecified function that maps $U(n) \to \mathbb{R}^m$). We choose the finite set (vector) of visibilities $\boldsymbol{\mathcal{V}}$ as such a function. The procedure is depicted in Fig 12.

We do so because it is because it is a function that is carrying information about the *quantum inference* and we are concerned about the quantum properties and quantum applications of the network.

Therefore, we propose the *self-consistency measure* as witness of the quality for the network tomography that had been applied to obtain $\mathbf{U}'$

$$\mathcal{Q} \coloneqq \left\| \tilde{\boldsymbol{\mathcal{V}}}(\mathbf{U}) - \boldsymbol{\mathcal{V}}(\tilde{\mathbf{U}}) \right\|_1 . \tag{106}$$

At first this definition seems a bit redundant and it is in fact only as reliable as the experimentally extracted visibilities $\tilde{\boldsymbol{\mathcal{V}}}(\mathbf{U})$ themselves are.[5] We are writing $\tilde{\boldsymbol{\mathcal{V}}}(\mathbf{U})$ to indicate that the extracted visibilities $\tilde{\boldsymbol{\mathcal{V}}}(\mathbf{U})$ are always an approximation. Its

[4] However, this scenario is not including numerical experiments, such as the one in section VIII, because in such a case we explicitly generate the underlying actual unitary $\mathbf{U}$.

[5] This is why one must have a description of the experiment that is as complete as possible which was object to the first part of this work

reliability depends on the respective model (consider (87)), whereas this is not the case for $\boldsymbol{\mathcal{V}}(\tilde{\mathbf{U}})$ since it can be obtained directly from the reconstructed unitary $\tilde{\mathbf{U}}$ via (85). In other words: $\mathcal{Q}$ is a proper measure to assess the quality of the reconstruction if we can extract a high quality of reconstruction if we can extract a highly reliable set of primary data. But, this might be a problem that is much better under control. And note that if the extraction is free of errors $\mathcal{Q}$ is a proper measure.

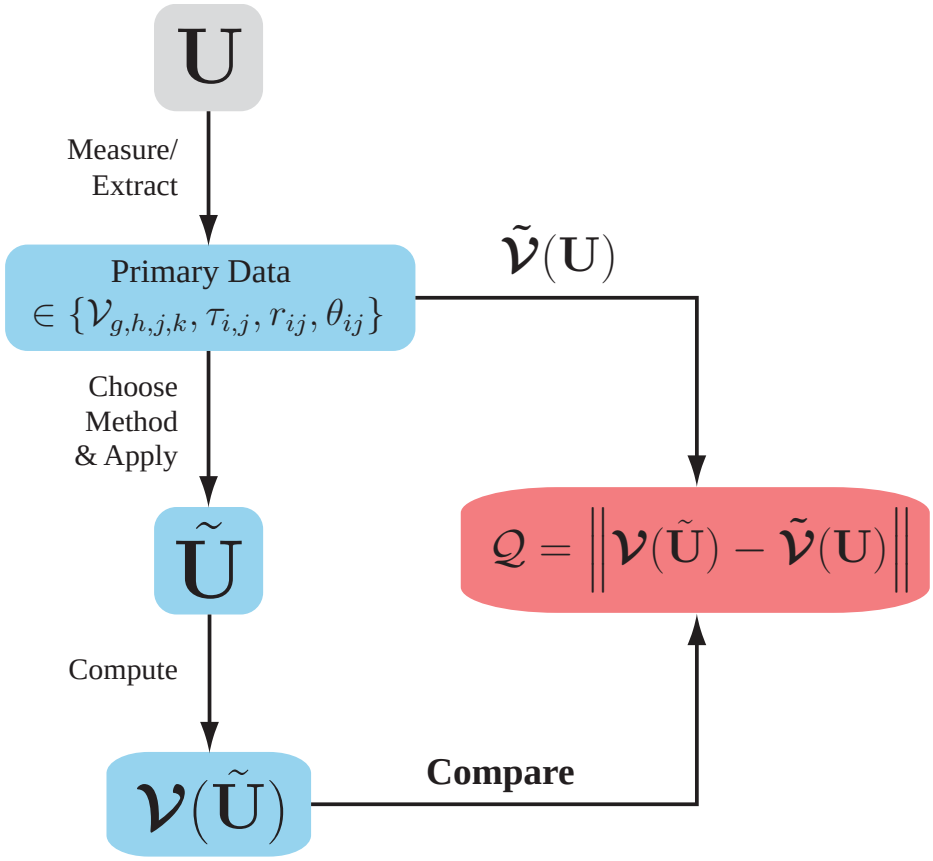


Figure 12: Flowchart representation for the self-consistency witness: After measuring the primary data for the respective method, e.g. $\mathcal{V}_{g,h,j,k}$ and $\tau_{i,j}$ for 'Bristol', this data is utilized for the network tomography that yields $\tilde{\mathbf{U}}$. Having $\tilde{\mathbf{U}}$ the hereby determined set of visibilities $\boldsymbol{\mathcal{V}}(\tilde{\mathbf{U}})$ can be computed. The self-consistency check is obtained by calculating $\mathcal{Q}$ and may serve as a witness for the quality of the reconstruction.

VII. (BACK)PROPAGATION OF ERRORS

The usual (forward) error propagation addresses the following scenario: Assume that we are interested in the properties of a system that can be described by the functional dependence $F(\mathbf{x})$, where $\mathbf{x}$ is the vector of parameters on which the systems properties depend. Now in the case of forward propagation one would ask for the perturbation of the system ΔF, given a certain perturbation of the parameter vector $\Delta\mathbf{x}$. To carry out an analytical error propagation of this kind, a necessary requirement is, that the functional dependance $F(\mathbf{x})$ is known. This is in principle the case for 'Bristol' and 'Brisbane' and we will carry out the respective error propagation in X B. This will give us a bound on $\Delta\mathbf{U}$, assuming some perturbation of the primary data. Therefore $\Delta\mathbf{U}$ provides a statement on the robustness and thus quality of the respective network tomography method. However, this is not possible for 'Vienna' because no closed analytical dependence can be given (at least no feasible). For this reason we need to follow another approach that we shall propose in the following subsection.

3. Backpropagation of Errors

Let us start with stating the problem we are facing. Our aim is to solve the following optimization problem

$$\min_{\mathbf{p}} D(\mathbf{p}), \quad \text{where } D(\mathbf{p}) \coloneqq \sum_{i=1}^{N} \frac{|\mathcal{E}_i - \mathcal{A}_i(\mathbf{U}(\mathbf{p}))|^2}{\sigma_i^2 \Gamma}, \tag{107}$$

where $\mathcal{E}_i$ are the visibilities, determined by measurement and $\mathcal{A}_i$ is a analytical express for the i^{th} visibility that depends on $\mathbf{p}$. Usually, no solution exists for (107). However, we obtain an approximate solution $\mathbf{p}'$ as described in V A. In analogy with the forward propagation the following question arises: Given a perturbation $\Delta\mathcal{E}_i$ what is the according perturbation of $\mathbf{p}$, namely $\Delta\mathbf{p}$. Note that finding a

bound for $\Delta\mathbf{p}$ and thus for $\Delta\mathbf{U}$ numerically is not sufficient in this case and one should consider an additional quality measure, such as the one introduced in VI. This is because we can not determine with certainty whether the perturbation are purely caused by $\Delta\mathcal{E}_i$ or maybe just by some numeric circumstances of the optimization routine, such as a bad choice of the algorithm, starting point, etc.[6] We are aware, that for similar problems one might make use of the inverse information matrix of $D(\mathbf{p})$, however these methods have failed and provided unphysical values in practice. One reason is the shape of the Hessian that is unsuitable for numerical inversion and the numerical approximations of the information matrix is equally bad. Furthermore the information matrix does only provide a lower bound and we are rather interested in an upper bound [52].

We propose a different approach that is based on imposing the theoretical error-distribution to the primary data $\tilde{\boldsymbol{\mathcal{V}}}(\mathbf{U})$ and performing the minimization in a Monte-Carlo like manner by drawing $\mathcal{E}_i$ from this distribution. The details are given in pseudocode in Alg. 1.

Some remarks:

- The optimization algorithm has not been specified in Algorithm1, for a discussion we refer to V A.
- The explicit form for $\boldsymbol{A}(\mathbf{p})$ can be obtained by (85) under the requirement that the underlying structure of the network is known (e.g. sectionXVII for the BosonSampling setup from Fig. 22). As we will show in the next section, obtaining $\boldsymbol{A}(\mathbf{p})$ can always be accomplished for any passive linear optical network that is known to behave unitary if its dimensionality, i.e. the number of modes M, is known. $\boldsymbol{\mathcal{E}}$ are the experimentally measured visibilities and $\boldsymbol{\sigma}$ are their standard deviations.

[6] Analyzing the optimization routine for such a case is a very broad problem and investigating it in the framework of this work is unrewarding.

Algorithm 1 Backpropagation

Require: $\mathbf{p_0}$, $\boldsymbol{\epsilon}_l$, $\boldsymbol{\epsilon}_u$, N_S, N_R

1: **procedure** SAMPLING($\boldsymbol{\mathcal{V}}$,$\boldsymbol{\sigma}$)
2: **for** $i \leftarrow 1, N_S$ **do**
3: $\delta\boldsymbol{\mathcal{V}} \leftarrow \mathcal{N}(0, \boldsymbol{\sigma})$
4: $\boldsymbol{\mathcal{V}} \leftarrow \boldsymbol{\mathcal{V}} + \delta\boldsymbol{\mathcal{V}}$
5: **if** $\boldsymbol{\mathcal{V}} > 1$ **then** $\boldsymbol{\mathcal{V}} \leftarrow 1$
6: **if** $\boldsymbol{\mathcal{V}} < -1$ **then** $\boldsymbol{\mathcal{V}} \leftarrow -1$
7: **for** $j \leftarrow 1, N_R$ **do**
8: ▷ draw $\mathbf{p}_0 \in [\mathbf{p}_0 - \boldsymbol{\epsilon}_l, \mathbf{p}_0 + \boldsymbol{\epsilon}_u]$ from the uniform distribution
9: $\mathbf{p}_{0,\,\mathrm{temp}} = \mathbf{p}_0 + \mathrm{random}(-\boldsymbol{\epsilon}_l, \boldsymbol{\epsilon}_u)$
10: $[\mathbf{p}_{\mathrm{temp}}, D_{\mathrm{temp}}] \leftarrow \min_{\mathbf{p}_{\mathrm{temp}}} \mathrm{D}(\mathbf{p}_{\mathrm{temp}}, \boldsymbol{\mathcal{V}}, \boldsymbol{\sigma}, \mathbf{p}_{0_{\mathrm{temp}}})$
11: subject to $\mathbf{p}_0 - \boldsymbol{\epsilon}_l \leq \mathbf{p}_{\mathrm{temp}} \leq \mathbf{p}_0 + \boldsymbol{\epsilon}_u$
12: **if** $j == 1$ **then**
13: $D[i] \leftarrow D_{\mathrm{temp}}$
14: $\mathbf{p}[i] \leftarrow \mathbf{p}_{\mathrm{temp}}$
15: **else if** $j > 1$ **then**
16: **if** $D_{\mathrm{temp}} < D[i]$ **then**
17: $\mathbf{p}[i] \leftarrow \mathbf{p}_{\mathrm{temp}}$
18: **else if** $D_{\mathrm{temp}} == D[i]$ **then**
19: $\mathbf{p}_{\mathrm{E}}[i, j] \leftarrow \mathbf{p}_{\mathrm{temp}}$
20:

Require: $\Gamma, \boldsymbol{\mathcal{A}}(\mathbf{p})$

21: **procedure** D($\boldsymbol{p}, \boldsymbol{\mathcal{V}}, \boldsymbol{\sigma}, \boldsymbol{p}_0$)
22: $D \leftarrow 0$
23: $\boldsymbol{\mathcal{E}} \leftarrow \boldsymbol{\mathcal{V}}$
24: $D \leftarrow \sum_{i=1}^{\mathrm{Length}(\boldsymbol{\mathcal{V}})} \frac{|\mathcal{E}[i] - \mathcal{A}(\mathbf{p})[i]|^2}{\sigma[i]^2 \Gamma}$

- Line 19 in Algorithm1 encapsulates both, that the photon statistic is invariant under the complex conjugation of the underlying network $\mathbf{U} \rightarrow \mathbf{U}^*$ (therefore the phase parameters $\mathbf{p}_\varphi$ and $-\mathbf{p}_\varphi$ are equivalent) and the fact that two different results might yield the same value for D. (Whether or not the first point has to be taken into account depends on the boundaries $\boldsymbol{\epsilon}_l$, $\boldsymbol{\epsilon}_u$ that have been impose.)

- The second loop is the random initialization that prevents the algorithm from converging into the wrong local minimum, as discussed in V A.

VIII. PROPOSAL: EXTENSION AND IMPROVEMENT

From the systematic discussions in V A, V B and V C two questions are arising: How to add a feedback technique and thus improve the passive schemes 'Bristol' and 'Brisbane'? And how to extend the active scheme 'Vienna' to black box devices?

As a matter of fact, we can achieve both aims with the same action and the improvement of 'Bristol' and 'Brisbane' as well as the extension of 'Vienna' can be accomplished by merging the passive schemes with an additional post-processing technique that is implementing active feedback in the manner of 'Vienna'.

First we propose the novel approach and thereafter an numerical experiment that might be utilized to provide numerical evidence of the benefit of the novel method and for which we shall present results in section X E.

As mentioned for all of the three methods we are concerned about contain major systematic drawbacks: 'Bristol' and 'Brisbane' are entirely passive and the unitary constraints are enforced by applying the polar decomposition. While 'Vienna' is implementing the unitary constraints in a natural manner throughout the reconstruction and it is an active scheme, it is only applicable to very certain problems for which an a priori knowledge of the physical network structure and an educated guess for the characterizing parameters $\mathbf{p}_0$ is present. I.e. the drawback of either of these two different groups of tomography schemes is the advantage of the other and in fact our novel approach can be seen as a fusion of the active and passive group of tomography schemes.

Now the novel approach shall be introduced:

Consider the completely unknown black box network $\mathbf{U}$. We have seen that it is possible to obtain a unitary approximation, say $\tilde{\mathbf{U}}$, by performing either of the two passive schemes 'Brisbane' or 'Bristol'.

By virtue of the unitary property, we can *always* find a natural mapping to the

set of parameters

$$\mathbf{p} = \{\omega_1, \omega_2, \ldots, \omega_n, \varphi_1, \varphi_2, \ldots, \varphi_n\}, \tag{108}$$

where $n = \binom{M}{2}$. These are the n-dimensional (unitary) spherical coordinates and are equivalent to the beam splitting ratios and phase shifting ratios from (22). Therefore, we have just obtained *one representation* of the actual physical network encoded in $\mathbf{p}$. Which is one of the requirements for 'Vienna'.

Given $\tilde{\mathbf{U}}$, we can compute the respective set of spherical coordinates $\tilde{\mathbf{p}} = \{\tilde{\omega}_1, \tilde{\omega}_2, \ldots, \tilde{\omega}_n, \tilde{\varphi}_1, \tilde{\varphi}_2, \ldots, \tilde{\varphi}_n\}$ by the algorithm proposed by Reck et al. (see II B 2 and note, that actually less than n phase shifting ratios are necessary as we do not have to compute the full set $\tilde{\mathbf{p}}_\varphi = \{\tilde{\varphi}_1, \tilde{\varphi}_2, \ldots, \tilde{\varphi}_n\}$ because the global phases do, as previously, not contribute and do consequently not need to be computed.) But by applying Reck et al.'s algorithm to the reconstructed unitary $\tilde{\mathbf{U}}$ we have just obtained the necessary educated guess for the *initial values* $\mathbf{p}_0$, encoded in $\tilde{\mathbf{p}}$, that is the last necessary requirement in order to apply 'Vienna'.

Although the actual physical structure might be unknown, we do possess *one* and we claim this to be sufficient for our purposes. We claim so because the advantage lays in the fact that we are now able to apply the unitary constraints in a natural way and obtain an active feedback technique (Fig. 9) by applying 'Vienna'. However, bare in mind that we also require the measured over-complete set of visibilities for this purpose. Depending on which method had been chosen to obtain $\tilde{\mathbf{U}}$ an extra set of measurement data might be required.[7]

Let us summarize: Applying 'Bristol' or 'Brisbane' provides us with a structure and therefore an initial value that is necessary to apply 'Vienna'. By doing so, we have just eliminated the drawbacks in both schemes.

Fig. 13 depicts the novel approach in a flowchart. Furthermore it presents a

[7] This is definitely the case for 'Brisbane' where no visibilities are collected. For 'Bristol' this requires to measure the set of all possible visibilities and not just a subset that is in general sufficient for the scheme.

scheme to numerically quantify the approach.

- Ignoring the dashed lines yields the experimental situation (red boxed) in which the corresponding primary set of data $\tilde{\boldsymbol{\mathcal{V}}}$, $\tilde{\boldsymbol{\tau}}$, $\tilde{\boldsymbol{\theta}}$ is determined by experimental measurements. The subsequent steps are described above.

