



SAPIENZA
UNIVERSITÀ DI ROMA

Certification of many-body bosonic interference in 3D photonic chips

Facoltà di Scienze Matematiche Fisiche e Naturali
Dottorato di Ricerca in Fisica – XXX Ciclo

Candidate

Niko Viggianiello
ID number 1246387

Thesis Advisor

Prof. Fabio Sciarrino

A thesis submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics

2014-2017

Thesis not yet defended

Certification of many-body bosonic interference in 3D photonic chips
Ph.D. thesis. Sapienza – University of Rome

© 2017 Niko Viggianiello. All rights reserved

This thesis has been typeset by L^AT_EX and the Sapthesis class.

Version: March 5, 2018

Author's email: niko.viggianiello@gmail.com

*Alla mia famiglia, ai miei amici, alle avventure e alle tragedie.
Grazie per avermi fatto crescere in questi 3 anni*

Contents

List of acronyms	5
Introduction	7
1 Classical and quantum information	12
1.1 Classical computation	12
1.1.1 The concept of algorithm	12
1.1.2 The Church-Turing thesis	14
1.1.3 Machine Learning	16
1.1.4 Computational complexity	18
1.2 Quantum computation	22
1.2.1 Qubit and density matrix	22
1.2.2 Entanglement	25
1.2.3 Circuit model of computation	26
1.2.4 Logic Quantum gates	28
1.3 Quantum algorithms	29
1.3.1 Overview on quantum algorithms	29
1.4 Quantum computation with linear optics	31
1.5 Discussion	34
2 Multi-bosons interference and Boson Sampling model	35
2.1 Indistinguishable Particles	35
2.2 Hong-Ou-Mandel effect	37
2.3 The Boson Sampling model	40
2.4 The computational complexity of Boson Sampling	43
2.4.1 Experimental requirements for Boson Sampling	45
2.5 Scattershot Boson Sampling	46
2.6 Discussion	49
3 Linear optics experiments with integrated photonics	50
3.1 Experimental realization of any unitary matrix	50
3.2 Two-dimensional waveguides	51
3.3 Integrated circuit fabrication techniques	52
3.3.1 Lithography	52
3.3.2 Femtosecond laser micromachining	54
3.4 Directional couplers properties	55
3.4.1 Tunable phases	56
3.5 Description of the experimental tools used for generation, manipulation and collection of single photons	57

4	The quest for the best photonic platform implementation	61
4.1	Rectangular decomposition	62
4.2	Fast decomposition	63
4.2.1	Hadamard matrices	66
4.3	Tomography of unitary matrices implemented with Fast decomposition	67
4.4	Benchmarking performances for different implementations	71
5	Quantum certification of many-body bosonic interference in photonic chips	81
5.1	Experimental Scattershot Boson Sampling	82
5.2	The problem of certification	85
5.3	Suppression law of quantum photonic states	88
5.4	Observation of Majorization Principle for quantum algorithms via 3-D integrated photonic circuits	89
5.5	Optimal photonic indistinguishability tests in multimode network	94
5.6	Suppression law with Sylvester interferometers	101
6	Exploiting machine learning techniques for validation protocols	105
6.1	Pattern recognition techniques for Boson Sampling validation	105
6.1.1	Adopted clustering techniques	107
6.1.2	Numerical results	108
6.1.3	Generalization for Scattershot Boson Sampling	108
6.2	Experimental statistical signature of many-body quantum interference	110
6.2.1	Generalizing the scheme with random forest classifiers	114
7	Quantum interference with deterministic single photon sources	118
7.1	Introduction to quantum dots	118
7.2	Structure of self-assembled quantum dots	119
7.3	Occupancy states of a quantum dot	119
7.4	Phenomenology of quantum dots in a cavity	121
7.4.1	Considering losses: the Purcell effect	122
7.5	Techniques of cavity QD coupling and fabrication	123
7.6	Brightness of quantum dots sources	124
7.7	Experimental properties and details to obtain photoluminescence (PL) emission	125
7.8	Merging deterministic single photon sources with integrated devices	127
7.8.1	Tritter specifications	127
7.8.2	Setup of quantum dot sources and analysis.	130
7.8.3	Tritter apparatus and analysis of experimental data	132
7.9	Discussion and future developments	139
	Conclusions	141
8	Appendices	144
	Appendices	144
8.1	Appendix Note 1: Multiphoton emission terms [5.3]	144
8.2	Appendix Note 2: Interferometer design in the 2^p -modes case [5.3]	145
8.3	Appendix Note 3: Discrimination from classical light [5.3]	146
8.4	Appendix Note 4: Classes for 2-photon output probabilities with Fourier interferometers [5.4]	146
8.5	Appendix Note 5: Majorization and fidelity of the operation [5.4]	148
8.6	Appendix Note 6: Optimal distinguishability for different n and m [5.5]	150
8.7	Appendix Note 8: Total variation distance for 3D interferometers with fast architecture	152

8.8	Appendix Note 9: Bayesian validation test	153
8.9	Appendix Note 10: Bayesian test on the experimental data	153
8.10	Appendix Note 11: Partial photon distinguishability in the binary Bayesian test	154
8.11	Appendix Note 12: Bayesian test on scattershot Boson Sampling experimental data	154
8.12	Appendix Note 13: Efficiency of the different pattern recognition techniques [6.1]	157
8.13	Appendix Note 14: Finding the optimal number of clusters for K-Means [6.1] . .	157
8.14	Appendix Note 15: Halting condition for hierarchical clustering [6.1]	158
8.15	Appendix Note 16: Reshuffling of the experimental data [6.1]	159
8.16	Appendix Note 17: Assigning new data with machine learning techniques [6.2] .	159
8.17	Appendix Note 18: Suppression laws for Sylvester interferometers [5.6]	161
	Bibliography	169
	List of publications	191

List of acronyms

AA	Aaronson and Arkhipov
APD	Avalanche Photodiode
BB	Barak and Ben-Aryeh
BBO	Beta-barium borate
BiBO	Beta-bismuth borate
BPP	Bounded error Probabilistic Polynomial time
BQP	Bounded error Quantum Polynomial time
BS	Beam splitter
CNOT	Controlled-Not
CTT	Church-Turing Thesis
CV	Coefficient of Variation
DC	Directional coupler
DFT	Discrete Fourier Transform
ECTT	Extended Church Turing Thesis
FBS	Fused fiber beam splitter
FFT	Fast Fourier Transform
FLM	Femtosecond laser micromachining
FWM	Four-wave mixing
HOM	Hong-Ou-Mandel
HWP	Half Wave Plate
KLM	Knill-Laflamme-Milburn
LOQC	Linear Optics Quantum Computation
MF	Mean Field
MOP	Multiplication of Pulses
MP	Majorization Principle
MZ	Mach-Zehnder interferometer

NM	Normalized Mean
NMR	Nuclear Magnetic Resonance
PBS	Polarizing beam splitter
PCTT	Physical Church-Turing thesis
PL	Photoluminescence
PPD	Passive Probabilistic Demultiplexer
PS	Phase Shifter
PTM	Probabilistic Turing Machine
QD	Quantum dot
QFT	Quantum Fourier Transform
qFFT	Quantum Fast Fourier Transform
QTM	Quantum Turing Machine
QWP	Quarter Wave Plate
RFC	Random Forest Classifier
RMT	Random Matrix Theory
RNE	Row Norm Estimator
SBS	Scattershot Boson Sampling
SCTT	Strong Church-Turing thesis
SoS	Silica-on-Silicon
SPDC	Spontaneous Parametric Down Conversion
TM	Turing Machine
TVD	Total Variation Distance
UTM	Universal Turing Machine

Introduction

Quantum information and quantum optics have reached several milestones during the last two decades. Starting from the 1980s, when Feynman and Deutsch [1,2] laid the foundations of quantum computation and information, in the last years there have been significant progresses both in theoretical and experimental aspects. A series of quantum algorithms has been proposed that promise computational speed-up with respect to its classical counterpart. If fully exploited, quantum computers are expected to be able to markedly outperform classical ones in several specific tasks. More generally, quantum computers would change the paradigm of what we currently consider efficiently computable, being based on a completely different way to encode and elaborate data, which relies on the unique properties of quantum mechanics such as linear superposition and entanglement. The building block of quantum computation is the *qubit*, which incorporates in its definition the revolutionary aspects that would enable overcoming classical computation in terms of efficiency and security. However recent developments in technologies claimed the realizations of devices with hundreds of controllable qubits, provoking an important debate of what exactly is a quantum computing process and how to unambiguously recognize the presence of a quantum speed-up [3]. Nevertheless, the question of what exactly makes a quantum computer faster than a classical one has currently no clear answer. Its applications could spread from cryptography, with a significant enhancement in terms of security, to communication and simulation of quantum systems. In particular, in the latter case it was shown by Feynman [1] that some problems in quantum mechanics are intractable by means of only classical approaches, due to the exponential increase in the dimension of the Hilbert space. Clearly the question of where quantum capabilities in computation are significant is still open and the hindrance to answer to these problems brought the scientific community to focus its efforts in trying to develop these kind of systems. As a consequence, significant progresses have been made in trapped ions [4,5], superconducting circuits [6,7], neutral atoms [8,9] and linear optics [10–12] permitting the first implementations of such devices.

Among all the scheme introduced, the approach suggested by linear optics, uses photons to encode information and is believed to be promising in most tasks. For instance, photons are important for quantum communication and cryptography protocols because of their natural tendency to behave as "flying" qubits. Moreover, with identical properties (energy, polarization, spatial and temporal profiles), indistinguishable photons can interfere with each other due to their boson nature. These features have a direct application in the task of performing quantum protocols. In fact they are suitable for several recent scheme such as for example graph- and cluster-state photonic quantum computation [10,11]. In particular, it has been proved that universal quantum computation is possible using only simple optical elements, single photon sources, number resolving photo-detectors and adaptative measurements (KLM scheme) [12], thus confirming the pivotal importance of these particles. Although the importance of linear optics has been confirmed in the last decades, its potentialities were already anticipated years before when (1) Burnham et al. [13] discovered the Spontaneous Parametric Down-Conversion (SPDC), (2) Hong, Ou and Mandel discovered the namesake effect (HOM) [14] and (3) Reck et al. showed how a particular combination of simple optical elements can reproduce any unitary transformation [15]. (1) SPDC consists in the generation of entangled photon pairs through a nonlinear crystal pumped with a strong laser and despite recent advancements in other

approaches [16–18], it has been the keystone of single photon generation for several years [19], due to the possibility to create entangled photon pairs with high spectral correlation. (2) The HOM effect demonstrated the tendency of indistinguishable photon pairs to "bunch" in the same output port of a balanced beam splitter, *de facto* showing a signature of quantum interference. Finally, (3) the capability to realize any unitary operation in the space of the occupation modes led to the identification of interferometers as pivotal objects for quantum information protocols with linear optics [15].

At this point, once recognized the importance of all these ingredients, linear optics aimed to reach large implementations to perform protocols with a concrete quantum advantage. Unfortunately, the methods exploited by bulk optics suffer of strong mechanical instabilities, which prevent a transition to large-size experiments. The need for both stability and scalability has led to the miniaturization of such bulk optical devices. Several techniques have been employed to reach this goal, such as lithographic processes and implementations on silica materials [20–23]. All these approaches are significant in terms of stability and ease of manipulation, but they are still expensive in terms of costs and fabrication time and, moreover, they do not permit to exploit the 3D dimension to realize more complex platforms. A powerful approach to transfer linear optical elements on an integrated photonic platform able to overcome these limitations has been recognized in the femtosecond laser micromachining (FLM) [24–26]. FLM, developed in the last two decades, exploits the mechanism of non-linear absorption in a medium with focused femtosecond pulses to design arbitrary 3D structures inside an optical substrate. Miniaturized beam splitters and phase shifters are then realized inducing a localized change in the refractive index of the medium. This technique allows to write complex 3D circuits by moving the sample along the desired path at constant velocity, perpendicularly with respect to the laser beam. 3D structures can also be realized either polarization sensitive or insensitive, due to the low birefringence of the material used (borosilicate glass), enabling polarization-encoded qubits and polarization-entangled photons to realize protocol of quantum computation [27, 28]. As a consequence, integrated photonics gives us a starting point to implement quantum simulation processes in a very stable configuration. This feature could pave the way to investigate larger size experiments, where a higher number of photons and optical elements are involved.

Recently, it has been suggested that many-particle bosonic interference can be used as a testing tool for the computational power of quantum computers and quantum simulators. Despite the important constraints that we need to satisfy to build a universal quantum computer [29] and perform quantum computation in linear optics [12, 30–32], bosonic statistics finds a new promising simpler application in pinpointing the ingredients for a quantum advantage. In this context, an interesting model was recently introduced: the Boson Sampling problem [33]. This model exploits the evolution of indistinguishable bosons into an optical interferometer described by an unitary transformation and it consists in sampling from its output distribution. The core behind this model is the many-body boson interference: although measuring the outcomes seems to be easy to perform, simulating the output of this device, is believed to be intrinsically hard classically in terms of physical resources and time, even approximatively. For this reason Boson Sampling captured the interest of the optical community, which concentrated its efforts to realize experimentally this kind of platforms. This phenomenon can be interpreted as a generalization of the Hong-Ou-Mandel effect of a n -photon state that interferes into an m -mode interferometer. In principle, if we are able to reach large dimensions (in n and m), this method can provide the first evidence of quantum over classical advantage and, moreover, it could open the way to the implementation of quantum computation based on quantum interference. Although the path seems promising, this approach has non-trivial drawbacks. First, **(a)** we need to reach large scale implementations in order to observe quantum advantage, so how can we scale them up? There are two roads that we can follow: (a1) to scale with the number of modes with the techniques developed in integrated photonics, trying to find the best implementation for our interferometers in terms of robustness against losses and choosing the best implementation, or (a2) to scale up the number of photons, identifying appropriate sources for this task. Second, **(b)** in order to

perform quantum protocols we should "trust" on the effective true interference that is supposed to occur the protagonist of the phenomenon. For large-scale implementations, simulating the physical behaviour by means of classical approaches, becomes quickly intractable. In this case the road that we chose is (1) to identify the transformation that are optimal in discriminating true photon interference and (2) to use classification protocols as machine learning techniques and statistical tools to extract information and correlations from output data. In Fig. 0.1 we report the conceptual scheme of the motivations and the problems which we analysed in this Thesis.

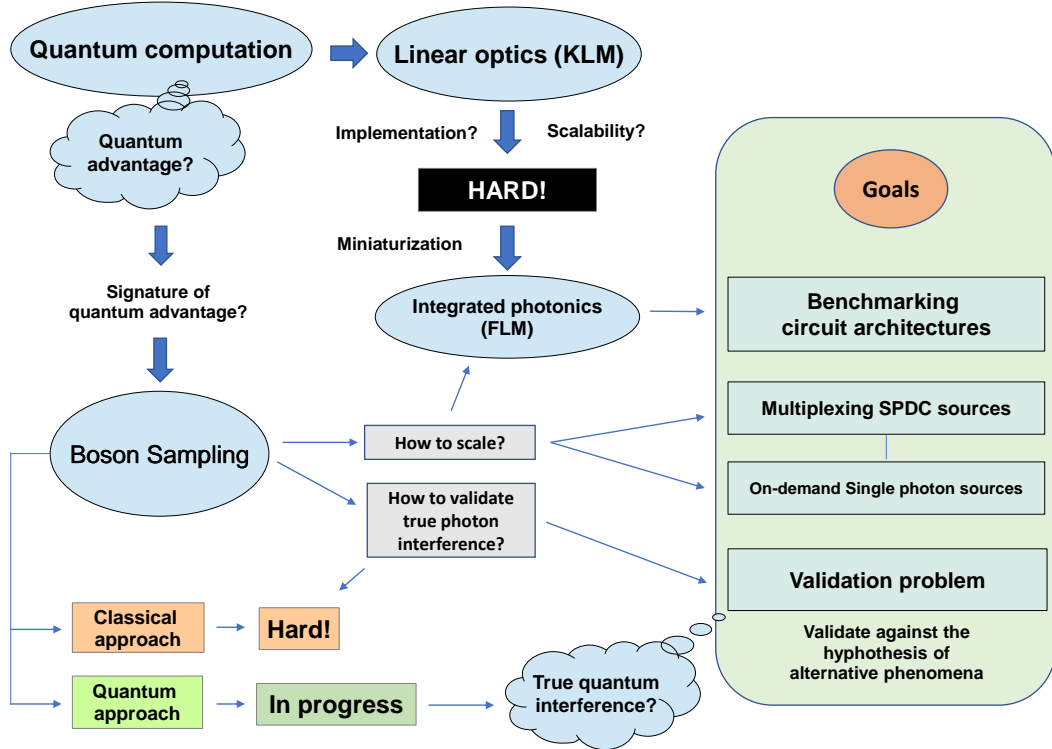


Figure 0.1. Conceptual scheme of the thesis: The demonstration the supposed quantum over classical advantage of quantum computers seems to be difficult to achieve. Boson Sampling is a promising model which attempts to show this advantage in a simpler experimental realization. In this thesis we performed our experiments on photonic platforms, focusing on the problem of "how we can scale up our systems" with optical devices by exploiting SPDC sources in parallel (multiplexing) and deterministic sources (on-demand single photon sources). Successively we focus on the problem "how to identify signatures of true photon interference in an efficient and unambiguous way, paying attention to the scalability of both the photonic implementation (benchmarking architectures) and computational resources (validation problem).

Following these premises, the main goal of this thesis is to address a solution to these problems by following the suggested paths. Firstly, we will give an overview of the theoretical and experimental tools used and, secondly, we will show the subsequent analyses that we have carried out. Regarding point (a1) we performed several analyses under broad and realistic conditions. We studied quantitatively the difference between the three known architectures to identify which scheme is more appropriate for the realization of unitary transformations in our interferometers, in terms of scalability and robustness to losses and noise. We also studied the problem comparing our results to the recent developments in integrated photonics. Regarding point (a2) we studied different experimental realizations which seem promising for scaling up both the number of photons and the performances of the quantum device. First, we used multiple SPDC sources

to improve the generation rate of single photons. Second, we performed an analysis on the performances of on-demand single-photon sources using a 3-mode integrated photonic circuit and quantum dots as deterministic single photon sources. This investigation has been carried out in a collaboration with the Optic of Semiconductor nanoStructures Group (GOSS) led by Prof. Pascale Senellart in Laboratoire de Photonique et de Nanostructures (C2N).

Finally, we focused on problem **(b)** trying to answer the question of how to validate genuine multi-photon interference in an efficient way. Using optical chips built with FLM we performed several experiments based on protocols suitable for the problem. We performed an analysis on finding the optimal transformations for identifying genuine quantum interference. For this scope, we employed different figures of merit as Total Variation Distance (TVD) and Bayesian tests to exclude alternative hypotheses on the experimental data. The result of these analysis is the identification of two unitaries which belong to the class of Hadamard matrices, namely the *Fourier* and *Sylvester* transformations. Thanks to the unique properties associated to the symmetries of these unitaries, we are able to formalize rules to identify real photon interference, the so-called *zero-transmission* laws, by looking at specific outputs of the interferometers which are efficiently predictable. Subsequently, we will further investigate on the validation problem by looking at the target from a different perspective. We will exploit two roads: retrieving signatures of quantum interference through machine learning classification techniques and extracting information from the experimental data by means of statistical tools. These approaches are based on choosing training samples from data which are used as reference in order to classify the whole set of output data accordingly, in this case, to its physical behaviour. In this way we are able to rule out against alternative hypotheses not based on true quantum interference.

In the following, I will present a summary of the work that I have carried out, giving a brief description of each chapter:

- In Chapter 1 we will introduce briefly the framework of this work and the basics that we need to introduce our works. In particular we will focus on the concept of computability, the concept of algorithm with a short description of machine learning. Then we will give the basics of quantum computation and quantum information describing the main developments.
- In Chapter 2 we will introduce the topic of multi-boson indistinguishability, discussing the importance of linear optics for quantum computation and introducing Boson Sampling as a pivotal model to study the quantum over classical advantage problem by means of photonic platforms.
- In Chapter 3 we will argue on the techniques we need to realize miniaturized interferometers, in particular femtosecond laser micromachining, and subsequently we will provide the experimental tools that we need to generate and manipulate photons for our experiments.
- In Chapter 4 we will focus on how to implement unitary transformations, identifying the parameters that are pivotal for a scalable implementation. Moreover, we will present a comparison between two universal schemes and a new, highly-specialized one that enables more robust and efficient implementations for specific relevant transformations.
- Finally in Chapter 5 and in Chapter 6 we will discuss the validation problem, starting with a little overview of the context and later on introducing the different models that I have investigated through these years where several validation protocols have been performed. In particular Chapter 5 focus on individuating and exploiting particular matrices with specific symmetries which highlights quantum interference, and in Chapter 6 we rather focus on pattern recognition techniques and statistical tools to cluster data according to particles behaviour. We conclude with a work done in collaboration with the group of Pascale Senellart in C2N, where we exploited quantum dots as on-demand single photons source, verifying with a tunable optical chip the indistinguishability of such sources (Chapter 7).

Furthermore, at the end of this manuscript I provide an Appendix Chapter in order to give to the reader, a deeper analysis of the models and techniques that we developed so far, to allow for a fully comprehension of the adopted procedures. Finally I want to acknowledge the groups that collaborated to our works, in particular Politecnico di Milano and IFN-CNR of Milano which provided us the optical circuits to perform our experiments with the technique of femtosecond laser micromachining, the Universidade Federal Fluminense and the Perimeter Institute for Theoretical Physics for carrying out part of numerical simulations, the GOSS group led by Pascale Senellart in C2N for the work related with deterministic single photon sources and last but not the least the Quantum Information Lab. of Rome where I worked for the whole period of my PhD project. My work has been supported by fundings QUCHIP and ERC 3D-QUEST.

Chapter 1

Classical and quantum information

In this chapter, we introduce the basic concepts of classical and quantum computation theory which are at the basis of this work, focusing on the points that will be useful in next chapters. We will describe the basic notions of classical information, defining the concept of algorithm and the formalism of the computational class theory. Then we briefly introduce the main topics of machine learning theory and its applications. Finally, we will discuss some useful quantum information and computation tools, which allow us to describe general properties of a quantum computer before moving to a more complete description of this device in the context of linear optics.

1.1 Classical computation

In the last century the theory of computation was separated from mathematics and it became an independent academic discipline. Starting from the pioneering works during the second World War by Alonzo Church [34] and Alan Turing [35], the improvement of classical computation led to a proliferation of works diffused in different areas and to new developments in technologies. The basic idea was that to translate mathematical operations into an automatic process, but there are problems regarding the introduction of a formal definition. Moreover, Gödel in 1930 enunciated the famous *incompleteness theorem* [36], stating that it is impossible to create a complete and coherent mathematical system. More formally, he proved that for any computable axiomatic system powerful enough to describe the arithmetic of the natural numbers: (1) if the system is consistent, it cannot be complete and (2) the consistency of the axioms cannot be proved within the system. The question that arise is: what can I define as computable? To answer the question we will now introduce the concept of algorithm in the next section.

1.1.1 The concept of algorithm

The goal of the theory of computation, is to find which kind of functions can be calculated through an automatic process, thus introducing the concept of sequential computation. Information and data are represented by means of dichotomic digits, the *bits*, with binary numeration (0,1) particularly suitable for technology. They can be seen from a hardware perspective as states ON/OFF of an electric device. Consequently, the procedure used to combine these states defines strings of instructions for specific purposes. It is thus important to define in a formal and rigorous way the intuitive idea of computable function. From one side, the approach is to focus on the concept of computability, trying to identify which classes of problems are theoretically solvable. From the other side, we want to determine which problems can be solved by a computer without specific constraints imposed by the size of the problem itself.

Thus, let us now introduce the concept of algorithm. An *algorithm* is a procedure, expressed in its most general and high-level form, to solve unambiguously, with a finite sequence of

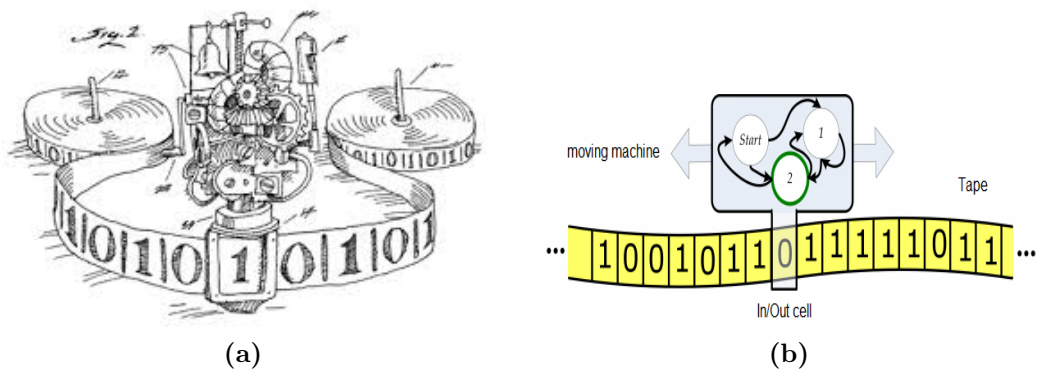


Figure 1.1. Turing machine. Schematic representation of a Turing machine. The calculation begins with the Start internal state and the head pointing to the first cell of the tape, with the input encoded in the tape status. Then, at each step, the internal status changes and the head moves, changing the tape content, according to the program lines.

instructions, a specific problem. A *computing machine* is a system which maps an input to an output, deterministically or governed by a probability distribution. Algorithms and computer machines work by performing a sequential number of operations which define a given computation. Nevertheless, for each algorithm in order to formally define the process, it must exist a model which describe the "behaviour" of the procedure. Basically, we need to define a model of computation: in computability theory and computational complexity theory, a model of computation is the definition of the set of allowable operations used and their respective costs. This is also useful to analyze the cost of such algorithms in order to measure the complexity of the procedure in execution time and memory space. In other words, by assuming a certain model of computation, it is possible to analyze the computational resources required or to discuss the limitations of algorithms or computers. There are several models of computation suitable for different kinds of algorithms and applications such as Markov approach, combinatorial logic, λ -calculus and register machines [37]. For the sake of brevity but completeness we focus only on the so-called *Turing machines* which is the basis of every model of computation. In fact, it is conjectured that every algorithm defined under a specific computational model can be mapped to a problem solvable with the formalism of a Turing machine with at most a polynomial overhead in the number of computational resources. For a more exhaustive description on the computational complexity we refer to the section 1.1.4. Let us thus introduce the Turing machine (TM) as the formal building block of the theory of computation.

A Turing Machine (TM) is constituted by four elements:

- An infinite **tape**, composed by cell. Each cell is filled with a final set of symbols $\{a_1, \dots, a_m\}$ belonging to a given alphabet Σ comprehensive of *blank* symbol for no active operation. For example $\Sigma = \{0, 1, \text{blank}\}$. A read-write head points at a specific cell during each calculation step. After each step it can read and change the symbols of the current cell, and move left or right to the adjacent tape cell.
- A **finite state control**, which is an internal state of the machine that can span over a finite set of states $S = \{s_1, \dots, s_f, H\}$, comprehensive of an initial state s_0 and a final state s_f . H is the *Halt* state which ends the execution.
- A **program**, i.e. a finite list of lines structured by a sequence of determined functional operations:

$$\bar{s} = f_s(s, a) \quad \bar{a} = f_a(s, a) \quad d = f_d(s, a) \quad (1.1)$$

In other words, given a specific set of inputs and configurations, the machine changes its state and move to another cell through simple discrete instructions set by the algorithm, until the *Halt* command.

$$t_i : s_i, a_i \implies t_{i+1} : \bar{s}_i, \bar{a}_i, d_i \quad \forall i \quad (1.2)$$

It is worth noting that this model represents the first formal theory on the sequential processing and accomplishes determined operations step by step following the path ruled by the algorithm. However, it is possible to define another type of **TM** which introduces a probabilistic approach to the problem and that is believed to solve specific problems even faster. This model is named *probabilistic Turing machine* (**PTM**) and differs from the previous one because the choice between available transitions at each point is performed according to some probability distribution:

$$\bar{s}_2, \bar{a}_2, d_2 \xleftarrow{1-p} s_i, a_i \xrightarrow{p} \bar{s}_1, \bar{a}_1, d_1 \quad (1.3)$$

One of the central questions of complexity theory is whether randomness adds power; that is, is there a problem which can be solved in polynomial time by a probabilistic Turing machine but not a deterministic Turing machine? Or can deterministic Turing machines efficiently simulate all **PTMs** with at most a polynomial slowdown? We will discuss more in detail the implication of this statement in section 1.2.

Let us introduce, in the next section, the motivations that brought to the need of finding a new approach for the computation.

1.1.2 The Church-Turing thesis

It is then with the introduction of a general model theory about **TMs** and the realization of the first transistors, that we lay the foundation of modern computability. Let us thus introduce the so-called *Universal Turing machine* **UTM** which incorporates the functioning of all the other **TMs**:

Definition 1.1: Universal Turing machine

Let T be a generic Turing machine and let n_T an integer number which identifies univocally T for a specific algorithm, we define Universal Turing machine U as:

$$U(n_T, x) = T(x). \quad (1.4)$$

Specifically, given an input x and an identifier n_T , the universal Turing machine returns a specific Turing machine T that will return its output.

The idea is that, the program, the alphabet and the set of states are fixed, while the first part of the initial state of the tape encodes the functioning of the **TM** to be simulated. This formalism, once extended to a bigger class of problems, leads to a rigorous approach to the theory of computation. This extension is embodied into the two independent seminal works of Alonso Church [34] and Alan Turing [35] to identify the class of algorithms which are "computable", which are usually combined in the conjecture of *Church-Turing thesis* (**CTT**):

"The class of functions computable by a Turing machine corresponds exactly to the class of functions we would naturally regard as being computable by an algorithm."

which means that all possible formalizations of the intuitive notion of computing are equivalent, and can be fully captured by the **TM** formalism. This strong statement constituted the basement

for all subsequent developments of computational complexity theory. But what exactly does computable mean?

There is still some ambiguity in the concept of what is computable and what is not. In 80s, David Deutsch tried to reinterpret this statement in a more *physical* way. Indeed, he switched the view from *computation* to *simulation*: in its own reformulation of the **CTT**, a computing machine is able to perfectly simulate a physical system if the two are computationally equivalent under a suitable choice of the program and the input-output labeling of the computing machine (or, in other words, if one cannot be able to distinguish a black box containing a physical system or its simulator just by looking at its inputs and outputs). So, he formulates the new version of the conjecture giving a wide breath on the possible implementations by physical means. The new conjecture, named "*physical Church-Turing thesis*" (**PCTT**) states that

"Every finitely realizable physical system can be perfectly simulated by a universal model computing machine operating by finite means".

At this point the attention started to focus on computer performances. The notion of efficient calculation, which is the main criteria for understanding the complexity behind a computational task, is of pivotal importance for modern computational theory. At the same time, it relies on a quite strong simplification: an algorithm is considered efficient if the required resources for its execution scale polynomially with the size of the input. The simplification lies obviously in the fact that nothing is said about the order of the polynomial; still, this distinction between "easy" (polynomial) and "hard" (exponential) computational problems seems to capture the main picture of the panorama of algorithmic complexity. A more detailed description on the field of computational classes will be introduced in section 1.1.4.

Moreover in 1960s and 1970s there were several studies to find possible alternatives to the Turing model, but it was observed that any computational model can be reproduced or "simulated" by a Turing machine with at least a polynomial overhead. This implication suggest an immediate change on **CTT** by introducing the *Strong Church-Turing thesis* (**SCTT**):

"Any algorithmic process can be simulated efficiently using a Turing machine."

It is worth noting that this strong implication relies on the word *efficiency*. Indeed, if the conjecture holds true, no matters what kind of model or device we use to perform our computation, that machine can be simulated efficiently by means of Turing machines, thus allowing to concentrate the analysis on computational efficiency of programs only to the **TM** model.

In the last 50 years there were several attacks on **SCTT**, but the first threat to the conjecture arrived from Robert Solovay and Volker Strassen in 70s, which exploited randomness as crucial part in the implementation of the algorithms [38] to retrieve the final result with a good approximation in an efficient way. Consequently, probabilistic programs pose a challenge to **SCTT** suggesting that there are problem that can be solved efficiently in probabilistic machines, while deterministic Turing machines cannot. Nevertheless the problem is easily solvable by introducing a little change in the **SCTT**. Combining the physical reinterpretation of the **PCTT** and adding a strong contribution due to the randomness, we can formulate the so-called *Extended Church-Turing thesis* (**ECTT**) which is comprehensive of both claims:

"Any physical system can be simulated on a probabilistic Turing machine with at most a polynomial overhead in the size of the system to be simulated"

Such statement has very strong computational and scientific implications, since it says that every aspect of Nature can be captured by a **UTM**. It was Feynmann's and Deutsch's works in early 80s that exposed how quantum mechanical processes can open the door to a new paradigm of computational which cannot be framed by the Turing machine methods. We will introduce, in

more details in the next sections, the quantum computation and the reason why this approach can fulfil the request to overcome the [ECTT](#).

Let's do for the moment a step back by discussing the principal discoveries of modern computation. In the last decades, the developments of this area both in theoretical and technological aspects, brought to a progressive proliferation of important topics. Computer science and telecommunications introduced the so-called "Society of Information", as a consequence of this development, there have been several needs to satisfy, as for example the enormous amount of data and information that spread quickly. The need to control this significant quantity of data, brought computer scientists to develop algorithms and techniques that aim to classify and reinterpret them in an efficient way. To mention the most diffused we can recognize techniques such as Bayesian inference, data mining and machine learning. Following this road, we will introduce the latter since it represents a frontier application of computer science which is believed to be pivotal for future applications in society. We give a non-exhaustive description of the *machine learning* field, and later we give more details on the computational classes, which represent the ingredients to understand how to frame quantum computation respect classical one from the point of view of computational efficiency.

1.1.3 Machine Learning

Computation theory developed extensively during the last six decades. Basic protocols have been widely used in order to elaborate and properly classify data in several fields of science. The next step is to evolve the way we think about computation, providing something which is beyond the simple concept of data processing. Is it possible to build algorithms which are able to learn from data itself? This question introduces the idea of *machine learning*: which aims to comprehend the behaviour hidden in raw data, to give a prediction on a subsequent process. Machine learning finds several applications in computer science, in the development of search engines and anti-spam control, in neural networks, in math economy, or in the analysis of large markets.

The advantage of machine learning with respect to fields as data mining and statistical analysis, is the possibility to avoid some statistical constraints imposed by the previous models. Hence, we have more freedom in analyzing complex real-life data sets, where samples are dependent, where there is excess noise, and where the distribution is entirely unknown or skewed. It was in 1997 that Mitchell introduced a formal definition of what is machine learning [39]:

Definition 1.2: Machine learning

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T , as measured by P , improves with experience E .

This definition is meant to be more operational rather than in cognitive terms. Basically it underlines the question: "Can machines do what we, as thinking entities, can do?". In the proposal he explores the various characteristics that could be possessed by a thinking machine. The second question is "How can we represent data in order to be 'understood' by learning algorithms?". The answer relies on a formal encoding of data sets through features. A feature is a measurable heuristic property of the phenomena. Learning algorithms work with features, therefore a careful selection of features will lead to a better model. [40] The multidisciplinary nature of machine learning is reflected in how features are viewed. There exist two kind of feature representations:

- **Geometric view:** we can represent features as tuples, vectors in a high dimensional space (the feature space). Thus features are grouped into a feature vector. Let d denote the

number of features. One vector of the canonical basis $\{e_1, e_2, \dots, e_d\}$ of \mathcal{R}^d is assigned to each sample j . Let x_{ij} be the weight of a feature i in a data sample j . Thus, a feature vector x_j for the object j is a linear combination of the canonical basis vectors:

$$x_j = \sum_{i=1}^d x_{ij} \mathbf{e}_i \quad (1.5)$$

By writing x_j as a column vector, we have $x_j = (x_{1j}, x_{2j}, \dots, x_{dj})$. For a collection of N data instances, the x_{ij} weights form a $d \times N$ matrix. Since the basis vectors of the canonical basis are perpendicular to one another, this implies the (often violated) assumption that the features are mutually independent. The assignment of features to vectors depends on user choice and thus it is arbitrary. Consequently, a feature may be assigned to any of the vectors of the canonical basis. With use of the geometric view, distance functions, norms of vectors, and angles help in the design of learning algorithms.

- **Probabilistic view:** In this case, each feature is a random variable, defined as a function that assigns a real number to every outcome of an experiment. A discrete or continuous random variable takes any of a specified finite or countable list of values which are bounded by a specific interval. Similarly to the representation of features group into a feature vector in the geometric view, in this case, the probabilistic view has a multivariate random variable for each data instance: (X_1, X_2, \dots, X_d) . A joint probability function or density function describes the distribution. The random variables are independent if and only if the joint probability decomposes to the product of the constituent distributions for every value of the random variables:

$$P(X_1, X_2, \dots, X_d) = P(X_1) P(X_2) \cdots P(X_d). \quad (1.6)$$

In both cases, collecting more features should enhance the discrimination power and thus the effectiveness in the learning process, however it has been shown that only relevant features helps the algorithm to converge.

Irrespective of the approach taken, we can distinguish two major group of machine learning algorithms (see Fig. 1.2):

- *Supervised learning:* In this group, algorithms work manipulating labeled data points. We denote a label by y . The training set is thus a collection of pairs of data points and corresponding labels: $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, where N is the number of training instances. Labels may have a discrete or continuous numerical range, leading respectively to a regression or classification problem. A supervised learner predicts the label of instances after training on a sample of labeled examples, the training set. At a high level, supervised learning is about fitting a predefined multivariate function to a set of points.
- *Unsupervised learning:* in this scenario the labels are missing. The task is to find structure in the samples on its own. Most algorithms belonging to this group, attempt to clustering data instances by means of discriminator function such as the Euclidean distance between data. The clusters may be embedded in one another, and the density of data instances often varies across the feature space; thus, clustering is a hard problem to solve in general.

We can also consider the merged case where input data is a mixture of labeled and unlabelled examples. In this case the model must learn the structures to organize the data as well as make predictions. Consequently, there are several branches that defines the widespread zoo of machine learning algorithms [39, 40]: *regression* is concerned with modelling the relationship between variables that is iteratively refined using a measure of error in the predictions made by the

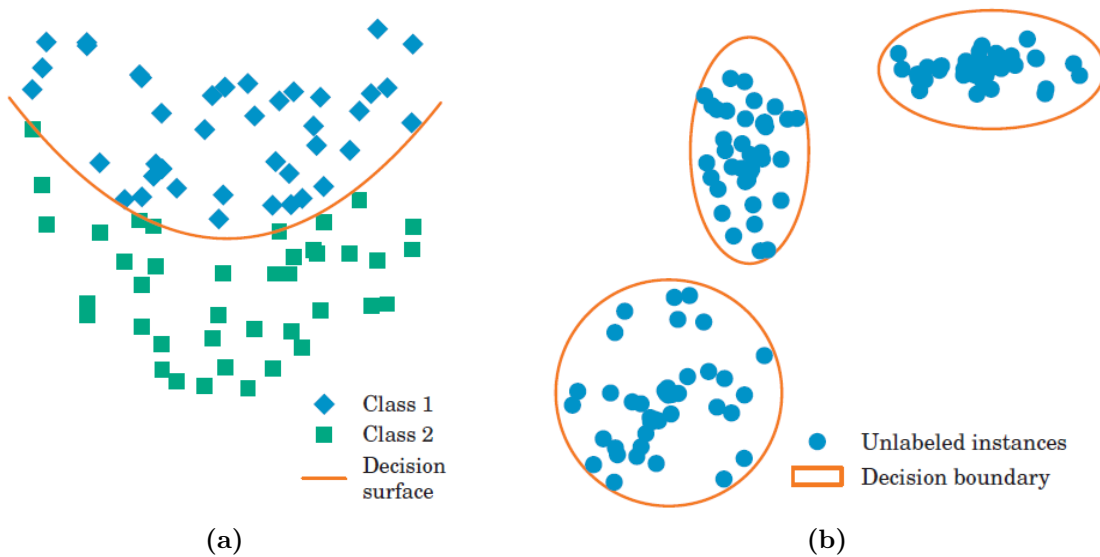


Figure 1.2. Unsupervised and supervised learning. a) In supervised scenario, we provide a labeled training set of data which identify a sample of input-output correlations. b) No labels are given to the learning algorithm, leaving it on its own to find structure in its input. Images taken from Ref. [40]

model. *Instance-based learning* models, as the k-Nearest Neighbour, are decision problem that typically compare new data using a similarity measure in order to find the best match and make a prediction. *Decision tree methods* construct a model of decisions based on actual characteristics hidden in the samples. *Bayesian methods* are those that explicitly apply Bayes' Theorem for problems such as classification and regression. *Artificial Neural Networks* are models that are inspired by the structure and/or function of biological neural networks in order to find structures in data. *Deep Learning methods* are a modern update to Artificial Neural Networks that exploit abundant cheap computation. They are concerned with building much larger and more complex neural networks. *Clustering methods*, like regression, describes the class of problem and the class of methods where data are classified in cluster according to their features. Clustering procedures are typically organized by the modelling approaches such as centroid-based and hierarchical. All methods are concerned with using the inherent structures in them to best organize the data into groups of maximum commonality. Honourable mentions for this group are K-Means, K-Medians and Hierarchical Clustering which will be discussed more deeply in Chapter 5, section 6.1.

Nevertheless, the fast development of computation led to an important and tricky question: how efficient can an algorithm be to perform a specific task? What are the resources necessary to implement such algorithms? The answer for all these questions are related to the so-called *complexity classes theory* which we introduce in the next section.

1.1.4 Computational complexity

Complexity theory aims to determine the number of resources required to solve a specific class of computational problems. Therefore, it is necessary to define a "hardness estimator" in order to compare different computational problems between them. First of all we must introduce two typologies of computational classes in which we can identify and divide the algorithms with respect to their hardness in terms of used sources:

- *Polynomial problems*: the time T_c and the number of resources needed to solve a problem is $T_c = O(t^k)$.
- *Superpolynomial problems*: the time T_c and the number of resources needed to solve a

problem grows faster of any power $T_c = o(t^k) \forall k$, and they are also called exponential problems.

For example, sum and multiplication belong to the polynomial class since we need only $O(N^2)$ number of resources, while factorization is a well known problem which is believed to belong to the superpolynomial class. In fact the best classic algorithm grows as $\tau = e^{o(n^{\frac{1}{3}}(\log n)^{\frac{2}{3}})}$. But what does it mean "efficient"? If we want to go to a more formal description of complexity we need to define its meaning. Thus we define efficient computation as:

Definition 1.3: Efficient computation

A computational process is *efficient* if it belongs to **P**, where P is the class of problems for which it exists a Turing machine M and a polynomial p such that $\forall x \in \{0, 1\}^n$ we have that:

1. $M(x)$ stops after a number of steps shorter than $p(|x|)$
2. $M(x)$ accepts x if and only if $x \in L$

Moreover, we can define **NP** (Non-Deterministic Polynomial), as the class of decision problems whose solution can be verified in polynomial resources. A well-known problem in this group is *factoring*, which aims to find non-trivial factors of an integer. It is immediate to see in this case the difference in terms of computational resources between finding a solution (exponential) and verifying at (polynomial).

If we want to determine univocally the efficiency of an algorithm we need to define a metric which does not depend on the development of technology. For this reason, the reference line is assigned to the concept of probabilistic Turing machines, following the **ECTT** enunciate. This means that if we are able to identify with the term "efficient" an algorithm which exploits polynomial resources, the investigation of complexity classes become independent from the model itself by mapping the hardness to a **PTM**.

The next step is to identify specific tools to study the problem associated to the complexity of algorithms. Consequently we introduce the *asymptotic notation*, a methodology used to estimate the principal resources involved. In this case we introduce two classifications of resources

- *Time*: given a **TM** T , we say that T solves a problem in time $g(n)$ if, given an input with length n , T produces an outcome in $g(n)$ steps.
- *Space*: given a **TM** T , we can say that T solves a problem in space $g(n)$ if, given an input with length n , T produces an outcome in $g(n)$ cells, neglecting input-output dimensions.

$g(n)$ must be a growing monotonic function to be a benchmark for the resources in time and space. The asymptotic notation reduces, shortly, to a study of three different quantities:

- **O**: we say that $g(n) = O(f(n))$ if $\exists c \geq 0, n_0 \in \mathbb{N}, : |g(n)| \leq c|f(n)| \forall n > n_0$ which provides an upper bound.
- **Ω** : we say that $g(n) = \Omega(f(n))$ if $\exists c \geq 0, n_0 \in \mathbb{N}, : |g(n)| \geq c|f(n)| \forall n > n_0$ which provides a lower bound.
- **Θ** : we say that $g(n) = \Theta(f(n))$ if $g(n) \in O(f(n))$ and $g(n) \in \Omega(f(n))$ which does not assume any difference in terms of performance in the best or worst case.

These tools provide a starting point to classify algorithms in specific classes defined by the runtime and space required. Before introducing the list of the main classes of computational complexity we have to describe in minimum terms what is the relation between them [41]. Therefore, we define **DSPACE** and **DTIME** as the classes of languages decided by **TM** of respectively of space and time complexity $t(n)$. In the case of Non-Deterministic Turing Machines we can associate the same statement for these two classes at the same way: **NSPACE** and **NTIME**.

We are ready to introduce the main classes. We have already shown the classes **P** and **NP**, and it is easy to show that the same definition holds in the context of **DSPACE** and **DTIME** [41].

$$\mathbf{P} = \cup_{k \geq 1} \mathbf{DTIME}[n^k] \quad \mathbf{NP} = \cup_{k \geq 1} \mathbf{NTIME}[n^k] \quad (1.7)$$

Consequently, we are able to construct a hierarchy of computational classes based on the concepts of time-space resources and boundary conditions, as previously done for the classes **P** and **NP**. Besides, the polynomial scaling we can identify also the **L** and **NL** which scale with a logarithmic number of resources. Finally, let us introduce shortly the main computational classes, which are reported also in Fig. 1.3a:

- **#P**: The class $\#P$ [42] is a class of counting problems of the form “how many solutions are there for a specific NP problem?”. To compare the class $\#P$ to decision problem classes like **P** or **NP** we have to define a decision problem class based on $\#P$. We can do this using the concept of *oracle*, which is defined as a black box able to solve, in a unit time, problems belonging to a specific class. We can define the class A^B as the class of problems which are in class A if we have access to an oracle for problems in class B; a class of this form is $P^{\#P}$, which identifies decision problems solvable in polynomial time by a **TM** with an oracle for $\#P$ problems.
- **PSPACE**: It is the class of decision problems, solvable with a polynomial-size memory (with no restriction on time).
- **EXP**: the class of problems solvable in exponential time.

A decision problem is said to be *A-hard* if every problem in the class A can be reduced to it using polynomial resources, and *A-complete* if it is both in A and in A-hard. There are several computational problems which are fundamental as critical boundaries for a robust definition of hierarchies. For example it exists:

- **NP – complete**: problems in this class are the most difficult of **NP**. It is not known whether the factoring decision problem is **NP**-complete, but the general belief is that this is not the case. This is because **NP**-complete problems in general have multiple solutions, while factoring is unique; another evidence is given by the fact that the fastest algorithms for **NP**-complete problems typically take order of e^n time, while the so-called *number field sieve* is able to factorize an n-bit integer with order of $e^{\frac{n}{3} \log \frac{2}{3}(n)}$ steps.

What is the relationship between classes, and how do we define the boundaries? For instance, $P \subseteq NP$, but up to now there is no formal proof on this statement and it remains one of the great questions in complexity theory. More in general, we know that:

$$\mathbf{P} \subseteq \mathbf{NP} \subseteq \mathbf{PSPACE} \subseteq \mathbf{EXP} \quad (1.8)$$

where at least one inclusion has to be strict, because it can be shown that $\mathbf{P} \neq \mathbf{EXP}$. Let’s individuate how **#P** problems are related to the previously described. The **#P** problems involve checking an exponential amount of different patterns in exponential time with polynomial memory, so that $\mathbf{P} \subseteq \mathbf{P}^{\#P} \subseteq \mathbf{PSPACE}$:

What about the problems that we can hope to solve by introducing probabilistic randomness with **PTMs** [43, 44]? In this case we introduce the class

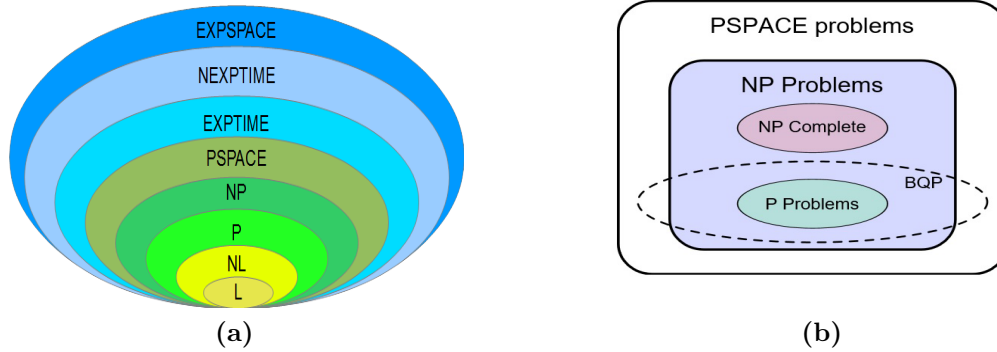


Figure 1.3. Relations between complexity classes. a) Relationship between some of the main complexity classes. b) Zoom on the supposed relations between **BQP** and other computational classes.

- **BPP** (Bounded Probabilistic Polynomial): the class of decision problem solvable by a **PTM** in polynomial time with probability $p \geq \frac{2}{3}$ (so that one can run the algorithm several times and take majority voting). This class is strictly linked to the “physically solvable problems” of the **ECTT**. It is not clear which is the relationship between **BPP** and **NP**, but we do know only the following demonstrated statements [43] $\mathbf{BPP} \subseteq \mathbf{NP}^{\mathbf{NP}}$ and $\mathbf{BPP} \subseteq \mathbf{P}^{\#\mathbf{P}}$.

It is possible to identify a relation between the classes by generalizing **P** and **NP** to oracle machines. This is called the *polynomial hierarchy* (**PH**) of complexity classes, and with **PH** we denote the union of all these infinite levels. It is strongly believed that each of the levels constitute a more powerful class than the lower levels (which justifies the term “hierarchy”) [44].

The quantum analogous of **BPP**, can be introduced by describing the quantum analogue of TM, that is substituted by a Quantum Turing machine **QTM**. We define a **QTM** considering the same road followed with the description of the TM, indeed we have an internal state s , a tape position x , I/O write device and a tape configuration c . The main differences are two: (1) the overall generic state is a superposition of the form $\sum_{i=1} \alpha_i |s, x, c\rangle$ (2) the program is substituted with a unitary operation. Thus, in the case of a **QTM**, after the measure, the superposition collapse to a particular state with probability $|\alpha_i|^2$. It is also possible to define a Universal **QTM** as done for the TM, further details are given in Ref. [46, 47].

Making a parallelism respect to the relation between TMs and **BPP** we say that . is called **BQP** [45]:

- **BQP** (bounded-error quantum polynomial time): is the class of decision problems solvable by a quantum computer in polynomial time, with an error probability of at most 1/3 for all instances. A decision problem is a member of **BQP** if there exists an algorithm for a quantum computer (a quantum algorithm) that solves the decision problem with high probability and is guaranteed to run in polynomial time. It is the computational class of problems solvable by a **TM** with the capability to associate a quantum circuit (see Section 1.25) of polynomial size with respect to the initial state [46].

How do this class compare to the others we mentioned? We can say that it is supposed to hold this relation:

$$\mathbf{P} \subseteq \mathbf{BPP} \subseteq \mathbf{BQP} \quad (1.9)$$

The inclusion with **BPP** depends from the fact that a **QTM** can exploit the intrinsic randomness of measurement outcomes. Nevertheless the best upper bound for the class describing **QTM** has been given in [47]

$$\mathbf{BQP} \subseteq \mathbf{P}^{\#\mathbf{P}} \subseteq \mathbf{PSPACE} \quad (1.10)$$

meaning that every problem solvable in poly-time by a QTM can be solved using poly-space with a classical TM. Combining equations (1.9) and (1.10), we can write:

$$\mathbf{P} \subseteq \mathbf{BPP} \subseteq \mathbf{BQP} \subseteq \mathbf{P}^{\#\mathbf{P}} \subseteq \mathbf{PSPACE} \quad (1.11)$$

Looking at the previous equation, one can see that we do not have any absolute result about the fact that quantum computers are more powerful than classical ones (it is believed that $\mathbf{BPP} \neq \mathbf{BQP}$, but it is still an open problem). In the last years, a great deal of attention has been devoted to investigate the relation between \mathbf{BQP} and \mathbf{NP} (see Fig. 1.3). The common belief is that quantum computers are able to provide an exponential speed-up for some specific \mathbf{NP} problems (like Factoring) by exploiting some internal structure of those problems, which is missing in \mathbf{NP} -complete problems, for which they can provide only a quadratic speed-up.

Let us now introduce formally in the next sections, the concept of quantum computation.

1.2 Quantum computation

Despite the promising development of classical computation, there is an intrinsic fabrication limit when we want to improve the calculation process, because at the state-of-art of technology, miniaturization of the devices is prohibitive due to the constraints imposed from the microscopical realm to the possibility to have deterministic and controllable stages. Therefore, these needs brought to search for alternative theories of computation that are more physical and near to nature behaviour. The field of quantum computing was initiated by the work of Paul Benioff [48], Yuri Manin in 1980 [49], Richard Feynman in 1982 [1] and David Deutsch in 1985 [2]. In particular Feynman in 1982 showed the limits of classical simulation to reproduce the behaviour of the quantum realm due to high difficulties in analysing the system for large scales. He introduced the basic concept of a quantum-like computing which uses quantum mechanics as a tool for the quantum simulation itself. Several theoretical developments have been made in this field, in the last three decade, the superior capabilities of the quantum computer for certain problems.

One could ask, at this point, what is the origin of the supposed computational superiority of quantum computing, where basically the strongest point is marked by the quantum parallelism procedure. In the next section we will introduce the basic tools which constitute the building blocks for quantum computation, superposition and entanglement. It has been argued that it is indeed entanglement, rather than just superposition to provide the basis of the computational complexity of simulating quantum systems [50]. The reason lies on the possibility to reproduce the phenomenon also with classical waves, consequently quantum mechanics is not necessary to reproduce these kind of phenomena. There is no quantum counterpart for entanglement and has been demonstrated [51] that it is necessary for a large class of computational problems to achieve a quantum-over-classical computational speedup. However, there are situations in which a quantum advantage (although small) is present without the use of entanglement [52,53]. In the next section we present the fundamental tools useful for a more detailed description of quantum computation.

1.2.1 Qubit and density matrix

Quantum information and computation are built on the concept of the quantum bit or *qubit*. Qubits are the expression of quantum superposition and they are described as two-level quantum systems $\{0,1\}$. There are several ways to implement a qubit by means of photons [54,55], ion traps [4], NMR [56], NV centers in diamonds [57], atoms [58,59] and superconducting models [6,7,60] This is due to its dichotomic natural predisposition for encoding, consequently it can be represented for example as ground-excited levels in an atom or vertical and horizontal polarizations. A general qubit is represented as a superposition of states in the computational

base $\{0, 1\}$:

$$|\Psi\rangle = \alpha |0\rangle + \beta |1\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad (1.12)$$

where α and β are complex numbers which represent the amplitudes of the quantum state, and for this reason they have to satisfy the normalization relation $|\alpha|^2 + |\beta|^2 = 1$. From (1.12), by means of the normalization condition, it is possible to show an alternative form, that is:

$$|\Psi\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ e^{i\phi} \sin \frac{\theta}{2} \end{pmatrix}. \quad (1.13)$$

The variables θ and ϕ contains the full description of the qubit on the Bloch sphere: θ represents the information regarding the possibility to find a state in a specific eigenstate and the azimuth angle ϕ describes the phase difference between the eigenstates of its computational basis. Any kind of qubit can be defined as a point in the Bloch sphere, even in the inner part. We define *pure* state a point on the surface, while a *mixed* state falls into the sphere itself. The Bloch sphere is formally analogous to the Poincaré sphere used to describe the polarization of radiation. It is pretty intuitive to find a correspondence between the degree of polarization of photons and the degree of purity of the qubits. A mixed state is different respect to the pure case because it is not possible to represent it as a superposition of states as in Eq. (1.12) and in a real experimental condition, the mixed state is the most realistic behaviour description of the phenomenon that we can consider. Nevertheless, to insert mixed states in the theory, we need to introduce the density matrix formalism.

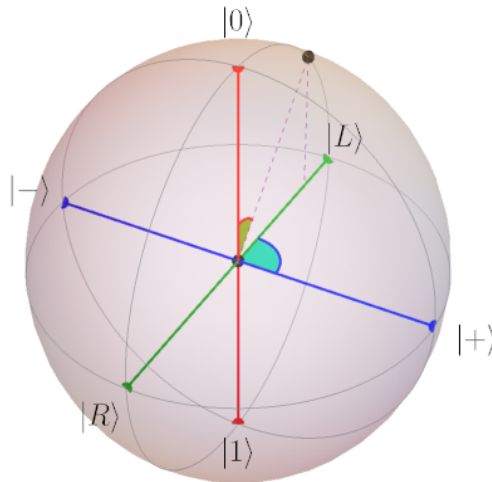


Figure 1.4. Bloch sphere: spherical representation of all possible combinations of a qubit in the computational bases (linear, diagonal and circled). The state lies on the surface of the sphere if it is a pure state. Any qubit built in a laboratory will be necessary an incoherent combinations of pure states (mixed states), thus they lies in the inner region, where the modulus of the vector is minor than 1. The poles represents the eigenstates, the polar angle θ contains the information on the probability to find the system in either one or the other eigenstate, while the azimuth angle ϕ is the phase difference between the two eigenstate components of the qubit.

Suppose to have N copies of a generic state $|\psi_k\rangle$ without any knowledge of the system and to know, instead, the probability that $|\psi_k\rangle$ will be an outcome. If $|\psi\rangle$ is a pure state, then the density matrix is described as:

$$\rho = |\psi\rangle \langle\psi| \quad (1.14)$$

that satisfy the relation $\rho^2 = \rho$. The importance of density matrices holds when we consider an incoherent blend of pure states $\{|\psi_k\rangle\}$, in fact in this case the formula reads:

$$\rho = \sum_k p_k |\psi_k\rangle \langle \psi_k| \quad (1.15)$$

where $0 < p_k \leq 1$ represents the associated probability for each state $|\psi_k\rangle$ and the normalization condition reads $\sum_k p_k = 1$. In this case, dealing with an incoherent ensemble of qubits, we have necessarily the property $\rho^2 \neq \rho$. Despite the purity of the system, there are several rules that we have to impose in order to associate a density matrix:

- $\rho = \rho^\dagger$. The matrix has to be hermitian (observable).
- $Tr(\rho) = 1$. The diagonal elements of ρ represent the populations of pure states in the ensemble, and the trace is the sum of the probabilities in the entire space.
- ρ not negative. All the eigenvalues have to be real.

If the elements ρ_{ii} on the diagonal represent the *population*, that is the probability for which the i^{th} -state has been measured, then the terms ρ_{ij} for $i \neq j$ describe the *coherence* between the i^{th} and the j^{th} state. Consequently coherence term exist if and only if exist both ρ_{ii} and ρ_{jj} . Density matrices fully characterize a closed quantum system and it is a powerful tool to retrieve the value of some observables measured on the system, in fact, given a generic observable A with eigenvalues a_i , we can compute the mean value through the ρ :

$$\langle A \rangle = Tr(\rho A) = \sum_{k=1}^n p_k \langle \psi_k | A | \psi_k \rangle \quad (1.16)$$

It is possible to diagonalize a density matrix in order to retrieve information regarding the purity of the system, through the number of non zero eigenvalues on the diagonal representation. Indeed after the diagonalization, a pure state is described as $\rho = diag(1, 0, \dots, 0)$ while in the mixed case there are several non-zero contributions $\rho = diag(\lambda_1, \dots, \lambda_n)$, where λ_i are the eigenvalues. It is worth noting that it is not possible to write a generic mixed state as a superposition of pure ones. In fact given for example without lose generality a state $\rho = \sum_{a=1}^2 p_a |\psi_a\rangle \langle \psi_a|$, it is always possible to choose another representation $|\phi_1\rangle = \sum_{a=1}^2 \sqrt{p_a} |\psi_a\rangle$ and $|\phi_2\rangle = \sum_{a=1}^2 \sqrt{p_a} |\psi_a\rangle$ where for particular k_a and g_a we get another similar description of the density matrix:

$$\rho = \sum_{a=1}^2 p_a |\psi_a\rangle \langle \psi_a| = \frac{1}{2} \sum_{a=1}^2 |\phi_a\rangle \langle \phi_a|. \quad (1.17)$$

Therefore, there is no unique way to decompose any density matrix into a set of pure states, thus providing the evidence for the need of this formalism.

We can also define the dynamics of a density matrix $\rho = \rho(t)$. Using the Schroedinger equation:

$$\frac{d\rho}{dt} = -\frac{1}{i\hbar} [H, \rho] \quad \implies \quad \rho(t) = U(t, t_0) \rho_0 U^\dagger(t, t_0). \quad (1.18)$$

we see that the generic evolution is provided by a unitary operator in the Heisenberg scheme, thus fully characterizing the system for each time t .

We are now able to use these tools to perform quantum computation. Qubits and density matrices play an important role, because with a proper combination of them in higher dimensions, we are able to realize supposed alternative kinds of computation which are not reproducible by a classical approach. The measurement at the end is retrieved quantum mechanically but the information is classical in the sense that we have to get access to results. From this point of view,

it seems that there is no advantage to explore a parallel computing through the superposition property. To understand this point it can be useful to focus on this example : let's think of a different kind of classical computer, which uses an ensemble of 4 bits as its elementary information cell. The amount of information in the cell would simply be one over $2^4 = 16$ possible states allowed to the 4-bit combination. In other words, to identify the state of n classical bits we would need a single integer $I \in [1, 2^n]$. For quantum states the situation is different: because of the superposition principle, the state of a "qudit", i.e. a D -dimensional quantum system, is described by $2D$ complex numbers which correspond to an amount of hidden information which is exponential with the number of qubits. This is strictly related to the fact that the state of two qubits cannot be in general expressed just as a single product state, but rather as a superposition of product states. There is also another genuine quantum property of fundamental importance in physics and quantum computation which is the *entanglement*. It has no counterpart in the classical world and it is believed that the core of the superiority of quantum computing in terms of processing is due not really at the parallelism but rather on entanglement. For this reason it is fundamental to introduce it.