- Considering the whole chart yields a procedure for a numerical experiment that can be employed to examine the improvement obtained by the novel approach: One can draw a random unitary $\mathbf{H}$ from the Haar measure (for it to be the unique measure invariant under the action of the unitary group [11]) and compute the exact set of $\boldsymbol{\mathcal{V}}$, $\boldsymbol{\tau}$, $\boldsymbol{\theta}$ analytically. Subsequently, one emulates the experimental situation by perturbing this set of data. Now one performs the new kind of network tomography, introduced above (inner dashed gray box). This is done repeatedly for different sets of $\tilde{\boldsymbol{\mathcal{V}}}$, $\tilde{\boldsymbol{\tau}}$, $\tilde{\boldsymbol{\theta}}$, drawn at random from the respective distributions (e.g. $\tilde{\boldsymbol{\mathcal{V}}} = \boldsymbol{\mathcal{V}} + \delta\boldsymbol{\mathcal{V}}$ and $\tilde{\boldsymbol{\tau}} = \boldsymbol{\tau}(1 + \delta\boldsymbol{\tau})$ with $\delta(\boldsymbol{\mathcal{V}}, \boldsymbol{\tau}) \sim \mathcal{N}(0, \boldsymbol{\sigma}_{(\boldsymbol{\mathcal{V}},\boldsymbol{\tau})})$). To ensure that the results are not just representing a special case, further sampling must be implemented by generating a representative set unitaries at random from the Haar measure and repeating the previous steps (outer dashed gray box).[8] At last the primal unitary $\mathbf{H}$ is compared to the unitaries reconstructed utilizing the different passive approaches $\tilde{\mathbf{U}}_\mu$ and the respective novel/extended approaches $\mathbf{U'}_\mu$.[9]

Some numeric results are presented in section X E.

---

[8] One could again implement additional random sampling for the initial values $\mathbf{p}_0$. However, this has not been done because we want to study the pure improvement accomplished by the new method.

[9] Here one can again make use of different measures to compare the results (as discussed in section VI).

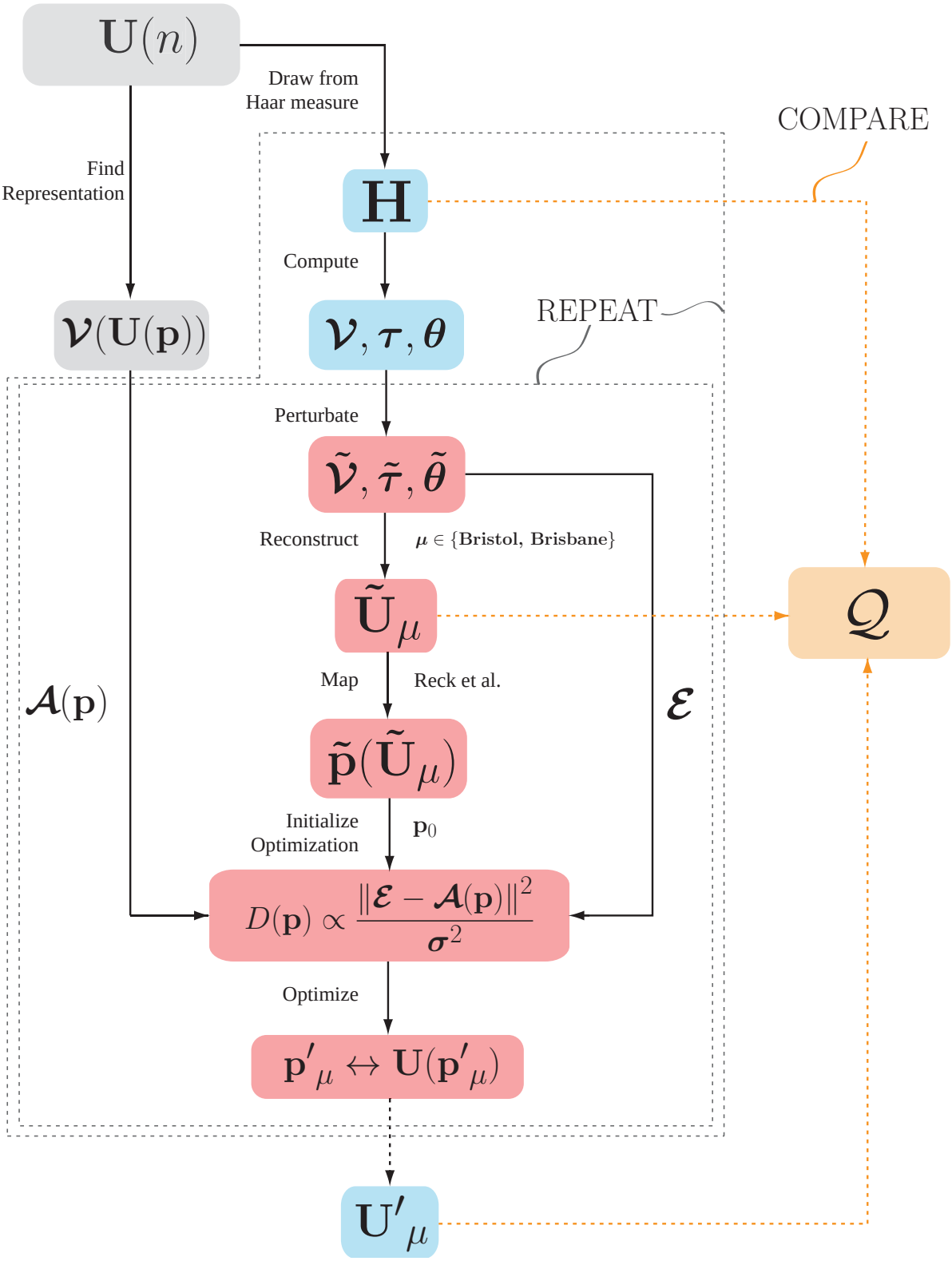


Figure 13: Flowchart representation for the proposed novel approach to network tomography (red) and a numerical approach to confirm the improvement (consider the whole chart).

Results

This chapter shall be devoted to results and will be structured into three main parts. In the first part we investigate the influence of correlations for the data extraction, necessary for the network tomography (see V) and present the extracted data.
Thereafter, in the second part, we present results regarding the network tomography. In particular, results for the proposed back-propagation approach for a given network that had been utilized for a BosonSampling experiment [15]. In the last part we present results for the proposed novel approach to network tomography.

IX. IMPROVING THE DESCRIPTION OF GENERALIZED TWO-PHOTON QUANTUM INTERFERENCE IN LINEAR OPTICAL WAVEGUIDE CIRCUITS

This section investigates the influence of correlations, namely, different interesting cases of $\mathbf{M}$, as well as different – experimentally relevant – parameter choices.
In the first part we will exclude all filtering processes as they are suppressing correlations as shown in III A 3. This enables us to investigate the pure correlation effects on the properties of a bi-photon state. We will find that fully correlated and fully anti-correlated JSA have very different impact on the distinguishability of non perfect photons, although they are carrying the same amount of entanglement.
Following this we will subsequently include filtering processes. As is well known the latter will turn out to have positive effects on the dip-visibility, but only for a relatively high trade-off due to negative effects of introducing filtering in experiments: such as alignment problems, tilt of the central wavelengths, decrease in brightness and so forth. We suggest that, for certain instances, it might even be beneficial to exclude filtering.

Besides other applications [22, 31, 53] it is vital to take the frequency correlations into account to find the most appropriate description for $G^{(2)}$.
In the last part of this section we employ the complete description of $G^{(2)}$, derived in IV A, to extract more reliable values - necessary for the network tomography - from experimental data.

A. Further Investigations on Correlation Measurements Excluding Filtering

Excluding filtering corresponds to setting $N_\mu^{(G)} = 1$ and $\sigma_\mu \to \infty$, i.e. $\mathbf{M}'$ vanishes. Taking account of this and renaming $G^{(2)} \equiv P_{s_1,s_2}(\Delta\tau)$ (82) becomes

$$\begin{aligned} P_{s_1,s_2}(\Delta\tau) \;=\; & |b_e|^2 + |b_o|^2 + \\ & (b_e^* b_o + \text{c.c.}) \cdot \Gamma(\sigma_l, \sigma_k, \theta) \cdot \chi(\Delta\omega, \theta) \cdot \underbrace{\exp\left[-\frac{\Delta\tau^2}{M_{1,1} + M_{2,2} - 2M_{1,2}}\right]}_{:=t(\theta,\Delta\tau)}. \end{aligned} \tag{109}$$

Where we have factorized the parts into three factors and set $\Delta\omega \equiv \omega_s^0 - \omega_i^0$. Each of which will be briefly investigated. All three depend on the angle θ that can be interpreted as correlation angle (see (43)). They do also depend on the major and minor axis of the ellipse, but let us assume them to be fixed for now. We have:

1. The correlation factor

$$\Gamma(\theta) = 2 \cdot \sqrt{\frac{\det(\mathbf{M})}{\det(\mathbf{M} + \boldsymbol{\sigma}_x \mathbf{M} \boldsymbol{\sigma}_x)}} = 2 \cdot \sqrt{\frac{M_{1,1}M_{2,2} - M_{1,2}^2}{(M_{1,1} + M_{2,2})^2 - 4M_{1,2}^2}}. \tag{110}$$

2. The frequency-mismatch

$$\chi(\theta) = \exp\left[-\frac{M_{1,1}M_{2,2} - M_{1,2}^2}{M_{1,1} + M_{2,2} + 2M_{1,2}} \Delta\omega^2\right]. \tag{111}$$

3. And the temporal factor

$$t(\theta) = \exp\left[-\frac{\Delta\tau^2}{M_{1,1} + M_{2,2} - 2M_{1,2}}\right]. \tag{112}$$

We can now investigate the effect of the above factors on the coincidence probability (109).

For this purpose we will first examine the three factors on their own to obtain an intuition for the way they influence the outcome. This will subsequently allow us to trace back certain behaviors to these specific terms.

Since the first two factors are decoupled from the temporal part, for a given frequency mismatch $\Delta\omega = \omega_s^0 - \omega_i^0$, they only affect the visibility of the HOM-Dip.

$$\begin{aligned} \mathcal{V} &= 1 - \frac{P_c(\Delta\tau = 0)}{P_c(\Delta\tau \to \infty)} \\ &= -\frac{b_e^* b_o + b_e b_o^*}{|b_e|^2 + |b_o|^2} \cdot \Gamma(\theta) \cdot \chi(\Delta\omega, \theta). \end{aligned} \tag{113}$$

1. The correlation factor $\Gamma(\theta)$

We start with the "correlation factor". Rewriting $\det(\mathbf{M} + \boldsymbol{\sigma}_x \mathbf{M} \boldsymbol{\sigma}_x) = \det(\mathbf{M} + \mathbf{M}) + (M_{1,1} - M_{2,2})^2$ unveils the bounds of Γ

$$0 < \Gamma(\theta) \leq 1. \tag{114}$$

Mathematically speaking the lower bond could be realized by $\sigma_\mu \to \infty$ with $\mu \in \{k, l\}$. However, since this is a non-physical case it is justified to assume that Γ is strictly greater than zero.

We can give a relatively simply expression for $\Gamma(\theta)$ in the framework of (43), that is

$$\Gamma(\theta) = 2\sqrt{\frac{1}{\left(\frac{\sigma_l^2}{\sigma_k^2} + \frac{\sigma_k^2}{\sigma_l^2}\right)\cos^2 2\theta + 2\left(1 + \sin^2 2\theta\right)}}. \tag{115}$$

For $\sigma_l = \sigma_k$ we get $\Gamma(\theta) = 1$ as expected. However, in the experimentally more realistic regime where $\sigma_l > \sigma_k$ - i.e. our photons get less indistinguishable - we find a very interesting behavior for $\Gamma(\theta)$. From (114) we know that $\Gamma(\theta)$ is bounded above by 1 and from (115) we can see that this bound is indeed realized for $\theta = \pm\frac{\pi}{4}$ (the maximal correlated $(+)$ and anti-correlated $(-)$ case). We would like to emphasis that this behavior is of great interest for PDC sources applications in optical quantum information experiments inasmuch as tuning the angle of correlation θ - which is effectively possible by tuning the pump field, crystal length and phase matching - we can *create photons with better visibility properties from non-optimal sources* (i.e. were $\sigma_l \neq \sigma_k$).

Bare in mind that $\Gamma(\theta)$ is $\frac{\pi}{2}$-periodic in θ, i.e. symmetric for the correlated and anti-correlated case, as can be seen from (115).

We depict the behavior of $\Gamma(\theta)$ for an exemplary value of $\sigma_l/\sigma_k \approx 3$ that corresponds to values, taken from a simulation for the experimentally utilized source ([54])

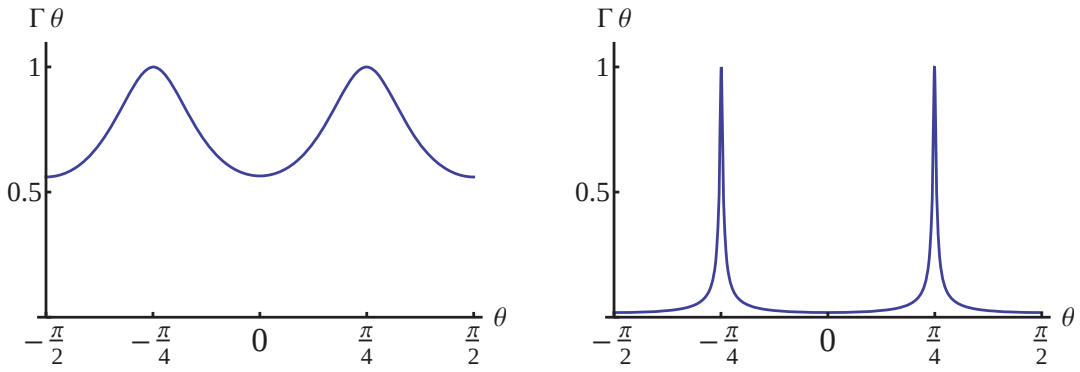


Figure 14: Two exemplary plots for the correlation factor $\Gamma(\theta)$. (a) for a parameter choice of the long axis σ_l and short axis σ_k of the joint spectral amplitude ellipse such that $\sigma_l/\sigma_k \approx 3$. This value was adapted from the experiment. And (b) depics the case of $\sigma_l \gg \sigma_k$ ($\sigma_l/\sigma_k \approx 100$).

Varying the parameter $\frac{\sigma_l}{\sigma_k}$ changes the quantitative behavior of $\Gamma(\theta)$ in the sense

that the peak is getting much sharper and narrower. Physically speaking this corresponds to a high sensibility to correlations, i.e. a non-perfect correlation in frequency space corresponds to a quick decrease in indistinguishability. This is certainly what one expects from the corresponding spectra (7b) where the projection on the ω_s- and ω_i-axis are identical for $\theta = \pm\frac{\pi}{4}$.

2. The frequency mismatch factor χ

We shall now proceed with the second factor, the "frequency mismatch factor".

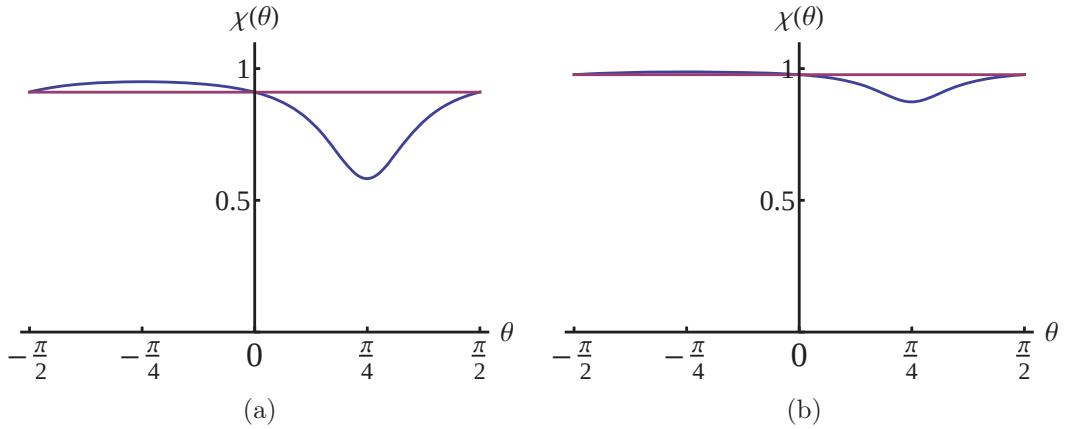


Figure 15: Two exemplary cases for the mismatch factor $\chi(\theta)$ for different ratios of the long and short axis of the joint spectral amplitude ellipse – σ_l and σ_k respectively – and different frequency mismatch $\Delta\omega$. (a) for $\sigma_l/\sigma_k \approx 3$, with $\Delta\omega \approx \sigma_k$. And (b) for the same ratio, but with $\Delta\omega \approx \frac{1}{2}\sigma_k$.

Plugging in our parametrization (43) for $\mathbf{M}$ we can rewrite the argument of the exponential in $\chi(\Delta\omega)$ as

$$\frac{-\ln\chi}{\Delta\omega^2} = \frac{1}{2}\left[\sigma_l^2\sin^2(\frac{\pi}{4}-\theta) + \sigma_k^2\sin^2(\frac{\pi}{4}+\theta)\right]^{-1}. \tag{116}$$

By means of (116) it is possible to investigate how the correlations entering in χ can affect the the coincidence probability (109).

We note that $-\ln\chi > 0$ for finite σ_μ and $\Delta\omega$. But also $-\ln\chi \leq 1$ as long as $\frac{\Delta\omega^2}{2\sigma_k^2} \leq 1$. Consequently, as long as the frequency mismatch is kept smaller than the one sigma interval of the minor axis, the mismatch causes a reduction in visibility by a factor that is greater than $1/\sqrt{e}$.

Studying the whole range of θ, assuming a significant difference in σ_l and σ_k, (116) reveals an asymmetric behavior:

If

$$0 < \theta < \frac{\pi}{2} \;\Rightarrow\; \frac{-\ln\chi}{\Delta\omega^2} > \left[\sigma_l^2 + \sigma_k^2\right]^{-1}$$

and the maximum is obtained for the maximally correlated case of $\theta = \frac{\pi}{4}$.

If

$$-\frac{\pi}{2} < \theta < 0 \;\Rightarrow\; \frac{-\ln\chi}{\Delta\omega^2} < \left[\sigma_l^2 + \sigma_k^2\right]^{-1},$$

obtaining the minimum for the maximally anti-correlated case of $\theta = -\frac{\pi}{4}$. However, the slope is less steep and the absolute distance at $\theta = \frac{\pi}{4}$ from the complete uncorrelated case ($\theta = 0$) is smaller than that at $\theta = -\frac{\pi}{4}$. Fig. 15b depicts two cases of the behavior characteristic for $\chi(\theta)$. Let us stress, that the above analysis reveals that the *anti-correlated case effectively suppresses frequency-mismatch*.