1.2.2 Entanglement

Let us introduce two systems, A and B , living respectively to two distinct Hilbert spaces \mathcal{H}_A and \mathcal{H}_B . A generic state of the two system is the vector:

$$|\psi\rangle_{AB} = \sum_{ij} a_{ij} |i\rangle_A |j\rangle_B \quad (1.19)$$

where $\{|i\rangle_A\}$ and $\{|j\rangle_B\}$ represent respectively two complete computational bases for the system A and B , and a_{ij} are complex amplitudes which satisfy the probability normalization $\sum_{ij} |a_{ij}|^2 = 1$. This new state belongs to a bigger Hilbert state of dimension $d = d_A \times d_B$ which is the tensor product of the separate one $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$.

It is worth noting that there are two different way to describe a quantum state, besides the density matrices formalism, which strictly depend on the link between degrees of freedom. We can identify two classes: *separable* and *entangled* states. A quantum state is separable when it is possible to decompose it as a product of a wave functions for the first system A and for the second B as

$$|\psi\rangle_{AB} = |\psi\rangle_A \otimes |\psi\rangle_B \quad (1.20)$$

otherwise it is called entangled, such as for instance:

$$|\psi\rangle_{AB} = \frac{|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B}{\sqrt{2}}. \quad (1.21)$$

From the experimental point of view, we can say that a separable state can be initialized locally for the two systems A and B . Conversely, an entangled state is forced to link the two systems and the correlation between the subspaces exists independently from the basis chosen to represent the whole system. It is interesting to show the simple case of the so-called Bell states:

$$|\Phi^+\rangle = \frac{|0\rangle |0\rangle + |1\rangle |1\rangle}{\sqrt{2}} \quad |\Phi^-\rangle = \frac{|0\rangle |0\rangle - |1\rangle |1\rangle}{\sqrt{2}}. \quad (1.22)$$

$$|\Psi^+\rangle = \frac{|0\rangle |1\rangle + |1\rangle |0\rangle}{\sqrt{2}} \quad |\Psi^-\rangle = \frac{|0\rangle |1\rangle - |1\rangle |0\rangle}{\sqrt{2}}. \quad (1.23)$$

which represent an entangled basis for a $d = 4$ system of two qubits. In the same way, we can define separable and entangled states through the density matrix formalism:

$$\rho_{AB} = \sum_i p_i \rho_i^{(A)} \otimes \rho_i^{(B)} \quad (1.24)$$

Therefore, they can be used to develop basic quantum tools for quantum computation. Entanglement is of fundamental importance since it highlights the inner correlations between quantum states. It is interesting to stress that this correlation seems to be in conflict with the general idea of classical locality. It was underlined in a historical work of Einstein, Podolsky and Rosen [61] that these states are in opposition to special relativity and leading to abandon the concept of localism and realism. Therefore to justify its existence, it was conjectured that quantum mechanics is incomplete and has to be ruled by a deterministic hidden theory in a deeper level. In 1964 John Bell formulates his famous inequality [62] to shed light on this aspect providing a mathematical tool which could be tested experimentally. Following this road, several experiments has been made [63–66] to verify such tests.

In the next section we introduce the description of circuit model of computation.

1.2.3 Circuit model of computation

Quantum computing aims to process quantum data to boost the computational process in terms of speed and resources. Nowadays, one of the principal problems regarding the implementation of a controllable quantum computer relies on the high requested level of control. In the last decades several efforts have been made to build a quantum device. For an incomplete list, we can mention quantum Turing machines [47], cluster states [67], quantum random walks [30], adiabatic [68], permutational [69] and one-clean-qubit [70] models. Moreover in the last decades strong improvements have been made for the implementation of quantum devices in the field of quantum annealing [71–73] and digital and analogic quantum computers and simulators with superconducting circuits [74,75], although criticisms arise in scientific community for the capability to fully control those devices in a quantum regime as for example the D-Wave systems [76].

Although the techniques developed so far are still not sufficient for a full description of a quantum computer, the building blocks of such devices are, at least half of models, the same from a conceptual point of view, because they lie on the idea of quantum circuits. We will briefly introduce the basics of the circuit model in quantum perspective. In particular, it is fundamental to list a set of *rules* or constraints in order to focus on the real targets for a near future implementation of a quantum device. A first step in this direction was done by the recognized DiVincenzo’s criteria [29] introduced in 2000, to identify a checklist of properties that a quantum computer must have:

1. A scalable physical system with well characterized qubits.
2. The ability to initialize the state of the qubits to a simple fiducial state, such as $|000\dots0\rangle$.
3. Long relevant decoherence times, much longer than the gate operation time.
4. A universal set of quantum gates.
5. A qubit-specific measurement capability.
6. The ability to interconvert stationary and flying qubits.
7. The ability to faithfully transfer qubits between distant locations.

The motivations behind the criteria are intuitive. We need a scalable and well characterize physical system because we have to assure a good computation without losing information during the manipulation process (1) as in the same way, we must have access to the initialization to the qubits in order to perform several new algorithms restarting the calculation independently (2). Nowadays, the decoherence of qubits is one of the most important problem in quantum platforms, since single qubits when interacting with the environment, lose information during the process (3). A figure of merit to estimate the efficiency has to be linked with the number of operations n_{op} executed inside the coherence time τ_Q of the system. The coherence time is defined as the

1.2.4 Logic Quantum gates

In the classical circuit model, it turns out that there is a set of single- and two-bit gates that are universal, meaning that every circuit can be realized from that finite set. The same holds true for the quantum circuit model where the only gates we can exploit to perform any kind of possible algorithm are *single-qubit gates*, *Hadamard gates*, *Phase gates* and *CNOT gates* [45, 77, 78]. A single-qubit gate, which performs the elementary operations on a single qubit, can be seen as 2×2 unitary matrices acting on the vector of the amplitudes describing the single quantum state (α, β) . It is important to mention the single-qubit Pauli gates

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (1.26)$$

and the Hadamard H and the Phase gate $\delta/2$

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \delta/2 = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{pmatrix}. \quad (1.27)$$

Pauli matrices are related to the rotation operator in the Bloch sphere by equation of the form $R_x(\theta) = e^{-i\theta \frac{X}{2}} = \cos\left(\frac{\theta}{2}\right)I - i \sin\left(\frac{\theta}{2}\right)X$. The action of the Hadamard matrix can be visualized in the Bloch sphere as a rotation of $\pi/2$ about the \vec{y} axis, followed by a reflection through the xy plane. It transforms qubits from a computational base to another as for example $H|0(1)\rangle = |+(-)\rangle$, where $\{|+\rangle, |-\rangle\}$ is the diagonal base. It is very simple to show that it is possible to connect different qubit gates with Hadamard gates through this relation that holds true

$$HXH = Z \quad HYH = -Y \quad HZH = X. \quad (1.28)$$

To make our set universal, we need to introduce also two-qubit gates and in particular the Controlled-NOT (**CNOT**) gate, which flips one of the qubits in input (*target*) if the other (*control*) is set to 1. The matrix representation in the bases $\{|i, k\rangle\}$ with $i, k \in [0, 1]$ is

$$CNOT = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (1.29)$$

we now briefly discuss why it represents an universal set of gates, describing the operation on a generic vector of the Hilbert space:

- Going from a point to another point on the surface of a sphere can be seen as a rotation under some axis (up to a phase factor) which in turn can be decomposed into three rotations $R_i R_j R_z$ around two orthogonal axis \vec{i} and \vec{j} .
- choosing $\delta = \pi/4$ we select a $\pi/8$ -gate. This gate corresponds, up to a phase, to $R_z(\pi/4)$. It can be shown that $\pi/8$ and H gates can give a rotation of an irrational multiple of 2π around the axis $\vec{n} = (\cos(\pi/8), \sin(\pi/8), \cos(\pi/8))$. Since the angle is irrational, several rotations of the same angle can approximate any angle to arbitrary accuracy.
- Finally, given the fact that we can compose a rotation around an axis i with the Hadamard, to give a rotation around an axis orthogonal to i , it follows that a single-qubit operation can be approximated with arbitrary accuracy using the Hadamard and $\pi/8$ gates. The Solovay-Kitaev theorem [79] states that this approximation can be done to an accuracy ϵ using $O(\log_2(\frac{1}{\epsilon}))$ gates from the discrete set.

These gates represent a universal set, since every multi-qubit unitary operation can be decomposed in a circuit composed of them [80]. Indeed, this is just one of the several possible universal sets. Still, it is important to note that the implementation of an arbitrary unitary by a quantum circuit is in general not efficient. In fact, quantum computation is basically limited to unitary transformations that can be realized by a polynomial sequence of 2-local unitaries [81].

1.3 Quantum algorithms

In this section we introduce a non-exhaustive list of the main quantum algorithms that have been theorized in the last two decades which highlight the hidden characteristics of quantum mechanics which manifest as a possible superior approach to the computation. We will introduce the logic behind a quantum algorithm, with a review on the different classes and typologies.

1.3.1 Overview on quantum algorithms

The idea of quantum computation was born three decades ago in the early 80's with the works of Feynmann and Deutsch. Since then, the field developed fast starting from the first theoretical quantum algorithm by Shor, Deutsch and Grover [2, 82, 83] which give the feeling for a real threat to ECTT. Nowadays several improvements have been made both theoretically and experimentally, and most algorithms improve the computational runtime providing a speed-up respect to its classical counterpart. Let us mention a few remarkable results of algorithms based on: *Quantum Fourier transform algorithms, amplitude amplification and quantum walks.*

Quantum Fourier Transform based algorithms: this class is based on the quantum counterpart of the discrete Fourier transform (DFT) which provides a significant speed-up in the computation [45]. The Fourier transform of a N vector $y = (y_1, y_2, \dots, y_N)$ acts on each component executing this transformation for each k :

$$x_k \longrightarrow \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} y_j e^{\frac{2\pi i j k}{N}} \quad (1.30)$$

while the quantum Fourier transform acts on a quantum basis $|j\rangle$ as:

$$|y\rangle \longrightarrow \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} e^{\frac{2\pi i j k}{N}} |k\rangle \quad (1.31)$$

which is a unitary transformation and can thus be implemented with the formalism of the circuit model. Consequently the QFT represents a rotation of the Hilbert space, thus a change of the computational basis. It is worth noting that the generic transformation of a generic state which is a linear superposition of the computational vectors will be

$$\sum_{j=0}^{2^N-1} c_j |j\rangle \longrightarrow \sum_{k=0}^{2^N-1} x_k |k\rangle \quad \text{with} \quad x_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} c_j e^{\frac{2\pi i j k}{N}} \quad (1.32)$$

where c_j are the amplitudes, for which it follows that the unitary matrix associated with the transformation has elements $U_{jk} = \frac{1}{\sqrt{2^N}} e^{2\pi i j k / 2^N}$. Using properly Hadamard and Phase shifters gates it is possible to retrieve the Fourier transform on N qubits as it shown in Fig. 1.6. It is worth noting that to transform 2^N amplitudes we require only $O(N^2)$ quantum gates instead of $O(N 2^N)$ gates in classical situation thus providing a computational boost. The best known upgrade of the QFT requires $O(N \log N)$ quantum gates [84].

Simon's algorithm solves with $O(N)$ queries so-called decision tree problems [85]. Shor's algorithm for which the classical problem is believed to belong to the NP complexity class, tries to factoring

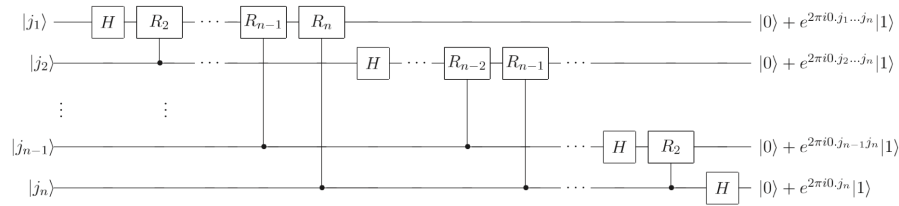


Figure 1.6. QFT circuit: circuitual representation of quantum Fourier transform acting on N qubits. Each node is properly transformed by a single-qubit gate in order to retrieve after the entire evolution, the discrete Fourier transform on the input qubits. Ref. [45].

an arbitrary long integer in polynomial time and for this reason belongs to the **BQP** class. Moreover there is a substantial difference between the classical and the quantum approach: in fact the best known classical algorithm to factorize large integers runs in time $e^{O((\log N)^{1/3}(\log \log N)^{2/3})}$, while the quantum approach show a strong reduction in computational run time $O((\log N)^3)$. It is interesting because the classical hardness of this system is widely exploited in RSA public-key cryptosystems [86] to protect bank codes. Another algorithm is the quantum phase estimation which is used to determine the eigenphase of an eigenvector of a unitary gate given a quantum state proportional to the eigenvector and access to the gate. The algorithm is frequently used as a subroutine in other algorithms.

Algorithms based on amplitude amplification: Amplitude amplification is a technique that allows the amplification of a chosen subspace of a quantum state. Applications of amplitude amplification usually lead to quadratic speed-ups over the corresponding classical algorithms. A honourable mention between the algorithms in this group is the Grover's algorithm [83]. The goal of this procedure is to find an element in an unstructured database with N entries with a number of resources which scales as $O(\sqrt{N})$ respect to its classical best known counterpart ($O(N)$). Although the purpose of Grover's algorithm is usually described as "searching a database", it may be more accurate to describe it as "inverting a function". Roughly speaking, if we have a function $y = f(x)$ that can be evaluated on a quantum computer, Grover's algorithm allows us to calculate x when y is given. Inverting a function is related to the searching of a database because we could come up with a function that produces one particular value of y ("true", for instance) if x matches a desired entry in a database, and another value of y ("false") for other values of x .

Grover's algorithm can also be used for estimating the mean and median of a set of numbers, and for solving the collision problem. The algorithm can be further optimized if there is more than one matching entry and the number of matches is known beforehand. Grover's algorithm can be used to crack passwords as well.

Algorithms based on quantum walks: Quantum walks [30] are the quantum analogue of classical random walks [87], i. e. the study of the dynamic of one or more particles (walkers) which can move in a random manner, continuously or at discrete time steps, through a discrete space. Classical random walks are of fundamental importance in the study of stochastic and Markov processes. The generalization to quantum walks implies the possibility for the walker to be in a superposition state of many positions. In continuous-time quantum walks [31], the state is uniquely determined by the position wavefunction of the particle on a graph, and its continuous evolution is determined by the Schrödinger equation $|\psi(t)\rangle = e^{-iHt} |\psi(0)\rangle$, where the Hamiltonian is proportional to the adjacency matrix of the graph. In discrete-time quantum walks [30] the system is described by the position of the walker and by an auxiliary entity named coin, which encodes the probability of taking the left or right direction at the upcoming step. In

the case of a discrete quantum walk on a line, the coin will be a two-level system, $\alpha |\rightarrow\rangle + \beta |\leftarrow\rangle$ and the unit step U of the walker will be given by the application of the coin flip operator C over the coin state (frequent is the case of the Hadamard coin) followed by the shift operator

$$S = |\rightarrow\rangle\langle\rightarrow| \otimes \sum_i |i+1\rangle\langle i| + |\leftarrow\rangle\langle\leftarrow| \otimes \sum_i |i-1\rangle\langle i| \quad (1.33)$$

If we introduce a measurement of the coin state between the application of operators C and S , we retrieve the classical random walk. Instead, keeping the quantum correlations between different positions will give rise to completely different phenomena, caused by quantum interference. In particular, the limiting distribution after many steps shows a completely different behaviour in the two cases, as the quantum case give rise to a strongly non-Gaussian distribution with variance larger than the linear trend typical of the classical case. More generally, the complex interference patterns (quantum carpets [32, 88]) and the revival of the wavepackets at certain periods [89, 90] typical of quantum walks have no analogue in classical random walks. One of the main point of interest in quantum walks is the investigation of quantum transport phenomena, with potential applications to understand processes such as energy transfer within photosynthetic complexes [91] or to connect distant quantum registers (for example by exploiting wavepacket revivals occurring in different positions [92]). From the point of view of quantum computing, quantum walks represents a computational model of their own, as they have been demonstrated to allow universal quantum computation [93]. Up to now, they have been successfully applied as algorithmic tools in problems such as estimating the volume of convex bodies [94], finding satisfying conditions for Boolean formulas [95] or approximating permanents [96]. In general, their key property is the ability to search and sample from an exponentially large set of combinatorial structures using only local operations. This point will be fundamental for the purpose of our work.

1.4 Quantum computation with linear optics

The realization of a scalable quantum device which is able to provide a superpolynomial speed-up over its classical simulation would be an outstanding evidence of the fact that $\mathbf{BQP} \neq \mathbf{BPP}$, with strong consequences over \mathbf{ECTT} 's tenability. Although the general thought about this realization is negative because of fundamental limitations such as influence of noise [45], in this scenario, several techniques started to develop and testing the first implementation with different platform. As we already mentioned, there are many possible physical schemes to implement qubits and quantum gates, each one with its own advantages with respect to the others. However, the choice of the scheme is strictly related to the way we are able to implement it experimentally. Despite ion traps or cold atoms implementations can make qubit interact in a highly controlled way, they have very low coherence times compared to photonic schemes. Photons can be well isolated from the environment, they have potentially infinite coherence time, but do not interact with each other. Because of this, while satisfying the first point is rather easy, the realization of quantum gates is demanding. For this reason, using photons implementations could provide a fascinating long-term development for this field. For example, the employment of schemes such as cluster-state quantum computation and graph quantum computation [10, 11] are believed to be promising because they are suitable for photonics chip implementations [97] and they are tolerant to noise [98]. This could be the key for realization of future large-size experiments.

However the use of photons as quantum information carriers, linear optical elements, such as fibers, beamsplitters and waveplates, and the capability to move them around and perform quantum operation on them, have found so far its development in the so-called *linear optical quantum computing* (LOQC) [12]. LOQC has been the subject of intensive research over the last 20 years, and it is still subject of deep analysis and a candidate for the development of quantum computing [99] despite the difficulties in scaling-up and noise control. A possible route to

implement two-qubit gates has been proposed in the seminal work of Knill, Laflamme and Milburn (KLM) [12], and consists in exploiting feed-forward mechanisms to get measurement-induced effective nonlinearities. This scheme is currently at the basis of the development of LOQC. However the experimental requirements for implementing universal quantum computing remain unfeasible with the current technologies, despite some improvements to the protocol [99, 100]. Photons remains one of the most feasible mean to achieve quantum computation thanks to its natural behaviour to transmit information along space. Although, they are non-interacting particles, the desired interaction is assured by the *interference* between photons. If we consider a state an indistinguishable photon state $|\psi\rangle = \sum_i \alpha_i |i\rangle$, and we inject it into a passive linear interferometer, the outcome probability is given by

$$Pr_{\psi \rightarrow \phi} = |\langle \phi | \psi \rangle|^2 = \left| \sum_i \beta_i^* \alpha_i \right|^2 = \sum_i |\beta_i|^2 |\alpha_i|^2 + \sum_{i \neq j} \beta_i^* \beta_j \alpha_j^* \alpha_i \quad (1.34)$$

It is intuitive understand the difference respect to the classical case. The additional term $\sum_{i \neq j} \beta_i^* \beta_j \alpha_j^* \alpha_i$ represents the interference among possible paths with $i \neq j$ and highlights the richness due to the quantum behaviour. This contribution is non-zero until decoherence provided by the interaction with the environment is negligible.

In this situation, the investigation of the power of passive linear optics with non-interacting photons naturally came along. In early times, there was general consensus over the fact that, regardless of the input state, no universal quantum computer could ever be built using only linear interferometers, nor even gaining any computational advantage. Later on, a series of results on hardness of simulation of linear interferometers started to come along: the output of a linear interferometer fed with coherent states is computable in polynomial time [101]. Any interferometer processing Gaussian states, even with added quadratic nonlinearities, can be simulated classically [102]. The probability of measuring a particular basis state in any n -photon linear optical circuit can be computed to within $\pm\epsilon$ additive error in a time which scales polynomially with n and $1/\epsilon$ (Gurvit's algorithm) [103]. The availability of adaptive measurements, i.e. the ability to modify the quantum operations in the network after intermediate measurements, makes linear optics universal for quantum computation either with single photons or Gaussian states in input (KLM scheme). If one have access to non-demolition measurement (which means that the post-measurement state is available after the measurement) this is true even with coherent states [12].

No one of the above results seemed to prefigure any possible quantum advantage from using passive linear optics alone. However several investigations on quantum communications, quantum key distribution and many-body quantum interference on this interferometric platforms seem to highlight some features on computation complexity and on the reliability of such particles to perform mid-term protocols.

The comprehension of the common operations of LOQC protocols pass trough the quantum field theory and the description of optical elements. We can describe the quantum electromagnetic field as a potential vector with wave-plane expansion by means of the creation and annihilation operators a and a^\dagger [104]:

$$a |n\rangle = \sqrt{n} |n-1\rangle \quad a^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle \quad (1.35)$$

where the operators follow the commutation relations provided by the bosonic statistic:

$$\left[a_k, a_l^\dagger \right] = \delta_{k,l} \quad \left[a_k, a_l \right] = \left[a_k^\dagger, a_l^\dagger \right] = 0. \quad (1.36)$$

This formalism identify, if we are in zero-losses condition, the so-called Fock states. we can assume the optical elements as for examples *beam splitters* (BS) and *phase shifters* (PS) are

unitary, thus providing the desired lossless condition. An optical circuit can be shortly described as an m - mode interferometer, that are m optical modes defined on the Fock space. In this case, the Hamiltonian will be

$$H = \sum_{j,k} A_{jk} a_j^\dagger a_k \quad (1.37)$$

that is a quadratic form on the creation operators. If we change the representation into Heisenberg form we can resume the behaviour of a passive linear interferometer as

$$a_i \longrightarrow \sum_j U_{ij} b_j \quad (1.38)$$

where U_{ij} are the element of the unitary matrix which identify the evolution. One of the fundamental components of linear optics, as shown before, are beam splitters (**BS**), which mix the optical modes respect to typology of the input injected. It can be seen as a 2×2 unitary matrix:

$$U_{BS} = \begin{pmatrix} \cos \theta & e^{-i\phi} \sin \theta \\ e^{i\phi} \sin \theta & \cos \theta \end{pmatrix}. \quad (1.39)$$

Angle θ corresponds to the transmitted and reflected beam, while the angle ϕ introduce a phase between the two beams. In the same way, we define another fundamental optical element: the phase shifter (**PS**). Usually, it consists in a material with refraction index different respect to the air, thus introducing as relative δ phase on the field.

$$U_{BS} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{pmatrix}. \quad (1.40)$$

These optical elements, together with other components can realize one- and two- (probabilistic) quantum gates [105–107]. Moreover, it is possible to realize universal circuital schemes by means of ancillar qubits [45], non-linear effects [108, 109] and ultrafast laser technologies [110, 111]. To perform protocols of quantum computation by means of linear optics elements, we can exploit the degrees of freedom of single photons. For this reason let's see in which way it is possible to represent the qubit through photon encoding. We can say that the state of each photon is expressible as a wave-packet

$$\hat{a}_i^\dagger(\tau_i) = \int_{-\infty}^{\infty} d\omega_i f_i(\omega_i) e^{-i\omega_i\tau_i} \hat{a}_i^\dagger(\omega_i) \quad (1.41)$$

where $i = 1, 2, 3$ identify the spatial mode k_i , τ_i is the delay due to the propagation and $f_i(\omega_i)$ is the wave function spectral function peaked to the frequency ω_0 and bandwidth $\delta\omega$

$$f_i(\omega_i) = \frac{1}{\sqrt[4]{2\pi\delta\omega^2}} e^{-\frac{(\omega_i - \omega_0)^2}{4\delta\omega^2}}$$

How can we represent the qubit by means of a wave-packet?

1. *temporal encoding*: we use a Mach-Zehnder (**MZ**) interferometer to encode the qubit. Path lengths must be chosen with a difference near to the coherence length of the photon. If we identify the path with $|0\rangle_{MZ}$ o $|1\rangle_{MZ}$, the photon state before the collapse is described by th superposition $|\psi\rangle = \alpha |0\rangle_{MZ} + \beta |1\rangle_{MZ}$, with

$$|0\rangle_{MZ} = \int_{-\infty}^{\infty} dz f\left(\frac{t - z/c}{\delta t}\right) e^{-i\omega_0(t-z/c)} \hat{a}^\dagger(z) |0\rangle \quad (1.42)$$

$$|1\rangle_{MZ} = \int_{-\infty}^{\infty} dz f\left(\frac{t - z/c + \tau}{\delta t}\right) e^{-i\omega_0(t-z/c+\tau)} \hat{a}^\dagger(z) |0\rangle \quad (1.43)$$

Its robustness suggests that temporal encoding is useful for communication protocols.

2. *Polarization encoding*: the encoding is performed through the exploitation of the polarization degree of freedom. In this case the state is described by the superposition $|\psi\rangle = \alpha |H\rangle + \beta |V\rangle$, with

$$|H\rangle = \int_{-\infty}^{\infty} dk f(k) e^{-i\omega_k t} \hat{a}_H^\dagger(k). \quad (1.44)$$

$$|V\rangle = \int_{-\infty}^{\infty} dk f(k) e^{-i\omega_k t} \hat{a}_V^\dagger(k). \quad (1.45)$$

3. *Occupation mode encoding*: the trick is to exploit different propagation modes, that is, the basis state is defined by occupation numbers. In this way, the superposition $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$ is realized distinguishing $|0\rangle$ e $|1\rangle$, respectively, the situation where we have 0 or 1 photons per mode.
4. *Transverse spatial-mode encoding*: we can encode through the orbital angular momentum on spatial modes. Any spatial mode of a laser beam can be written in terms of solutions of the paraxial Helmholtz's equation [110]. If the system has cylindrical symmetry, solutions of the equation are the Laguerre-Gauss modes

$$u(r, \phi, z) = \sqrt{\frac{2p}{\pi(l+p)}} \frac{1}{\omega_z} \left(\frac{r\sqrt{2}}{\omega_z}\right)^{|l|} L_p^{|l|} \left(\frac{2r^2}{\omega_z^2}\right) e^{-\frac{r^2}{\omega_z^2} + ik\frac{r^2}{2R_z}} e^{i(2p+|l|+1)\zeta_z} e^{il\phi} \quad (1.46)$$

where $\omega_z = \omega_0 \sqrt{1 + \left(\frac{\lambda z}{\pi\omega_0^2}\right)^2}$ is the spot size, $R_z = z \left(1 + \left(\frac{\pi\omega_0^2}{\lambda z}\right)^2\right)$ the radius of curvature, $\zeta(z) = \arctan(z/z_R)$ is the Gouy's phase and $\phi = \arctan(y/x)$. The fundamental property of these kind of encoding relies on the capability to carry the angular orbital momentum with a phase factor $e^{il\phi}$. To build a computational basis, we can choose to represent states with orbital momentum, as for example $l_1 = +\hbar \rightarrow |0\rangle$ e $l_{-1} = -\hbar \rightarrow |1\rangle$.

1.5 Discussion

Linear optics and photons are one of the most promising candidates to perform basic operation of quantum computations and communication. Currently, the continuously increasing number of entangled qubits with ions and superconducting qubits makes them the best systems for local information processing. Nevertheless, to connect two of such local stations, photons appear to be the most natural candidates. Indeed, the speed of transmission and the low-noise properties of photons constitute extremely important characteristics to account for quantum technologies. Encoding information on a single photon is relatively easy and moreover is secure because any not controlled measurement that affects the system corrupts inevitably the information, thus protecting from eavesdropping. Recently, several ideas arose to ascertain the advantage to use many-body bosonic interaction in order to test a computational advantage in quantum computation. In next chapter we will introduce the most recent development in this field from the perspective of many-body bosonic interference phenomenon.

Chapter 2

Multi-bosons interference and Boson Sampling model

In this Chapter we introduce the main topic of this work, that is the use of many-body bosonic interference to implement a non-universal model that promises a quantum computational advantage with relatively low experimental effort. We start from a theoretical background on bosonic coalescence, then we will introduce the Boson Sampling problem that is believed to be hard to compute for classical computers. Moreover, we will describe an improved version of Boson Sampling, which exploits an advanced application of SPDC sources in order to significantly boost the quantum computation power in terms of generation rate. All these topics represent the justification of the research that I have carried out and they are a necessary introduction for the problem of quantum certification of genuine quantum interference.

2.1 Indistinguishable Particles

In nature there are many examples of systems made of particles of the same species, *i.e.* with same mass, spin and charge. The electrons in an atom, for instance, are identical particles. In classical physics it is also possible to keep track of single particles even though they look identical. In fact we can always, in principle, follow the trajectory of each particle at each instant of time. In quantum mechanics, however, identical particles are truly indistinguishable: the concept of trajectory does not exist and we cannot specify more than a complete set of commuting observables for each of the particles.

For simplicity, consider a system with just two particles. Suppose k' is a collective index of quantum numbers describing the system. Since we are dealing with two identical particles, the wavefunctions $\Psi(k_1, k_2)$ and $\Psi(k_2, k_1)$, *i.e.* after a particles exchange, must represent the *same* quantum state, and therefore they must differ only by a phase factor [112]:

$$\Psi(k_1, k_2) = e^{i\alpha} \Psi(k_2, k_1). \quad (2.1)$$

If we repeat the exchange of the two particles, obtaining again $\Psi(k_1, k_2)$, we find:

$$e^{2i\alpha} = 1 \quad \implies \quad e^{i\alpha} = \pm 1, \quad (2.2)$$

that is, the wave function of a system of two identical particles must be either symmetric or antisymmetric under the exchange of the two particles. As a result, we have the *spin-statistics* theorem [112]:

Theorem 2.1: Spin-Statistics

Systems of identical particles with integer spin ($s = 0, \pm 1, \dots$), known as **bosons**, have

wave functions symmetric under exchange of any pair of particle labels. The wave function is said to obey the Bose-Einstein statistics.

Systems of identical particles with half-odd-integer spin ($s = \pm 1/2, \pm 3/2, \dots$), known as **fermions**, have wave functions antisymmetric under exchange of any pair of particle labels. The wave function is said to obey the Fermi-Dirac statistics.

This fundamental law was discovered by Wolfgang Pauli [113] and it has been largely supported by experimental evidences.

Moreover it has been observed that in two-dimensional systems, we can individuate quasiparticles that can be observed to obey statistics ranging continuously between Fermi-Dirac and Bose-Einstein statistics. These particles are named *anyons* and are represented, given two identical particles an the action of the exchange operator, as

$$\Psi(x_1, x_2) = e^{i\theta} \Psi(x_2, x_1) \quad (2.3)$$

this particles can be seen as a generalization of bosons and fermions and they were introduced in the seminal work of Leinaas and Myrheim in 1977 and independently from Wilczek in 1982 [114, 115]. These particles have found interest in quantum information field because it is believed that those particle can provide a powerful tool for topological quantum computing [116]. In parallel experimental works with photonics architectures demonstrated the importance of this quantum state in governing the behaviour of quantum walks [27].