3. The temporal factor t

Studying the temporal factor will give indications about the temporal information encoded in the PDC-state.

In contrast to $\Delta\omega$ which is a non dynamical mismatch arising from experimental imperfections, $\Delta\tau$ is a tunable dynamic parameter in the experiment. Here $\Delta\tau/\Delta\omega$ should not be confused with the temporal/frequency uncertainties which correspond to $\sigma_{k/l}$. However, it is still worth noting that $-\frac{\ln t}{\Delta\tau^2} = 2 \cdot \sigma_l^2\sigma_k^2 \cdot \frac{-\ln\chi}{\Delta\omega^2}$, i.e. we can "map" a certain frequency mismatch to a temporal mismatch.

Whereas the previous two factors were only influencing the absolute value of the visibility - as they are constant in the experiment w.r.t. $\Delta\tau$ - the temporal factor

is determining the dynamical behavior w.r.t. $\Delta\tau$. Therefore we shall focus on the full half width τ_{FWHM} of (109) that is characteristic for these dynamics. By definition τ_{FWHM} can be extracted from the equation $t(\theta, \tau_{\text{FWHM}}) = \frac{1}{2}t(\theta, 0) = \frac{1}{2}$, which yields

$$\tau_{\text{FWHM}} = \sqrt{\ln 2} \cdot \frac{\sqrt{\sigma_k^2 \sin^2(\frac{\pi}{4} - \theta) + \sigma_l^2 \sin^2(\frac{\pi}{4} + \theta)}}{\sigma_l \sigma_k}. \tag{117}$$

Extracting the width of the dip from (117) highlights again how vital it is to take frequency correlations in the JSA into account.

Fig. 16 depicts the crucial influence of the spectral correlations in just one picture. We find that correlations for which $0 < \theta < \frac{\pi}{2}$ the width of the $G^{(2)}$-function (HOM-Dip) becomes broader, whereas it becomes narrower for $-\frac{\pi}{2} < \theta < 0$. This very interesting behavior has also been described in [55] for states with no frequency mismatch whilst preparing this script. However, it should be stressed that it is not just an interesting physical behavior, but also of great importance for extracting visibilities from experimental HOM-Dip data and thus has to be taken into account to get reliable results.

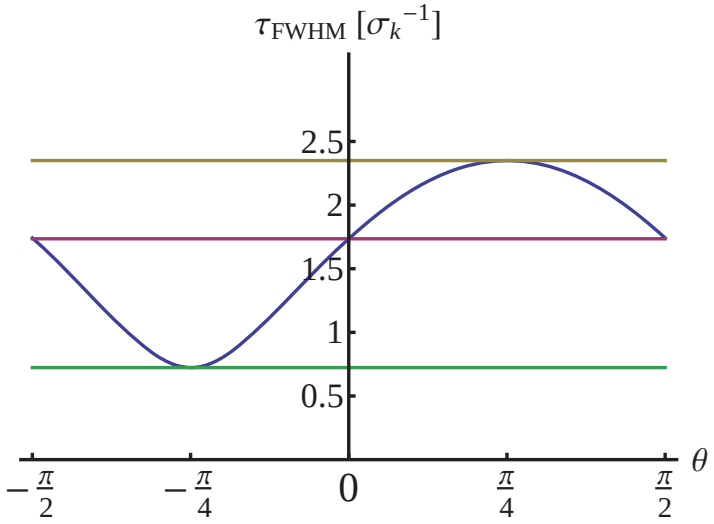
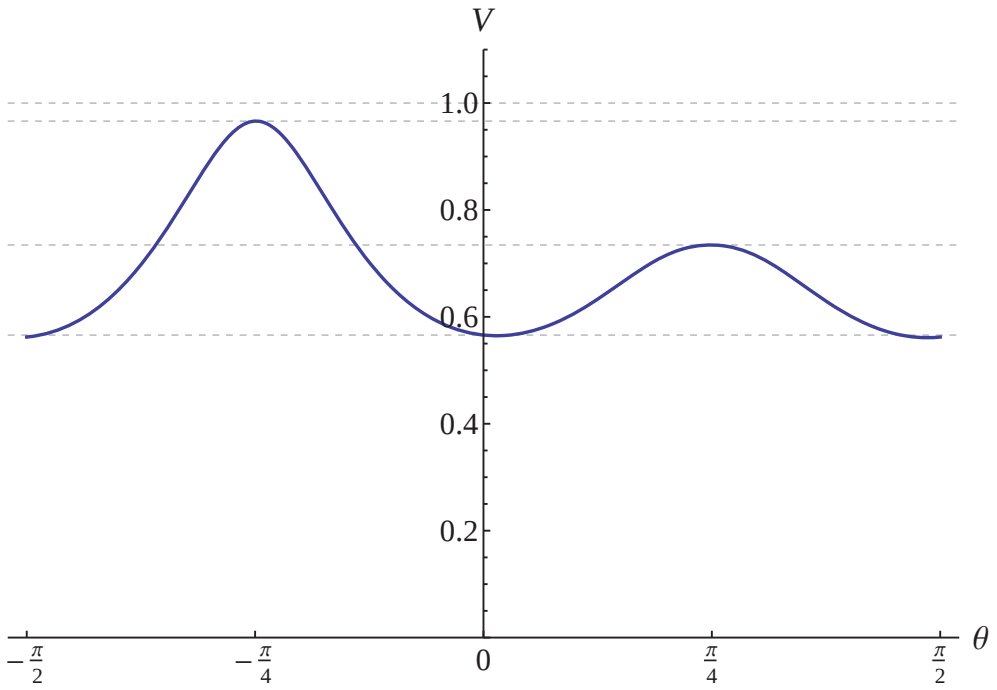
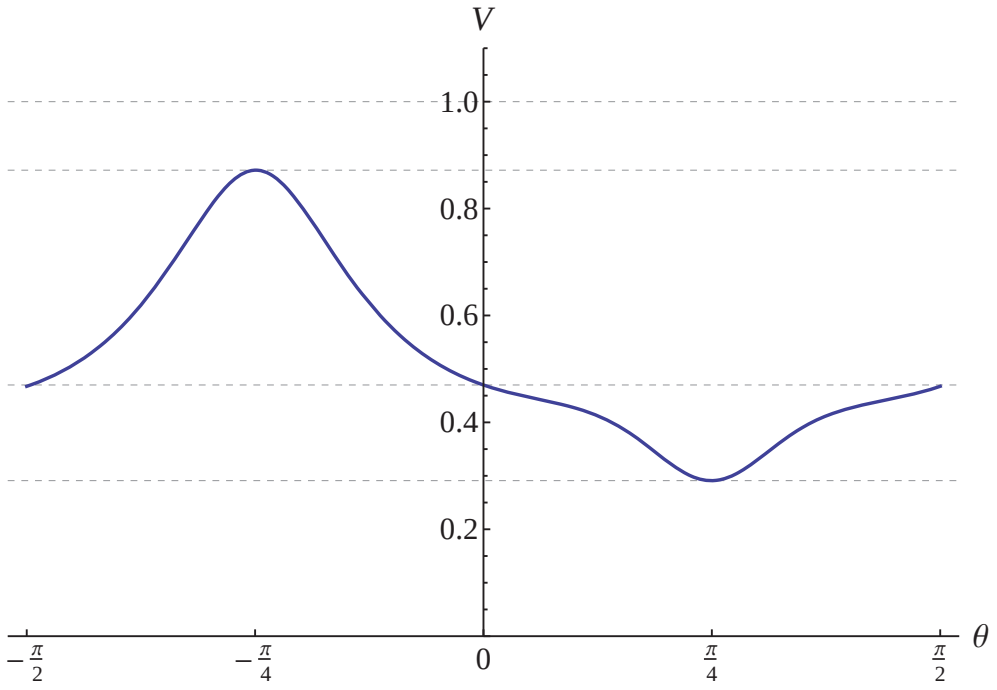


Figure 16: Width of the coincidence probability (109) against the correlation parameter θ. For the correlated ($\theta > 0$) case the dip is broadening, while for the anti-correlated ($\theta < 0$) it is narrowing. Note that the function is asymmetric w.r.t. $\theta = 0$. Here we have chosen $\sigma_l/\sigma_k \approx 3$.

B. Impact on the Visibilities



(a)



(b)

Figure 17: Exemplary behavior of the Visibilities w.r.t. θ. Here for a 50 : 50 BS. The parameters describing the JSA are taken from experimental values (see Tbl. V). The major axis of the joint spectral intensity corresponds to a value of about 11 nm and the minor axis is about a third of the major axis. In (a) the frequency mismatch corresponds to a mismatch in wavelength of about 1.7 nm. In (b) it is twice as large.

The modification of the visibility due to frequency correlations is the key point for the network tomography and distinguishability of the photons. Fig. 17 depicts how (110) and (111) alter the measured visibilities (113).
In fact that *anti-correlated state results in a higher dip-visibility* than the correlated state. This is due to the asymmetry in $\chi(\theta)$ that had been discussed in IX A 2. Furthermore *the anti-correlated state seems to be more robust against frequency mismatch* which is attributed to the less steep slope of $\chi(\theta)$ in the anti-correlated regime. Compare the change in the values at the local Extrema in in the two regimes $\theta \gtrless 0$ in Fig. 17a and Fig. 17b. If $\Delta\omega$ overcomes a certain threshold the completely correlated case ($\theta = \frac{\pi}{4}$) is even worse than the uncorrelated case.

C. Consequence of Neglecting Correlations

If correlations are neglected the JSA factorizes into the product of the marginals. The marginal corresponds to the partially traced out system of either idler or signal. Or in a more picturesque way, the marginal $\beta_\mu(\omega_\mu)$, with $\mu \in \{s, i\}$ it is the projection of the JSA on the ω_μ-axis: $\beta_\mu(\omega_\mu) = \int \mathrm{d}\omega_\nu \, \beta(\omega_\mu, \omega_\nu)$. Hence

$$\beta(\omega_s, \omega_i) \to \beta_s(\omega_s) \cdot \beta_i(\omega_i). \tag{118}$$

And if we identify $s \leftrightarrow 1$ and $i \leftrightarrow 2$ the marginal reads

$$\beta_\mu(\omega_\mu) = \frac{1}{M_{1,1}\sqrt{2\pi}} \exp[-\frac{1}{2} M_{1,1} \cdot (\omega_\mu - \omega_\mu^0)^2]. \tag{119}$$

Resulting in

$$\beta(\omega_s, \omega_i) \to \frac{1}{2\pi M_{1,1} M_{2,2}} \exp[-\frac{1}{2} \left(M_{1,1} \cdot (\omega_s - \omega_s^0)^2 + M_{2,2} \cdot (\omega_i - \omega_i^0)^2\right)]. \tag{120}$$

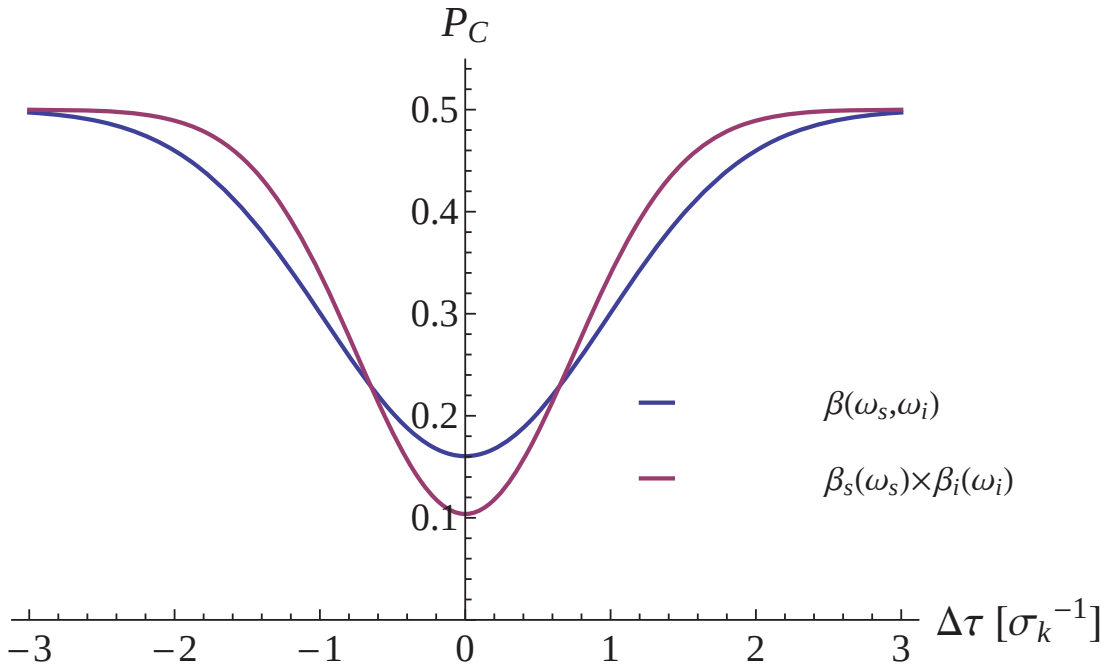
(Depending on the notation, one might need to change $\mathbf{M} \to \gamma\mathbf{M}$. This is because

we project β and not $|\beta|^2$ in our notation. If we were projecting $|\beta|^2$, γ would take a value of two.) Carrying out the coincidence probability yields

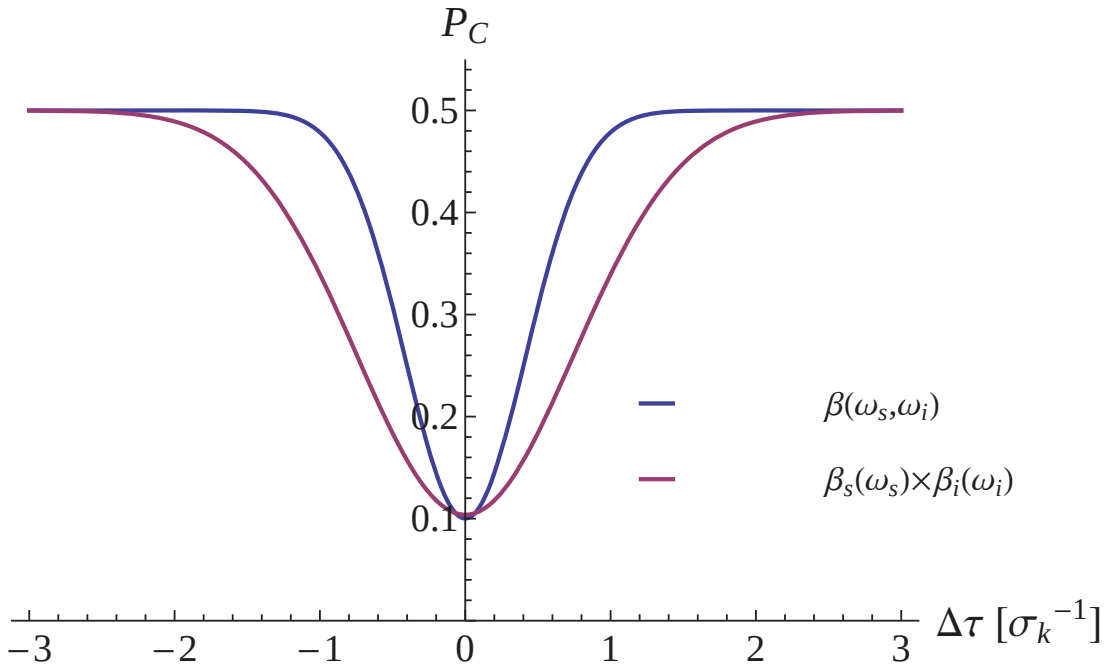
$$P_{1,2}(\Delta\tau) = |b_e|^2 + |b_o|^2 + (b_e^* b_o + \text{c.c.}) \frac{\sigma_1\sigma_2}{\sigma_1^2+\sigma_2^2} \exp\left[-\frac{\Delta\omega^2}{2(\sigma_1^2+\sigma_2^2)}\right] \cdot \exp\left[-2\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}\Delta\tau^2\right], \tag{121}$$

were we have defined $2\sigma_i^2 = M_{i,i}^{-1}$. From (43) we infer that the correlated and anti-correlated case behave identical in contrast to (121). This is conditioned by the fact that all appearing parts are symmetric under exchange of σ_1 and σ_2 and $M_{1,1}(\pi-\theta) = M_{2,2}(\theta)$. I.e. changing from the correlated regime ($\theta > 0$) to the corresponding anti-correlated regime is equivalent to exchanging σ_1 with σ_2, but since (121) is invariant under such an exchange, $P_{1,2}(\Delta\tau)$ takes the same form in the correlated regime ($\theta > 0$) as in the corresponding anti-correlated regime ($\pi - \theta$).

Fig. 18 compares the behavior for values taken from Tbl. V and the corresponding anti-correlated case.



(a)



(b)

Figure 18: Comparison of the coincidence probability of a 50 : 50 BS, where values from Tbl. V had been considered. In (18a) we depict the HOM-Dip for a correlated JSA with $\theta = +60°$. (18b) depicts the anti-correlated case in which $\theta = -60°$. In both cases, the blue line corresponds to the case in which the correlations are taken into account ($\beta(\omega_s, \omega_i)$), whereas the red line corresponds to a JSA that is set together as a product of its marginal distributions ($\beta_s(\omega_s) \cdot \beta_i(\omega_i)$) and therefore ignoring the correlations.

Indeed this is giving us a vital feedback for the visibility extraction for the reconstruction. The visibilities are differing decisively. Though it should be noted that we have not considered filtering up to now. However, we expect from the Schmidt decomposition that the correlations will still affect the outcome, although the effect might differ. But bare in mind that this is a very interesting phenomenon on its own and therefore investigated here.

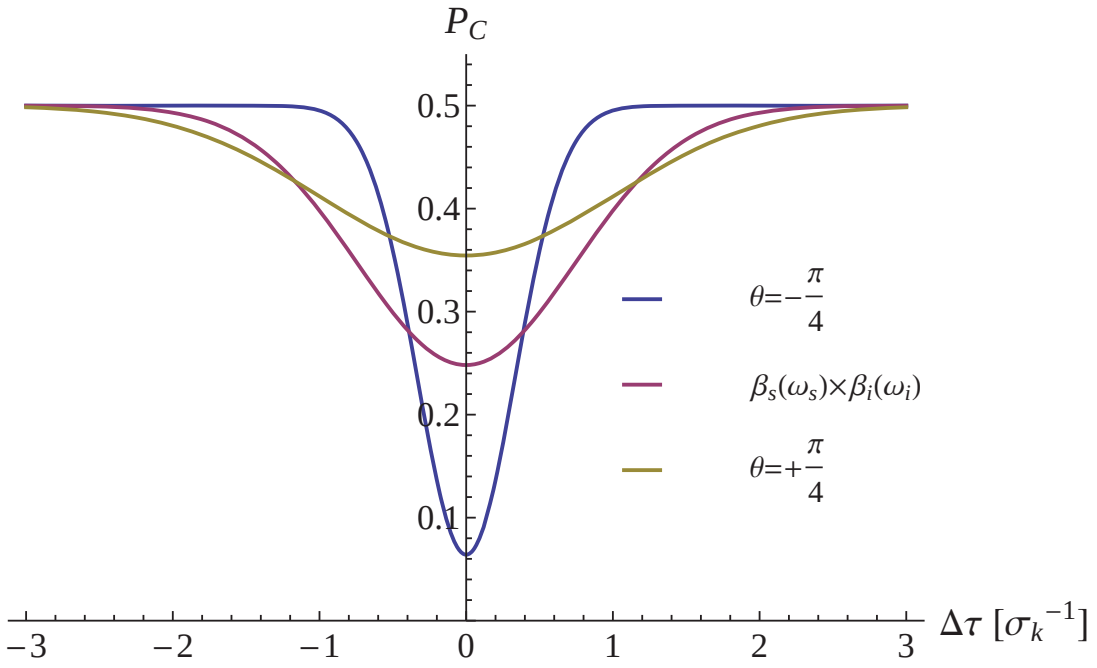


Figure 19: Another exemplary comparison of the coincidence probability of a $50:50$ BS where we have considered the same values for the minor and major axis of the JSA as we did in the previous figure (Fig. 18). The frequency mismatch is twice as large and the correlation angles can be drawn from the figure.

D. Including Filtering Processes

Filters are present in most experimental setups and most generally speaking they lead to higher indistinguishability. For instance in the BosonSampling setup they are employed to suppress additional distinguishability from longitudinal walk-off [15, 56–58]. The idea is that narrow-band filtering tailors the coherence time of the photons τ_c such that $\tau_c > \delta T$, where δT is the temporal longitudinal walk-off of the photons after they leave the BBO crystal [59, 60]. However, they also suppress filtering [39] as can be inferred from Fig. 20.

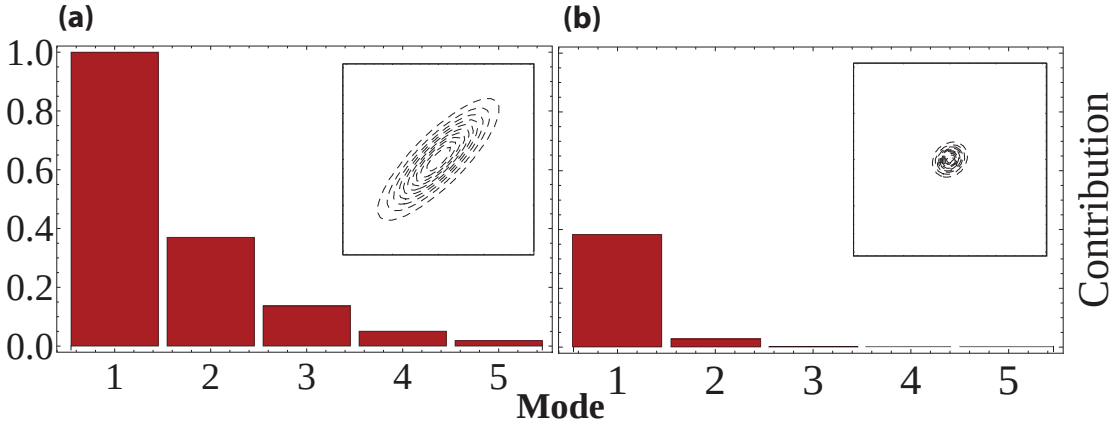


Figure 20: Schmidt modes of a JSA that corresponds to the experimental values from Tbl VII and VIII. (a) before filtering and (b) after filtering.

Obviously, well aligned filters provide a positive effect in terms of indistinguishability, but this accomplishment is coinciding with a trade-off in brightness and correlations. (The previous paragraph suggests that this is not necessary in every situation. One could think of tuning the correlations rather than including filters to accomplish better indistinguishability with the advantage of higher brightness and one potential source of errors less.)

We will not repeat the previous investigations again. Instead, we conclude from the behavior of the Schmidt modes (Fig. 20) that filtering strongly suppresses the influence of correlations, but does not completely eliminate them. Let us restrict

our investigations to the specific case from Tbl. VIII and VII. See also XIII for details.

The functional interaction of the different factors in (84) add up to a more complex behavior than previously. One difference is a symmetry-breaking within the correlated and anti-correlated regimes. Another difference is that the anti-correlated case does no longer generally improve visibilities by suppressing frequency mismatch (Fig. 21). This is caused by the fact that $\omega_s^F \neq \omega_s^0$ and $\omega_i^F \neq \omega_i^0$ or $\omega_s^F \neq \omega_i^F$ which causes additional change in the effective frequency mismatch after the filter. I.e. unlike before, where we could assume $\Delta\omega = \omega_s^0 - \omega_i^0$ to be fixed and rotate the correlation angle θ (IX A), we do now effectively alter the centroid frequencies of the ellipse after filtering, depending on the correlation angle θ before filtering. (A code allowing to vary the parameters in (84) can be found in [61])

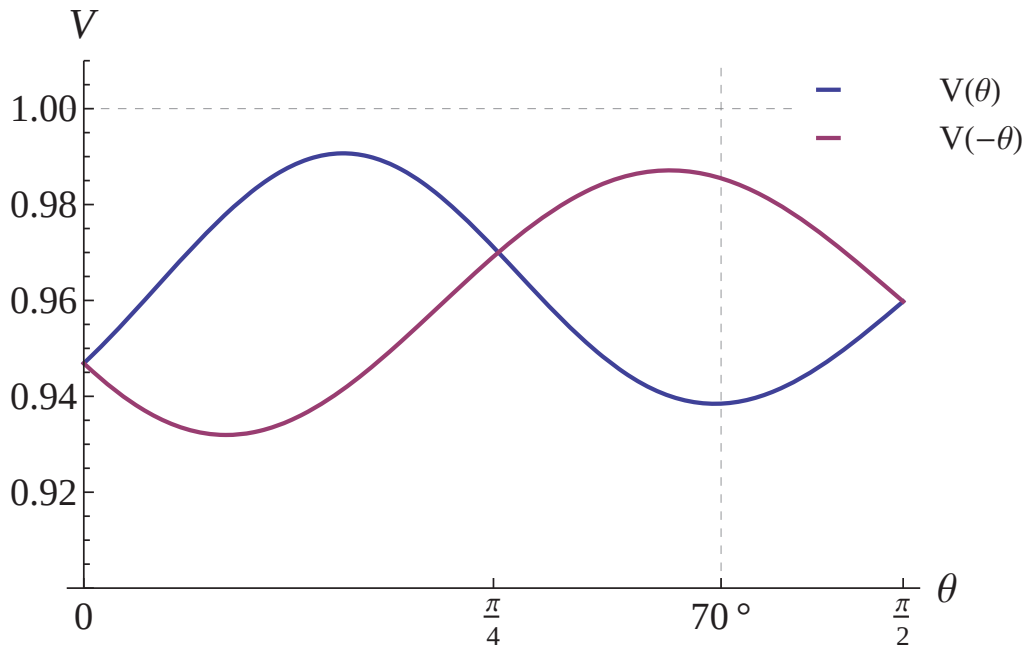


Figure 21: The visibility as a function of the correlation angle. All parameters, except from θ are fixed according to Tbl. VII and VIII. The asymmetry of the centroid filter frequency causes the asymmetric behavior of the visibility and it is also responsible for the fact that $\mathcal{V}(\theta)$ takes smaller values, as well as larger values than $\mathcal{V}(-\theta)$, unlike in the previous section.

With this in mind let us now move on to investigate the case from Tbl. VIII and VII. Fig. 21 depicts this case for all possible angles, including the one where $\theta = 70^\circ$ that had probably been used to carry out the reconstruction for [15], as we describe in XIII.

However, the original motivation to study all these different influences was to find an improved description for the extraction of visibilities. After the significance of such a more complete description has now been discussed, it should now be applied in the subsequent section.

E. Visibility Extraction from Experimental Data

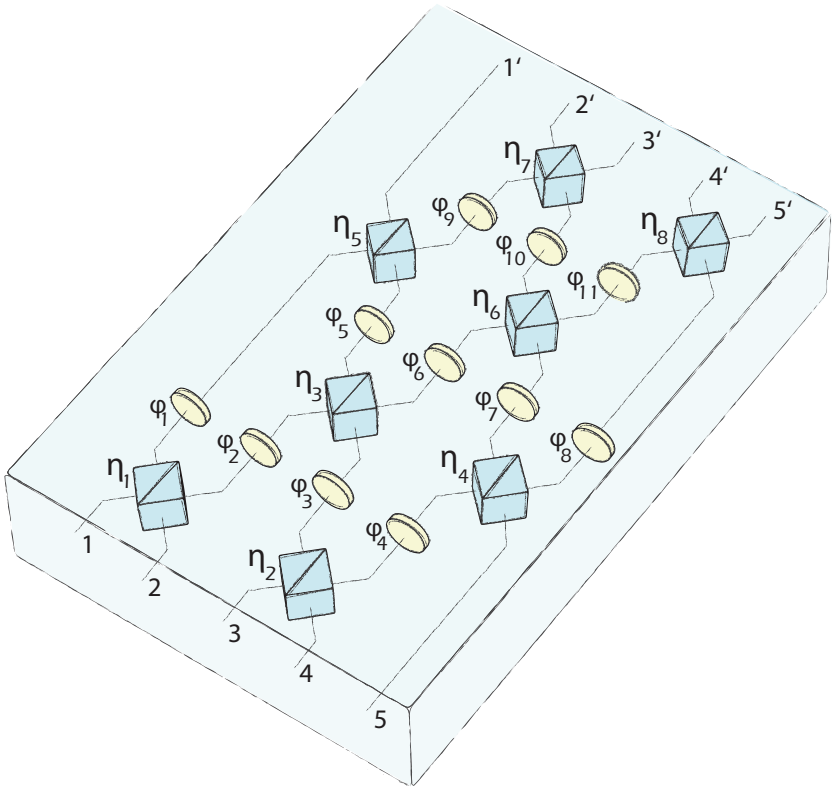


Figure 22: 5×5-Network, utilized to carry out the experimental BosonSampling in [15]. Tbl. I lists the results of visibilities, characterizing this network. Figure taken from [26].

The procedure introduced in IV B 1 is employed to extract the visibilities. The results obtained via. (for 87) are listed in Tbl. I and Tbl. II contains the corresponding errors. Extracted values for the case in which the correlations are neglected are listed in Tbl. IX and X. (Carried out with the marginal coincidence count rate (121)) The mean deviation between the two data sets is 0.0265. The data is collected for a 5×5 unitary, that corresponds to the network, depicted in Fig. 22. The routines and experimental data are provided in [62].

Table I: Extracted Visibilities for the 5×5-network from Fig. 22. The extraction had been carried out as described in IX E.

(t_1, t_2) \ (s_1, s_2)	$(4,5)$	$(3,5)$	$(2,5)$	$(1,5)$	$(3,4)$	$(2,4)$	$(1,4)$	$(2,3)$	$(1,3)$	$(1,2)$
$(1,2)$	-1.01714	-0.268516	0.594285	0.0865994	-0.254499	0.609374	0.0936144	-0.0677409	0.299734	0.212145
$(1,3)$	-0.676692	0.212803	0.240638	0.345146	0.490329	0.665138	0.354358	-0.470308	-0.314639	0.347222
$(1,4)$	0.582587	0.564828	-0.416501	-0.823415	-0.956158	0.242011	0.445519	0.255417	0.411079	-0.0275681
$(1,5)$	0.817264	-0.296532	0.441006	0	0.18503	-0.377536	0	0.681214	0	0
$(2,3)$	-0.681008	0.163607	-0.0161495	0.0803231	0.12563	-0.195876	0.259363	-0.749796	0.261005	0.759891
$(2,4)$	0.58342	0.552903	0.156447	-0.116548	-0.238325	-0.330316	0.527473	0.620812	-0.303252	0.439315
$(2,5)$	0.821944	-0.382812	-0.0838709	0	0.191731	0.0937749	0	0.407255	0	0
$(3,4)$	0.900126	-0.5933	0.254115	0.612399	0.500349	-0.204998	-0.522286	0.658694	1.00188	-0.608981
$(3,5)$	0.466946	0.850836	-0.882292	0	-0.633199	0.658326	0	0.856185	0	0
$(4,5)$	-0.396626	0.845285	0.388788	0	0.233151	0.0896736	0	-0.627071	0	0

Table II: The Errors, associated with Tbl. I.

(t_1, t_2) \ (s_1, s_2)	$(4,5)$	$(3,5)$	$(2,5)$	$(1,5)$	$(3,4)$	$(2,4)$	$(1,4)$	$(2,3)$	$(1,3)$	$(1,2)$
$(1,2)$	0.0291723	0.0101141	0.00903224	0.00653225	0.00822143	0.0109054	0.00468338	0.00442748	0.00395376	0.00558262
$(1,3)$	0.008063	0.00504786	0.00636622	0.00626547	0.00752015	0.00979642	0.00733547	0.00977201	0.00843022	0.0126068
$(1,4)$	0.0111414	0.00705219	0.00947389	0.0145284	0.0183785	0.00568282	0.0131073	0.00461	0.0061197	0.0064363
$(1,5)$	0.00734576	0.00505715	0.00764286	0	0.00410907	0.00621354	0	0.00787464	0	0
$(2,3)$	0.0155012	0.00470242	0.00300846	0.00355508	0.0122164	0.00614345	0.0040153	0.0125312	0.00508491	0.00806716
$(2,4)$	0.0176236	0.0127206	0.00843719	0.00667298	0.0144222	0.0069789	0.0102752	0.0057799	0.00577611	0.00539616
$(2,5)$	0.0103131	0.010123	0.00676895	0	0.0054029	0.00392052	0	0.00984262	0	0
$(3,4)$	0.00914779	0.00640217	0.00357984	0.0050244	0.00856178	0.0047252	0.0106177	0.00613756	0.00844412	0.00793212
$(3,5)$	0.00314808	0.00537998	0.0124291	0	0.00955853	0.00615956	0	0.0121062	0	0
$(4,5)$	0.00742234	0.00948506	0.00603427	0	0.00461111	0.00364311	0	0.00803362	0	0

X. NETWORK TOMOGRAPHY FOR PASSIVE LINEAR OPTICAL NETWORKS

This section shall be devoted to the different approaches to network tomography. First we present the results for the backpropagation of errors - introduced in VII - for the obtained experimental datasets from IX E with the method introduced in V A. Additionally we will impose a self-consistency check for how good the recovered network predicts the results in IX E. This will be done for the new dataset containing the correlations, as well as for the old one that had not taken them into account.
Thereafter we perform analytic and numeric error propagation for V B and V C which enables us to compare these three methods subsequently.
In the last section the results for a new approach to network tomography following VIII will be presented.

A. Backpropagation of Errors

In VII we have proposed a Monte-Carlo-like approach to error propagation whose results are now presented.
Let us first present the resulting raw data for the 5×5-network (depicted in Fig. 22). Fig. 23 depicts the respective beam-splitting and phase-shifting ratios, resulting from a Monte-Carlo like backpropagation for the network tomography described in V A. We have employed visibilities with a Gaussian error of variance according to Tbl. II and generated 10000 samples for each of which a backpropagation had been carried out. For each sampling run we have run the tomography for ten different random starting points. These starting points were uniformly picked from the intervall $[\mathbf{p}_0 - 0.1, \mathbf{p}_0 + 0.1]$. Subsequently the run that minimized the process had been selected and contributes as one of the 10000 data points. The code can be found in [63].