Let us now focus on the previous statistics introduced: bosonic and fermionic in order to highlight why they are important and how they are connected to this manuscript work. In general, given a quantum system of N identical particles, the global state of the system, both symmetric (S) and antisymmetric (A), is described by a linear superposition of all possible permutations p of product states [117]:

$$|k_1, k_2, \dots, k_N\rangle_S = \sqrt{\frac{\prod_j n_j!}{N!}} \sum_p |k_{p(1)}, k_{p(2)}, \dots, k_{p(N)}\rangle, \quad (2.4)$$

$$|k_1, k_2, \dots, k_N\rangle_A = \sqrt{\frac{1}{N!}} \sum_p \text{sgn}(p) |k_{p(1)}, k_{p(2)}, \dots, k_{p(N)}\rangle, \quad (2.5)$$

with n_j the number of repetitions in the exchange. Now, to completely characterize the quantum state, we have to take in account the signs due to particle exchange, thus we need to compute two quantities, for asymmetric and symmetric state. Thus, we can define the following definitions, given a $N \times N$ matrix:

$$\text{Det}(U) := \sum_{\sigma \in S_N} \prod_{i=1}^N \text{sgn}(\sigma) a_{i\sigma(i)} \quad \text{Per}(U) := \sum_{\sigma \in S_N} \prod_{i=1}^N a_{i\sigma(i)}, \quad (2.6)$$

respectively as the determinant and the permanent of a matrix, where S_N contains all the permutations σ relative to the set $\{1, 2, \dots, N\}$, and $\text{sgn}(\sigma)$ is +1 or -1 whether the permutation is even or odd. Let us introduce an example of an application of these two quantities to build a system with two identical fermion particles. The simplest way to approximate the wave function of a many-particle system is to take the product of properly chosen orthogonal wave functions of the individual particles. For the two-particle case with coordinates x_1 and x_2 , we have $\Psi(x_1, x_2) = \psi_1(x_1) \psi_2(x_2)$ as follows by the Hartree-Fock approximation in the case of fermionic particles [112]. In order to take in account the statistics we have to consider a linear superposition which contain the action of the exchange. Therefore, if we build a matrix with the single terms we can retrieve the wave function:

$$\frac{1}{\sqrt{2}} \text{Det} \begin{pmatrix} \psi_1(x_1) & \psi_1(x_2) \\ \psi_2(x_1) & \psi_2(x_2) \end{pmatrix} \implies \Psi(x_1, x_2) = \frac{1}{\sqrt{2}} (\psi_1(x_1) \psi_2(x_2) - \psi_1(x_2) \psi_2(x_1)) \quad (2.7)$$

The choice between determinant and permanent is pivotal in order to distinguish the behaviour of the quantum system. In particular, as a generalization, the wave function of a N -particle fermionic system is given by the *Slater determinant*, by considering a $N \times N$ matrix U :

$$\Psi_{k_1, \dots, k_N}^{(A)}(x_1, \dots, x_N) = \frac{1}{\sqrt{N!}} \text{Det} \begin{pmatrix} \psi_{k_1}(x_1) & \psi_{k_1}(x_2) & \cdots & \psi_{k_1}(x_N) \\ \psi_{k_2}(x_1) & \psi_{k_2}(x_2) & \cdots & \psi_{k_2}(x_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{k_N}(x_1) & \psi_{k_N}(x_2) & \cdots & \psi_{k_N}(x_N) \end{pmatrix}. \quad (2.8)$$

Otherwise, for a globally symmetric wave function, that is, a bosonic quantum system, we need to compute the permanent:

$$\Psi_{k_1, \dots, k_N}^{(S)}(x_1, \dots, x_N) = \frac{1}{\sqrt{N!}} \text{Per} \begin{pmatrix} \psi_{k_1}(x_1) & \psi_{k_1}(x_2) & \cdots & \psi_{k_1}(x_N) \\ \psi_{k_2}(x_1) & \psi_{k_2}(x_2) & \cdots & \psi_{k_2}(x_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{k_N}(x_1) & \psi_{k_N}(x_2) & \cdots & \psi_{k_N}(x_N) \end{pmatrix}. \quad (2.9)$$

These two expressions could appear very similar, but they are extremely different in terms of computational complexity. The calculation of the determinant scales polynomially with the input size, so it is in the computational class \mathbf{P} and therefore tractable on a classical computer. For the permanent, instead, there is no efficient algorithm able to compute it efficiently. Indeed, it scales exponentially with the input size and, consequently, it is intractable for a classical computer. In particular, it was pointed out by Valiant [42] that the permanent calculus belongs to the $\#\mathbf{P}$ -hard class. This difference in computational complexity relies on the permutation signs, which alternate in determinants allowing some terms to cancel out. For permanents this is not possible and the computation becomes much harder. We will provide further details on the computational complexity of evaluating permanents in Section 2.4.

In the following section we describe how quantum interference between photons can be described using permanents, even for the simple case of only two photons impinging on a beam-splitter. Moreover, this approach will be useful to understand the Boson Sampling problem.

2.2 Hong-Ou-Mandel effect

The Hong-Ou-Mandel (**HOM**) effect is a quantum phenomenon due to two-photon interference on a beam-splitter (**BS**) [14]. It occurs when the two photons are *indistinguishable* and impinge on a 50:50 **BS** from different input modes (a, b). There are four possibilities for the photons at the output:

1. The photon entering from mode a is reflected and the other one entering from mode b is transmitted.
2. The photon entering from mode a is transmitted and the other one entering from mode b is reflected.
3. Both photons are transmitted.
4. Both photons are reflected.

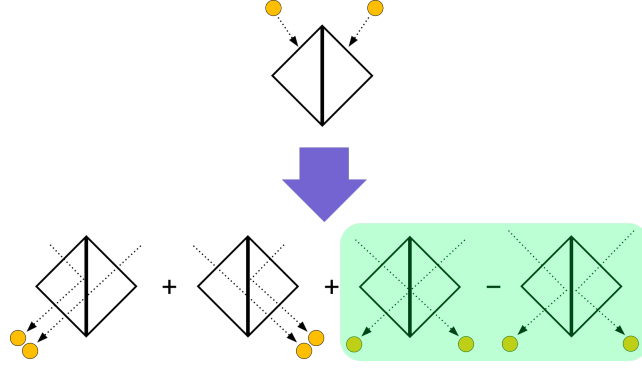


Figure 2.1. Possible scenarios for two photons impinging on a 50:50 BS: "from left to right" the photon entering from mode a is reflected and the other one is transmitted, the photon entering from mode a is transmitted and the other one is reflected, both photons are transmitted, both photons are reflected. If the photons are indistinguishable, the last two cases cancel each other, due to destructive interference process. This effect is known as Hong-Ou-Mandel effect.

The beam-splitter reflectivity (\mathcal{R}) and transmissivity (\mathcal{T}) determine the relative probabilities of reflection and transmission. When the photons are perfectly indistinguishable, a destructive interference occurs between the possibilities 3. and 4., making them exit *always* from the same output mode, c or d according to scenarios (1) and (2) (2.1).

Formally, we can describe the HOM effect using the second quantization formalism, that is exploiting the Fock states and the creation and annihilation operators, which obey the bosonic commutation rules:

$$[\hat{a}_i^\dagger, \hat{a}_j^\dagger] = [\hat{a}_i, \hat{a}_j] = 0, \quad [\hat{a}_i^\dagger, \hat{a}_j] = \delta_{ij}. \quad (2.10)$$

The action of a BS is described by the matrix [118]:

$$\hat{B} = \begin{pmatrix} t & ir \\ ir & t \end{pmatrix}, \quad (2.11)$$

where t and r are the transmission and reflection coefficients, related to \mathcal{T} and \mathcal{R} by the expression¹:

$$t^2 + r^2 = \mathcal{T} + \mathcal{R} = 1. \quad (2.12)$$

Having indicated the input modes as (a, b) , the two-photons initial state is described by the Fock state: $|\psi\rangle_{a,b}^{(in)} = \hat{a}_a^\dagger \hat{a}_b^\dagger |0, 0\rangle_{a,b} = |1, 1\rangle_{a,b}$. In order to determine the output probabilities, we have to find out how the input modes transforms after the BS. This transformation is easily carried out thanks to the matrix equation:

$$\begin{pmatrix} \hat{a}_a^\dagger \\ \hat{a}_b^\dagger \end{pmatrix} = \hat{B} \begin{pmatrix} \hat{a}_c^\dagger \\ \hat{a}_d^\dagger \end{pmatrix} \implies \begin{cases} \hat{a}_a^\dagger = t \hat{a}_c^\dagger + ir \hat{a}_d^\dagger \\ \hat{a}_b^\dagger = ir \hat{a}_c^\dagger + t \hat{a}_d^\dagger \end{cases} \quad (2.13)$$

Hence, we can rewrite the input state obtaining:

$$\begin{aligned} |\psi\rangle_{a,b}^{(in)} &= |1, 1\rangle_{a,b} = \hat{a}_a^\dagger \hat{a}_b^\dagger |0, 0\rangle_{a,b} = (t \hat{a}_c^\dagger + ir \hat{a}_d^\dagger)(ir \hat{a}_c^\dagger + t \hat{a}_d^\dagger) |0, 0\rangle_{c,d} \implies \\ &= irt |2, 0\rangle + irt |0, 2\rangle + (t^2 - r^2) |1, 1\rangle \end{aligned} \quad (2.14)$$

where the coefficients are the probability amplitude for each two-photon Fock state. Hence, requiring our beam-splitter to be 50:50, *i.e.* $t^2 = r^2 = 1/2$, we see that the last term in (2.14)

¹We are referring to a lossless beam-splitter.

vanishes due to destructive interference.

Thus, under the indistinguishability hypothesis, two photons interfering on a symmetric BS can only exit from the same output port:

$$|\psi\rangle_{a,b}^{(in)} = \frac{1}{\sqrt{2}} (|2, 0\rangle + |0, 2\rangle)_{c,d}. \quad (2.15)$$

Now, let us consider the general case in which a state $|n, m\rangle$ enters a general BS. Such beam-splitter is described by the unitary matrix \hat{B}_G :

$$\hat{B}_G(\theta, \phi) = \begin{pmatrix} \cos\theta & e^{-i\phi}\sin\theta \\ e^{-i\phi}\sin\theta & \cos\theta \end{pmatrix}, \quad (2.16)$$

with $\cos\theta$ and $\sin\theta$ being the transmission and reflection amplitudes respectively. Assuming $\phi = \pi/2$, and supposing the photon number conservation, we expect at the output of the beam-splitter:

$$|n, m\rangle \longrightarrow \sum_{p=0}^{n+m} f_p |p, n+m-p\rangle. \quad (2.17)$$

It is worth noting that, if we let the state $|N, N\rangle^{(in)}$ enter a symmetric beam-splitter, the output state will be only occupied by even photons, that is:

$$|N, N\rangle^{(in)} = \frac{i^N}{2^N N!} \sum_{p=0}^N \binom{N}{p} \sqrt{2p!(2N-2p)!} |2p, 2N-2p\rangle^{(out)}. \quad (2.18)$$

Experimental evidences of this phenomenon was first obtained in Ref. [14]. In 2.2 it is reported an example of **HOM dip**: coincidence counts are plotted with respect to the relative delay between the photons. We can observe that the minimum, corresponds to a relative delay shorter than their coherence time. Theoretically the coincidences should vanish, however the actual *experimental* indistinguishability can be only partial, this is why they do not.

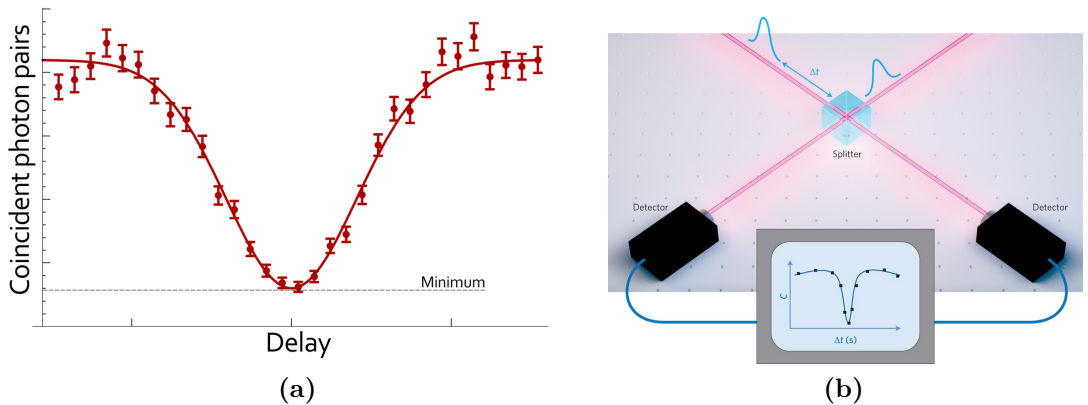


Figure 2.2. Coincidences number trend with respect to the relative two-photon delay. When the relative delay is shorter than their coherence time, the coincidences number drops due to destructive interference. The minimum value corresponds to the case of bunched photons exiting from the same output.

Experimental information about the degree of indistinguishability is given by an important figure of merit of the Hong-Ou-Mandel effect, the dip **visibility**. It is defined as:

$$\mathcal{V} = 1 - \frac{Q}{C}, \quad (2.19)$$

where C are the coincidences between distinguishable photons, *outside* of the dip, while Q are the coincidences between indistinguishable photons, that is *inside* the dip.

Let us show how the output probabilities are related to permanents. In detail, the initial state transforms with the matrix \hat{B} , thus, if we identify the definition matrix \hat{U} (2.6) with \hat{B} , we see that the probability to have $|1, 1\rangle$ is proportional to the matrix permanent:

$$B_{11} = \begin{pmatrix} t & ir \\ ir & t \end{pmatrix}, \quad (2.20)$$

that is, $\text{Per}(B_{11}) = (t^2 - r^2)$. Similarly, the output probabilities for the states $|2, 0\rangle$ and $|0, 2\rangle$ are proportional, respectively, to the permanents of the matrices

$$B_{20} = \begin{pmatrix} t & ir \\ t & ir \end{pmatrix}, \quad B_{02} = \begin{pmatrix} ir & t \\ ir & t \end{pmatrix}. \quad (2.21)$$

To sum up, one can calculate the output probabilities starting from the permanent of the *sub-matrices* a given unitary transformation \hat{U} . These sub-matrices are built by the cross selection among the columns and rows relative to the occupied input and output modes respectively. If a bunched input/output arrives on the BS, the corresponding column/row has to be taken n -times, where n is the number of bunched photons on that mode. For example, the sub-matrix B_{20} has been obtained by selecting the two columns of the matrix B (input $|1, 1\rangle$) and twice the first row (output $|2, 0\rangle$).

2.3 The Boson Sampling model

The *boson sampling* model, introduced by Aaronson and Arkhipov (AA) in 2011 [33], can be interpreted as a generalised Hong-Ou-Mandel effect, that is, a n -photons through an arbitrarily complex linear optical circuit (See Chapter 3).

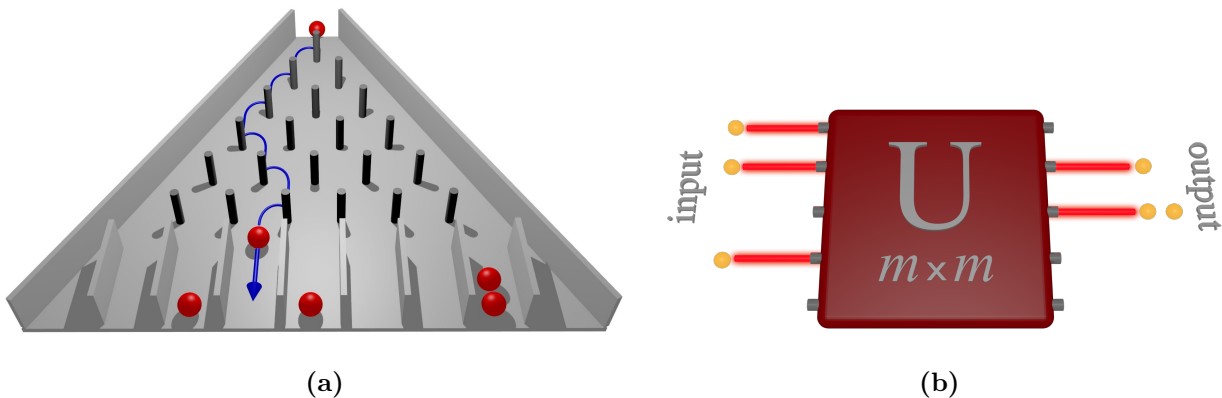


Figure 2.3. Analogy between a Galton's board and a Boson Sampling platform. (a) The Galton's board is composed by k identical balls dropped one by one into a lattice made of pins, each of which randomly scatters each incoming ball onto other pins and, at the end, the balls are collected in different containers (b) Case of a 5-mode boson sampling device with 3 input photons. Note the analogy between the two protocols, where instead of balls and pins, we use n -boson states and m -mode interferometers.

The Boson Sampling represents a *non-universal* model for quantum computation and consists of sampling from the output probability distribution of indistinguishable bosons evolving in a linear optical interferometer.

It is the quantum analogue of the *Galton's board*, a device used to illustrate the binomial distribution. (see Fig. 2.3a). Boson Sampling is very similar to the Galton's board, but it uses identical bosons, instead of classical spheres, sent from different starting points. The main differences between the two processes is referred to how we calculate the output probability distribution. For a Boson Sampling device, the dimension of the output Hilbert space scales exponentially with the number of bosons n and modes m , as:

$$\dim(\mathcal{H}) = \binom{m+n-1}{n}, \quad (2.22)$$

This is the core of the problem, because from one side, enlarging the system, the Hilbert space grows exponentially, from the other, calculate permanents is not efficient in n . It is worth noting that the Boson Sampling problem is not thought as a calculation problem, or in other words, these devices does not calculate permanents because in order to do so, we have to wait an infinite amount of time to obtain the full output distribution. Instead the problem is focused on sampling from the output distribution.

A *sampling problem* S is a collection of probability distributions $(\mathcal{D}_x)_{x \in \{0,1\}^*}$, input string composed of $x \in \{0,1\}^n$. To “solve” S means to sample from \mathcal{D}_x , given x as input [33, 119]. In other words, we would like to sample the outcomes of n indistinguishable photons entering in a m -modes interferometer, represented as an $m \times m$ unitary matrix \hat{U} . To this purpose, we need the output probability distribution, which can be obtained classically by computing the permanent, by extracting the submatrices related to each input-output combination considering fixed the input. Due to the intrinsic algorithmic complexity of the permanent, one cannot predict the outcomes with a classical computer without an exponential overhead in time or physical resources. Why do we need indistinguishable particles? If we consider the same problem, but instead, injecting distinguishable bosons, the matrix elements are non-negative real numbers. This makes us able to approximate the permanent in probabilistic polynomial time, by the Jerrum, Sinclair and Vigoda algorithm [96]. In the quantum case, by contrast, the matrix elements are complex numbers and given a general matrix $A \in \mathbb{C}^{n \times n}$, and even approximating $\text{Per}(A)$ to within a constant factor is a $\#\mathbf{P}$ -complete task. Hence, sampling from the output distribution of non-interacting bosons, appears a task solvable only by quantum simulators. In particular, integrated photonics devices are currently considered as the most promising tool for a scalable implementation of this model and moreover because photons solve the problem in a natural way. Let us now proceed with a formal description of the model, how permanents are linked to the output distributions. The privileged experimental condition for the Boson Sampling is the *collision-free* case, in which *only* one photon per mode at the device's output is collected, therefore the Hilbert space dimension \mathcal{H}_{CF} , to:

$$\dim \mathcal{H}_{CF} = \binom{m}{n}. \quad (2.23)$$

Nevertheless, it was demonstrated by Arkhipov [120], that n photons entering from the same input mode of an $m \gg n^2$ interferometer, tend to exit from separate output modes. Consider now an input state, whose representation in the Fock basis is given by:

$$|\psi\rangle_{in} = \hat{a}_1^\dagger \dots \hat{a}_n^\dagger |0_1, \dots, 0_m\rangle, \quad (2.24)$$

where \hat{a}_i^\dagger is the photon creation operator in the i -th mode. The input state evolves according to a unitary matrix \hat{U} , which acts as a map on the creation operators:

$$\hat{U} : \hat{a}_i^\dagger \rightarrow \sum_{j=1}^m \hat{U}_{i,j} \hat{b}_j^\dagger, \quad (2.25)$$

where m is the number of modes and \hat{b}_j^\dagger is the creation operator in the j -th output mode. Thus, the state will transform as:

$$\hat{U} |t_1, \dots, t_n\rangle = \prod_{i=1}^n \frac{1}{\sqrt{n_i!}} \left(\sum_{j=1}^m \hat{U}_{i,j} \hat{b}_j^\dagger \right) |0\rangle. \quad (2.26)$$

Since we have to take in account all the possible output configurations, the output state is:

$$|\psi\rangle_{out} = \sum_S \gamma_S |s_1, \dots, s_m\rangle, \quad (2.27)$$

where s_j is the number of photons in each j th output mode² and γ_S is probability amplitude, relative to the configuration S .

If we identify the input state with $|T\rangle$ and the output one with $|S\rangle$, the output probability associated to a specific configuration S is:

$$P_S = |\langle S | \hat{U} |T\rangle|^2 = |\gamma_S|^2. \quad (2.28)$$

It was demonstrated by Scheel [121] that the probability amplitudes for indistinguishable bosons, and thus the probability, are related to the permanents by:

$$\gamma_S = \frac{\text{Per}(U_{S,T})}{\sqrt{s_1! \dots s_n!} \sqrt{t_1! \dots t_n!}}, \quad (2.29)$$

where $U_{S,T}$ is the $n \times n$ sub-matrix of U , obtained by repeating s_i times its i th column and t_j times its j th row. It is now more clear, the difference between considering indistinguishable rather than distinguishable particles. In fact, in the case of distinguishable particles, the probability is

$$|\gamma_S|^2 = \frac{\text{Per}(|U_{S,T}|^2)}{s_1! \dots s_n!}, \quad (2.30)$$

which is computed starting from non-negative real numbers. Thus it is possible apply the Jerrum's algorithm and simplify the problem.

Now, to better understand the relationship between the permanent and the submatrices, let's consider an example. Suppose two photons enter the first two inputs modes (1,2) with the remaining modes in the vacuum state and let's assume to be in the case of *collision-free* condition. Let us estimate the amplitude of measuring the two photons exiting from the output modes (2,3). There are two ways in which this could occur: either the first photon comes out from mode 2 and the second from mode 3, or vice-versa, they swap and the first photon reaches mode 3 and the second mode 2:

$$\begin{array}{ccc} 1 & \longrightarrow & 2 \\ & & \searrow \\ 2 & \longrightarrow & 3 \end{array} \quad , \quad \begin{array}{ccc} 1 & & 2 \\ & \searrow & \nearrow \\ 2 & & 3 \end{array}, \quad (2.31)$$

therefore there are $2! = 2$ ways in which the photons can exit the outputs. Hence, the amplitude probability can be written as:

$$\gamma_{\{1,2\}} = \underbrace{U_{12}U_{23}}_{\text{photons don't swap}} + \underbrace{U_{13}U_{22}}_{\text{photons swap}}. \quad (2.32)$$

Now, looking at the matrix representation of the unitary operator \hat{U} :

$$U = \begin{pmatrix} U_{11} & U_{12} & U_{13} & U_{14} & U_{15} \\ U_{21} & U_{22} & U_{23} & U_{24} & U_{25} \\ U_{31} & U_{32} & U_{33} & U_{34} & U_{35} \\ U_{41} & U_{42} & U_{43} & U_{44} & U_{45} \\ U_{51} & U_{52} & U_{53} & U_{54} & U_{55} \end{pmatrix}, \quad (2.33)$$

²Since we are dealing with the collision-free case, s_j can only be 0 or 1.

we see that $\gamma_{\{1,2\}}$ is the 2×2 permanent, corresponding to the submatrix built by the cross selection among the input columns and the output rows.

In general, having fixed the input/output modes, with n photons there are $n!$ way in which they can reach the outputs and the associated probability amplitude will be related to an $n \times n$ submatrix permanent.

As already mentioned, in terms of computational complexity, the permanent is strongly believed to be intractable by means of a classical computer. In the following section, we are going to present with more details the connections between the permanent and the complexity classes.

2.4 The computational complexity of Boson Sampling

The boson sampling strength, in showing real evidences of the quantum-over-classical computation advantage, lies in the hardness to calculate permanents with classical algorithms.

The evaluation of the permanent belongs to the $\#\mathbf{P}$ complexity class, which is related to counting problems. We have introduced in previous chapters the classes \mathbf{P} and \mathbf{NP} , that, instead, are related to decisional problems. Thus, to compare the former class with the latter two, we introduce another class based on decisional problems but referred to the $\#\mathbf{P}$ one. To this scope, as mentioned in Section 1.1.4, we can use the concept of *oracle*, which is defined as a black box able to solve, in a unit time, problems belonging to a specific class. Supposing \mathbf{A} and \mathbf{B} are two complexity classes, we can define the class $\mathbf{A}^{\mathbf{B}}$ as the class of problems which are in class \mathbf{A} if we have access to an oracle for problems in class \mathbf{B} . With this definition we define $\mathbf{P}^{\#\mathbf{P}}$, as the class that corresponds to decision problems solvable in polynomial time by a \mathbf{TM} , given an oracle for $\#\mathbf{P}$ problems. $\#\mathbf{P}$ problems can require checking an exponential amount of different patterns, which can be done in an exponential amount of time with polynomial memory, so that the following relation holds:

$$\mathbf{P} \subseteq \mathbf{P}^{\#\mathbf{P}} \subseteq \mathbf{PSPACE}. \quad (2.34)$$

It is possible to identify a subgroup with an internal hierarchy among the different complexity classes encountered up to now. It is composed by the sequence $\mathbf{P}, \mathbf{NP}, \mathbf{NP}^{\mathbf{NP}}, \mathbf{NP}^{\mathbf{NP}^{\mathbf{NP}}}, \dots$, and it is called *Polynomial Hierarchy* (\mathbf{PH}). It is strongly believed that each of the levels constitutes a class more powerful than the lower levels (which justifies the term “hierarchy”). The widely accepted conjecture that $\mathbf{P} \neq \mathbf{NP}$. Since the ability to count solutions to \mathbf{NP} problems allows to climb all the way up to the top of the \mathbf{PH} , the classes $\mathbf{P}^{\#\mathbf{P}}$ and \mathbf{PH} are related as follows:

$$\mathbf{PH} \subseteq \mathbf{P}^{\#\mathbf{P}} \subseteq \mathbf{PSPACE}. \quad (2.35)$$

In this context, the Boson Sampling problem is efficient since the required physical resources scale polynomially with m and n , conversely with the classical case. The core of the problem relies on the fact that Boson Sampling is a solution devised to overcome the limit of Troyansky and Tishby in [122], where we move from the calculation of the permanent to the problem of sampling. In the following we briefly discuss the proof of the hardness for the exact Boson Sampling and its more difficult approximated version [33].

Exact Boson Sampling: The hardness of the problem relies on the computational effort to evaluate permanents. A formal proof is reported in Ref. [33], based on precedent remarkable results in the study of computational complexity:

1. *Stockmeyer’s approximate counting:* estimating to within a multiplicative error the probability that a polynomial-time randomized algorithm accepts a certain input is in $\mathbf{BPP}^{\mathbf{NP}}$. This is true even if the probability to be estimated is exponentially small [123].

2. It is always possible to efficiently write an $n \times n$ complex matrix A as a submatrix of an $m \times m$ unitary matrix (with $m > 2n$), modulo a rescaling of A by a factor of $\epsilon < 1/\|A\|$ [122].

Now, suppose to have an efficient classical algorithm A for exact Boson Sampling; using point (2) one could insert any $n \times n$ matrix X in the upper-left corner of a unitary matrix U by rescaling it of a factor $1/\|X\|$, and run A setting U as the unitary and $(1_1, 1_2, \dots, 1_n, 0_{n+1}, \dots, 0_m)$ as the input Fock state. The number of times that A will output the state with all photons in the first n modes will be proportional to the squared permanent of the upper-left $n \times n$ scattering matrix, *i.e.* the rescaled X . But then, we could also use Stockmeyer's counting algorithm to estimate the same probability to within multiplicative error, even if the rescaling of X decreases exponentially this probability. This implies that, if A exists, then estimating the square modulus of a permanent is in $\mathbf{BPP}^{\mathbf{NP}}$. But then, using this fact, AA have shown how to calculate the permanent of a matrix in $\mathbf{BPP}^{\mathbf{NP}}$ with a binary search. In fact, consider the matrix X_r of which we want to compute the permanent with the x_{11} entry replaced by $x_{11} + r$. By searching over r and using A to approximate the permanent square, one can exactly find such that $|Per(X_{r*})| = 0$. With this value, by solving a linear equation, the calculation of $Per(X)$ can be reduced to the calculation of the permanent of the bottom-right $(n \neq 1) \times (n \neq 1)$ submatrix of X . This procedure can be iterated to find efficiently the permanent of X , which is in conflict with the proof from Valiant [42] of the $\#\mathbf{P}$ -hardness of calculating the permanent. More in particular, this result would imply that a $\#\mathbf{P}$ -complete problem (and, as a consequence, every $\#\mathbf{P}$ problem) can be solved in $\mathbf{BPP}^{\mathbf{NP}}$, which would mean that [43]:

$$\mathbf{BPP} \subseteq \mathbf{NP}^{\mathbf{NP}}, \quad (2.36)$$

$$\mathbf{BPP} \subseteq \mathbf{P}^{\#\mathbf{P}}. \quad (2.37)$$

For the \mathbf{BPP} quantum analogue, that is \mathbf{BQP} , the following relation with $\mathbf{P}^{\#\mathbf{P}}$ as described in Ref. [47] holds :

$$\mathbf{BQP} \subseteq \mathbf{P}^{\#\mathbf{P}} \subseteq \mathbf{PSPACE} \quad (2.38)$$

which would mean that the entire polynomial hierarchy is contained in its third level. This kind of collapse would be a very striking result in computational complexity theory (even more than proving that $\mathbf{BQP} = \mathbf{BPP}$!), and it is considered extremely unlikely, concluding the proof that an efficient algorithm for the exact Boson Sampling cannot exist.

Approximate Boson Sampling: The hardness of approximating Boson Sampling cannot be regarded as rigorously demonstrated by AA, because is based on two unproven conjectures [101]. Still, these two "weak points" are considered very weak assumptions by the computational complexity community, and are well believed to hold [119]; indeed, they can be regarded as much "safer" assumptions than believing that factoring is not in \mathbf{BPP} . First, we need to introduce the Haar-random unitaries and Gaussian matrices. The Haar measure [124] is the best way to define the analogous of the uniform measure on unitary matrices. An operative procedure to pick a Haar-random $m \times m$ unitary is to draw the first row as a random unit vector in C_m , then the second as another random unit vector orthogonal to the first, and so on. In their work, AA prescribed Haar-random unitaries to be encoded in the m -mode interferometer. Since a random choice of the unitary is necessary to avoid any simplification/spoiling of the computational complexity. Next, we can define a complex Gaussian matrix X as a matrix whose entries are complex random values following a Gaussian distribution with a fixed mean and sigma (in our case, 0 and 1 respectively). The reason behind the interest in Gaussian matrices is that they naturally come as submatrices of random unitaries, as demonstrated by AA [33]. Because of this fact, the hardness demonstration of approximate BS is essentially based on proving the complexity of *Gaussian Permanent Estimation* (GPE).