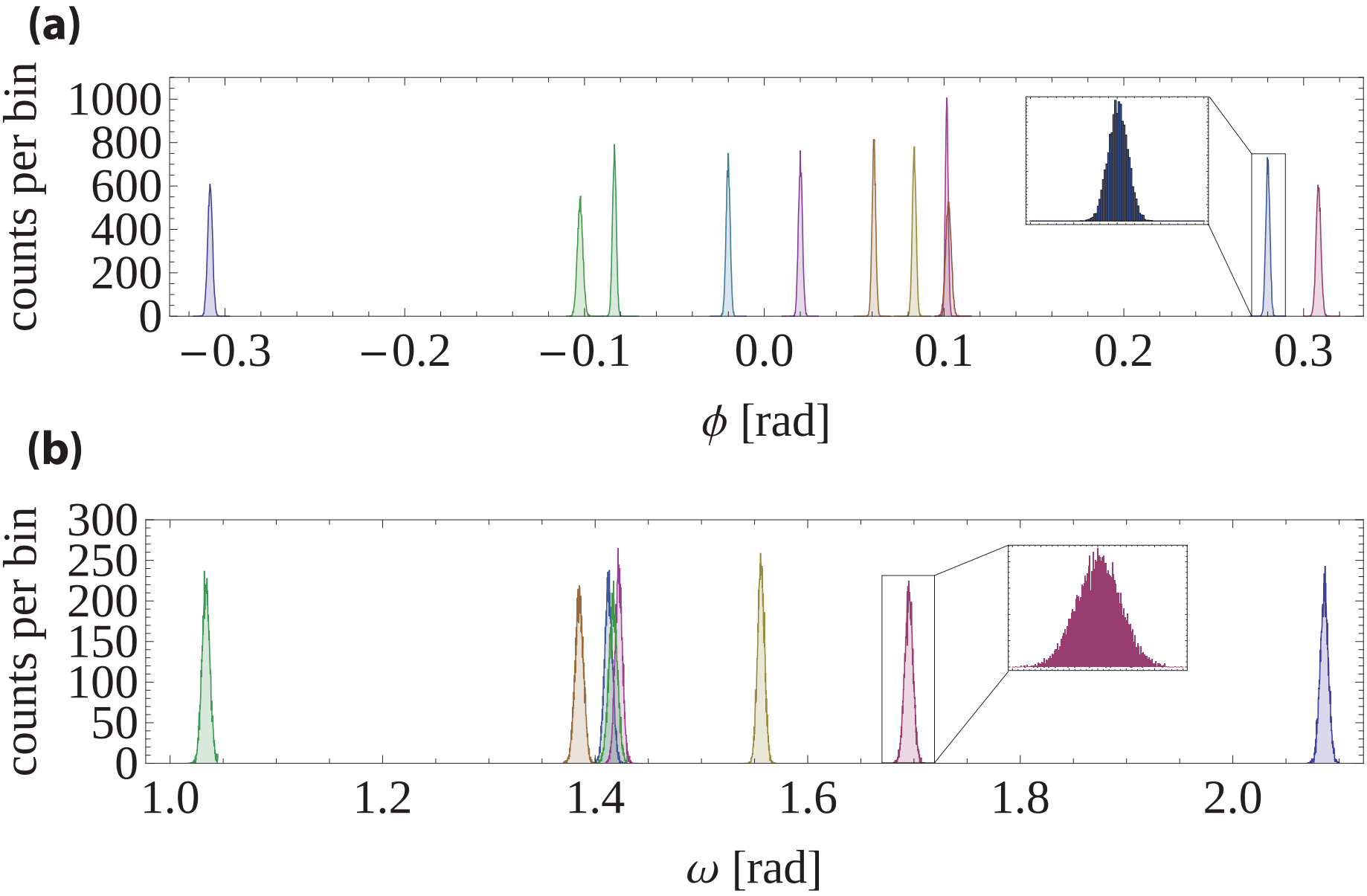


Figure 23: Parameter distributions, resulting from a Monte-Carlo like backpropagation of a Gaussian error with variances according to Tbl. II. The histograms in (a) depict the eleven phase shifts ϕ for the network from Fig. 22. And the histograms (b) depict the eigth beam splitting ratios ω. For both histograms we have binned values that fall into an interval of 0.0002 rad.

Table III: Expectation values and standard deviation for the 19 parameters characterizing the network that had been used to carry out the experiment resulting from the backpropagation.

	Expectation Value
ω_1	2.0863±0.0039
ω_2	1.6953 ± 0.0039
ω_3	1.5562 ± 0.0032
ω_4	1.0339 ± 0.0037
ω_5	1.4125 ± 0.0036
ω_6	1.4220 ± 0.0034
ω_7	1.3849 ± 0.0039
ω_8	1.4167 ± 0.0042
φ_1	-0.3082 ± 0.0014
φ_2	0.3082 ± 0.0014
φ_3	0.0834 ± 0.0010
φ_4	-0.0834 ± 0.0010
φ_5	0.2801 ± 0.0012
φ_6	0.1015 ± 0.0008
φ_7	0.0610 ± 0.0010
φ_8	-0.1025 ± 0.0015
φ_9	-0.0201 ± 0.0011
φ_{10}	0.0201 ± 0.0011
φ_{11}	0.1025 ± 0.0015

Calculating the mean and standard deviation of the dataset results in the parameters listed in Tbl. III. Where the error corresponds to the unbiased standard deviation estimator of the dataset produced by the backpropagation (Fig. 23)
For each sampling run we carry out a reconstruction of the unitary, that is composed of the basic building blocks (22) that depend on the reconstructed parameters according to the decomposition in XVII. This yields the following estimator

for the matrix $\mathbf{U}_{\mathrm{V}}$ including error bounds $\Delta\mathbf{U}_{\mathrm{V}}$.

$$\mathrm{Re}\,\mathbf{U}_{\mathrm{V}} = \begin{pmatrix} 0.0333 & 0.0703 & -0.2817 & 0.1214 & 0 \\ 0.0095 & -0.3869 & -0.1386 & -0.7828 & 0.0167 \\ -0.7709 & -0.2951 & -0.2707 & 0.0233 & -0.2474 \\ 0.1473 & -0.1527 & -0.1408 & -0.1330 & 0.0116 \\ 0.2244 & 0.0735 & -0.7887 & 0.0875 & 0.4164 \end{pmatrix}, \tag{122}$$

$$\mathrm{Im}\,\mathbf{U}_{\mathrm{V}} = \begin{pmatrix} -0.3379 & 0.8200 & -0.1075 & -0.3192 & 0 \\ 0.2751 & 0.1809 & 0.2078 & -0.1558 & -0.2052 \\ -0.2410 & -0.0047 & -0.0191 & 0.3532 & -0.0201 \\ -0.2620 & -0.0859 & 0.3597 & -0.0092 & 0.8432 \\ 0.1262 & -0.1308 & -0.0304 & 0.3093 & -0.1072 \end{pmatrix}$$

$$\mathrm{Re}\,\Delta\mathbf{U}_{\mathrm{V}} = \begin{pmatrix} 0.0012 & 0.0015 & 0.0009 & 0.0007 & 0 \\ 0.0010 & 0.0014 & 0.0011 & 0.0007 & 0.0003 \\ 0.0009 & 0.0012 & 0.0010 & 0.0013 & 0.0009 \\ 0.0007 & 0.0006 & 0.0016 & 0.0008 & 0.0006 \\ 0.0008 & 0.0004 & 0.0007 & 0.0011 & 0.0010 \end{pmatrix}. \tag{123}$$

$$\mathrm{Im}\,\Delta\mathbf{U}_{\mathrm{V}} = \begin{pmatrix} 0.0013 & 0.0009 & 0.0006 & 0.0011 & 0 \\ 0.0010 & 0.0009 & 0.0010 & 0.0006 & 0.0007 \\ 0.0012 & 0.0009 & 0.0007 & 0.0010 & 0.0004 \\ 0.0008 & 0.0004 & 0.0010 & 0.0007 & 0.0006 \\ 0.0007 & 0.0007 & 0.0008 & 0.0009 & 0.0015 \end{pmatrix}. \tag{124}$$

Given (122) the self-consistincy check can be performed and compared to the set of visibilities from (148). The self-consistency check is performed by extracting the set of visibilities $\mathcal{V}_{\mathrm{P}}$ from $\mathbf{U}_{\mathrm{NBS}}$ via (85) and comparing there deviation by means of the quality witness defined in (106) to the set of visibilities extracted $\mathcal{V}_{\mathrm{E}}$

in IX E. This procedure yields the following value for the self-consistency witness

$$\mathcal{Q}_0 = 0.0145. \tag{125}$$

The relative deviation $\mathcal{Q}_{\mathrm{R}}$, i.e. the $\mathcal{Q}$-value as fraction of the averaged absolute experimentally extracted and predicted visibilities, $\boldsymbol{\mathcal{V}}_{\mathrm{E}}$ and $\boldsymbol{\mathcal{V}}_{\mathrm{P}}$ respectively, yields correspondingly

$$\mathcal{Q}_{\mathrm{R}} = \frac{\mathcal{Q}_0}{\left\| \boldsymbol{\mathcal{V}}_{\mathrm{E}} + \boldsymbol{\mathcal{V}}_{\mathrm{P}} \right\| /2} = 0.0327. \tag{126}$$

Meaning that the deviation is about 3% which should be compared to the respective value that results from the old extraction procedure $\mathcal{Q}^* = 0.0756$ and correspondingly $\mathcal{Q}_R^* = 0.1743$ which is about 17%. Thus we conclude that we have indeed improved the extraction procedure by taking into account correlations. Note also that the value of $\mathcal{Q}_{\mathrm{R}}$ is very close to the average standard deviation of the Visibilities (0.0184, i.e. around 2%) which provides further evidence that the obtained data (Tbl. I as well as (122)) is indeed trustworthy and much more reliable than the previously used data.

The self-consistency witness had been collected for each sampling run.

The respective values in each run yield the following minimal and maximal values for $\mathcal{Q}$

$$\mathcal{Q}_{\mathrm{min}} = 0.0145 \quad \text{and} \quad \mathcal{Q}_{\mathrm{max}} = 0.0159. \tag{127}$$

It is worth focussing the attention to the errors of the reconstructed parameters in Tbl. III and those for the unitary $\Delta\mathbf{U}_{\mathrm{V}}$. First of all we note that the resulting standard deviations for the phase shifter φ_μ seem to be smaller than those for the beam splitting ratios ω_μ. Indeed, the beam splitting ratios are larger than the phase shifting ratios and hence the relative error is not smaller. However, from Fig. 23 we know that the distributions are rather Gaussian and we should therefore not consider the relative error. Our interpretation of this behavior is that the phase

shifting ratios in (89) do only enter as a sum in the cosine and thus, when small error are imposed on them, they have a smaller impact on the visibilities than the beam splitting ratios. A handwavy interpretation for this behavior might be that some of the phases can be seen as embedded in a Mach-Zehnder-Interferometer which can again be represented by an efficient beam splitter [30]. Thus the error might be propagated back to such a beam splitting ratio rather than the original phase itself. Further investigations should be made and alternative interpretations are possible.

However, carrying out a forward propagation for the set of visibilities with the errors found for the characterizing parameters - p_μ, where $p \in \{\omega, \varphi\}$ - and for the entries of the unitary we find that they are in good agreement with the errors of the initial data (Tbl. II) from which had been sampled to perform the backpropagation. This provides further evidence for the stability of the algorithm. The average error for the visibilities obtained from (122) is $\bar{\epsilon}_{\mathrm{BP}} = 0.0053$, whereas the average error in Tbl. II yields $\bar{\epsilon}_{\mathrm{E}} = 0.0081$. Note that the sum of these two values are roughly $\mathcal{Q}$.

The explanation we propose for this behavior has been the starting point for the improved approach for reconstructing the network's unitary via tomography methods that we have suggested in VIII. We propose that the significant drop in standard deviations from the inital set of data to the set of parameters p_μ (the beam splitting and phase shifting angles) that remain after the backpropagation is caused by forcing unitary constraints throughout the tomography process. As described in V A this is accomplished because the structure of the network is known and by applying this structure in (88) the optimization routine we efficiently impose additional constraints, i.e. we reduce the degrees of freedom in our system. This leads to the reduced standard deviation of $\mathbf{p}$. Similar to the inverse behavior were we wish to calculate the estimator of the standard deviation s for n data points under ν additional constraints for the data points: $s = \frac{1}{\sqrt{n-\nu}} \sqrt{\sum_i (x_i - \mu)^2}$

. Therefore we conclude that applying unitary constraints throughout the network tomography might be beneficial.

B. Forwardpropagation of Errors

In this section the susceptibility of the network tomography introduced in V B and V C shall be investigated. This will be done in two steps. First we give some analytical results obtained in the manner of a stability analysis, i.e. the error is propagated forward. Then we follow a further numerical Monte-Carlo like approach, similar to what has been done in the previous section. The term "forwardpropagation" might be slightly confusing and we mainly use it to clarify that it is different from the previous approach.

1. Brisbane

For the method introduced in V C an error propagation is straight forward. Each complex matrix entry of the reconstructed unitary u_{ij} is obtained as the polar decomposition of the matrix m_{ij} build as the product of the modulo r_{ij} and the respective phase-factor $e^{i\theta_{ij}}$ $m_{ij} = r_{ij} \cdot e^{i\theta_{ij}}$. The measured observables are r_{ij} and θ_{ij}. Therefore we impose errors on these quantities: $r_{ij} \to r_{ij} + \delta r_{ij}$ and $\theta_{ij} \to \theta_{ij} + \delta\theta_{ij}$. Thus we obtain $m_{ij} \to m_{ij}\,(e^{i\delta\theta_{ij}} + \frac{\delta r_{ij}}{r_{ij}})$. The phase-factor occurring in the brackets rotates each entry of the initial matrix slightly away from its true value. Additionally the second part in the brackets imposes a rescaling for each entry. Expanding the exponential $e^{i\delta\theta_{ij}} \approx 1 + i\delta\theta_{ij} + \mathcal{O}(\delta\theta_{ij}^2)$ yields

$$m_{ij} \to m_{ij} + \underbrace{m_{ij}(i\delta\theta_{ij} + \frac{\delta r_{ij}}{r_{ij}})}_{\equiv \delta m_{ij}}. \tag{128}$$

However, another step is performed, namely the polar decomposition of $\mathbf{M}$ from

which we obtain the desired unitary $\mathbf{U}$. If $\epsilon = \|\delta\mathbf{M}\| / \|\mathbf{M}\|$ satisfies $\kappa_F(\mathbf{M}) \cdot \epsilon < 1$, performing a the polar decomposition of $\mathbf{M} \to \mathbf{M} + \delta\mathbf{M}$ yields [64]

$$\mathbf{M} + \delta\mathbf{M} = (\mathbf{U} + \delta\mathbf{U})\,(\mathbf{H} + \delta\mathbf{H})\,, \tag{129}$$

where

$$\frac{\|\delta\mathbf{U}\|_F}{\|\mathbf{U}\|_F} \leq (1+\sqrt{2})\kappa_F(\mathbf{M}) \cdot \epsilon + \mathcal{O}(\epsilon^2). \tag{130}$$

$\|.\|_F$ denotes the Frobenius norm and $\kappa_F(\mathbf{M})$ the condition number of $\mathbf{M}$ in the Frobenius norm. Therefore we conclude that, in the case that $\delta m_{ij} \backsim \mathcal{N}(0, \sigma^2)$ with $\sigma \approx \epsilon$, the polar decomposition only causes small variations as $\kappa_F(\mathbf{M})$ would be expected to be close to one.

It is worth noting though that another consequence of (130) is that this approach is very sensitive to corrupt input (r_{ij}/θ_{ij}) because it propagates straight through the reconstruction procedure and hence raises ϵ which can quickly yield a violation of the bound $\kappa_F(\mathbf{M})\epsilon < 1$. This gives raise to the question if an improvement of the procedure is possible by introducing some kind of feedback loop. The proposed approach from VIII can be seen as such an additional feedback loop.

Let us now proceed with the analysis of the approach introduced in V B.

2. Bristol

For this approach the experimental observables are one photon transmission amplitudes τ_{ij} and two-photon visibilities $\mathcal{V}_{ij}$.

Perturbing these observables results in the following first order transformation for (92): $\mathbf{M}_1 \to \mathbf{M}_1 + \delta\mathbf{M}_1$. If the perturbations $\delta\tau_{ij}$ and $\delta\mathcal{V}_{ij}$ are small, they cause the following explicit transformation for each element x_{ij} and α_{ij} in $\mathbf{M}_\mu$:

$x_{ij} \to x_{ij} + \delta x_{ij}$, where in first order approximation

$$\delta x_{ij} = x_{ij} \left(\frac{\delta\tau_{ij}}{\tau_{ij}} + \frac{\delta\tau_{kl}}{\tau_{kl}} - \frac{\delta\tau_{kj}}{\tau_{kj}} - \frac{\delta\tau_{il}}{\tau_{il}} \right), \tag{131}$$

and $\alpha_{ij} \to \alpha_{ij} + \delta\alpha_{ij}$, where in first order approximation

$$\begin{aligned} \delta\alpha_{ij} &= -\frac{1}{2\sqrt{1 - (\frac{1}{2}\mathcal{V}_{ij} \cdot (x_{ij} + x_{ij}^{-1}))^2}} \cdot \\ &\quad \left[(x_{ij} + x_{ij}^{-1})\delta\mathcal{V}_{ij} + (x_{ij} - x_{ij}^{-1}) \left(\frac{\delta\tau_{ij}}{\tau_{ij}} + \frac{\delta\tau_{kl}}{\tau_{kl}} - \frac{\delta\tau_{kj}}{\tau_{kj}} - \frac{\delta\tau_{il}}{\tau_{il}} \right) \right] \\ &= -\frac{(x_{ij} + x_{ij}^{-1})}{2\sqrt{1 - (\frac{1}{2}\mathcal{V}_{ij} \cdot (x_{ij} + x_{ij}^{-1}))^2}} \cdot \left[\delta\mathcal{V}_{ij} + \frac{(x_{ij}^2 - 1)}{(x_{ij}^2 + 1)} \cdot \frac{\delta x_{ij}}{x_{ij}} \right]. \end{aligned} \tag{132}$$

And thus $m_{\mu,ij} \to m_{\mu,ij} + \delta m_{\mu,ij}$ with

$$\delta m_{\mu,ij} = m_{\mu,ij} \left(\delta x_{ij} + i\delta\alpha_{ij} \right). \tag{133}$$

As before we are interested in finding a bound for the error on the recovered unitary $\frac{\|\delta\mathbf{U}\|}{\|\mathbf{U}\|}$ caused by $\delta\mathbf{M}_\mu$, i.e. the errors imposed on the input data for the network tomography. Similar to the previous procedure, a polar decomposition needs to be applied to $\mathbf{M}_1$ in order to perform the next steps of the tomography algorithm (see V B). As above, this results in

$$\mathbf{M}_1 + \delta\mathbf{M}_1 = \left(\mathbf{U} + \delta\mathbf{U} \right)_1 \left(\mathbf{H} + \delta\mathbf{H} \right)_1, \tag{134}$$

where

$$\frac{\|\delta\mathbf{U}_1\|_F}{\|\mathbf{U}_1\|_F} \leq (1 + \sqrt{2})\kappa_1(\mathbf{M}_1) \cdot \epsilon + \mathcal{O}(\epsilon^2). \tag{135}$$