Let us define two versions of this computational problem: estimating the square modulus of the permanent with additive error, or estimating the permanent with multiplicative error. More formally, given a $n \times n$ GM X ,

1. *Permanent anti-concentration*: $\forall X \sim \mathcal{N}_C^{m \times m} \exists$ a polynomial $Q \mid \forall (m, d) > 0$:

$$\text{prob}\left(\text{Per}(X) < \frac{\sqrt{m!}}{Q(m, \frac{1}{\delta})}\right) < \delta \quad (2.39)$$

2. *Gaussian permanent estimation (GPE_X)*: Approximating $\text{Per}(X)$ for a matrix $X \sim \mathcal{N}_C^{m \times m}$ of independent and identically distributed random Gaussian elements to within a multiplicative (\times) error is a $\#\mathbf{P}$ -hard problem.

The proof of the hardness then goes through the following results [33]:

Theorem 1 (GPE_± and Boson Sampling). If an efficient classical algorithm exists to sample from a distribution ϵ -close to the Boson Sampling distribution in total variation distance, then $|GPE|^2$ with an additive (\pm) error is in $\mathbf{BPP}^{\mathbf{NP}}$.

Theorem 2 (Haar-unitary hiding). : Any $n \times n$ submatrix of an $m \times m$ Haar random unitary is close in variation distance to a gaussian matrix if $m > n^5 \log(2n)$.

The loop closes when we take in account the Haar-random measure, randomly selecting the unitary matrix in order to avoid reductions of the computational complexity. Using the above conjectures, Aaronson and Arkhipov finally state that a classical polynomial-time algorithm for the approximate Boson Sampling would imply the collapse of \mathbf{PH} .

2.4.1 Experimental requirements for Boson Sampling

At this point we should stress the requirements that we need to "disprove"³ the Extended Church Turing Thesis (ECTT). The experiment will require the generation of n -photon Fock states, with a high degree of indistinguishability, to be injected in stable and low-loss m -mode linear optical interferometers, and finally collected and sent to an array of m photodetectors. Let us now identify the "boundaries" respect to the number of photons and modes, showing the scalability problems related to post-selection. Concerning the number of photons n , one could roughly identify three different regimes:

$3 \leq n \leq 10$	Regime where there are the first experimental attempts exploiting n -bosons interference with a medium size interferometer.
$10 \leq n \leq 50$	In this regime, computers start to suffer, since, the best known classical algorithm to compute permanents calculation (Ryser's algorithm [125]) requires $O(2^n n^2)$ operations. This is the region where a quantum device will start surpassing a classical computer.
$n \geq 50$	In this region the classical simulation becomes intractable. In principle we do not want to go beyond for technological reasons but also because: 1) since the ECTT is an asymptotic statement, it could never be formally disproved by any experiment, which could only provide strong evidence against it. 2) in the high-computational-complexity regime, certifying the correctness of the quantum device's output becomes a hard task.

³It is worth noting that a real proof cannot exist because, the ECTT itself is a conjecture. Consequently we can only see an evidence of the declared superiority of these quantum machines.

Moreover, the problem is not only circumscribed to photon generation but it also referred to the collection. In fact, because of experimental losses and non-unity detector efficiencies, any real experiment will require post-selection to obtain a fair sample of n -fold coincidence at the output of the interferometer. Since the post-selected rate will suffer of an exponential slowdown η^n (due to the probability to observe a n -fold coincidence event), it becomes very hard to perform an experiment with a very large number of photons, which means, that the ECTT could never be challenged in the very-high complexity regime. What can we say about the number of modes? There are two lower bounds that have to be fulfilled, according to AA [33]:

$\mathbf{m} > \mathbf{n}^2$	The reason is basically to remain in the collision-free regime, where the events with more than one photon per mode occurs with low probability. Even if AA expressed the belief that hardness-of-simulation still holds even in the $m \sim n$ regime, there would be less theoretical support in this condition.
$\mathbf{m} > \mathbf{n}^5 \log(\mathbf{n})^2$	(often reported in literature as $m > n^6$ or $m > n^{5.1}$): this is required for the Haar-Unitary Hiding Theorem to hold. AA believes that the bound can be lowered down to a much more feasible $m > n^2$.

In practice, $m > n^2$ has become the standard bound for the implementations of Boson Sampling. There is also an experimental reason not to lower this bound further: in the collision-free regime, it is possible to use non-number-resolving detectors (which give a “click” if one or more photons hit the sensitive area) which are much more compact and cheap than number-resolving ones [126]. One could say that, since the Boson Sampling complexity scales exponentially with the number of photons because of the hardness of calculating the permanent, it should scale exponentially as well in m because of the increase in the number of possible output configurations $\binom{m}{n}$. This would be true if a classical simulation of Boson Sampling would necessarily imply reconstructing the whole probability distribution ($\binom{m}{n}$ permanents of $n \times n$ matrices) and then sampling from that distribution. This is not true, since we are dealing with a sampling problem, that does not necessarily imply a full simulation of the system.

Following this road, further investigations have been performed to explore the scalability of its implementations and in order to provide bounds to satisfy with large scale demonstrations. In Ref. [127], AA analyzed Boson Sampling in the case of photon losses, while the level of error tolerance allowed in practical implementations, fabrication imperfections in the quantum circuit and investigations of alternative protocol have been analyzed extensively in Ref. [128–130], Ref. [131, 132] and Ref. [133], respectively. The definition of a real bound to prove a concrete advantage of quantum machines over classical computers is still an open problem and this attack can be addressed to further developments of hardware architectures [134] and sampling methods [135]. Moreover Clifford and Clifford [136] have been introduced recently a new sampling algorithm for Boson Sampling which requires $O(n2^n + \text{poly}(m, n))$ time resources plus an additional $O(m^2)$ space resources, clearly enhancing the number of photons and mode required to find evidences of a real quantum advantage. In Table 2.1 we list the Boson Sampling experiments performed in the last five years.

2.5 Scattershot Boson Sampling

The Boson Sampling problem aims to find concrete evidences that quantum protocols can be more efficient than the classical ones. However, in order to show the differences in terms of resources employed we have to discuss the scaling of n and m 2.4.1. Experimentally, one of the main difficulties in scaling up the complexity of Boson Sampling devices is the simultaneous generation of a large number of indistinguishable photons. From the first proof-of-principle demonstrations for small size Boson Sampling experiments [137–140], several theoretical and

Year	Authors	m	n	Platform	Source
2013	Spagnolo et al. [146]	5	3	FLM	SPDC
2013	Broome et al. [137]	6	3	in-fiber	SPDC
2013	Spring et al. [138]	6	3	SoS	SPDC
2013	Tillmann et al. [139]	5	3	Si	SPDC
2013	Crespi et al. [140]	5	3	FLM	SPDC
2014	Spagnolo et al. [198]	13	3	FLM	SPDC
2014	Carolan et al. [199]	21	5	Si_2N_3	SPDC
2015	Carolan et al. [22]	6	6	SoS	SPDC
2015	Tillmann et al. [220]	5	3	FLM	SPDC
2015	Bentivegna et al. [149]	13	3	FLM	SPDC
2016	Crespi et al. [182]	8	2	FLM	SPDC
2017	Loredo et al. [143]	6	4	in-fiber	QD
2017	He et al. [142]	8	5	bulk	QD
2017	He et al. [145]	9	2	fused-quarz	QD

Table 2.1. List of Boson Sampling experiments performed from 2013 to nowadays.

experimental works have reported a significant advance in number of modes and photons, by means of different techniques [141] and sources [142, 143]. However, despite significant improvements in technologies, currently the most widely used source remains the Spontaneous Parametric Down Conversion (SPDC). In this context, the scalability of these sources as the relevance of the detectors' efficiency have been studied in [144]. Its advantages are the high photon indistinguishability, collection efficiency and relatively simple experimental setups. Beyond these features, the SPDC suffers two important drawbacks:

- since the process is spontaneous, the phenomenon is *non-deterministic*;
- the laser pump power, and hence the source brilliance, has to be kept low to prevent unwanted higher-order terms in the generation process.

These disadvantages place strict limitations on the implementations of Boson Sampling experiments. In order to use the SPDC sources at their best possibilities and to increase the generation rate, Lund et al. have introduced a new scheme named by Aaronson *Scattershot*

Boson Sampling (SBS) [147, 148]: a generalization of the Boson Sampling problem that exploits the full capabilities of SPDC to get an exponential enhancement in the generation rate with respect to the the standard version.

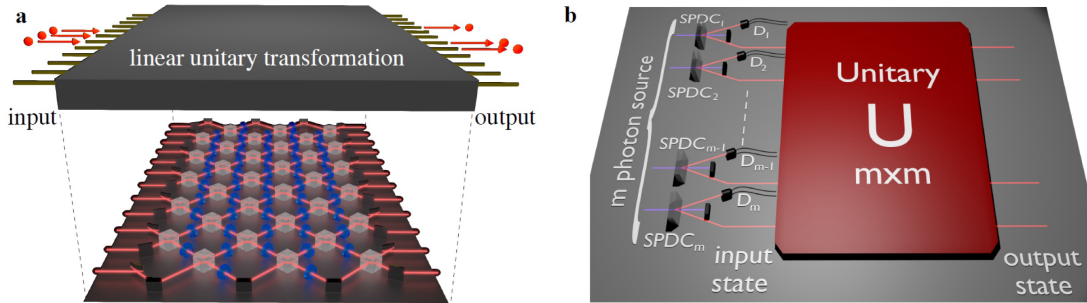


Figure 2.4. a), Standard Boson Sampling experiment. A n -photon Fock state is injected into a linear interferometer. The outcome depends on the probability distribution retrieved by their evolution b), Scattershot configuration. In this case we have m heralded single-photon sources, one for each input port, that are coupled to the interferometer. n photons ($n < m$) are probabilistically generated from SPDC sources and injected into the interferometer. Each detected n -photon event at the interferometer's output can be assigned to its corresponding input state by the heralding detectors, Ref. [149].

The idea is the following: we know that the photon generation is due to SPDC sources and it is not efficient, but very simple to use. Consequently this scheme, involves the connection of k , with $k > n$, parametric down conversion sources to different input ports of the interferometers in order to boost the generation rate while exploiting their individual case of implementation (see Fig. 2.4).

Suppose each SPDC crystal yields a single photon with probability ϵ per pulse. By pumping all k sources with simultaneous laser pulses, n photons will be simultaneously generated in a random, but *heralded* [149] set of input ports with probability equal to:

$$P = \binom{k}{n} \epsilon^n, \quad (2.40)$$

which for $k \gg n$ represents an exponential improvement in generation rate with respect to the usual, fixed-input Boson Sampling.

Thus, the task of the Scattershot Boson Sampling is to sample from the output probability distribution given a random set of input configurations. In other words, it is the problem of sampling from the input/output joint probability distribution. All of these features suggest that the SBS approach to Boson Sampling will be decisive in future, larger experiments designed to reach the quantum advantage regime. Note that the pump laser power does not need to be increased k -fold, as the laser can sequentially pump each SPDC source with very little loss to down-converted photons. In this way, the ratio between one-pair production rate and higher order terms can be kept low, as in the standard space-multiplexed heralded photon sources.

What about computational the complexity of SBS? Intuitively, one could say that increasing the number of possible inputs should (at least) keep the same complexity of the original problem. Further details on Scattershot Boson Sampling computational complexity, refer to [147].

We will now discuss how SBS may bring within reach an experimental regime approaching quantum advantage. Let us consider experiments with $N = 2000$ events. In this regime, with high probability each recorded event is sampled from a different input state, provided that $\binom{m}{n} \gg N$ and assuming the use of one SPDC source per input mode. To get an insight into the hardness

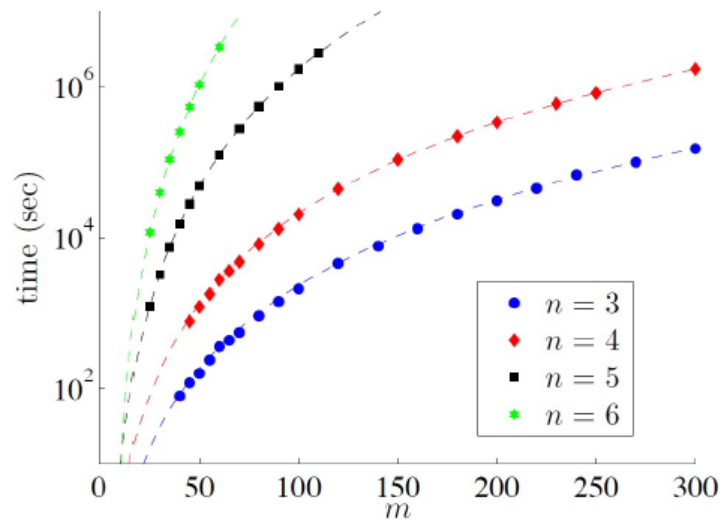


Figure 2.5. Simulation of SBS experiment. Time required with a laptop to calculate $N = 2000$ Boson Sampling probability distributions, each one corresponding to a different input configuration, as a function of the number of modes m , for different number of photons n . The simulation is performed by using a brute-force approach: for each event, an input is drawn uniformly and the full probability distribution is calculated to sample a single event. Ref. [149].

of calculating the whole output distribution corresponding to a SBS experiment, we illustrate the required computational time on a standard laptop in Fig. 2.5. The main advantage of the SBS approach is to decrease the experimental run-time with respect to the usual, fixed-input setup. We get an estimated runtime of ~ 108 seconds for a 2000-event, fixed-input BS experiment with $n = 4$, $m = 100$. The corresponding SBS experiment uses $k = 100$ SPDC sources in parallel, resulting in a simulated quantum runtime of ~ 50 seconds. These estimates clearly illustrate the boost in computational speed provided by the Scattershot approach. Moreover, it has been shown that the computational resources for the Scattershot implementations, reduce the bound for quantum advantage to nearly 7 photons [150].

In Chapter 5 we will show our work reported in Ref. [149], introducing the first experimental Scattershot approach to the Boson Sampling.

2.6 Discussion

The study of the phenomenology associated to many-body evolution of bosonic states in an interferometer, as pointed out, if we compare the quantum approach versus the classical computation, the latter is believed to be hard to compute, due to the hardness of the problem. At this point the important question is: "*how can we be sure that the output distribution corresponds to a genuine outcome of many-body quantum interference of indistinguishable states?*". For short sizes it is not evident the problem because we can still deal with a reasonable little calculus to retrieve the output distribution and verify the sampling through the probabilities. Nevertheless, if we want to study the protocol for larger sizes, at this regime we cannot be sure (considering the fact that we are dealing with a same problem) if the outcomes corresponds to the indistinguishable behaviour or not. This is due to the exponential increase in m and n of the number of possible output states, in addition to the permanent calculus. In Chapter 5 we will deal with this problem, focusing on the so-called *validation* problem or, what protocol we can use to certify true quantum interference in an efficient way. This is important not only for Boson Sampling itself but also to certificate quantum interference which is at the core of LOQC.

Chapter 3

Linear optics experiments with integrated photonics

In this Chapter we introduce the field of integrated photonics, describing the most diffused techniques whose permit the implementation of any unitary transformation, using only phase shifters and beam splitters. We firstly discuss the linkage between unitary transformation and interferometry showing the algorithms involved for the fabrication of such structures. As following, we will focus on the description of the most known techniques, focusing in particular to the femtosecond laser writing which lies at the core of the platforms used in this PhD project.

3.1 Experimental realization of any unitary matrix

We start by describing the motivations behind the development of integrated photonics. Remarkably, the starting point is the question: "how can I implement any unitary transformation on the optical modes?" The answer is positive. This question was answered by Reck, Zeilinger, Bernstein and Bertani [15], who demonstrated that any discrete finite-dimensional unitary operator acting on m optical modes can be realized in the laboratory using optical devices, more precisely using a sequence of two-dimensional beam-splitters and phase shifters. Indeed, given a set of linear optical elements, we know that they can be always described by a unitary matrix. If we assign to each optical device a specific matrix (*e.g.*, the BS matrix), the whole system unitary representation can be obtained by the matrix product among all the devices.

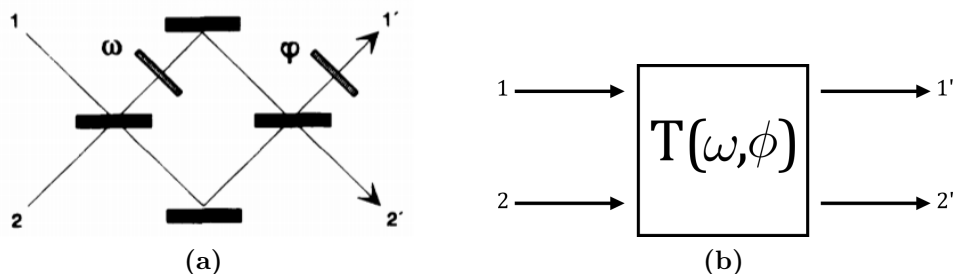


Figure 3.1. (a) A Mach-Zehnder interferometer made with 50 : 50 beam-splitters, phase shifters and mirrors. (b) Corresponding transformation dependent on the parameters (ω, ϕ) . Ref. [15].

The experimental realization of the most general element belonging to $U(2)$ is already known. It uses a lossless beam-splitter matrix with a phase shifter at one output port, and transforms an input state with modes (k_1, k_2) into the output state with modes (k'_1, k'_2) :

$$\begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} e^{i\phi} \sin \omega & e^{i\phi} \cos \omega \\ \cos \omega & -\sin \omega \end{pmatrix} \begin{pmatrix} k'_1 \\ k'_2 \end{pmatrix}, \quad (3.1)$$

where ω is a parameter describing the reflectivity $\sqrt{\mathcal{R}} = \sin(\omega)$ and the transmittance $\sqrt{\mathcal{T}} = \cos(\omega)$ of the beam-splitter. The parameter ϕ can be realized by an external phase shifter after the BS. Varying the input port phases, the beam-splitter matrix performs any transformation in $U(2)$ [151]. A varying reflectivity beam-splitter can be substituted by a Mach-Zender (MZ) interferometer with 50 : 50 BSs (Eq. 3.1).

Following this approach, one can perform an experiment equivalent to any unitary matrix $U(N)$ by using beam-splitters and phase shifters devices, such that the full N -dimensional Hilbert space can be factorized in successive $U(2)$ transformations, acting on its two-dimensional subspaces.

By this approach any *arbitrary* unitary operator can be built *only* with beam-splitters and phase shifters. It is possible to verify that the maximum number of beam-splitters needed to build a general N -dimensional matrix is [15]:

$$\binom{N}{2} = \frac{N(N-1)}{2}. \quad (3.2)$$

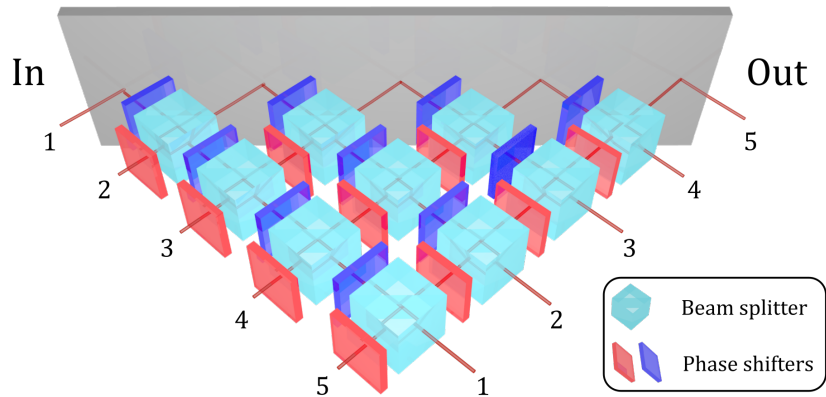


Figure 3.2. A triangular array of beam-splitters implementing a 5×5 unitary matrix. Each diagonal row reduce the effective dimension of the Hilbert space by one.

Since this number grows quadratically, *i.e.* $O(N^2)$, it represents an efficient scaling law for experiments in higher-dimensional Hilbert spaces. The practical implementation of this method involves a triangular array of beam-splitters: each diagonal row reduce the effective dimension of the Hilbert space by one (3.2). However, its physical realization with *conventional* optical techniques results to be out of reach. Indeed, with a rather high number of modes, the size of the interferometric network grows too much, reducing the overall stability and, finally increasing the losses.

This problem can be overcome thanks to the technology of *integrated photonics* through optical confinement of light in small size optical chips. Let us introduce in more details this technology in next sections.

3.2 Two-dimensional waveguides

The basic concept of optical confinement is quite intuitive. A medium of refractive index n_2 embedded in a medium of lower refractive index n_1 acts as a light “trap” within which optical rays remain confined by multiple total internal reflections at the boundaries if $n_2 \gg n_1$. The light is transported through the inner medium without radiating into the surrounding medium. The most widely used of these waveguides is the optical fiber, made of two concentric cylinders

of low-loss dielectric material such as glass. Integrated optics is the technology of integrating various optical devices and components all into a single chip. Optical waveguides provide the connections between these components and such chips represent the optical version of electronic integrated circuits. In the following we give a brief description of two-dimensional waveguides and of the coupling between them (for an in-depth analysis see e.g. [110]).

As before, we can consider a rectangular dielectric waveguide which is a cylinder of refractive index n_1 with square cross section and width d , embedded in a medium of slightly lower refractive index, namely n_2 . The components of the wavevector $k = (k_x, k_y, k_z)$ of guided light must satisfy the condition

$$k_x^2 + k_y^2 \leq n_1^2 k_0^2 \sin^2 \theta_c \quad (3.3)$$

where $\theta_c = \cos^{-1}(n_2/n_1)$ and k_0 is the wave vector in the vacuum. Depending on the waveguide size and on its refractive indices, the number M of guided modes satisfying equation (3.3) is

$$M = \frac{1}{4\pi} (d n_1 k_0 \sin \theta_c)^2 \sqrt{n_1^2 - n_2^2}. \quad (3.4)$$

M is proportional to the cross section as d , thus, having fixed n_1 and n_2 , single mode waveguides can be realized by simply adopting a proper cross section¹. If two waveguides are sufficiently close such that their fields overlap, light can be coupled from one into the other. Optical power can be transferred between the waveguides, an effect that can be used to make optical couplers and switches. Power is transferred between two waveguides in a periodic way: the light jumps back and forth from one waveguide to the other until their fields overlap. In the simple case of two identical waveguides - labeled 1 and 2 - brought close together for a length z , by injecting light into the first waveguide, the output power $P_i(z)$ with $i = 1, 2$, will be:

$$\begin{aligned} P_1(z) &= P_1(0) \cos^2(Cz) \\ P_2(z) &= P_1(0) \cos^2(Cz + \phi) \end{aligned} \quad (3.5)$$

where C is the coupling coefficient and ϕ is a phase difference between the two fields.

3.3 Integrated circuit fabrication techniques

Several techniques have been developed to build waveguides into a small size chip. Such approaches to largely improve performance of quantum optics experiments since spatial mode matching, which is crucial for classical and quantum interference, since nearly perfect in such an architecture. The first experiment on integrated quantum photonics was reported in 2008, on silica-on-silicon (!SoS) waveguide quantum circuits that are fabricated and used to achieve quantum logic gates with high fidelity [152]. Coupling between waveguides, to realize BS-like operations, may be achieved when two waveguides are brought sufficiently close together that the evanescent fields overlap: this system is known as *directional coupler* (Fig. 3.3) [110]. The coupling depends on the separation between the two (or more) waveguides and on the interaction length. In the following, let's thus introduce the most recent techniques: lithography and femtosecond laser micromachining Femtosecond laser micromachining (FLM), which are nowadays widely adopted.

3.3.1 Lithography

Lithography is described as the transfer of a pattern to a photosensitive material by selective exposure to a radiation source. A photosensitive material is a material that changes its physical properties when exposed to a radiation source. If we selectively expose a photosensitive material

¹The diameter is $\sim 4 \mu m$ as for single mode fibers

²Ref.https://www.rp-photonics.com/fiber_couplers.html

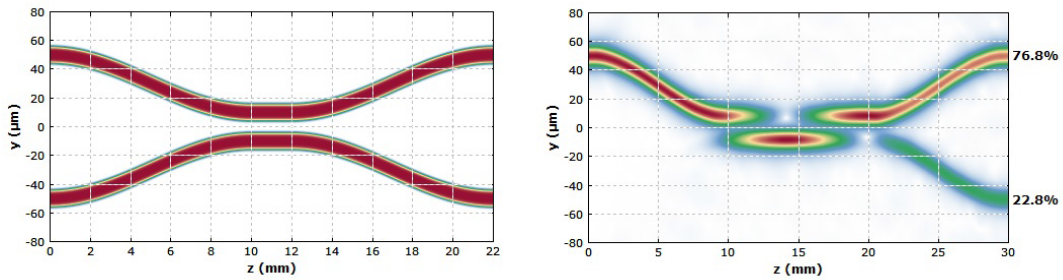


Figure 3.3. Left: refractive index profile of a directional coupler. Right: simulation of the amplitude ²field behaviour in a directional coupler, where the interaction length is chosen to reach a specific value of the reflectivity. A full description of the methodology used is provided in Section 3.4.

to radiation (e.g. by masking some of the radiation) the pattern of the radiation on the material is transferred to the material exposed, as the properties of the exposed and unexposed regions differs [20, 21]. In lithography for micromachining, the photosensitive material used is typically a photoresist. When the photoresist is exposed to a radiation source of a specific wavelength, the chemical resistance of the resist changes. On the other hand, if the resist is placed in a developer solution after selective exposure to a light source, it will etch away one of the two regions (exposed or unexposed). If the exposed material is etched away by the developer and the unexposed region is resilient, the material is considered to be a positive resist. If the exposed material is resilient to the developer and the unexposed region is etched away, it is considered to be a negative resist. By simply varying the pattern of the exposed region it is possible to realize different circuits in which the surviving resist acts as a waveguide. This type of waveguide can only be arranged in 2-dimensional configurations and present squared cross sections. In particular, since the squared geometry is not isotropic, this waveguide does not allow the optimal propagation of gaussian modes thus preventing their use for manipulation of polarized light.

Silicon waveguides: These circuits are realized through lithographic techniques, with all the related disadvantages, but the possibility to exploit the huge technology expertise and supply chain from electronic industry, as well as the full compatibility with CMOS technology, are strong incentives. Indeed, silicon circuits are typically realized using electron beam or optical lithography and reactive ion etching, mainly using CMOS manufacturing tools [153]. The silicon crystal is transparent in the $1.3 \mu\text{m} - 1.6 \mu\text{m}$ region. The most used configurations involve rectangular shape silicon waveguides ($n = 3.47$) surrounded by a SiO_2 bottom cladding ($n = 1.44$) and a top cladding of SiO_2 or air. The high index contrast provides all the advantages of a strong mode confinement, such as the possibility to have waveguides with sub-micrometric sections, tight bends (with radius down to $3 \mu\text{m}$) and very dense and compact circuits. Rectangular section are preferable to minimize the scattering losses induced by surface imperfections, which are much higher at the sidewalls with respect to top and bottom surface, due to characteristic fabricative errors of processes like electron beam lithography. The strong light confinement in silicon waveguides makes them possible candidates as integrated probabilistic single photon sources. Even if a silicon crystal lacks $\chi(2)$ nonlinearity due to its centrosymmetric structure, the high field intensity allows to exploit the $\chi(3)$ nonlinearity to generate photon pairs through spontaneous four-wave mixing [154, 155]. The use of optical resonators [156] to enhance the power density has proven its efficiency in boosting photon generation. The large thermo-optic response of the silicon has advantages and disadvantages. To maintain the stability of the circuits a control in temperature is indeed necessary, and high-power effects happening in microring resonators has to be kept under control. At the same time, thermal tuning can be easily achieved in such devices, thanks to the very strong dependence of the resonance wavelength from the effective optical path.

3.3.2 Femtosecond laser micromachining

Femtosecond laser micromachining (FLM) is a fabrication technique in which a femtosecond laser beam is used in order to induce structural changes in the transparent material. The very first demonstration of FLM was done in 1994, when it was employed to create surface ablation features on silica and silver substrates [157, 158]. Subsequently, Davis et al. [159] managed to write for the first time optical waveguides in the bulk of various transparent materials. Nowadays, just two decades after these very first demonstrations, FLM has evolved into a very powerful and versatile experimental tool for several applications [24–26]. The inscription of integrated devices is achieved by focusing the femtosecond laser beam with the use of a common microscope objective into the bulk of a transparent material. This gives rise to a series of nonlinear absorption processes that lead to a permanent, local modification of the refractive index of the material. By translating the sample, using a system of three dimensional translational stages, it is possible to realise various different structures such as simple optical waveguides, or more complex waveguide circuits. The motivation to adopt this method, over other integration techniques, is a number of very important advantages as presented below:

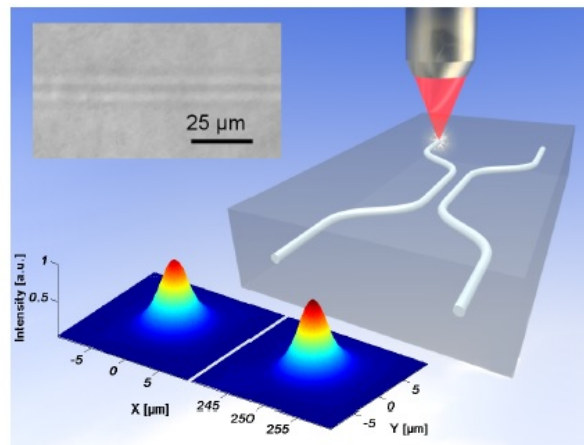


Figure 3.4. Scheme of femtosecond laser writing: directional couplers are written on a borosilicate glass, simply translating a focused laser beam along the chip. Non-linear absorption of the medium, change permanently the refraction index. Then, we can choose the desired coupling between waveguides setting properly distance among them and orientation through space. Ref. [28].

1. Three-dimensional capabilities. It is possible to inscribe different 3D geometries at different depth of the glass: this is a unique feature of this technique.
2. Rapid device prototyping (the device pattern can be easily changed by a simple software control).
3. Simpler and less expensive production plants: we do not need clean rooms or very expensive machines to write circuits and the designing process is done by a simple software, thus avoiding all the consuming steps required by other techniques such as mask prototyping in photolithography.

We shall focus on the femtosecond micromachining of bulk transparent materials, *i.e.*, materials that do not have any linear absorption at the wavelength of the femtosecond laser as for example *borosilicate glass*. There are unique advantages in favour of femtosecond laser micromachining of transparent materials over other photonic-device fabrication techniques. *First*, the nonlinear nature of the absorption confines any induced change to the focal volume. This spatial confinement, combined with laser-beam scanning or sample translation, makes it possible to model geometrically complex structures in three dimensions. *Second*, the absorption process

is independent of the material, enabling optical devices to be fabricated in compound substrates of different materials. *Third*, **FLM** can be used for the fabrication of an ‘optical motherboard’, where all interconnects are fabricated separately, before (or even after) bonding several photonic devices to a single transparent substrate. When a femtosecond laser pulse with a high enough peak pulse intensity is focused into a material, an optical breakdown is observed. The laser pulse energy is partially transferred to the electrons in the short duration of the pulse. The highly excited electrons thermalize with the ions and alter the material permanently. Depending on the degree of excitation, cracking, void formation or localized melting occurs. In the absence of impurities, carriers are generated initially by multiphoton absorption, promoting electrons from the valence to the conduction band. Several photons must be incident on an electron at the same time for the process to occur with a high probability. The micromachined feature size will depend on many experimental parameters: E (the pulse energy), τ (the pulse duration) and NA (the focusing numerical aperture). However, under special conditions, exposure to multiple pulses can further change the feature size.