Here κ_μ is again the condition number of $\mathbf{M}_1$ in the Frobenius norm. In order to obtain the unitary $\mathbf{U}$ that is characterizing the network another step needs to be applied. Namely the system of equations $\mathbf{M}_\mu \boldsymbol{\tau} = \mathbf{n}$ needs to be solved in order

to obtain the single photon transition amplitudes. This corresponds to inverting $\mathbf{M}_\mu$. But with (133) we have $\mathbf{M}_\mu \to \mathbf{M}_\mu + \delta\mathbf{M}_\mu$. I.e. the inversion performed undergoes the following transformation: $\mathbf{M}_\mu^{-1} \to (\mathbf{M}_\mu + \delta\mathbf{M}_\mu)^{-1}$.
Now, in order to obtain an upper bound similar to (130) we differentiate the equation $\mathbf{M}_\mu^{-1}\mathbf{M}_\mu = \mathbf{I}$

$$d\left(\mathbf{M}_\mu^{-1}\right)\mathbf{M}_\mu + \mathbf{M}_\mu^{-1} d\mathbf{M}_\mu = 0. \tag{136}$$

Rearranging and taking the norms yields

$$\frac{\left\|d\mathbf{M}_\mu^{-1}\right\|}{\|\mathbf{M}_\mu\|} \leq \underbrace{\|\mathbf{M}_\mu\|\left\|\mathbf{M}_\mu^{-1}\right\|}_{=\kappa_\mu(\mathbf{M}_\mu)} \frac{\|d\mathbf{M}_\mu\|}{\|\mathbf{M}_\mu\|}. \tag{137}$$

Combining (137) with (135) then yields the following bound

$$\frac{\left\|\delta\mathbf{M}_\mu^{-1}\right\|}{\|\mathbf{M}_\mu\|} \leq \kappa_\mu(\mathbf{M}_\mu) \cdot \kappa_1(\mathbf{M}_1) \cdot (1+\sqrt{2})\epsilon. \tag{138}$$

Therefore the fractional change can be $\kappa_\mu(\mathbf{M}_\mu)$ times larger than (130). To obtain actual values numerical experiments have been performed and the results are subject of section X E.
Furthermore, and as mentioned in the previous section, the approach is sensitive to corrupt inputs $(\tau_{ij}/\mathcal{V}_{ij})$ as it does not contain any kind of feedback loop or (unitary) constraints for $\mathbf{M}_\mu$.

C. Experimental Results for 'Bristol' and 'Brisbane'

This section will provide experimental results obtained for the 'Bristol' and 'Brisbane' method for the BosonSampling network (Fig. 22). In the spirit of X A we shall compare the resulting $\mathcal{Q}$-values as a measure for the quality of the respective network tomography. Furthermore we argue that permuting the matrix

prior, as proposed in [24], does not improve the 'Bristol' approach. Instead we propose a rudimentary feedback scheme for both approaches that can lead to an improvement.

1. Brisbane

It seems suitable to present the experimental results, carried out for the same 5×5 network (Fig. 22). The raw data is attached in XV (At this point we like to mention that the obtained data had been taken for the version of V C that is not considering loss). So let us proceed with the resulting quality measure $\mathcal{Q}$ obtained from the reconstructed unitary. The result reads as follows

$$\mathcal{Q}_0 = 0.0232 \quad \text{and} \quad \mathcal{Q}_R = 0.0527. \tag{139}$$

Which is slightly worse than (125).
Both results (139) and (125) are very close to the errors of the extracted data and deviations might be explained by results that have not been taken into account (e.g. higher order and for (139) also loss as we mentioned above). Which, again, raises the question if we could implement some additional post processing feedback-loop that improves the raw data obtained. We like to mention that we have also carried out the the reconstruction for 10000 Monte-Carlo propagation sampling runs and taken the mean of $\mathbf{U}$ over all these samples to calculate $\mathcal{Q}$. However, because of 128 one expects this to yield the same as carrying out the reconstruction for $\mathbf{M}$ if δm_{ij} is symmetric. And indeed both yield the same result.
However, performing a backwards propagation and collecting the respective self consistency witnesses in each run yields the following minimal and maximal values for the self-consistency witness

$$\mathcal{Q}_{\text{min}} = 0.0192 \quad \text{and} \quad \mathcal{Q}_{\text{max}} = 0.0353. \tag{140}$$

Table IV: Numerically obtained values for the self-consistency witness for the BosonSampling network for the three different approaches 'Brisbane', 'Bristol' and 'Vienna'.

	$\mathcal{Q}_{\min}$	$\mathcal{Q}_0$	$\mathcal{Q}_{\max}$
Brisbane	0.0192	0.0232	0.0353
Bristol	0.0173	0.0362	0.0948
Vienna	0.0145	0.0145	0.0159

Here we have performed 20000 sampling runs to obtain the maximal and minimal bounds.

2. Bristol

Performing the reconstruction with measurement data from Tbl. XIII and Tbl. I and without any sampling (the unitary is given in (149)), yields the following value for the self-consistency witness

$$\mathcal{Q}_0 = 0.0362 \quad \text{and} \quad \mathcal{Q}_R = 0.0821. \tag{141}$$

This value should be compared to (125) in order to judge the performance of the two approaches (for this particular network). We will compare and discuss all results in X D.

Performing 10000 Monte-Carlo runs yields the minimal and maximal values for the self-consistency witness yield

$$\mathcal{Q}_{\min} = 0.0173 \quad \text{and} \quad \mathcal{Q}_{\max} = 0.0948. \tag{142}$$

D. Summary of the Results

This section summarizes and discusses consequences of the numerical results obtained for the BosonSampling network. We have evaluated the self-consistency

witness as a quality measure. The results are listed in Tbl. IV. Some remarks:

1. For 'Vienna' the mean value $\mathcal{Q}_0$ coincides with the minimum value $\mathcal{Q}_{\min}$. This is because $\mathcal{Q}_0$ is a function of the expectation value of the recovered 19 parameters $\mathbf{p}_0$ which is the value that minimizes (88). The fact that $\mathcal{Q}_0 = \mathcal{Q}_{\min}$ is basically just stating that the procedure is self consistent and that the extracted visibilities $\mathcal{V}_{ij}$ are indeed the best choice (within the accuracy of our statistics).

2. The $\mathcal{Q}$-values obtained by 'Vienna' are always smaller than the $\mathcal{Q}$-values obtained for any of the other methods. We are interpreting this as a consequence of the fact that we apply unitary constraints throughout the network tomography. Therefore, as long as the errors are not to large, the parameters in each sampling run $\mathbf{p}_S$ always locate close to the expected parameter vector $\mathbf{p}_0$. Further analytical investigations should investigate this.

3. The $\mathcal{Q}$-value es obtained for 'Bristol' and 'Brisbane' are well explained by our analytical investigations from X B. The average and extreme cases for 'Bristol' seems to be slightly worse than those for 'Brisbane', as expected and explained by comparing the bounds (130) and (138).

We would like to add one more remark. As mentioned in [24] one could think about permuting the matrix prior or employing additional subsets to improve the robustness to noise. If we consider the statistical ensemble of the noise imposed primary data this should average out though. I.e. reconstructing for each possible subset of visibilities should not bring any advantage as the errors are not assumed to behave biased in any way.
However, one could improve the procedure by applying a self-consistency check in the following manner. Experimental measurement provide us with the primary data $\tilde{\tau}_{ij}$ and $\tilde{\mathcal{V}}_{ij}$ (accordingly for $\tilde{r}_{ij}$ and $\tilde{\theta}_{ij}$). As previously one imposes the respective errors on this set of data. Now one samples from the respective error

distributions and reconstructs $\mathbf{U}$ for the corresponding set of data $\tilde{\mathcal{V}}_{ij} + \delta\mathcal{V}_{ij}$ and $\tilde{\tau}_{ij} + \delta\tau_{ij}$. This is done for S sampling runs. For each of these runs the corresponding self-consistency witness is calculated as $\mathcal{Q}_S = \left\| \tilde{\boldsymbol{\mathcal{V}}} - \boldsymbol{\mathcal{V}}_S(\mathbf{U}_S) \right\|$. Subsequently one chooses the sampling run S that is minimizes $\mathcal{Q}_S$ and hence obtains $\mathbf{U}_S$.
It should be emphasized that the utility of this self-consistency check does strongly depend on the accuracy with which the primary data $\tilde{\boldsymbol{\tau}}$ and $\tilde{\boldsymbol{\mathcal{V}}}$ had been extracted. And in fact, we have propose a physically more sensible approach to improve the tomography in VIII that is taking the natural unitary structure of network into account. In the following section we provide numeric evidence for this result.

E. Extension and Improvement

In VIII we have proposed a new approach to improve the quality of network tomography. Besides the arguments given in that section we do now provide further numerical evidence that this approach is indeed accomplishing the desired improvements. Furthermore, we give numerical evidence for the error scaling found in (X B).
For that purpose we have carried out the numerical experiment proposed in VIII. Fig. 24 depicts the results for dimensions $D = 4, \ldots, 7$. For the numerical experiment we have created 20 random Haar unitaries. For each of these unitaries we have drawn 20 ($S = 1, \ldots, 20$) sampling sets of measurement data ($\tilde{\mathcal{V}}_{ij} + \delta\mathcal{V}_{ij}$ and $\tilde{\tau}_{ij} + \delta\tau_{ij}$ ($\tilde{\theta}_{ij} + \delta\theta_{ij}$ and $\tilde{r}_{ij} + \delta r_{ij}$)) from the respective error distributions. These error distributions are characterized by the variance of the underlying normal distribution $\delta \in [0.01, \ldots, 0.1]$ (note though that for instance $\delta\tau_{ij} \neq \delta$ as described in the methods). We have chosen $\delta \in [0.01, \ldots, 0.1]$ as higher values seem to be unrealistic because if the measurement precision was worse it would not be worth to do an experiment in the first place.
The sampling sizes (20 sampling runs for each of the 20 Haar unitaries) should

be increased further to provide better statistics. Due to the relatively high time requirements, especially for higher dimensions, we have restricted ourselves to this relatively small sample size (400 runs for each of the ten width of the underlying normal distribution and each approach $\Rightarrow$ 8000 optimization runs for each dimension). We would like to stress that they are qualitatively in agreement with the analytical investigations conducted in sections X B and X E.

The numerical results support our findings from section X B and arguments given in section X E. Considering Fig. 24 and comparing the results obtained for the different approaches to network tomography provides further evidence for both: The fact that we expect 'Bristol' to be less robust to errors than 'Brisbane', as well as for the improvement due to the post processing optimization. 'Brisbane' seems to scale linearly as one expects from (128), whereas 'Bristol' exhibits a less obvious behavior (see (133)) and more fluctuations. We are interpreting the fluctuations as a consequence of the bad response to corrupt values of this approach. The scaling behavior does depend on the condition number of the underlying matrices as we have shown in X B and these condition numbers can differ significantly for the different matrices $\mathbf{M}_\mu$. Additionally the errors do not enter directly as they do for 'Bristol'. They first undergo additional transformations which makes them behave rather "unpredictable" (see also (131) and (132)).

Let us now drag the attention to the new version (or improved versions of these two approaches).

First, we note that, without exception, improvement is obtained for both methods (see Fig. 24). Note that we converge into different minima though. Second, the improvement obtained increases with the errors imposed on the primary data. This is interesting as it means that we obtain an improvement that is better than $\mathcal{O}(1)$ by applying "unitary optimization". Finally we would like to emphasize that the size of the error bars in Fig. 24 is related to how strong the reconstructed unitaries scatter around the true one. It can be seen as a measure of how robust the

reconstruction is to the imposed perturbations. And indeed the results agree with what we expect from our analytical arguments derived in section X B. Namely the results obtained via 'Bristol' show worse behavior which is caused by the complex functions and the additional matrix inversion that has to be carried out.

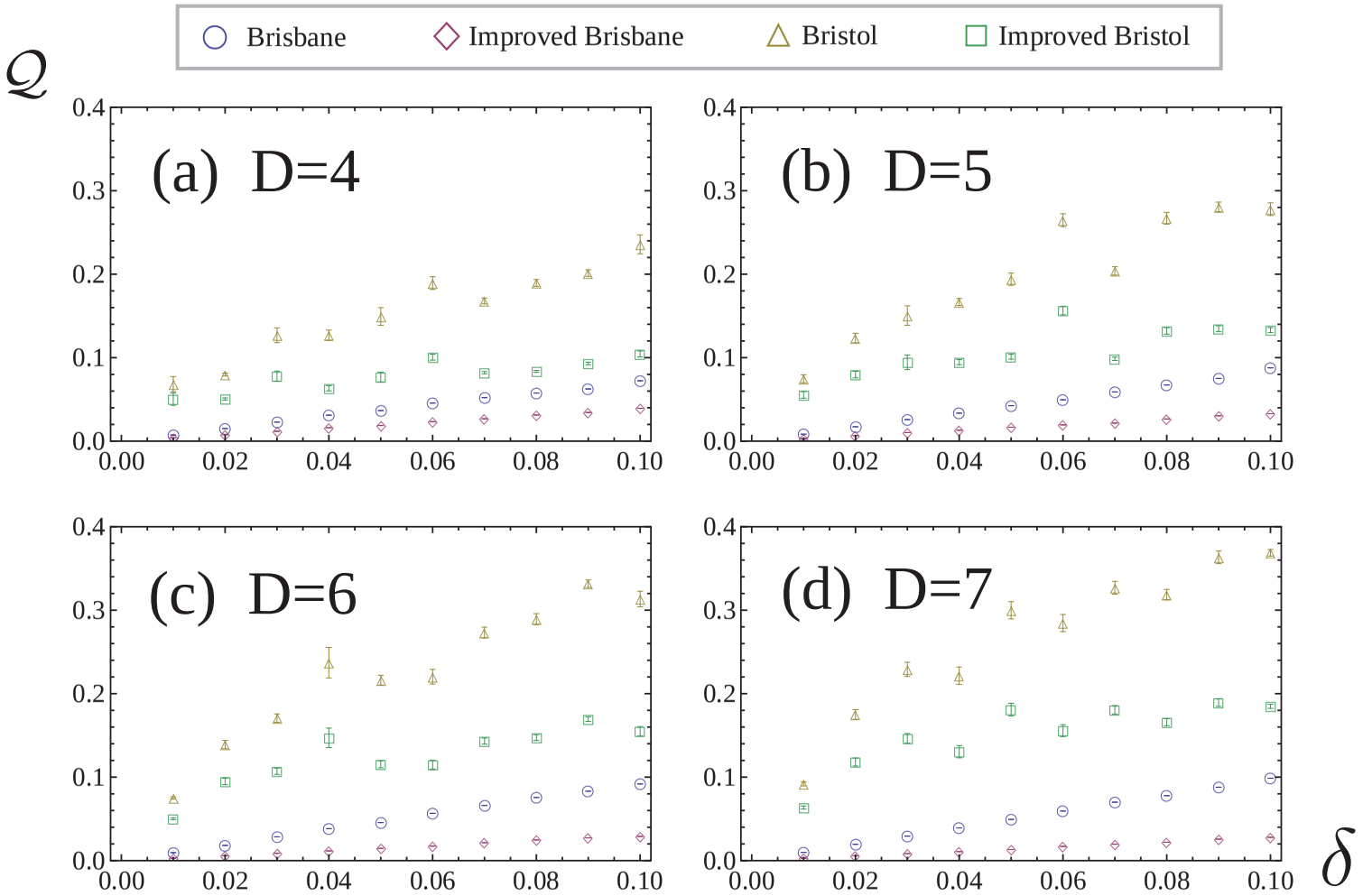


Figure 24: Behavior of the network tomography, measured by Q, with increasing error size, characterized by δ, for network dimensions $D = 4, \ldots, 7$. Fig (a) - (d) do provide evidence for both conjectures made. Namely, the conjecture that the 'Bristol' approach to network tomography exhibits poorer error scaling behavior than the 'Brisbane' approach (X B), as can be seen by comparing the results obtained for 'Bristol' (circles) with the results for 'Brisbane' (triangles). The second conjecture made states that 'Bristol' as well as 'Brisbane' can be improved by applying post processing optimization (VIII). Evidence for this conjecture is provided by comparing them (circles/triangles) with the respective improved version (diamonds/squares) respectively in Fig. (a) - (d). The data points are the estimator for the expectation values (i.e. the mean value of all sampling values). The error bars are the respective variance estimator (a consequence is that actual values from the numerical experiments can also lay outside of these error bars).

XI. CONCLUSION & OUTLOOK

In conclusion, the description of two-photon quantum interference in passive linear optical circuits was successfully extended by including frequency correlations. Those frequency correlations have an huge influence on the shape and visibility of the quantum interference feature and can both increase and decrease the dip-width and -visbility. This is an important finding as narrowband interference filters can typically reduce the frequency correlations, but they can never totally suppress them. This improved description was then applied in the modeling of experimental data generated with a 5×5 BosonSampling network. The thereby generated primary data set of 100 dip-visibilities was used to carry out a network tomography using two different techniques [24, 26], expanded by a numeric error propagation. A self-consistency check was implemented to provide an evaluation of the different approaches 'Bristol' [24], 'Brisbane' [25] and 'Vienna' [26]. A Monte-Carlo like backpropagation of errors for the approach 'Vienna' was implemented which yielded a confidence region for the BosonSampling network used in the experiment. The obtained results suggest that active approaches to network tomography containing feedback routines, such as the 'Vienna' approach, result in better error robustness than passive schemes like the 'Bristol' or 'Brisbane' method. Additional analytical arguments support this interpretation. This was done by performing an analytical error propagation for the two approaches 'Brisbane' and 'Bristol'. This analysis revealed that especially in the case of 'Bristol' errors propagate unfavorably and can affect the outcome in a negative way.

Finally, we conclude that 'Vienna' provides the best approach to network tomography under the condition that the initial points are close to the actual parameters. We consider this condition to be the primary drawback of this active method and resolve it by providing the initial guess with a passive scheme (preferable 'Brisbane'). The approach 'Vienna' can be applied to the two passive schemes and extends them by an additional post-processing optimization, forcing unitarity in

a natural manner. Using these techniques it was possible to outperform all of the previously known approaches which we could verify numerically.
Consequently, this approach improves current network tomography techniques which are essential for present and future applications of integrated linear optical waveguide circuits and is the best choice for network tomography that is currently known. An open question is the scalability of our methods for higher dimensions.

Appendix

XII. NORMALIZATION OF THE JSA AND FILTER FUNCTION

Before applying filtering the JSA is given by (45), i.e.

$$\beta(\omega_{s,}\omega_i) = \frac{1}{\sqrt{N}} \exp\left[-\frac{1}{2}(\boldsymbol{\omega} - \boldsymbol{\omega_0})^T \mathbf{M} (\boldsymbol{\omega} - \boldsymbol{\omega_0}) + i\boldsymbol{\omega}^T \boldsymbol{\tau}\right].$$

The normalization condition reads

$$\int\int d\omega_s d\omega_i \left|\beta(\omega_s, \omega_i\right|^2 \stackrel{!}{=} 1. \tag{143}$$

Or equivalently

$$N = \int\int d\boldsymbol{\omega} \exp\left[-\frac{1}{2}(\boldsymbol{\omega} - \boldsymbol{\omega_0})^T \mathbf{X} (\boldsymbol{\omega} - \boldsymbol{\omega_0})\right]. \tag{144}$$

To carry out the integral we can exploit the symmetry of $\mathbf{X} = 2\mathbf{M}$. Since $\mathbf{X}$ is symmetric it can be diagonalized by the orthogonal transformation $\mathbf{X} = \mathbf{O}^T\mathbf{D}\mathbf{O}$, where $\mathbf{O}$ is an orthogonal matrix and $\mathbf{D}$ a positive diagonal matrix. Now we substitute $\boldsymbol{\nu} = (\boldsymbol{\omega} - \boldsymbol{\omega_0})$ and then $\boldsymbol{\nu} \to \mathbf{O}\boldsymbol{\nu}$. Both transformation have unit Jacobian, $\det \mathbf{O} = 1$. After the substitution we end up with a Gaussian integral with exponent $-\frac{1}{2}\boldsymbol{\nu}^T\mathbf{D}\boldsymbol{\nu}$. But since $\mathbf{D}$ is diagonal the integral factorizes into two Gaussian integrals. Each of them contributes $\sqrt{\frac{2\pi}{d_{i,i}}}$, thus for an N dimensional integral we have $\sqrt{\frac{(2\pi)^N}{\prod_i d_{i,i}}}$. But, $\prod_i d_{i,i} = \det \mathbf{D} = \det \mathbf{X}$. Therefore (144) becomes

$$N = \sqrt{\frac{(2\pi)^2}{\det \mathbf{X}}} = \sqrt{\frac{(2\pi)^2}{\det 2\mathbf{M}}} = \sqrt{\frac{(2\pi)^2}{2^2 \det \mathbf{M}}} = \frac{\pi}{\sqrt{\det \mathbf{M}}}. \tag{145}$$

A similar procedure yields the normalization for the Filter function (48)

$$T_\mu(\omega_\mu) = \sqrt{N_\mu^{(G)}} \exp\left[-\frac{(\omega_\mu - \omega_\mu^F)^2}{2\sigma_\mu^2}\right].$$

The normalization condition in this case reads

$$\int d\omega_\mu \, |T_\mu(\omega_\mu)|^2 \stackrel{!}{=} \mathscr{T}_\mu. \tag{146}$$

Where $\mathscr{T}_\mu$ is the transmissivity of the respective filter over the whole spectrum. Thus we obtain for the normalization constant

$$N_\mu^{(G)} = \frac{\mathscr{T}_\mu}{\pi \cdot \sigma_\mu^2} \tag{147}$$

XIII. EXPERIMENTAL VALUES FOR THE JSA AND FILTER

Table V: Assumed values for the JSA.

λ_s^0 [nm]	λ_i^0 [nm]	σ_l [nm]	σ_k [nm]	θ
790.63	788.90	$\frac{9}{2\sqrt{\ln 2}}$	$\frac{2.77}{2\sqrt{2\ln 2}}$	$\pm 60°$

Table VI: Measured values for the utilized filters. The subscript is not uniquely determined because the arm in which the filter was mounted was uncertain

$\lambda_{s/i}^F$ [nm]	$\lambda_{i/s}^F$ [nm]	$\sigma_{s/i}^F$ [nm]	$\sigma_{i/s}^F$ [nm]	$\mathscr{T}_{s/i}$ [%]	$\mathscr{T}_{i/s}$ [%]
789.13	788.09	$\frac{3.05}{2\sqrt{\ln 2}}$	$\frac{3.41}{2\sqrt{\ln 2}}$	84	75

Table VII: Assumed values for the JSA.

λ_s^0 [nm]	λ_i^0 [nm]	σ_l [nm]	σ_k [nm]	θ
790.63	788.90	$\frac{11}{2\sqrt{\ln 2}}$	$\frac{11/3}{2\sqrt{2\ln 2}}$	$\pm 70°$

Table VIII: Assumed values for the filters.

λ_s^F [nm]	λ_i^F [nm]	σ_s^F [nm]	σ_i^F [nm]	$\mathscr{T}_s$ [%]	$\mathscr{T}_i$ [%]
788.45	789.05	$\frac{3.0}{2\sqrt{\ln 2}}$	$\frac{3.0}{2\sqrt{\ln 2}}$	84	75

The aim of this section will be to bring the three parts from IX A together in a more vivid picture. To do so, we will investigate (109) for the whole temporal range.

From simulations of the PDC source [54] we extract probable values for σ_l and σ_k, as well as θ. It should be emphasized that these values might differ from the real values to which we do not have access (though there are techniques which allow obtaining this data [65]). Furthermore the marginal signal- and idler- spectrum had been measured respectively to obtain the center frequencies [66]. Tbl. V lists the values that were utilized.

Additionally, the filter intensity distributions [67] and marginals *after* applying filtering [66] had been measured. The resulting values are listed in Tbl. VI.

However, these measurements were made after the original experiment had been moved, i.e. deviations are very likely. To circumvent these inaccuracies we have applied an additional optimization in which we have have minimized the deviation between the marginal values that had been measured after the filters and (49). Where we have restricted the parameters to values laying in a close region around those from Tbl. VI and Tbl. V.

This procedure yields.

XIV. EXTRACTED EXPERIMENTAL VISIBILITIES WHEN CORRELATIONS ARE NEGLECTED

Table IX: Visibilities, extracted via. (121).

(t_1, t_2) \ (s_1, s_2)	$(4,5)$	$(3,5)$	$(2,5)$	$(1,5)$	$(3,4)$	$(2,4)$	$(1,4)$	$(2,3)$	$(1,3)$	$(1,2)$
$(1,2)$	-0.99774	-0.26075	0.57953	0.08355	-0.24927	0.59421	0.090739	-0.06633	0.29434	0.20671
$(1,3)$	-0.66160	0.20828	0.23463	0.33708	0.47983	0.64939	0.34450	-0.45859	-0.30735	0.33854
$(1,4)$	0.56952	0.55389	-0.40494	-0.80636	-0.93579	0.23848	0.43712	0.25225	0.40415	-0.02325
$(1,5)$	0.84515	-0.30556	0.45332	0	0.19145	-0.39030	0	0.70383	0	0
$(2,3)$	-0.65884	0.15801	-0.01569	0.07748	0.12132	-0.18852	0.25136	-0.72639	0.25251	0.74049
$(2,4)$	0.56604	0.54032	0.15386	-0.1130	-0.23101	-0.32213	0.51476	0.60992	-0.29580	0.43273
$(2,5)$	0.80072	-0.37040	-0.08037	0	0.18541	0.09046	0	0.39417	0	0
$(3,4)$	0.88345	-0.58549	0.25030	0.60248	0.49074	-0.20129	-0.51563	0.64727	0.98364	-0.60233
$(3,5)$	0.45957	0.83641	-0.86624	0	-0.62193	0.64541	0	0.83560	0	0
$(4,5)$	-0.38701	0.82482	0.37927	0	0.22698	0.08721	0	-0.61351	0	0

Table X: The to Tbl. IX associated errors.

(t_1, t_2) \ (s_1, s_2)	$(4,5)$	$(3,5)$	$(2,5)$	$(1,5)$	$(3,4)$	$(2,4)$	$(1,4)$	$(2,3)$	$(1,3)$	$(1,2)$
$(1,2)$	0.03376	0.01009	0.01036	0.00540	0.00870	0.00912	0.00468	0.00307	0.00230	0.00419
$(1,3)$	0.00814	0.00222	0.00463	0.00451	0.00448	0.00860	0.00705	0.00667	0.00501	0.00807
$(1,4)$	0.01206	0.00565	0.00675	0.01275	0.01336	0.00480	0.01052	0.00247	0.00451	0.00331
$(1,5)$	0.00714	0.00444	0.00871	0	0.00215	0.00613	0	0.00736	0	0
$(2,3)$	0.01177	0.00419	0.00263	0.00160	0.00791	0.00618	0.00361	0.01046	0.00408	0.00642
$(2,4)$	0.02326	0.01063	0.00604	0.00433	0.01744	0.00911	0.01005	0.00467	0.00463	0.00244
$(2,5)$	0.01019	0.00981	0.00544	0	0.00431	0.00278	0	0.00789	0	0
$(3,4)$	0.00852	0.00689	0.00245	0.00498	0.00892	0.00440	0.01130	0.00548	0.00809	0.00716
$(3,5)$	0.00261	0.00555	0.01091	0	0.00786	0.00600	0	0.01233	0	0
$(4,5)$	0.00614	0.00823	0.00417	0	0.00400	0.00172	0	0.00793	0	0

Table XI: Measured intensities. For the sampling we have imposed poissonian errors.

output \ input	1	2	3	4	5
1	0.1141	0.6969	0.0961	0.1181	0
2	0.0915	0.1740	0.0626	0.6308	0.0452
3	0.6480	0.0805	0.0711	0.1337	0.0637
4	0.0838	0.0280	0.1519	0.0148	0.7209
5	0.0626	0.0207	0.6183	0.1026	0.1702

$$\mathrm{Re}\,\mathbf{U}_{\mathrm{V}_0} = \begin{pmatrix} 0.0320 & 0.0724 & -0.2780 & 0.1228 & 0 \\ 0.0114 & -0.3863 & -0.1353 & -0.7842 & 0.0124 \\ -0.7757 & -0.2937 & -0.2677 & 0.0267 & -0.2476 \\ 0.1444 & -0.1518 & -0.1392 & -0.1327 & 0.0203 \\ 0.2225 & 0.0715 & -0.7929 & 0.0871 & 0.4123 \end{pmatrix}. \tag{148}$$

$$\mathrm{Im}\,\mathbf{U}_{\mathrm{V}_0} = \begin{pmatrix} -0.3370 & 0.8203 & -0.1060 & -0.3220 & 0 \\ 0.2751 & 0.1860 & 0.2073 & -0.1502 & -0.2036 \\ -0.2328 & 0.0018 & -0.0162 & 0.3517 & -0.0151 \\ -0.2611 & -0.0840 & 0.0839 & -0.0092 & 0.8449 \\ 0.1231 & -0.1293 & -0.0268 & 0.3067 & -0.1121 \end{pmatrix}.$$

XV. DATASET OBTAINED FOR "BRISBANE"

$$\mathrm{Re}\,\mathbf{U}_{\mathrm{BA}} = \begin{pmatrix} 0.3344 & 0.8287 & 0.2985 & 0.33497 & -0.0032 \\ 0.2953 & -0.0988 & -0.0921 & 0.0087 & -0.1459 \\ 0.8068 & -0.2863 & 0.1149 & -0.1956 & 0.1137 \\ 0.2912 & -0.0358 & -0.1305 & -0.0850 & -0.1482 \\ 0.2551 & -0.0323 & -0.4806 & 0.2677 & -0.0183 \end{pmatrix}. \tag{149}$$

Table XII: Mean values of the measured phases in rad. The mean values and errors are a rough estimate obtained by measuring the over complete set of visibilities. I.e. the relative phases have been measured for each possible reference mode and not just for one.

output \ input	$(1,1)$	$(1,2)$	$(1,3)$	$(1,4)$	$(1,5)$
$(1,1)$	0	0	0	0	0
$(1,2)$	0	-1.8380	1.9262	1.5697	2.3603
$(1,3)$	0	3.0500	1.1161	-2.1475	-1.0974
$(1,4)$	0	1.8003	-1.9295	-2.2440	1.7501
$(1,5)$	0	1.7939	-2.2142	0.5677	-1.6230

Table XIII: Measured one photon transition intesities. Note also the difference, compared to Tbl. XI that indicates that errors are indeed present.

output \ input	1	2	3	4	5
1	0.1358	0.7056	0.1265	0.1311	0.0002
2	0.0970	0.1831	0.0826	0.6674	0.0606
3	0.6499	0.0763	0.0871	0.1132	0.0761
4	0.0671	0.0200	0.1388	0.0130	0.6811
5	0.0503	0.0150	0.5651	0.0753	0.1820

$$\mathbf{U}_{\mathrm{BA}} = \begin{pmatrix} 0.0061 & 0.0022 & -0.0066 & -0.0055 & -0.0031 \\ 0.0011 & 0.4086 & -0.2379 & -0.7934 & -0.1412 \\ -0.0002 & -0.0356 & -0.2348 & 0.3124 & 0.2205 \\ -0.0032 & -0.1712 & 0.3629 & 0.1057 & -0.8334 \\ -0.0052 & -0.1471 & 0.6305 & -0.1791 & 0.4236 \end{pmatrix}. \tag{150}$$

Note that the the boarders are not real and $\left(\mathbf{U}_{\mathrm{BA}}\right)_{1,5} \neq 0$. This is caused by the polar decomposition.

XVI. DATASET OBTAINED FOR "BRISTOL"

$$\mathrm{Re}\,\mathbf{U}_{\mathrm{B}} = \begin{pmatrix} 0.3173 & -0.8428 & 0.1449 & -0.1754 & -0.0095 \\ 0.2851 & 0.053718 + 0.39242i & 0.1724 & -0.6974 & -0.2049 \\ 0.8167 & 0.2783 & 0.27285 & 0.3529 & 0.2377 \\ 0.2966 & 0.047279 - 0.15226i & -0.4313 & 0.1311 & -0.7739 \\ 0.2511 & 0.042948 - 0.12674i & -0.7570 & -0.2893 & 0.3856 \end{pmatrix}. \quad (151)$$

$$\mathrm{Im}\,\mathbf{U}_{\mathrm{B}} = \begin{pmatrix} 0 & -0.1100 & -0.2366 & 0.2629 & -0.0101 \\ 0 & 0.3924 & 0.1893 & -0.4019 & -0.0800 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -0.1523 & -0.0524 & -0.0310 & 0.2843 \\ 0 & -0.1267 & 0.1459 & 0.1608 & -0.2578 \end{pmatrix}. \quad (152)$$

Note that $\left(\mathbf{U}_{\mathrm{B}}\right)_{1,5} \neq 0$ which is caused by the polar decomposition. The fact that the third row is real instead of the first is originating from permuting the first row with the third in order to apply the network tomography (no vanishingly small elements in the first two row are permitted).