In the next section we will introduce in more details the main features of this technique, in particular regarding the malleability to realize very complex circuits inside and the choice of a good material to implement quantum information tasks without significant losses.

3.4 Directional couplers properties

Integrated optical circuits are built upon architectures of directional couplers (**DC**), with additional specific geometries to account for the deformation-induced phase shifts. Despite the clear mathematical structure, and differently from the development of miniaturized electronic circuits, linear optics does not have a unique manufacturing technique. The most common material adopted for the fabrication of integrated photonic circuits, as well as for optical fibers, is fused silica. There are indeed several advantages in considering silica as the medium for the propagation:

- low birefringence ($B = (7 \pm 1) 10^{-5}$, and order of magnitude lower than silica) and low propagation losses (~ 0.01 dB).
- low temperature dependence (no random changes with phases).
- very large spectral region of operations (UV to infrared).

Using a pulsed laser with high repetition rate, it is possible to induce a isotropic thermal diffusion that provides a fusion around the focal point, that creates a waveguide with a circular cross section³. In this way it is possible to avoid to choose a specific shape of the laser beam. The writing process is based on controlling phase shifters and transmissivities by changing its optical path during the translational movement. According to the equation (3.6) we can individuate three features: (a) Controlled deformation of the S-bent waveguide at the input of each directional coupler and coupling geometry allow independent control over the phase shift and transmissivity. (b) The deformation is given by a non-linear coordinate transformation, which is function of a deformation coefficient d . The graph shows the undeformed S-bend together with a deformed one; the experimental dependence of the induced phase shift on the deformation parameter d . (c) Control over the transmissivity of the directional coupler is performed by modulating the coupling coefficient; this is achieved by changing the waveguide spacing in the coupling region by rotating one arm of the directional coupler out of the main circuit plane. Let us now discuss in more details, how **DC** are modulated.

³At wavelength of 800 nm, the waveguide supports a Gaussian mode with $8 \mu\text{m}$ diameter to $1/e^2$

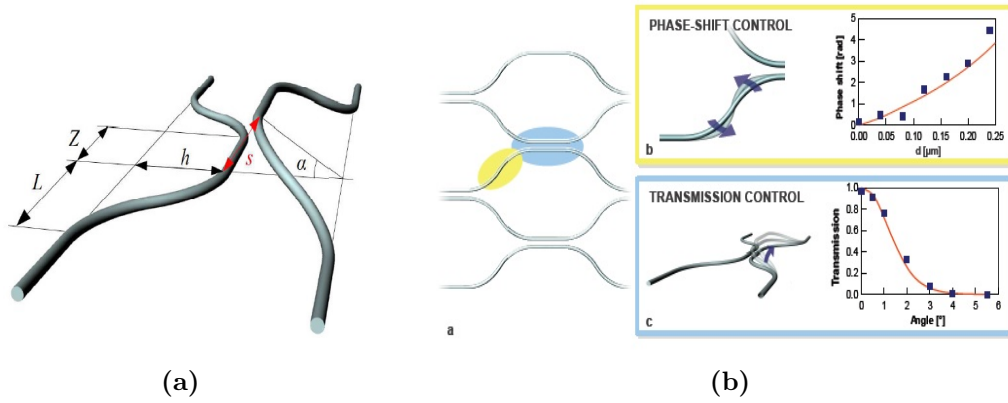


Figure 3.5. a) Scheme of a 3D directional coupler, where we define all the relevant geometric parameters. b) Controlled deformation and rotation of the S-bent waveguide at the input of each directional coupler allow independent control over the phase shift and transmissivity. The graphs show the variation of phase and transmissivity with the two parameters. Ref. [138].

Controlling phase shifters: to control phase shifters Ref. [138] proposed to extend the optical path by performing an S-bend with this coordinate transformation (see Fig. 3.5)

$$y(x) = -\frac{h}{2} \cos\left[\frac{2\pi}{L}x\right] \longrightarrow y(x + \delta(x)) = -\frac{h}{2} \cos\left[\frac{2\pi}{L}(x + d \sin\frac{2\pi}{L}x)\right] \quad (3.6)$$

with particular precision during the bending in order to avoid losses as possible.

Controlling transmissivities: to control transmissivities and phase shifter at the same time, it is used the 3D geometry capability, rotating the waveguide around its axis. In this way, we do not modify the S-bending. The transmissivity in directional couplers have the form $T = \sin^2(k(s)Z)$. The transmissivity depends on the distance between the two arms at the interaction length which decays exponentially through its coupling constant ($k = k_0 e^{-\frac{s}{s_0}}$). Thus, displacing the waveguides in a 3D arrangement, by changing the angle α it is possible to obtain any desired transmissivity as shown in the following equation:

$$T = \sin^2\left(e^{\log(k_0 L) - \frac{\sqrt{C_1 - C_2 \cos \alpha}}{s_0}}\right) \quad (3.7)$$

where s is the interaction length between beam splitters, C_1 and C_2 are expressions that take in account parameters h and L (see Fig. 3.7).

3.4.1 Tunable phases

One of the key advantages of FLM is that the fabricated devices are extremely stable and robust, and after the writing process, the interferometer is permanent without any chance to control inner phases. Recent works showed how it is possible to tune inner phases even with the case of borosilicate glasses [160, 161].

Flamini et al. [161], demonstrated recently a thermally reconfigurable MZ for 1550 nm in alumino-borosilicate glass, that operated based on thermo-optic modulation. In their approach, the MZ consisted of two symmetric directional couplers with balanced splittings, fabricated $70 \mu\text{m}$ beneath the glass surface. The arms of the interferometer were made 12 mm long and were placed at different depths, one at $14 \mu\text{m}$ and the other at $115 \mu\text{m}$. Above the

shallow arm a gold resistor was fabricated by sputtering a thin film of gold and then patterning it with FLM ablation as shown in Fig. 3.6. By applying a small voltage to the electrode, phase modulation was achieved due to thermal index change. The phases change according to a linear realization:

$$\Phi = \phi_1 - \phi_2 = \Phi_0 + \sum_i \alpha_i P_i \quad (3.8)$$

where P_i are the powers dissipated from the resistors and α_i coefficients linked to the heaters. In this way, the devices were able to achieve a full 2π phase shift and fringe visibilities greater than 95%. This method can be used not only to modulate an interferometer phase, but also to impose arbitrary phases to waveguide devices in general.

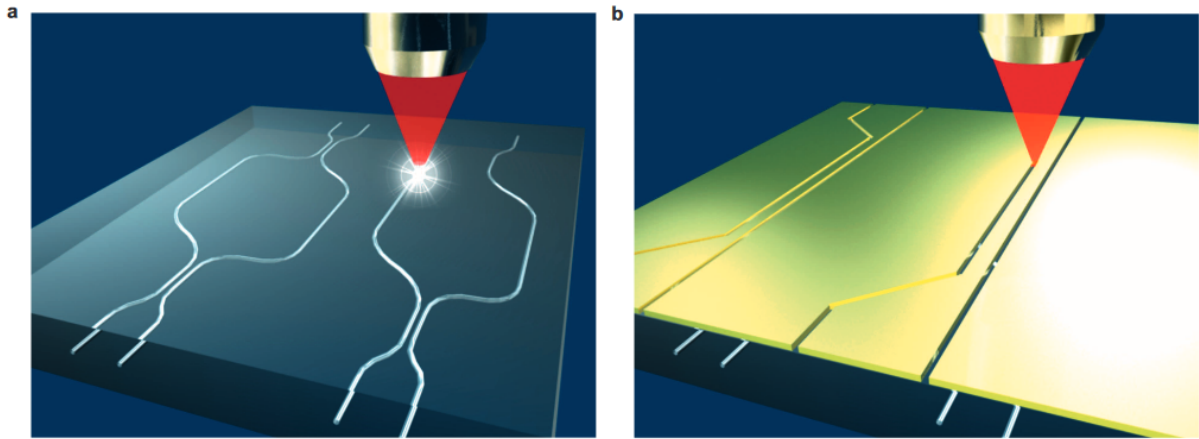


Figure 3.6. Illustrative image that shows the two procedures to realize a photonic chip with a tunable phase. a) The interferometer is written into the glass following the standard procedure as shown in previous sections. b) After gold coating the sample top surface, the resistors are built by ablation using the same femtosecond laser. Accurate alignment ($\sim 1 \mu m$) between the resistors and the waveguides arms is achieved by using the reference markers inscribed on the glass surface together with the interferometers. Ref. [161].

Tunable optical chips are believed to be a keystone for further developments in integrated photonics. In Table 3.1 let us report a table with the state-of-art of photonic chip developed so far

3.5 Description of the experimental tools used for generation, manipulation and collection of single photons

In this section we will describe briefly the tools that we exploited to perform our experiments. Photons generation and manipulation have been carried out through a probabilistic source of kind χ^2 : the *Spontaneous Parametric Down Conversion* (SPDC). In SPDC [13,171,172] the pump field excite the second order polarization of a nonlinear crystal, generating outcoming radiations at frequencies E_s and E_i , with momentum k_s and k_i . There are two possible configurations for the linear polarizations of outcoming photons, dictated by material dispersion and crystal orientation with respect to the pump beam polarization: in type I SPDC, the two outcoming photons have the same polarization, which is orthogonal to the polarization of the pump, while in type II they have orthogonal polarizations. Also, in type I the two photons are emitted at the opposite point of the same cone, while in type II the two cones (ordinary and extraordinary) are slightly non-centered. The phase matching condition requires:

$$\omega_p = \omega_i + \omega_s \quad (3.9)$$

Year	Authors	Platform	N_{phas}	(m,n)	λ
2009	Smith et al. [162]	SoS	1	(2,2)	830
2009	Matthews et al. [163]	SoS	1	(2,4)	780
2011	Shadbolt et al. [164]	SoS	8	(6,2)	808
2012	Bonneau et al. [165]	Ti:LN	1	(2,2)	1550
2012	Bonneau et al. [166]	SoI	1	(2,2)	1550
2014	Jin et al. [167]	LN	1	(4,2)	1560
2014	Metcalf et al. [168]	SoS	1	(6,3)	830
2015	Flamini et al. [161]	FLM	1	(2,2)	1550
2015	Carolan et al. [22]	SoS	30	(6,6)	808
2016	Pitsios et al. [175]	FLM	2	(5,2)	785
2017	Harris et al. [23]	SoS	176	(26,1*)	1570

Table 3.1. Non-exhaustive list of experiments performed with tunable integrated circuits. First demonstrations of reconfigurable circuits for quantum information tasks have been reported in the last years. While SoS-based devices can benefit of the know-how and the facilities developed for the electronics components, the unique advantages in versatility, mode-matching and compatibility with telecom optical fibers offered by FLW photonic circuits make this technology promising for larger scale implementations. Here, n is the number of photons, m is the number of modes of the interferometer, λ is the work wavelength and is the number of phases reconfigurable. The circuit is tunable also in polarization. The circuit * inject coherent laser in each input

$$\vec{k}_p = \vec{k}_i + \vec{k}_s \quad (3.10)$$

which, in practice, translates in conditions over the propagation directions of the outcoming photons with respect to the pump direction and the crystal axis, for specific signal and idler wavelengths. Concerning the photon number, the general output of SPDC is a squeezed state [173] of the form:

$$|\Psi\rangle = \sqrt{1 - \lambda^2} \sum_{n=0}^{\infty} \lambda^n |\psi\rangle_i |\psi\rangle_s \quad (3.11)$$

where λ is the squeezing parameter, and n is the number of photons. Consequently, no genuine Fock state can be produced via SPDC. The most followed approach when needing single-photon Fock states consists in having $\lambda \ll 1$, so that most of the times no downconversion happens, but the ratio between the numbers of single pair generations and the number of double pair generation (which is equal to λ) is low, and the heralding of an idler from a non-number-resolving detector implies with high probability a single-photon signal. This characteristic, of course, comes at the cost of a reduced source brilliance. In practice, one has to compromise between $g^{(2)}$ and generation rate. Number-resolving detectors can of course help in discriminating higher-order terms, but they are much less compact and cheap. SPDC sources have been, and still are, the main tool of quantum optics experiments. Other possible sources exploited are non-linear generation of χ^3 as for example *Four-Wave Mixing* [16, 17]. Basically, this effect is present in every medium that exhibits non linearity with the constraint of phase-matching and the most promising material that provides strong χ^3 are silicons. Usually this kind of sources needs of long interaction region and they are very useful when we deal with fully integrated sources.

Nevertheless, if we provide a high power peak and if we keep the integrated power low (to avoid pairs unwanted), SPDC remains the most diffused source in the optical community (as we showed in chapter 2). This is due to its natural predisposition to the collection, high indistinguishability of photon pairs (nearly 98% of visibility) and because it is possible to generate (thanks to the phase matching conditions) couple of entangled photons (Bell states) with same of orthogonal polarization, depending on the type. The only drawbacks are the non-efficient generation and the probabilistic behaviour which do not permit full control of the source and very long experimental runtime.

In this thesis we will use mostly this kind of sources for the generation of 2- and 3- photon states. At this point, if we want to obtain indistinguishable photons, we have to manipulate them in a controllable way. In Fig. 3.7 it is shown the generation and manipulation stage that acts before the injection in the integrated device. Here we give a brief description of the setup: single photons were generated at 785 nm with a type II parametric down-conversion process pumping a crystal (2-mm long BBO) with a 392.5 nm wavelength Ti:Sa pulsed laser (pump laser for SPDC generated previously by BiBO with a second harmonic generation). The two photons are spectrally filtered by means of 3 nm interferential filters (F), and coupled into single mode fibers. The walk-off effect is compensated by the usage of HWP and BBO/2 arranged properly in order to correct deviation of the group velocity and spatial momentum [110]. The indistinguishability of the photons is then reached by a polarization compensation stage (PC), and by propagation through delay lines for each possible path (used to adjust the degree of temporal distinguishability) before injection into the interferometer via a single mode fiber array. The generalization to a multiple input configuration has been employed in the same way and it is discussed in details at the beginning of Chapter 5. After the evolution through the integrated devices, photons are collected via a multimode fiber array and counted with photo-detectors. A reliable single photon detector is a device able to detect energies of the order of 10^{-19} joules, with high efficiency, low dark counts (false positives), a small timing jitter (i.e. variation in temporal response) and a high maximum count rate. An additional characteristic is the ability to resolve the number of incoming photons: non-photon-number-resolving detectors are on-off devices that can only distinguish between zero and non-zero photons (apart from efficiency).

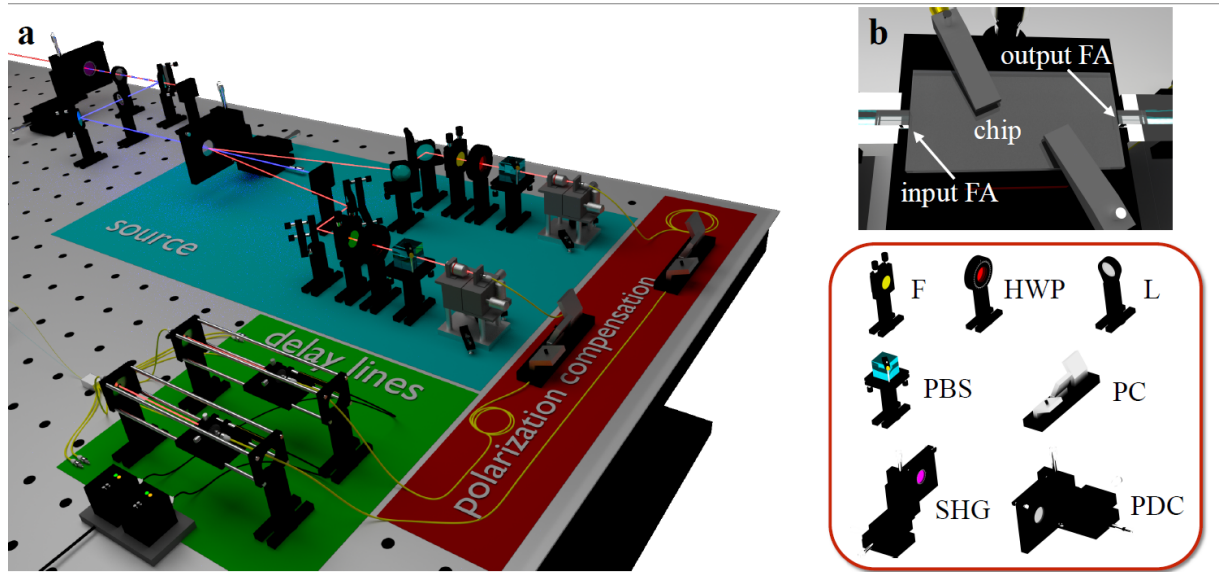


Figure 3.7. The photon source (IF, HWP, PBS, PC, PDC, DL and SMF). (b) Photon injection (extraction) before (after) the evolution through the interferometer. DL, delay lines with motorized stages; FA, fibre array; HWP, half-wave plate; F, interferential filter; PBS, polarizing beam splitter; PC, polarization compensator; PDC, parametric down-conversion; SMF, single-mode fibre

Avalanche photodiodes (APD) typically use, as active area, a semiconductor lattice with an applied voltage greater than the diode breakdown voltage. An incoming photon with an energy greater than the separation between valence and conduction band, if absorbed, will be able to create an electron-hole pair, which will be accelerated creating a detectable avalanche discharge. This so called “Geiger mode” operation allows to reach very good efficiencies in the visible range (between 50% and 85%, with tens or few hundreds Hz of dark count rates). Data are finally post-selected by an electronic device which is able to retrieve n -fold coincidences between the clicks of the APDs. To measure full-bunching outputs, we used to put in cascade fiber beam splitters (FBS) to each output mode. One FBS for 2-photon contributions and two FBS for 3-photon contributions in order to emulate number resolving detectors [126].

Discussion

In this Chapter we discussed how to translate our optical transformations in a more suitable and compact platform in order to provide scalability to our systems. In fact, the procedures adopted in bulk optics suffer of this important drawbacks which does not permit so far to realize large size experiments. With the technique of femtosecond laser micromachining we are able both to realize any unitary operation (supported by the milestone work of Reck [15]) by exploiting the 3D geometry capability and the considerably easier approach to the writing compared to other techniques as for example lithography. Moreover, we saw that the recent developments and the current state-of-art give us the possibility to explore a dynamical approach on the implementation of our devices by means of thermo-optical resistors which provide tunable phases. First notable experimental realization with the latter approach are described in Ref. [160] where they explored for the first time HOM interference in a tunable beam splitter at Telecom wavelength and furthermore two remarkable applications in contextuality [174] and quantum simulation [175].

In the next chapter we will focus on one of the main work which deals with identifying and characterize the best architectures that we can use to easily retrieve good signature of many-body quantum interference with photonic chips. These results will be directly used to validate this phenomenon in Chapter 5.

Chapter 4

The quest for the best photonic platform implementation

In this context, given suitable techniques to implement miniaturized interferometers into little-size chip, it is essential to identify suitable schemes to implement interferometric networks within an integrated platform. As showed in Ref. [15], Reck *et al.* proposed a universal algorithm to implement an arbitrary unitary transformation by decomposing it in a suitable network of only beam splitters and phase shifters. However its triangular layout lacks a perfect symmetry, thus making it highly sensitive to internal losses that lower the adherence of the implemented transformation to the ideal one. Recently, two different architectures have been proposed for the implementation of linear optical networks. A first scheme has been reported in Ref. [178], which ultimately corresponds to a rearrangement of the scheme of Ref. [15] in a symmetric layout. While keeping the same number of optical elements and the capability of implementing an arbitrary unitary transformation, this scheme presents reduced sensitivity to losses within the interferometer. A second scheme, inspired by the classical algorithm of Cooley and Tukey [179] for the fast Fourier transform, has been proposed recently in Refs. [180, 181] and implemented experimentally in Refs. [170, 182] by FLM. This layout, though not supporting arbitrary unitary evolutions, allows to implement a significant class of linear optical networks with a substantial reduction in the number of necessary optical elements. Such class of matrices includes the Hadamard ones, with a notable example provided by the Fourier transformation that is widely employed in a large set of quantum information protocols [45, 77]. In the next sections we will introduce two new architectures, the Rectangular and Fast schemes, which have been analysed extensively in comparison with Triangular one, showing the difference in terms of performance with a large number of modes and the behaviour with respect of internal losses. The sequent works reported have been carried out during my research project, focusing on search for the best implementation to design our devices.

- Flamini, F., Spagnolo, N., **Viggianiello**, N., Crespi, A., Osellame, R. and Sciarrino, F., Benchmarking integrated linear-optical architectures for quantum information processing. *Scientific Reports*, 7, 15133 (2017) [176]. In this work I contributed to the discussion and the development of the simulation which I partially carried out.
- Flamini, F., **Viggianiello**, N., Bentivegna, M., Spagnolo, N., Mataloni, P., Crespi, A., Ramponi, R., Osellame, R. and Sciarrino, F. Generalized Quantum Fast Transformations via Femtosecond Laser Writing Technique. *Interdisciplinary Information Sciences*, 23, 115-118, (2017) [177]. In this work I contributed to the development of the theory and the proofs.

4.1 Rectangular decomposition

Two decades after the original paper, an improvement respect Reck's decomposition has been proposed by Clements *et al.* [178]. In their work, they demonstrate a new design for universal multiport interferometers based on an alternative arrangement of beam splitters and phase shifters. This new scheme outperforms that by Reck *et al.* [15] in terms of compactness and optical losses.

Specifically, Reck's unitary decomposition can be summarized by the formula:

$$\hat{U} = D \left(\prod_{(m,n) \in S} T_{m,n} \right), \quad (4.1)$$

where S defines a specific ordered sequence of two-mode transformations, $T_{m,n}$ is the sub-unitary extracted for the decomposition (see Fig. 3.1b for an analogy) and D is a diagonal matrix of complex elements with modulus equal to one. The unitary decomposition introduced by Clements moves from the intuition that any element in row n or m of U can also be nulled by multiplying U with a $T_{m,n}^{-1}$ matrix. The protocol consists of nulling successive diagonals of U , in such a way that the sequence of $T_{m,n}$ and $T_{m,n}^{-1}$, multiplying U from left and right respectively, guarantees that nulled elements of U are not affected by subsequent operations.

Exploiting this decomposition, we obtain the following expression:

$$\left(\prod_{(m,n) \in S_L} T_{m,n} \right) \hat{U} \left(\prod_{(m,n) \in S_R} T_{m,n}^{-1} \right) = D, \quad (4.2)$$

where S_L and S_R are the pairs of (m,n) indices for the $T_{m,n}$ or $T_{m,n}^{-1}$ matrices given by the decomposition. The last expression can be rewritten as:

$$\hat{U} = \left(\prod_{(m,n) \in S_L^T} T_{m,n}^{-1} \right) D \left(\prod_{(m,n) \in S_R^T} T_{m,n} \right). \quad (4.3)$$

An illustration of the order in which the elements of a 5-dimensional \hat{U} are nulled is represented below [178]:

$$\begin{array}{cccccc} 15 & & & & & \\ 7 & 14 & & & & \\ 6 & 8 & 13 & & & \\ 2 & 5 & 9 & 12 & & \\ 1 & 3 & 4 & 10 & 11 & \end{array}, \quad (4.4)$$

where the bottom left matrix element is the first to be nulled, while the following elements are nulled in consecutive diagonals. Black elements located in column i are those cancelled by the matrix $T_{i,i+1}^{-1}$, the blue elements located in row i , instead, are those cancelled by the matrix $T_{i-1,i}$. After these steps, U is a lower triangular matrix (see (4.4)), which by virtue of its unitarity must be diagonal. It can be demonstrated that, if D consists of single-mode phase-shifts, then for any $T_{m,n}^{-1}$ matrix one can find a matrix D' of single-mode phases and a matrix $T_{m,n}$ such that:

$$T_{m,n}^{-1} D = D' T_{m,n}. \quad (4.5)$$

In this way we can put equation (4.3) in the Reck's form (4.1), that is:

$$\hat{U} = D' \left(\prod_{(m,n) \in S} T_{m,n} \right). \quad (4.6)$$

By construction, equations (4.3) and (4.6) physically corresponds to the multiport interferometer shown in 4.1.

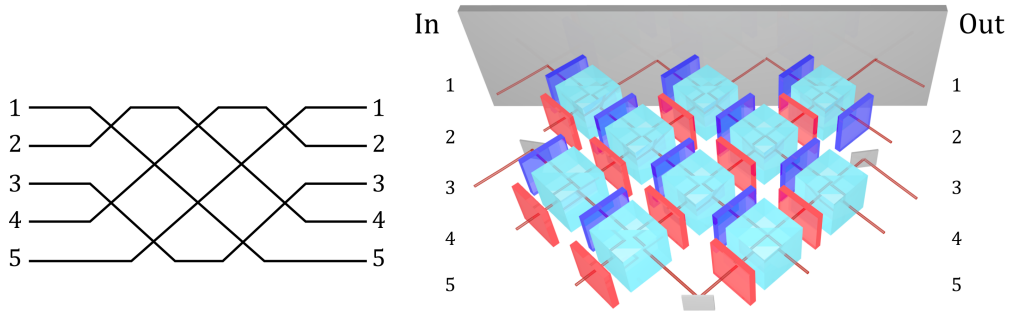


Figure 4.1. Clements design for a 5×5 interferometer. The left-hand side reports a two-dimensional representation, while the right-hand side is the equivalent bulk representation of the circuit.

Comparing the two interferometric architectures presented we see that, while both have the same number of beam-splitters, the Clements' scheme is quite smaller than the Reck's one of a factor two in the circuital depth. In practice, what changes is the **depth of the interferometer**, that is the longest path through the interferometer, enumerated by counting the number of beam-splitters traversed by that path. Precisely, for an $N \times N$ interferometer, the Reck design [15] has an optical depth of $2N - 3$, whereas the Clements one has an optical depth of N . The minimization of the optical depth of an interferometer is a fundamental task, since area of the circuit is an important constraint for the fabrication of planar waveguide structures. Furthermore, propagation losses are reduced in interferometer with small optical depth. Hence interferometers built according to last design will suffer from only about half the propagation loss of an interferometer built according to the Reck architecture.

4.2 Fast decomposition

The general method to realize an arbitrary unitary transformation using passive linear optical devices, was introduced by Reck *et al* [15] and later improved by Clements *et al* [178]. They provided a decomposition of a unitary of dimension m as a sequence of $m(m-1)/2$ beam-splitters and phase shifters, that is $O(m^2)$ elements.

However, in the special case of the Discrete Fourier Transform (DFT) matrices, a more efficient method has been proposed [180,181], which takes advantage of the DFT symmetries to significantly reduce the number of linear optical elements required. Based on the classical algorithm of Cooley and Tukey [179], who first introduced the Fast Fourier Transform (FFT) as a more efficient way to calculate the DFT, Barak and Ben-Aryeh [181] developed a quantum analogue in the linear optics domain.

Before providing a formal description of the Barak and Ben-Aryeh decomposition method, let us observe that a Discrete Fourier transform implemented by using quantum states is called Quantum Fourier Transform (QFT). For 2^n computational quantum states, representing n

qubits, the $2n$ -dimensional DFT matrix is represented by the unitary matrix [181]:

$$QFT(n) = \frac{1}{2^{n/2}} \begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & u & u^2 & u^3 & \dots & u^{2^n-1} \\ 1 & u^2 & u^3 & u^4 & \dots & u^{2(2^n-1)} \\ 1 & u^3 & u^4 & u^5 & \dots & u^{3(2^n-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & u^{2^n-1} & u^{2(2^n-1)} & u^{3(2^n-1)} & \dots & u^{n(2^n-1)} \end{pmatrix}, \quad (4.7)$$

where $u = e^{2\pi i/2^n}$. It has been shown that any n qubit unitary transformation can be implemented by at most $2^{n-1}(2^n - 1)$ BS and phase shifter combinations. The Cooley–Tukey algorithm improves this relation breaking down the $2^n \times 2^n$ unitary matrix into $n2^{n-1}$ qubit transformations. The algorithm uses n stages of 2^{n-1} transformations, where in each stage the input is given by the output of the previous one. This algorithm leads to the concept of Quantum Fast Fourier Transform qFFT. This approach, valid for $2p$ -dimensional Fourier matrices, requires only $\frac{m}{2} \log m$ beam-splitters and phase shifters, thus, compared with the Reck and Clements schemes, it enhances the compactness and scalability of the platform for a more reliable experimental realization.

The algorithm demonstrated by Barak and Ben-Aryeh, for implementing Fourier matrices with passive optical circuits, is based on the following steps [181]:

1. The algorithm decompose the action of the matrix into n discrete transformations. Each partial transformation consists of 2^{n-1} interactions between pairs of optical modes;
2. At the stage k all modes are distributed in 2^{k-1} groups, containing 2^{n-k+1} input modes each. Within each group, the mode j is coupled through a beam splitter with the input $j + 2^{n-k}$ belonging to the same group, after a multiplication by a phase factor $e^{\frac{f(k,m)}{2^k}}$ with $f(k, m) = \sum_{s=0}^{k-1} m_{k-s} 2^s$;
3. Successively it perform a permutation of the outputs,
4. The fast implementation of the QFT can be assembled recursively out of two 2^{n-1} -dimensional QFT circuits,
5. The unitary matrix $U(\alpha, \beta, \gamma, \delta)$ in the action which links input/output modes at each step, is defined as

$$e^{i\alpha} \begin{pmatrix} \cos \beta e^{i(\gamma+\delta)} & \sin \beta e^{i(-\gamma+\delta)} \\ -\sin \beta e^{i(\gamma-\delta)} & \cos \beta e^{i(-\gamma-\delta)} \end{pmatrix} \quad (4.8)$$

To retrieve the QFT in the BB scheme, all the beam splitters are set balanced (50:50) while the phase shifts vary in discrete steps multiples of $\pi/2^n$. The general transformation has the form ($j = 0, 1, \dots, 2^{n-1}$)

$$U\left(\pi \frac{2^{n-1} - j + 1}{2^n}, \frac{\pi}{4}, \pi \frac{2^{n+1} - j + 1}{2^n}, \frac{\pi}{2}\right) \quad (4.9)$$

For a large number of qubits, the number of optical elements required with the BB scheme is reduced by a factor of $2^n/n$, which makes the algorithm suitable for practical realizations. In the case of a n -qubit– 2^n -dimensional discrete Fourier transform, we obtain the following decomposition:

$$QFT(n) = R^{(n)} S_n^n S_{n-1}^n, \dots, S_2^n, S_1^n, \quad (4.10)$$

where $R^{(n)}$ is a permutation matrix rearranging the outputs.

The algorithm uses only $n2^n$ optical elements. Thus, for a large number of qubits, we find that the number of beam-splitters and phase shifters needed is significantly reduced by a factor of $2^n/n$.

The procedure to design interferometers implementing the fast Fourier transform is now described in the general case of $m = 2^n$ modes, where n is an integer number.

Each optical mode, $k \in [1, m]$, can be numbered accordingly to its binary representation and can be associated to the set $(b_1, \dots, b_n)_k$. The $\{b_i\}$ are ordered from the most significant, b_1 , to least significant, b_n .

Once obtained the n -dimensional vector (b_1, \dots, b_n) , its components can be thought as the coordinates of the vertices of a n -dimensional hypercube in \mathbb{R}^n . An example with $n = 3$ is provided in Fig. 4.2 [182]. The algorithm consists of n steps. The j -th step connects all pairs of modes that differ only for the j -th bit. Each waveguide position in the cross-section plane can be defined by suitably projecting in two dimensions the vertices of the hypercube. At every FFT step, the pairs of modes that have to interact are connected by directional couplers. Such connections correspond to all the edges of the hypercube with a given direction. The scheme of an 8-mode interferometer is showed in Fig. 4.3 [182]. The inset shows which modes are connected by directional couplers at each step k . At each step, only the modes corresponding to the vertices connected by parallel edges interact. The phase shifters $\{P_i\}$, instead, specify the amount of phase shift to introduce on specific modes. For a full description of the integrated implementation see Appendix Note 8.2. A good strategy for an integrated implementation of a QFT matrix exploits the peculiar features of the femtosecond laser-writing technique. Indeed, Reck and Clements schemes are efficiently built with planar structures. In comparison, the complex arrangement of pairwise interactions, necessary for the FFT and Quantum Fast Fourier Transform (qFFT) algorithm, cannot be easily implemented using a planar architecture.

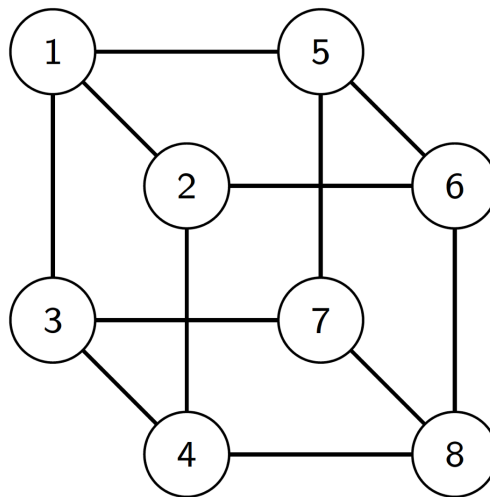


Figure 4.2. Planar projections of an hypercube of dimension $p = 3$. The numbered vertices correspond to the possible position of the waveguides in the cross-section of a 3D interferometer implementing the FFT. Each connection between the vertices corresponds to a directional coupler in a given section of the interferometer. Geometrical scheme taken from [182].

The essential structure of such fast architectures presents some interesting advantages with respect to the other universal schemes. First, the depth of the circuit scales only logarithmically with the size of the interferometer, i.e. with the number of optical modes, leading to an even more compact operation area and to reduced propagation losses. Moreover, the layout is fully symmetric and naturally fits a description in terms of qubit states, thanks to the binary interactions between the modes.

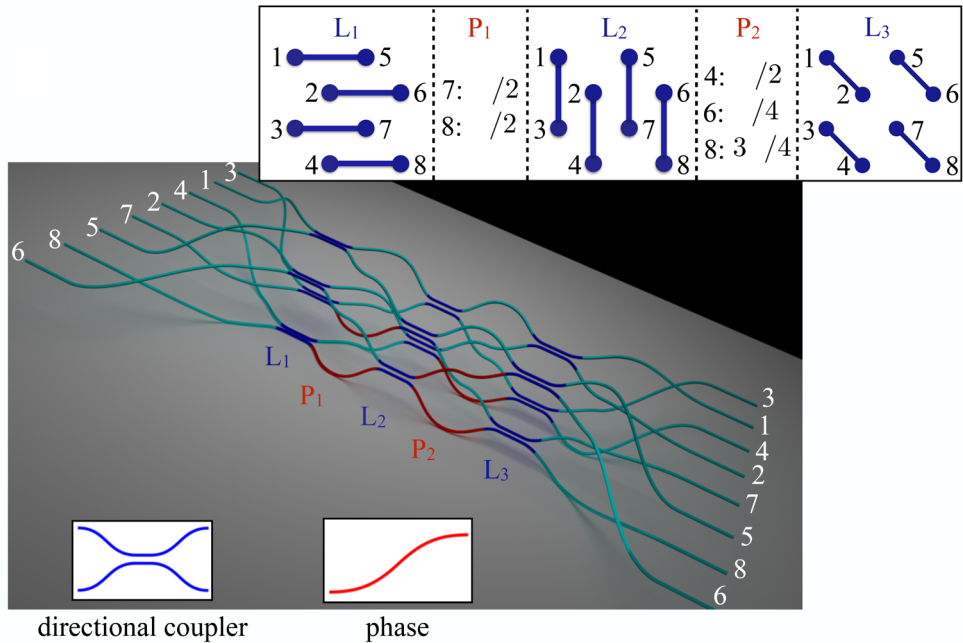


Figure 4.3. Internal structure of an 8-mode interferometer implementing the qFFT. Input and output modes have been relabeled in order to obtain the effective Fourier transformation and the to minimize the bending losses. The insets shows the waveguides disposition and coupling among the modes. Figure taken from Ref. [182].

4.2.1 Hadamard matrices

Fast decomposition allow to realize a QFT by exploiting the symmetries of these kind of matrix in order to cross specific modes, reducing significantly the circuit depth. However, this architecture can be used to implement also unitaries which belong to a bigger class of transformations [177]: the class of Hadamard transformations. A matrix belonging to this class is defined as an orthogonal matrix of complex numbers, in which all the elements have a unitary modulus. Indeed, the Fourier matrix is a particular case of this general class. Notable representatives of the Hadamard class are the Sylvester (U_S) transformations, which are unitaries where all the elements are real and are equal to ± 1 . Sylvester matrices are described as 2^n -dimensional unitary matrices given respectively by

$$U^H(2^n) = H(2^n) = \begin{pmatrix} H(2^{n-1}) & H(2^{n-1}) \\ H(2^{n-1}) & -H(2^{n-1}) \end{pmatrix} \quad (4.11)$$

being $H(2^0) = (1)$, $H(2) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ and n any positive integer. The unitary matrix is finally retrieved by multiplying for a normalization factor $S(m) = \frac{1}{\sqrt{m}} H(m)$ where m is the matrix dimension. The scheme proposed by Barak and Ben-Aryeh [180, 181], is particular suitable for these kind of matrices. It is possible to generalize their scheme to span a whole class of generalized Hadamard transformations [177] by keeping fixed the interferometric structure and tuning the parameters of the beam splitters and phase shifters. Let us show how can we retrieve a general closed-form expression [177] of the element $U_{a,b}$ for the most general 2^n -dimensional fast unitary transformation. We will show how can we identify Fourier and Sylvester from this formula.

$$U_{a+1,b+1}^{(n)} = e^{i b_r \phi_r^{\xi_r^{(a,b)}} + i \frac{\pi}{2} a_r^{(n)} \oplus b_r^{(n)}} \prod_{s=1}^n \cos \chi_{a,b}^{(n,s)} \quad (4.12)$$

where $a, b \in [0, 2^n - 1]$ label the input/output modes and some shorthand notations have been used, following the Einstein summation convention

$$\chi_{a,b}^{(n,s)} = \theta_{s,f_s^{(a,b)}} - \frac{\pi}{2} |a_s^{(n)} - b_s^{(n)}| \quad \xi_r^{(a,b)} = 1 + 2^{n-r} + b \bmod 2^{n-r} + \left\lfloor \frac{a}{2^{n-r+1}} \right\rfloor 2^{n-r+1} \quad (4.13)$$

where $f_s^{(a,b)} = 1 + b + (a_r^{(n)} - b_r^{(n)}) 2^{n-r}$ and $m_r^{(n)}$ equals the r -th digit of the n -bit binary representation of m , being $[x]$ the integer part of the real number x .

In general, any m -dimensional photonic architectures can be described in terms of consecutive layers s of optical elements L_s , made up of a network of phase shifters and beam splitters mixing a subset of modes no more than once each. Specifically, each matrix L_s consists in turn of a layer $B^{(s)}$ of beam splitters, coupling a set of pairs of modes (k_1, k_2) , and $\frac{m}{2}$ phase shifters $e^{i\phi_{s,k_1}}$ placed for each pair on one of the two interacting modes (k_1, k_2) . The particular sequence of mode interactions $\{(k_1, k_2)\}$ depends on the triangular, square or fast designs. While clearer for the first two, the geometry of the third scheme for a 2^n -dimensional unitary transformation is slightly more complex and arises from the binary representations of the optical modes. Using $\tau_{s,k}$ for the beam splitters transmissivities on mode k , the beamsplitters layer is described by the matrix

$$B_{k_1, k_2}^{(s)} \equiv \begin{cases} \tau_{s, k_1} & k_1 = k_2 \\ i(1 - \tau_{s, k_1}^2)^{1/2} & (k_1, k_2) \in \{(\alpha, \beta)\}^{(n,s)} \\ 0 & otherwise \end{cases} \quad (4.14)$$

where $\{(\alpha, \beta)\}^{(n,s)} = \{(a + 2^s b, a + 2^s b + 2^{s-1})\}$ are the pairs of modes interacting in the layer s , with $a \in \{1, \dots, 2^{s-1}\}$, $b \in \{0, \dots, 2^{n-s} - 1\}$. For example, the 8-dimensional quantum Fourier transform is obtained, modulo a relabeling of the output modes [181], by choosing $\tau_{s,k} = \sqrt{2^{-1}}$ and $\phi_{2,7} = \phi_{2,8} = \phi_{3,4} = \frac{\pi}{2}$, $\phi_{3,6} = \frac{\pi}{4}$ and $\phi_{3,8} = \frac{3\pi}{4}$. Similarly, the Sylvester transformation corresponds to the choice $\tau_{s,k} = \sqrt{2^{-1}}$ and $\phi_{s,k} = 0$.

In the next sections we will introduce the procedure to reconstruct unitary matrices by means of single photon measurements.

4.3 Tomography of unitary matrices implemented with Fast decomposition

In this section we will focus on the procedure chosen to reconstruct a unitary matrix implemented with the Fast scheme by means of single-photon and two-photon measurements. In particular we will show reconstructed matrices in the case of Fourier and Sylvester implementations. In Fig. 4.4 are reported the scheme of Sylvester and Fourier matrices using Fast decomposition.

Several algorithms have been developed to perform the reconstruction of a unitary linear transformation [183–186]. One first difference among the various procedures resides in the amount of exploited knowledge about the physical structure of the interferometer. The reconstruction can be performed considering the device as a black box, whose only known information is the n -photon input-output relation [185–187], considering classical resources [188] or considering the device as a complex structure whose elementary constituents' arrangement is partially or completely known. Here, we adopt the second approach for the reconstruction of our matrices. With this approach, the retrieval of the unitary transformation is based on the characterization of the constituent optical elements, that is, beam-splitters and phase shifters.

The knowledge of the inner structure can in principle provide advantages of the algorithm, which are:

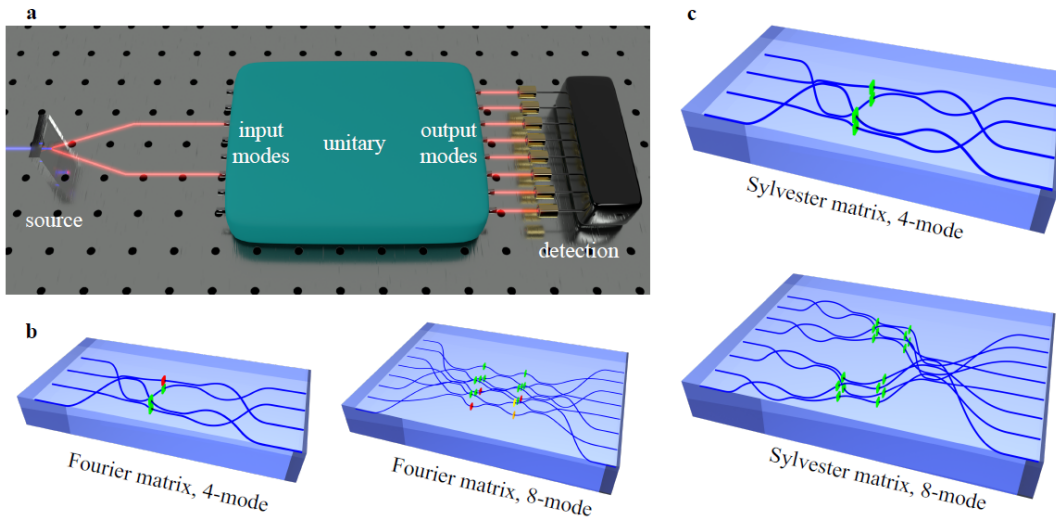


Figure 4.4. Conceptual scheme of reconstruction process. **a**, Scheme of the experimental approach for the reconstruction. A photon source based on parametric down-conversion produces pairs of photons that can be used to measure HOM visibilities to retrieve phases from any possible input combination of a unitary transformation, with photo-detection at the output. **b**, Internal structure for the 4- and 8-mode laser-written Fourier and Sylvester interferometers. Known the architecture, we can parametrize inner phases and reflectivities and apply the protocol.

Significance: Given two different reconstructed matrices having the same fidelity with the ideal one, a unitary transformation reconstructed by adopting this approach is closer to the implemented one. Indeed, the constraints on the architecture can highlight possible internal symmetries otherwise ignored by other methods.

Unitarity: Other algorithms which treat the system as a black box may in general give a non-unitary matrix as output of the reconstruction [183]. The subsequent need for recovering a unitary transformation may thus alter, to some extent, the result of the reconstruction process.

Characterization: Ultimately, further information is obtained by focusing on the properties of the elementary optical components. Specifically, this reconstruction method permits to investigate the tolerance to their unavoidable fabrication imperfections, by comparing their retrieved characteristics with those of the ideal ones.

However, the formal proof of this advantages is still in study. To perform the reconstruction, *i.e.* to characterize the single components within the integrated architecture, it is necessary to gain information on both the transmittivities $\{\tau\}$ and the relative phases $\{\phi\}$ associated to each element. Since the inner structure for implementation is similar both for Fourier and Sylvester matrices, we will use this information to apply the same algorithm for the reconstruction of the unitaries. In the following we will illustrate the procedure adopted.

The reconstruction algorithm is made up of two independent stages, corresponding to the separate retrieval of the transmittivities $\{\tau\}$ and the phases $\{\phi\}$. The estimate of the two sets of parameters is performed via a maximum-likelihood method. In Fig. 4.5 we report the schemes of the internal structure of the implemented interferometers.

Estimate of the transmittivities

Information on the $|U_S^{(d)}|_{i,j}$ can be obtained directly with single-photon measurements, recalling that the scattering probability from modes i to j is indeed exactly $|U_S^{(d)}|_{i,j}^2$. Thus, by experi-

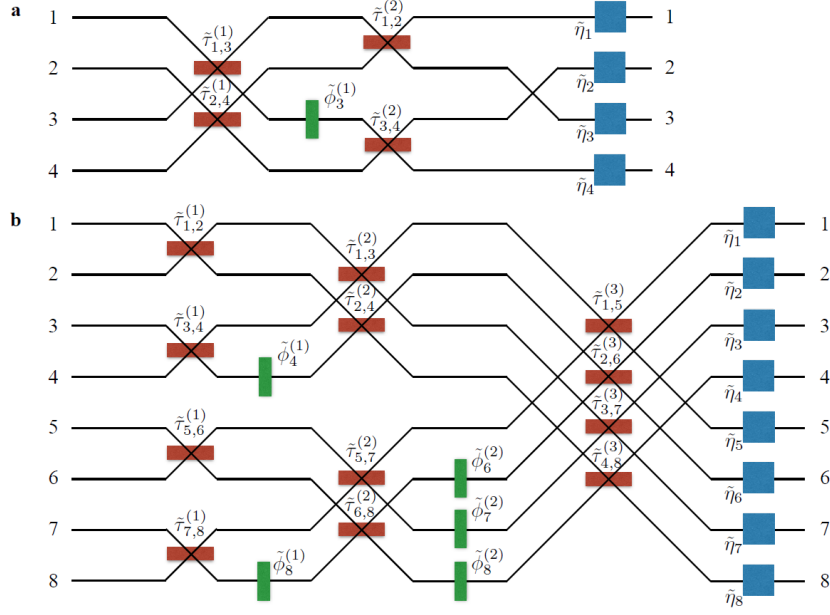


Figure 4.5. Internal structure of the 4-mode device **(a)** and the 8-mode device **(b)** for Sylvester implementation. A different notation for $\{\tau\}$ and $\{\phi\}$ is adopted with respect to Eqs. of 4.2.1. Red boxes: beam-splitters. Green boxes: fabrication phases relevant for the algorithm. Blue boxes: relative output losses. For the Sylvester transformations, all beam-splitters have transmittivities $\tau_{id} = 2^{-1/2}$ and all phases are 0.

mentally measuring all the d^2 scattering probabilities $[\tilde{P}_S^{(d)}]_{i,j}$, it is possible to look for the set $\{\tau\}$ that best reproduces the values registered. This can be done, via a maximum likelihood procedure, by numerically minimizing the χ_τ^2 quantity

$$\chi_\tau^2 = \sum_{i,j=1}^d \left(\frac{[\tilde{P}_S^{(d)}]_{i,j} - |U_S^{(d)}|_{i,j}^2}{[\sigma_{\tilde{P}_S^{(d)}}]_{i,j}} \right)^2, \quad (4.15)$$

where $\sigma_{\tilde{P}_S^{(d)}}$ are the experimental errors associated to the measured probabilities $\tilde{P}_S^{(d)}$. However, the set $\{\tau\}$ found in this way is insensitive to possible external losses in the output of the device (see Fig. 4.5 for the Sylvester case). Such unavoidable relative losses modify the transformation actually implemented, entailing biases in the distribution of the moduli. To overcome this issue it is sufficient to introduce a second, non-unitary diagonal transformation $D(\eta_1, \eta_2, \dots, \eta_d)$ acting after the device. Thus, defining

$$M_S^{(d)} = D(\eta_1, \eta_2, \dots, \eta_d) U_S^{(d)} = \begin{pmatrix} \eta_1 & 0 & \dots & 0 \\ 0 & \eta_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \eta_d \end{pmatrix} U_S^{(d)}, \quad (4.16)$$

and including the $\{\eta\}$ in the parametrization of the unitary, it is possible to obtain an estimate of the actual relative losses $\{\tilde{\eta}\}$ involved, as well as a set of transmittivities $\{\tilde{\tau}\}$ which properly takes them into account.

The retrieved values of the output losses for the 4-mode and the 8-mode Sylvester devices, as well as the transmittivities $\{\tilde{\tau}\}$, are reported in Table 4.1. Note that since the method permits reconstruction of relative losses only, one of the parameters must be taken as a reference (for instance, η_1).

4-mode device: $\{\tilde{\eta}\}$	4-mode device: $\{\tilde{\tau}\}$	8-mode device $\{\tilde{\eta}\}$	8-mode device $\{\tilde{\tau}\}$
$\tilde{\eta}_1 = 1$	$\tilde{\tau}_{1,3}^{(1)} = 0.6879 \pm 0.0003$	$\tilde{\eta}_1 = 1$	$\tilde{\tau}_{1,2}^{(1)} = 0.669 \pm 0.008$
$\tilde{\eta}_2 = 1.134 \pm 0.002$	$\tilde{\tau}_{2,4}^{(1)} = 0.7195 \pm 0.0003$	$\tilde{\eta}_2 = 0.94 \pm 0.02$	$\tilde{\tau}_{3,4}^{(1)} = 0.621 \pm 0.007$
$\tilde{\eta}_3 = 1.147 \pm 0.002$	$\tilde{\tau}_{1,2}^{(2)} = 0.7139 \pm 0.0003$	$\tilde{\eta}_3 = 1.07 \pm 0.02$	$\tilde{\tau}_{5,6}^{(1)} = 0.639 \pm 0.008$
$\tilde{\eta}_4 = 0.961 \pm 0.002$	$\tilde{\tau}_{3,4}^{(2)} = 0.7031 \pm 0.0003$	$\tilde{\eta}_4 = 0.93 \pm 0.02$	$\tilde{\tau}_{7,8}^{(1)} = 0.630 \pm 0.008$
		$\tilde{\eta}_5 = 0.96 \pm 0.02$	$\tilde{\tau}_{1,3}^{(2)} = 0.750 \pm 0.007$
		$\tilde{\eta}_6 = 1.05 \pm 0.02$	$\tilde{\tau}_{2,4}^{(2)} = 0.715 \pm 0.007$
		$\tilde{\eta}_7 = 1.01 \pm 0.02$	$\tilde{\tau}_{5,7}^{(2)} = 0.755 \pm 0.007$
		$\tilde{\eta}_8 = 1.02 \pm 0.02$	$\tilde{\tau}_{6,8}^{(2)} = 0.729 \pm 0.006$
			$\tilde{\tau}_{1,5}^{(3)} = 0.748 \pm 0.008$
			$\tilde{\tau}_{2,6}^{(3)} = 0.723 \pm 0.007$
			$\tilde{\tau}_{3,7}^{(3)} = 0.774 \pm 0.006$
			$\tilde{\tau}_{4,8}^{(3)} = 0.755 \pm 0.007$

Table 4.1. Relative losses $\{\tilde{\eta}\}$ and transmittivities $\{\tilde{\tau}\}$ reconstructed from the first step of the algorithm for the 4- and 8-mode devices in the case of Sylvester implementations.

The values of $\{\tilde{\tau}\}$ are to be compared with the expected one ($\tau_{\text{id}} = 2^{-1/2}$). In all estimates, the experimental errors are obtained by a Monte Carlo approach on the reconstruction process.

Let us now define $U_{\tilde{\tau}}^{(d)}$ as the temporary matrix obtained by inserting the $\{\tilde{\tau}\}$ in the decomposition of $U_S^{(d)}$.

Estimate of the phases

Once we have the estimate of the set $\{\tilde{\tau}\}$, it is possible to apply the same procedure to find the values of the phases $\{\tilde{\phi}\}$ that best reproduce further experimental measurements. This is done by minimizing a second χ_{ϕ}^2 quantity, which compares the measured visibilities ($\tilde{V}_{i,j}^{m,n}$) of the two-photon HOM dips/peaks with the ones ($V_{i,j}^{m,n}$) expected from the unitary $U_{\tilde{\tau}}^{(d)}$. Note that the expected values are parametric in the unknown phases $\{\tilde{\phi}\}$. The visibilities are defined as $V_{i,j}^{m,n} = (D_{i,j}^{m,n} - Q_{i,j}^{m,n})/D_{i,j}^{m,n}$, being $D_{i,j}^{m,n}$ the transition probability from inputs (i, j) to outputs (m, n) with distinguishable particles and $Q_{i,j}^{m,n}$ the one with indistinguishable photons. The visibilities $V_{i,j}^{m,n}$ are measured by recording the output coincidences as a function of the relative delay $\Delta\tau$, and by fitting the experimental interference pattern with the function

$$C_{i,j}^{m,n}(\Delta\tau) = B_{i,j}^{m,n} [1 - V_{i,j}^{m,n}] \exp\left(-\Delta\tau^2/(2\sigma_{\tilde{\tau}}^2)\right). \quad (4.17)$$

The χ_{ϕ}^2 quantity to be minimized for the reconstruction algorithm is defined as

$$\chi_{\phi}^2 = \sum_{i,j,m,n}^d \left(\frac{\tilde{V}_{i,j}^{m,n} - V_{i,j}^{m,n}}{\sigma_{\tilde{V}_{i,j}^{m,n}}} \right)^2 \quad (4.18)$$

where the sum is extended over all the measured input-output combinations. Note that visibilities are independent of the output losses. Here, for each visibility, the error $\sigma_{\tilde{V}_{i,j}^{m,n}}$ is obtained from the fitting procedure. Conveniently choosing the two-photon input states to inject, whose visibilities depend on all the relevant phases, it is in principle possible to complete the reconstruction using only one state. However, to improve the reconstruction stability, for both devices we measured an overcomplete set of data. For Sylvester interferometers, in the 4-mode case, we recorded all possible 36 collision-free two-photon visibilities. In the 8-mode case, all collision-free

visibilities for 8 two-photon input states $\{(2, 5), (2, 6), (2, 8), (3, 7), (3, 8), (5, 7), (5, 8), (6, 8)\}$ have been recorded, for a total number of 224 visibilities.

The estimated values for the phases are reported in Table 4.2.

4-mode device: $\{\tilde{\phi}\}$	8-mode device $\{\tilde{\phi}\}$
$\phi_3^{(1)} = (+0.229 \pm 0.004)$ rad	$\phi_4^{(1)} = (+0.38 \pm 0.02)$ rad
	$\phi_8^{(1)} = (+0.36 \pm 0.01)$ rad
	$\phi_6^{(2)} = (-0.19 \pm 0.03)$ rad
	$\phi_7^{(2)} = (+0.43 \pm 0.02)$ rad
	$\phi_8^{(2)} = (-0.04 \pm 0.03)$ rad

Table 4.2. Relevant phases $\{\tilde{\phi}\}$ reconstructed from the second step of the algorithm for the 4- and 8-mode devices.

Note that in the case of Sylvester matrices, thanks to the internal symmetries in the evolution through the interferometer, it is possible to describe the action of this unitary with fewer phases than those actually present in the devices (see scheme of Fig. 4.5). First, input phases are irrelevant when the system is injected with Fock states, and thus can be all set to 0. In the case of the 4-mode device, for instance, only one phase is relevant for the algorithm, and we can thus set $\phi_4^{(1)} = 0$ for the sake of simplicity. In the case of the 8-dimensional Sylvester transformation, the functional dependence on all the 4+4 phases of the two internal layers can be compressed in that on only 5 phases ($\phi_3^{(1)} = \phi_7^{(1)} = \phi_5^{(2)} = 0$), eventually involving differences between them.

Once all the parameters $\{\tilde{\tau}\}$ and $\{\tilde{\phi}\}$ are known, the actual implemented unitary transformations $\tilde{U}_{\text{IS}}^{(d)}$ can be obtained by inserting the obtained values in the decomposition.

The real and imaginary parts of the reconstructed transformation are compared with the Sylvester transformations in Fig. 4.6 (for the 4-mode device) and in Fig. 4.7 (for the 8-mode device). The overall data obtained from the reconstruction process for both Fourier and Sylvester matrices, is shown in Fig. 4.8. In the final step, in order to verify that we performed the correct implementation, we calculated the fidelities

$$F^{(d)} = \frac{1}{d} \left| \text{Tr}(U^{(d) \dagger} \tilde{U}^{(d)}) \right| \quad (4.19)$$

for both Fourier and Sylvester transformations. With the ideal 4- and 8- dimensional Sylvester transformation, the fidelities are respectively $F_{U_S}^{(4)} = 0.99807 \pm 0.00005$ and $F_{U_S}^{(8)} = 0.9813 \pm 0.0005$. In the case of Fourier matrices the fidelity between the reconstructed unitary and the theoretical one is $F_{U_F}^{(4)} = 0.9822 \pm 0.0001$, the fidelity between the reconstructed unitary and the ideal 8-mode is $F_{U_F}^{(8)} = 0.9527 \pm 0.0006$, thus confirming the high quality of the fabrication process in both implementations.

4.4 Benchmarking performances for different implementations

Here we focus on the analysis of the robustness of our implementations, making a comparison between different schemes. This work is reported in [176]. A crucial requirement towards the identification of optimal architectures for the implementation of large size interferometers is a detailed knowledge of the tolerance to fabricative errors, namely propagation losses and imperfections in the parameters of the optical elements. Indeed, in all quantum information applications the applicability of a given experimental platform is limited by the maximum amount of noise tolerable in the interferometric networks.

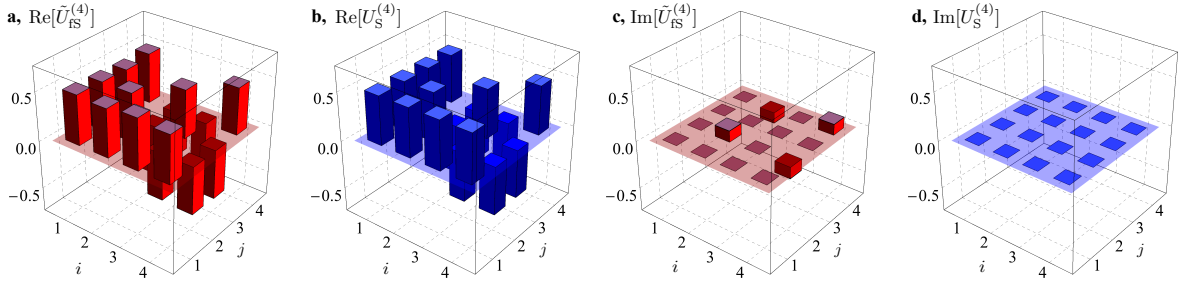


Figure 4.6. Results of the unitary reconstruction process for the 4-mode device. **a** Real part and **c** imaginary part of the implemented interferometer $\tilde{U}_{\text{fs}}^{(4)}$. Lighter regions in the bars represent the error in the reconstruction process. **b** Real part and **d** imaginary part of the Sylvester matrix $U_{\text{S}}^{(4)}$.

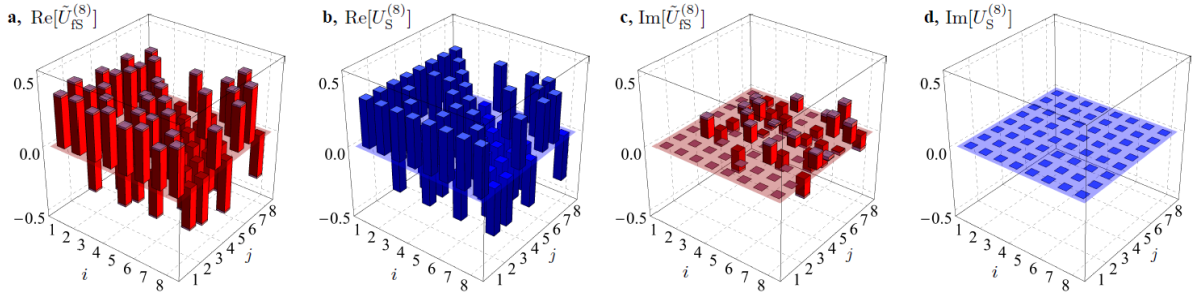


Figure 4.7. Results of the unitary reconstruction process for the 8-mode device. **a** Real part and **c** imaginary part of the implemented interferometer $\tilde{U}_{\text{fs}}^{(8)}$. Lighter regions in the bars represent the error in the reconstruction process. **b** Real part and **d** imaginary part of the Sylvester matrix $U_{\text{S}}^{(8)}$.

In this section we focus on this topic, presenting our work on the analysis of the performance of the main photonic architectures under imperfect operational conditions presented in the previous sections. First, we compare briefly the performances of the first two schemes, which were shown to be universal for unitary evolutions, for the implementation of Haar-random transformations of increasing size (for further details refer to our paper [176]). Second, we compare the operation of the three schemes for the implementation of Fourier and Sylvester [169] interferometers, which represent the fundamental building blocks in a significant number of relevant quantum information protocols. For our numerical benchmark we employ as figures of merit the fidelity [178] and the total variation distance, as good estimators of the distance between ideal and imperfect implementations in relevant applications. Our analysis highlights the advantages and limitations of each scheme, providing an essential reference point for the design of future larger-scale photonic technologies, whose optimal configurations may well benefit from a joint integration.