XVII. BOSONSAMPLING UNITARY

$$
\begin{aligned}
U(1,1) &= e^{i\Phi_1}\sin(\Omega_1)\sin(\Omega_5) - e^{i\Phi_2+i\Phi_5}\cos(\Omega_1)\cos(\Omega_5)\sin(\Omega_3)\\
U(2,1) &= e^{i\Phi_2}\cos(\Omega_1)\left(-e^{i\Phi_6+i\Phi_{10}}\cos(\Omega_3)\cos(\Omega_7)\sin(\Omega_6) - e^{i\Phi_5+i\Phi_9}\sin(\Omega_3)\sin(\Omega_5)\sin(\Omega_7)\right)\\
&\quad -e^{i\Phi_1+i\Phi_9}\cos(\Omega_5)\sin(\Omega_1)\sin(\Omega_7)\\
U(3,1) &= e^{i\Phi_2}\cos(\Omega_1)\left(e^{i\Phi_6+i\Phi_{10}}\cos(\Omega_3)\sin(\Omega_6)\sin(\Omega_7) - e^{i\Phi_5+i\Phi_9}\cos(\Omega_7)\sin(\Omega_3)\sin(\Omega_5)\right)\\
&\quad -e^{i\Phi_1+i\Phi_9}\cos(\Omega_5)\cos(\Omega_7)\sin(\Omega_1)\\
U(4,1) &= -e^{i\Phi_2+i\Phi_6+i\Phi_{11}}\cos(\Omega_1)\cos(\Omega_3)\cos(\Omega_6)\sin(\Omega_8)\\
U(5,1) &= e^{i\Phi_2+i\Phi_6+i\Phi_{11}}\cos(\Omega_1)\cos(\Omega_3)\cos(\Omega_6)\cos(\Omega_8)
\end{aligned}
$$

$$
\begin{aligned}
U(1,2) &= -e^{i\Phi_2+i\Phi_5}\cos(\Omega_5)\sin(\Omega_1)\sin(\Omega_3) - e^{i\Phi_1}\cos(\Omega_1)\sin(\Omega_5)\\
U(2,2) &= e^{i\Phi_1+i\Phi_9}\cos(\Omega_1)\cos(\Omega_5)\sin(\Omega_7)\\
&\quad +e^{i\Phi_2}\sin(\Omega_1)\left(-e^{i\Phi_6+i\Phi_{10}}\cos(\Omega_3)\cos(\Omega_7)\sin(\Omega_6) - e^{i\Phi_5+i\Phi_9}\sin(\Omega_3)\sin(\Omega_5)\sin(\Omega_7)\right)\\
U(3,2) &= e^{i\Phi_1+i\Phi_9}\cos(\Omega_1)\cos(\Omega_5)\cos(\Omega_7)\\
&\quad +e^{i\Phi_2}\sin(\Omega_1)\left(e^{i\Phi_6+i\Phi_{10}}\cos(\Omega_3)\sin(\Omega_6)\sin(\Omega_7) - e^{i\Phi_5+i\Phi_9}\cos(\Omega_7)\sin(\Omega_3)\sin(\Omega_5)\right)\\
U(4,2) &= -e^{i\Phi_2+i\Phi_6+i\Phi_{11}}\cos(\Omega_3)\cos(\Omega_6)\sin(\Omega_1)\sin(\Omega_8)\\
U(5,2) &= e^{i\Phi_2+i\Phi_6+i\Phi_{11}}\cos(\Omega_3)\cos(\Omega_6)\cos(\Omega_8)\sin(\Omega_1)
\end{aligned}
$$

$$
\begin{aligned}
U(1,3) &= -e^{i\Phi_3+i\Phi_5}\cos(\Omega_3)\cos(\Omega_5)\sin(\Omega_2)\\
U(2,3) &= -e^{i\Phi_4+i\Phi_7+i\Phi_{10}}\cos(\Omega_2)\cos(\Omega_6)\cos(\Omega_7)\sin(\Omega_4)\\
&\quad -e^{i\Phi_3}\sin(\Omega_2)\left(e^{i\Phi_5+i\Phi_9}\cos(\Omega_3)\sin(\Omega_5)\sin(\Omega_7)-e^{i\Phi_6+i\Phi_{10}}\cos(\Omega_7)\sin(\Omega_3)\sin(\Omega_6)\right)\\
U(3,3) &= e^{i\Phi_4+i\Phi_7+i\Phi_{10}}\cos(\Omega_2)\cos(\Omega_6)\sin(\Omega_4)\sin(\Omega_7)\\
&\quad -e^{i\Phi_3}\sin(\Omega_2)\left(e^{i\Phi_5+i\Phi_9}\cos(\Omega_3)\cos(\Omega_7)\sin(\Omega_5)+e^{i\Phi_6+i\Phi_{10}}\sin(\Omega_3)\sin(\Omega_6)\sin(\Omega_7)\right)\\
U(4,3) &= e^{i\Phi_3+i\Phi_6+i\Phi_{11}}\cos(\Omega_6)\sin(\Omega_2)\sin(\Omega_3)\sin(\Omega_8)\\
&\quad +e^{i\Phi_4}\cos(\Omega_2)\left(e^{i\Phi_8}\cos(\Omega_4)\cos(\Omega_8)+e^{i\Phi_7+i\Phi_{11}}\sin(\Omega_4)\sin(\Omega_6)\sin(\Omega_8)\right)\\
U(5,3) &= e^{i\Phi_4}\cos(\Omega_2)\left(e^{i\Phi_8}\cos(\Omega_4)\sin(\Omega_8)-e^{i\Phi_7+i\Phi_{11}}\cos(\Omega_8)\sin(\Omega_4)\sin(\Omega_6)\right)\\
&\quad -e^{i\Phi_3+i\Phi_6+i\Phi_{11}}\cos(\Omega_6)\cos(\Omega_8)\sin(\Omega_2)\sin(\Omega_3)
\end{aligned}
$$

$$
\begin{aligned}
U(1,4) &= e^{i\Phi_3+i\Phi_5}\cos(\Omega_2)\cos(\Omega_3)\cos(\Omega_5)\\
U(2,4) &= e^{i\Phi_3}\cos(\Omega_2)\left(e^{i\Phi_5+i\Phi_9}\cos(\Omega_3)\sin(\Omega_5)\sin(\Omega_7)-e^{i\Phi_6+i\Phi_{10}}\cos(\Omega_7)\sin(\Omega_3)\sin(\Omega_6)\right)\\
&\quad -e^{i\Phi_4+i\Phi_7+i\Phi_{10}}\cos(\Omega_6)\cos(\Omega_7)\sin(\Omega_2)\sin(\Omega_4)\\
U(3,4) &= e^{i\Phi_4+i\Phi_7+i\Phi_{10}}\cos(\Omega_6)\sin(\Omega_2)\sin(\Omega_4)\sin(\Omega_7)\\
&\quad +e^{i\Phi_3}\cos(\Omega_2)\left(e^{i\Phi_5+i\Phi_9}\cos(\Omega_3)\cos(\Omega_7)\sin(\Omega_5)+e^{i\Phi_6+i\Phi_{10}}\sin(\Omega_3)\sin(\Omega_6)\sin(\Omega_7)\right)\\
U(4,4) &= e^{i\Phi_4}\sin(\Omega_2)\left(e^{i\Phi_8}\cos(\Omega_4)\cos(\Omega_8)+e^{i\Phi_7+i\Phi_{11}}\sin(\Omega_4)\sin(\Omega_6)\sin(\Omega_8)\right)\\
&\quad -e^{i\Phi_3+i\Phi_6+i\Phi_{11}}\cos(\Omega_2)\cos(\Omega_6)\sin(\Omega_3)\sin(\Omega_8)\\
U(5,4) &= e^{i\Phi_3+i\Phi_6+i\Phi_{11}}\cos(\Omega_2)\cos(\Omega_6)\cos(\Omega_8)\sin(\Omega_3)\\
&\quad +e^{i\Phi_4}\sin(\Omega_2)\left(e^{i\Phi_8}\cos(\Omega_4)\sin(\Omega_8)-e^{i\Phi_7+i\Phi_{11}}\cos(\Omega_8)\sin(\Omega_4)\sin(\Omega_6)\right)\\
U(1,5) &= 0\\
U(2,5) &= e^{i\Phi_7+i\Phi_{10}}\cos(\Omega_4)\cos(\Omega_6)\cos(\Omega_7)\\
U(3,5) &= -e^{i\Phi_7+i\Phi_{10}}\cos(\Omega_4)\cos(\Omega_6)\sin(\Omega_7)\\
U(4,5) &= e^{i\Phi_8}\cos(\Omega_8)\sin(\Omega_4)-e^{i\Phi_7+i\Phi_{11}}\cos(\Omega_4)\sin(\Omega_6)\sin(\Omega_8)\\
U(5,5) &= e^{i\Phi_7+i\Phi_{11}}\cos(\Omega_4)\cos(\Omega_8)\sin(\Omega_6)+e^{i\Phi_8}\sin(\Omega_4)\sin(\Omega_8)
\end{aligned}
$$

[1] M. Planck, in *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin.* (Verl. d. Kgl. Akad. d. Wiss., 1899), pp. 440–480.

[2] R. P. Feynman, International Journal of Theoretical Physics **21**, 467 (1982).

[3] P. Benioff, Journal of Statistical Physics **29**, 515 (1982).

[4] D. Deutsch, Proceedings of the Royal Society A **400**, 97 (1985).

[5] P. W. Shor, in *Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on* (IEEE, 1994), pp. 124–134.

[6] D. Divincenzo, in *Mesoscopic Electron Transport* (Springer Netherlands, 1997), pp. 657–677.

[7] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge Series on Information and the Natural Sciences (Cambridge University Press, 2010).

[8] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O/'Brien, Nature **464**, 45 (2010), ISSN 0028-0836.

[9] E. Knill, R. Laflamme, and G. J. Milburn, Nature **409**, 46 (2001).

[10] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001).

[11] S. Aaronson and A. Arkhipov, Theory of Computing **4**, 143 (2013).

[12] P. Kaye, R. Laflamme, and M. Mosca, *An Introduction to Quantum Computing* (Oxford University Press, 2007).

[13] A. Politi, J. C. F. Matthews, and J. L. O'Brien, Science **325**, 1221 (2009).

[14] J. C. F. Matthews, A. Politi, StefanovAndre, and J. L. O'Brien, Nat Photon **3**, 346 (2009).

[15] M. Tillmann, B. Dakic, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, Nat Photon **7**, 540 (2013), ISSN 1749-4885.

[16] B. J. Metcalf, J. B. Spring, P. C. Humphreys, N. Thomas-Peter, M. Barbieri, W. S.

Kolthammer, X.-M. Jin, N. K. Langford, D. Kundys, J. C. Gates, et al., Nat Photon **8**, 770 (2014).

[17] W. H. Louisell, A. Yariv, and A. E. Siegman, Phys. Rev. **124**, 1646 (1961).

[18] D. C. Burnham and D. L. Weinberg, Phys. Rev. Lett. **25**, 84 (1970).

[19] B. R. Mollow and R. J. Glauber, Phys. Rev. **160**, 1076 (1967).

[20] B. R. Mollow, Phys. Rev. A **8**, 2684 (1973).

[21] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, Rev. Mod. Phys. **79**, 135 (2007).

[22] A. B. U'Ren, K. Banaszek, and I. A. Walmsley, Quantum Info. Comput. **3**, 480 (2003).

[23] W. P. Grice and I. A. Walmsley, Phys. Rev. A **56**, 1627 (1997).

[24] A. Laing and J. L. O'Brien, eprint arXiv:1208.2868 (2012).

[25] S. Rahimi-Keshari, M. A. Broome, R. Fickler, A. Fedrizzi, T. C. Ralph, and A. G. White, Opt. Express **21**, 13450 (2013).

[26] M. Tillmann, S.-H. Tan, S. E. Stoeckl, B. C. Sanders, H. de Guise, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, eprint arXiv:1403.3433 (2014).

[27] R. J. Glauber, Phys. Rev. **130**, 2529 (1963).

[28] K. J. Blow, R. Loudon, S. J. D. Phoenix, and T. J. Shepherd, Phys. Rev. A **42**, 4102 (1990).

[29] M. S. Moslehian and E. W. Weisstein, *Haar measure* (2014), from MathWorld–A Wolfram Web Resource., URL `http://mathworld.wolfram.com/HaarMeasure.html`.

[30] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Phys. Rev. Lett. **73**, 58 (1994).

[31] A. Christ, Ph.D. thesis, Universität Paderborn (2013).

[32] R. Loudon, *The Quantum Theory of Light* (Oxford University Press, 2010), 3rd ed.

[33] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics* (Wiley, 2007).

[34] S. Barnett and P. Radmore, *Methods in Theoretical Quantum Optics*, Oxford science publications (Clarendon Press, 1997).

[35] L. Mandel and E. Wolf, *Optical coherence and Quantum Optics* (Cambridge University Press, 2008).

[36] M. Bensimhoun.

[37] A. M. Branczyk, T. C. Ralph, W. Helwig, and C. Silberhorn, New Journal of Physics **12**, 063001 (2010).

[38] K. N. Cassemiro, K. Laiho, and C. Silberhorn, New Journal of Physics **12**, 113052 (2010).

[39] A. Christ, C. Lupo, M. Reichelt, T. Meier, and C. Silberhorn, Phys. Rev. A **90**, 023823 (2014).

[40] E. Schmidt, Mathematische Annalen **65**, 370 (1908), ISSN 0025-5831.

[41] C. K. Law, I. A. Walmsley, and J. H. Eberly, Phys. Rev. Lett. **84**, 5304 (2000).

[42] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).

[43] R. Hanbury Brown and R. Q. Twiss, Nature **178**, 1046 (1956).

[44] C. K. Hong, Z. Y. Ou, and L. Mandel, Phys. Rev. Lett. **59**, 2044 (1987).

[45] R. Byrd, J. C. Gilbert, and J. Nocedal, Mathematical Programming **89**, 149 (2000).

[46] R. A. Waltz, J. Morales, J. Nocedal, and D. Orban, Mathematical Programming **107**, 391 (2006).

[47] R. Byrd, M. Hribar, and J. Nocedal, SIAM Journal on Optimization **9**, 877 (1999).

[48] A. Angermann, M. Beuschel, M. Rau, and U. Wohlfarth, *MATLAB - Simulink - Stateflow: Grundlagen, Toolboxen, Beispiele* (Oldenbourg Wissenschaftsverlag, 2011).

[49] C. Audet and J. J. Dennis, SIAM Journal on Optimization **13**, 889 (2003).

[50] K. Fan and A. Hoffman, Proceedings of the American Mathematical Society **6**, 111 (1955).

[51] J. Demmel, I. Dumitriu, and O. Holtz, Numer. Math. **108**, 59 (2007).

[52] B. Frieden, *Physics from Fisher Information: A Unification* (Cambridge University Press, 1998).

[53] V. Giovannetti, S. Lloyd, and L. Maccone, Nature **412**, 417 (2001).

[54] M. Tillmann, *spdc geometry*, URL `http://ul.to/t0847m8d`.

[55] V. Ansari, B. Brecht, G. Harder, and C. Silberhorn, eprint arXiv:1404.7725 (04/2014).

[56] A. Crespi, R. Osellame, R. Ramponi, D. J. Brod, E. F. Galvao, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino, Nat Photon **7**, 545 (2013).

[57] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White, Science **339**, 794 (2013).

[58] J. B. Spring, B. J. Metcalf, P. C. Humphreys, W. S. Kolthammer, X.-M. Jin, M. Barbieri, A. Datta, N. Thomas-Peter, N. K. Langford, D. Kundys, et al., Science **339**, 798 (2013).

[59] P. G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A. V. Sergienko, and Y. Shih, Phys. Rev. Lett. **75**, 4337 (1995).

[60] D. Bouwmeester, J.-W. Pan, M. Daniell, H. Weinfurter, and A. Zeilinger, Phys. Rev. Lett. **82**, 1345 (1999).

[61] C. Schmidt, *Coincidence probability and visibilities for full treatment (correlations&filter)*, URL `http://ul.to/0qx1avdm`.

[62] C. Schmidt, *Visibility extraction routine including experimental data*, URL `http://ul.to/0cae3tci`.

[63] C. Schmidt, URL `http://ul.to/mobcgbml`.

[64] N. Higham, SIAM Journal on Scientific and Statistical Computing **7**, 1160 (1986).

[65] M. Avenhaus, A. Eckstein, P. J. Mosley, and C. Silberhorn, Opt. Lett. **34**, 2873 (2009).

[66] M. Tillmann, *Measured source data*, URL `http://ul.to/gtkaw7hr`.

[67] M. Tillmann, *Filter intensity distributions*, URL `http://ul.to/q0o1va3f`.

[68] C. Schmidt, URL `http://ul.to/6gyuooe7`.

[69] C. Schmidt, URL `http://ul.to/avy8daob`.

[70] C. Schmidt, URL `http://ul.to/mb478e01`.

[71] A. Altland and B. D. Simons, *Condensed Matter Field Theory* (Canbridge University Press, 2010).

[72] J. L. O'Brien, G. J. Pryde, A. Gilchrist, D. F. V. James, N. K. Langford, T. C. Ralph, and A. G. White, Phys. Rev. Lett. **93**, 080502 (2004).

[73] J. L. O'Brien, G. J. Pryde, A. G. White, T. C. Ralph, and D. Branning, Nature **426**, 264 (2003).

[74] A. Peres, Nuclear Physics B **6**, 243 (1989).

[75] E. T. Whittaker and G. Robinson, *The Calculus of Observations: An Introduction to Numerical Analysis* (Dover Publications, 1967).

[76] A. Zeilinger, American Journal of Physics **49**, 882 (1981).

Hinnerk Christian Schmidt

M.Sc. Student, Physics

University of Vienna
Boltzmanngasse 5, 1090 Vienna, Austria
christian.schmidt@univie.ac.at

Research

Optical Quantum Computation: I am currently working on a theoretical approach to the characterization of quantum interference in linear optical networks. In particular I am studying the robustness, extension and new approaches of tomography algorithms for such networks. Additionally I work on the influence of correlations in parametric down conversion sources.

Interests: Statistical Physics and its applications to Complex & Dynamical Systems, Cellular Automata, Spin Glasses, Constraint Satisfaction Problems, Machine Learning, (Quantum) Random Walks, Quantum Computation, Quantum Simulation

Education

University of Vienna, Vienna, Austria, 2013 - current
M.Sc. in Physics, Expected 2014, Supervisor: Philip Walther

University of Münster, Münster, Germany**,** 2009 - 2011 & 2012 - 2013
B.Sc. in Physics, Distinguished Performance, 2012

University of York, York, UK, 2011-2012
B.Sc. Project, Supervisor: Irene D'Amico

Awards

DAAD Scholarship, German Government(German Academic Exchange Service), 2013 - current

Positions

Research Assistant, M.Sc. Project in the Group of Photonic Quantum Computation and Quantum Simulation, University of Vienna, 2013 - 2014, Supervisor: Philip Walther

Research Intern, Project on non-linear optimization in natural science, Applied Mathematics Münster, University of Münster, Winter Term 2012/2013, Supervision: Michael Schaefer, Mario Ohlberger

Undergraduate Research, B.Sc Project on the properties of spin chains for transferring information and the potential for entangling qubits, University of York, 2011 - 2012, Supervisor:Irene D'Amico

Teaching

Tutor, Physics for Chemists, University of Münster, 2012 - 2013

Mentor, Physics I-II, Teaching Program by the Principles of Teaching Submission of the Physics Department, University of Münster, 2010 - 2011

Experience

Spin Glasses: An old tool for new problems, School on Spin glasses and its applications in Cargèse, August 2014

Quantum [Un]Speakables II: 50 Years of Bell's Theorem, Conference on Bell's Theorem in Vienna, June. 2014

Complex Quantum Systems Summer School 2013, Summer School on Quantum Foundations and Manipulations in Vienna, Sept. 2013

Erasmus Program, University of York, 2011 - 2012

Founder of the Engineering Project Group, Fridjof-Nansen School, Flensburg, 2004-2008