Modelling non-ideal unitary implementations

In the following we will refer to the *C*- (Clements *et al.* [178]), *R*- (Reck *et al.* [15]) and *F*- (Fast [181]) designs looking at their fixed, solid photonic architectures. This aspect is especially relevant if we are to choose the layout of a fully reconfigurable quantum circuit [22], which is designed to be multi-purpose and optimal when averaging over all its applications of interest. In general, the two main factors affecting the implementation of photonic quantum circuits are propagation losses and imperfect settings of the parameters describing the optical elements.

- *Losses*: Coupling losses at the input/output of any circuits remain relevant aspects in practical situations; however, their effect can be regarded as independent of the internal photonic architecture adopted and, thus, they will not be included in our study. Propagation

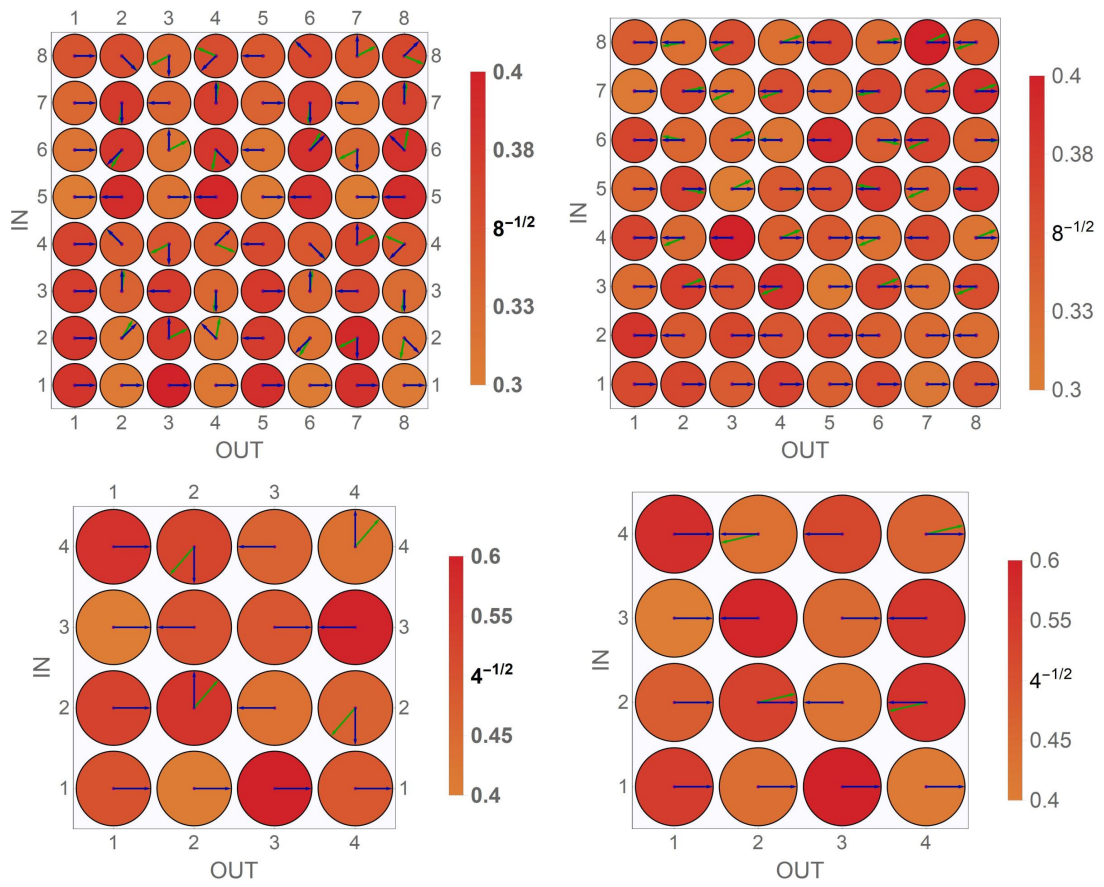


Figure 4.8. Representation of the reconstructed experimental transformation U_F^d and U_S^d . Coloured disks represent the moduli of the reconstructed matrix elements (all ideally equal to $d^{-1/2}$ for both matrices, where d is the matrix dimension). Arrows represent the phases of the unitary matrix elements (green: reconstructed unitary, blue: ideal matrix). Left: results for Fourier U_F^8 and U_F^4 . Right: results for Sylvester U_S^8 and U_S^4 .

	Reck <i>et al.</i>	Clements <i>et al.</i>	Fast
Symmetry	✗	✓	✓
Depth	$2m - 3$	m	$\log_2 m$
# elements	$\frac{m}{2}(m - 1)$	$\frac{m}{2}(m - 1)$	$\frac{m}{2} \log_2 m$
Haar	✓	✓	✗
Hadamard	✓	✓	✓

Figure 4.9. Table summarizing the main features of the three architectures. Image taken from [176]

losses, occurring in both straight and bent waveguides, play instead the main role in spoiling multipath interference. Their impact was shown to be mitigated by a more compact and symmetric interferometric structure [178], where optical modes interact with each other in a balanced way. Though photon losses unavoidably occur all along the circuit, a simple and effective way to analyse their effect is that of inserting constant losses at the output of each two-mode unit cell.

- *Fabricative noise:* Imperfect control over the fabricative parameters naturally leads to deviations from the ideal evolution on the circuit while keeping the unitarity of the process. The source of noise, arising from imperfect fabrication of the beam splitters and phase shifters associated to the $T_{m,n}$ has a different nature depending on the practical realization. Recent technological achievements have enabled the realization of reconfigurable quantum circuits [22, 23], where generic beam splitters are implemented as Mach-Zehnder interferometers with two cascaded symmetric beam splitters and one tunable phase shift. In this case, which will be at the core of our analysis, noise arises from imperfect fabrication of the fixed symmetric beam splitters and from a non-perfect control over the thermo-electric or electro-optic phase manipulation [22, 161, 174, 175], which becomes non-negligible for large-scale circuits or with high-speed tunings.

To focus our study around realistic values for the above mentioned sources of imperfections, we will consider femtosecond laser writing as the reference fabrication technology, as it is the only one capable to realize all the three considered architectures. In this case, typical bending losses with respect to straight waveguides are of the order of 0.5 dB/cm, which corresponds to 0.2 dB per directional coupler. Realistic values for fabrication errors at each directional coupler are of 0.21 rad for the phase shifters [265] and about 1-2% for the beam splitter transmissivities. These values do not change significantly considering a 2D or 3D architecture and, concerning phase shifts, a quite higher accuracy down to about 0.01 rad can actually be reached with an active control through thermo-optic phase shifters [160]. Therefore, we can consider as realistic fabrication tolerances 0.01 for the beam splitter transmissivity and 0.01 rad for the active phase control.

Our investigation on the performances of the three architectures was carried out by introducing various levels of noise on the optical elements and losses after each unit cell. Our Monte Carlo simulations proceed through the following steps:

1. sample a unitary transformation U according to the Haar measure;
2. apply C or R algorithms to retrieve the parameters (ω_i, ϕ_i) according to Eq.(4.20);

$$T_{k,k+1}^{(i)} = \begin{pmatrix} \not\leftarrow & & & 0 \\ -\sin \omega_i & e^{-i\phi_i} \cos \omega_i & & \\ \cos \omega_i & e^{-i\phi_i} \sin \omega_i & & \\ 0 & & & \not\leftarrow \end{pmatrix} \quad U_{R,C} = \prod_{i=1}^{\frac{m(m-1)}{2}} T_{k,k+1}^{(i)} \quad (4.20)$$

3. implement each $T_{k,k+1}^{(i)}$ as a Mach-Zehnder with input phase shift ϕ_i and $\tau_{i,1} = \tau_{i,2} = 2^{-1/2}$:

$$T^{(i)} \rightarrow -i \begin{pmatrix} \tau_{i,2} & i\sqrt{1-\tau_{i,2}^2} \\ i\sqrt{1-\tau_{i,2}^2} & \tau_{i,2} \end{pmatrix} \begin{pmatrix} e^{-i\omega_i} & 0 \\ 0 & e^{i\omega_i} \end{pmatrix} \begin{pmatrix} \tau_{i,1} & i\sqrt{1-\tau_{i,1}^2} \\ i\sqrt{1-\tau_{i,1}^2} & \tau_{i,1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi_i} \end{pmatrix} \quad (4.21)$$

4. introduce Gaussian noise on the four parameters $(\tau_{i,1}, \tau_{i,2}, \omega_i, \phi_i)$, by sampling new values from a normal distribution centered on the ideal ones and with widths σ_{BS} and σ_{PS} respectively for (τ_1, τ_2) and for (ω, ϕ) ;
5. generate new unit cells from the noisy values, adding possible losses $\text{diag}(\eta_i)$ at the output, and rebuild the noisy U .

This simple procedure allows us to investigate the effect of noise and losses on the C - and R -designs for the implementation of Haar random unitaries. Incidentally, our numerical simulations confirm that predictions would be remarkably different if, instead of generating every time a new unitary according to the Haar measure, we directly generated sets of uniformly distributed random parameters (ω_i, ϕ_i) to -mistakenly- speed up the calculation. Subsequently, the same analysis is repeated for all three architectures (C , R and F) focusing on the implementation of the Fourier and the Sylvester transformation.

Benchmarking Haar-random interferometers

The results of our first analyses are shown in Fig.4.10. The figure of merit adopted to compare the two architectures is the fidelity [178]

$$\mathcal{F} = \left| \frac{\text{Tr}(U_{imp}^\dagger U)}{\sqrt{m \text{Tr}(U_{imp}^\dagger U_{imp})}} \right|^2 \quad (4.22)$$

being U a $m \times m$ unitary transformation and U_{imp} its imperfect implementation. Thanks to the rescaling factor in the denominator, accounting for the contribution of lossy implementations, this fidelity is particularly suitable to characterize non-unitary transformations. Fig. 4.10a,b show the deterioration of \mathcal{F} for increasing values of losses η per unit cell and size m of the circuit for the C - and R - designs. In this simulation, optical elements are assumed to be immune to fabrication noises in order to isolate the η -contribution. As pointed out in Ref. [178], the C -design is more resilient to internal losses than the R - design, thanks to the almost total balance between the mode interactions. Moreover, as the heuristic non-linear fit suggests (see Supplementary Note 1), also the scaling of the fidelity is much more favourable for the former architecture, features that pushes C as a promising candidate for large-scale platforms or where, due to technical issues inherent to the specific implementation, it is not possible to guarantee a low level of losses in each optical element. We observe that, as already pointed out in Ref. [178, 192], inserting additional lossy elements in the R -scheme allows to compensate for the imbalance, at the price of significantly increasing the total amount of losses. Fig. 4.10c,d show instead the average effect of a noisy implementation of the optical components over the fidelity. Similarly to the previous analysis, our simulation is carried out for different levels of noise $\sigma = 0, \dots, 0.020$ and different sizes $m = 4, \dots, 128$, aiming to retrieve a more complete feeling of its scaling. Noise is assumed to be of equal intensity at this stage on both beam splitters transmissivities (σ_{BS}) and phase

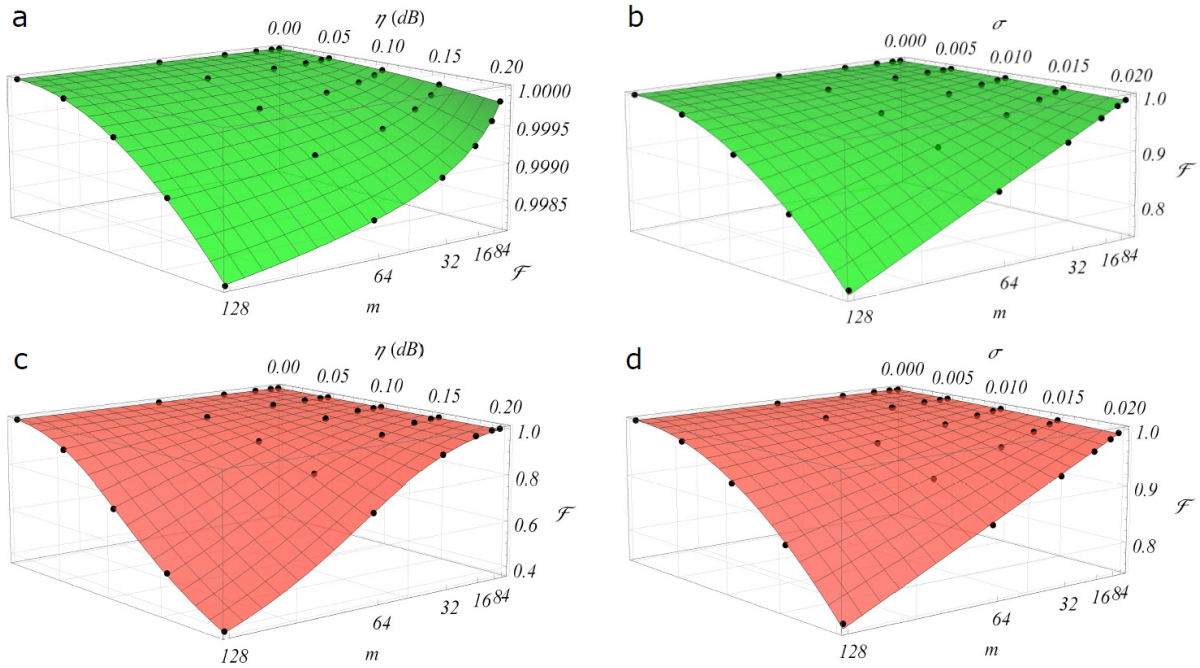


Figure 4.10. Haar-random with losses and noise. Noise and losses affect the implementation of Haar-random unitaries in different ways in the C - and R -designs. **a, c**, average fidelity F for different levels of loss η per beam splitter and size m of the interferometer in the C (**a**) and R (**c**) designs. Note the difference in vertical scale, due to the more balanced structure of C where the dependency on η arises from the slight path asymmetry of the outer waveguides. **b, d**, average fidelity F for different levels of noise σ in the optical elements and size m , averaged over 500 noisy unitaries. Here noise is treated equally on both beam splitters and phase shifters, namely $\sigma = \sigma_{BS} = \sigma_{PS}$. Note that the scaling is identical in (**b**) and (**d**) since, averaging over the unitaries, the deterioration of F is due to the number of noisy elements, which is the same in the two schemes. Surfaces: heuristic non-linear fits of the data.

shifts (σ_{PS}), i.e. $\sigma = \sigma_{BS} = \sigma_{PS}$, in order to capture the scaling of the performance in a unique three-dimensional plot. Our simulations confirm previous qualitative estimates [178] concerning the similarity of the scalings of the average fidelity in the two architectures, providing a clear and quantitative picture of its dependency on the fabrication noise over the optical elements. Such a similar trend is observed also when investigating their performance in the case of crosstalk between thermal shifters.

After this preliminary stage, more general investigations have been carried out by considering different levels of noise over the fabrication transmissivities and phases, i.e. studying the practical situation where $\sigma_{BS} \neq \sigma_{PS}$. Results of this analysis are shown in the contour plots of Fig. 4.11 for different sizes of the interferometers, highlighting a number of interesting features. First, the qualitative dependency on the noise seems to remain fixed while increasing the dimension of the circuit, though the average fidelity rapidly drops to low values already at $m = 64$, for intensities of noise that are within the tolerances of current technology, in particular for reconfigurable circuits. Moreover, we observe that the two sources of noise affect the fidelity in a similar way, with an intensity approximately double for the one on the transmissivities in the Mach-Zehnder. Thus, while dynamic control over the phases can in principle mitigate the effects of noisy implementations, non-ideal values of the transmissivities of the symmetric beam splitters remain critical when looking at large-scale implementations. To focus our analysis on the circuit size allowed by current technology, we have considered as a reference current top-level implementations for quantum applications. Circuits with up to $m = 16$ have been reported for non-reconfigurable femtosecond laser written circuits [265], $m = 6$ for universally reconfigurable lithographic circuits [22] and $m = 26$ for reconfigurable lithographic circuits that

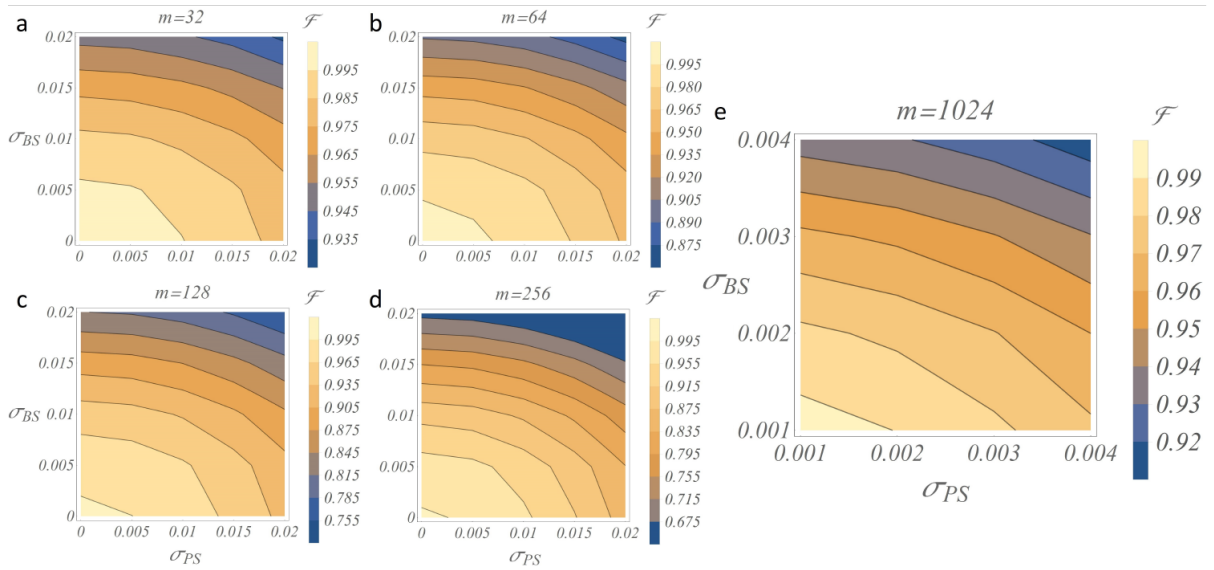


Figure 4.11. Haar-random with non-ideal transmissivities and phase shifts. Fabrication imperfections and errors in real-time control over the phases are in general on different scales, depending for instance if we are considering (un)balanced beam splitters or (not-)reconfigurable phase shifters. Here noisy Haar-random implementations are investigated under different values of σ_{BS} and σ_{PS} , to analyze the separate contribution of each source of error to the final unitary transformation. Since the average robustness against noise of the C - and R -designs is equivalent (see Fig.4.10), here the contour plots are shown for the C scheme for $m = 32$ (a), 64 (b), 128 (c), 256 (d) and 1024 (e) (see Supplementary Note 2 for the optimized decomposition adopted for these simulations).

do not implement a fully arbitrary unitary matrix [23]. Reasonable sizes achievable in the near future thus can be $m = 32$ or $m = 64$ depending on the level of reconfigurability required. However, looking at future large-scale implementations, we see that for $m = 1024$ even a very low level of noise makes the average fidelity drop to values as low as ~ 0.90 . Note the shift in the axes scales between the four left contour plots ($m \leq 256$) and the one on the right ($m = 1024$). Thus, for an efficient scaling a technology leap will be required not only in the circuit size but also on the level of fabrication tolerances.

So far, we adopted as figure of merit the fidelity of the quantum process. Though perfectly suitable to benchmark noisy transformations, this quantity fails in general to highlight more complex many-particle phenomena. Partial deviation from their ideal symmetric structures can then spoil the interference effects in the output distributions. For this reason, we investigated the performance of noisy architectures also in the scope of multiphoton output probability distributions, employing as figure of merit the total variation distance (TVD) between the ideal (P) and actual (\tilde{P}) n -photon distributions:

$$TVD(P, \tilde{P}) = \frac{1}{2} \sum_i |p_i - \tilde{p}_i|. \quad (4.23)$$

Indeed, the TVD plays a relevant role in testing statistical hypotheses since the quantity $1 - TVD$ is a lower bound on the sum of Type I and Type II error rates [222]. From a practical point of view, $TVD(P, \tilde{P})$ is also a simple and natural metric to quantify the discrepancy between the two probability distributions P and \tilde{P} . Results for this analysis are shown in Fig. 4.12: again, noise affects almost equally the C and R architectures when averaging over all the input/output states. The slight difference in favour of the R -design may not be practically appreciable in real experimental conditions and it rapidly becomes negligible for higher values of (n, m) . Thus, the two architectures behave equally as far as lossless multiphoton investigations are concerned.

We conclude from our analysis that the C - and R -designs are ultimately equivalent in terms of resilience to noise when averaging over all input/output configurations. On one side, the single

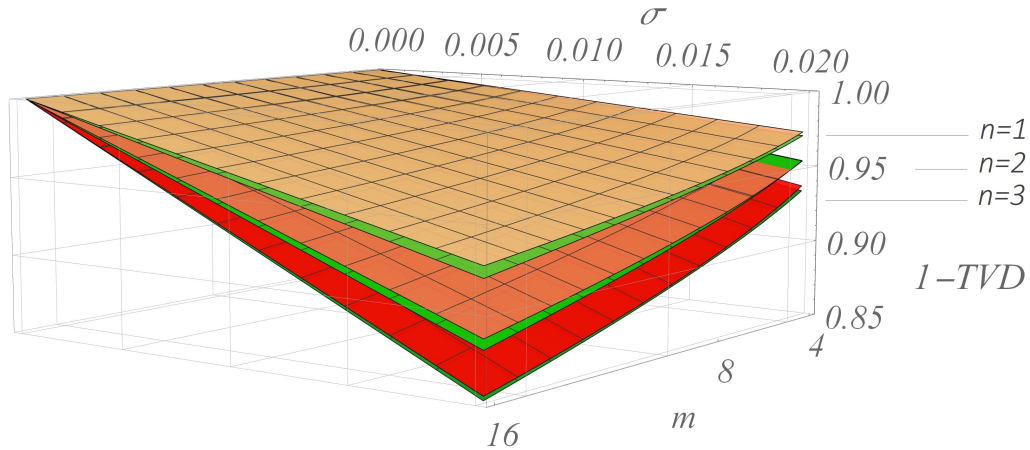


Figure 4.12. Multiphoton interference in noisy Haar-random interferometers. Several applications require to evolve many-photon states in large-scale quantum circuits. The total variation distance (TVD) is in this sense a good estimate of the goodness of an experimental implementation, being it a natural measure of the distance between two probability distributions. Here, plots compare the ideal and noisy output probability distributions relative to $n = 1, 2, 3$ -photon collision-free states, averaged over all the inputs. For each n , 100 unitaries are sampled and implemented according to the C - (green) and R - (red) designs setting $\sigma_{BS} = \sigma_{PS} = \sigma$. Our analysis confirms on one hand the prediction of Fig.4.10c,d, where the R -design is found to be slightly more robust against noise, but on the other it shows how this difference seems to become negligible for increasing values of n . Surfaces: heuristic non-linear fits of the data.

optical elements affect in a different way the elements of the unitary transformation in the two schemes, being the C and R parameters more localized respectively in the upper corner and central part of the unitary matrix. However, for practical noise levels and general multi-input applications, this difference does not give rise to any effective deviation between the two schemes. In contrast, the two architectures behave differently when it comes to lossy implementations, where the symmetric C -design outperforms the R scheme.

Analysis of Hadamard matrices

Here we will focus on the performances of such architectures in the specific cases of Fourier and Sylvester transformation which will be important in the next Chapter. We use the fidelity in Eq. (4.22) as a figure of merit to compare the two architectures:

being U a $m \times m$ unitary transformation and U_{imp} its imperfect implementation. Thanks to the rescaling factor in the denominator, accounting for the contribution of lossy implementations, this fidelity is particularly suitable to characterize non-unitary transformations. Now, though the C - and R -designs are universal for unitary evolutions, often it is desirable to have circuits optimized for specific relevant tasks. It is the case of the quantum Fourier [82, 182, 201] or Sylvester [169] transforms, which have importance on their own in different scopes of quantum information processing.

We investigate the performance of the Fast (F) scheme [181, 182] for the implementation of generalized Hadamard transformations [177] and compare it with the average performance of the two universal designs. In Fig. 4.13 is shown the comparison. In the Fast scheme, being the circuit completely symmetric, the structure is totally immune to constant propagation losses, beating even the highly resilient C -design. The F -design is also more resilient to noise, thanks to the reduced depth of the circuit which lowers the number of noisy optical elements from $\frac{m}{2}(m-1)$ to $\frac{m}{2} \log m$. A benchmark summary of this comparison is reported in Table 4.3 for both the quantum Fourier transform and the Sylvester interferometers.

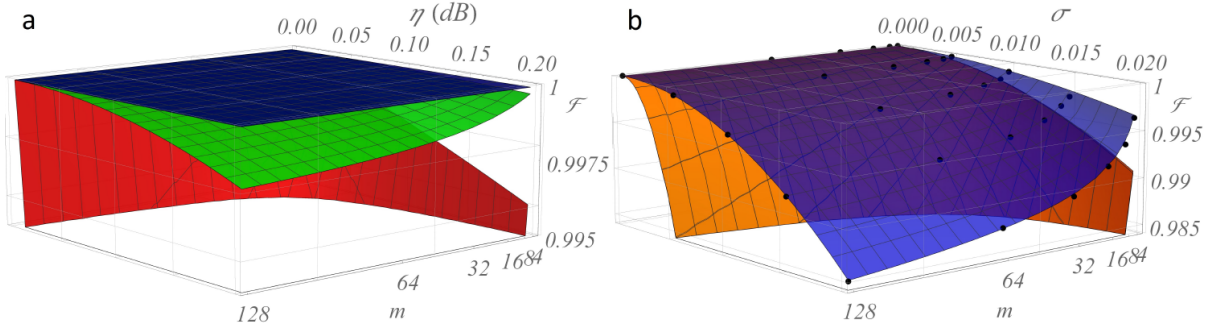


Figure 4.13. Generalized Hadamard transformations with C -, R - and F designs. While universal for unitary decompositions, the C - and R - architectures offer suboptimal solutions for the implementation of specific classes with higher symmetries. The Fast architecture is optimized for the implementation of the quantum Fourier transform and the class of Hadamard transformations. **a)**, Average fidelity as a function of losses per unit cell and network size (blue, green and red surfaces for F -, C - and R -designs respectively). Full symmetry between the optical paths cancels out the effect of constant losses per unit cell in the F - design. **b)**, Fast architectures (blue surface) are also more resilient to fabrication noise, thanks to the reduced depth of the circuit. Here, only one (orange) surface is shown for both C - and R -designs, assuming equal resilience to noise. Surfaces: heuristic non-linear fits of the data. Image taken from [176].

Quantum Fourier transform and Sylvester									
	$m = 64$			$m = 128$			$m = 256$		
$\sigma_{BS} \rightarrow$	0.005	0.01	0.02	0.005	0.01	0.02	0.005	0.01	0.02
$\sigma_{PS} = 0.001$	0.999	0.995	0.981	0.998	0.994	0.976	0.998	0.992	0.972
	0.994	0.975	0.904	0.987	0.950	0.817	0.975	0.903	0.667
$\sigma_{PS} = 0.01$	0.998	0.994	0.980	0.997	0.993	0.975	0.997	0.992	0.971
	0.984	0.966	0.895	0.969	0.933	0.803	0.938	0.870	0.644
$\sigma_{PS} = 0.02$	0.995	0.991	0.976	0.994	0.989	0.971	0.993	0.988	0.968
	0.956	0.939	0.870	0.917	0.882	0.759	0.838	0.770	0.575

Table 4.3. Benchmarking Hadamard transformations. Generalized Hadamard transformations benefit from optimized architectures, obtained from the efficient algorithms [179, 181] developed for the fast quantum Fourier transform. Here, average Fidelities (in each row, F : up; average of C and R : down) are reported for each combination of size m and noises σ_{BS}, σ_{PS} , averaging over 100 noisy unitaries. Data for the Fourier and Sylvester transformations are displayed in a single table since the values in the two cases are equal within a discrepancy lower than ~ 0.001 .

Discussion

Photonic technologies promise to enable the application of several quantum information protocols, ranging from fundamental research to quantum computation and optical quantum networks. In this chapter we have provided a description of the most diffused integrated photonics techniques and we performed a brief analysis of the performance of the three main interferometric schemes, namely the triangular [15] and square [178] designs and the Fast [181] architecture, under realistic conditions of losses and noise.

Fast architectures represent a specialized design to implement a significant class of unitary evolutions, highly optimized for the realization of Fourier and Sylvester (Hadamard class) quantum transformations. Our results quantitatively highlight the improved performance of this scheme

with respect to the universal ones, thus making it the preferred choice in this context. According to this analysis we expect that Fast architectures can play a key role among photonic platforms in synergy with the universal schemes, to fully benefit from the unique advantages of both designs. In Chapter 5 we will see deeply the usage of this platform by exploiting its resilience and robustness to imperfections to validate genuine quantum interference of indistinguishable boson states.

Chapter 5

Quantum certification of many-body bosonic interference in photonic chips

In this chapter we will present different approaches and tools to generate many-body bosonic interference and subsequently we will focus on how to identify a signature of indistinguishability by employing different validation protocols. We begin the description focusing on the task of "how can I scale up my system" by describing the experimental realization of the Scattershot Boson Sampling using three SPDC sources to generate random inputs. Subsequently we will focus on how to validate our data performing different protocols with our integrated platforms. We will perform an analysis on the optimal unitaries which are suitable for addressing genuine quantum interference. Following these roads, we targeted two unitaries belonging to the class of Hadamard matrices, namely Fourier and Sylvester matrices, which allow us to introduce validation protocols based on the so-called *suppression law* which is an intrinsic property of these matrices. Finally we will investigate alternative protocols based on machine learning techniques and statistical tools. In Chapter 4 we discussed the importance to search for the best implementation in order to avoid significant losses to perform correctly our protocols. Here we show how using **FLM** with combination of the Fast architecture we can find alternative models in order to generate and individuate a good signature of true particle interference without employing the calculus of the permanent and thus avoiding to reconstruct the output distribution of these devices. This Chapter represents the second part of PhD work and we report our results respectively in these papers, in which I contributed mainly in the development of the experimental setup, characterizations and data analysis and secondly to the simulations of our model, except for where I listed as first author where I contributed equally likely in all fields:

- M. Bentivegna, N. Spagnolo, C. Vitelli, F. Flamini, N. Viggianiello, L. Latmiral, P. Mataloni, D. J. Brod, E. F. Galvao, A. Crespi, R. Ramponi, R. Osellame and F. Sciarrino. *Experimental scattershot boson sampling*. *Sci. Adv.*, **1**, e1400255 (2015) [149].
- A. Crespi, R. Osellame, R. Ramponi, M. Bentivegna, F. Flamini, N. Spagnolo, N. Viggianiello, L. Innocenti, P. Mataloni and F. Sciarrino. *Suppression law of quantum states in a 3D photonic fast Fourier transform chip*, *Nat. Comm.*, **7**, 10469 (2016) [182].
- F. Flamini, N. Viggianiello, T. Giordani, M. Bentivegna, N. Spagnolo, A. Crespi, G. Corrielli, R. Osellame, M. A. Martin-Delgado and F. Sciarrino, *Observation of Majorization Principle for quantum algorithms via 3-D integrated photonic circuits*, preprint at arXiv:1608.01141, (2016) [170].
- N. Viggianiello, F. Flamini, M. Bentivegna, N. Spagnolo, A. Crespi, D. J. Brod, E. F.

Galvao, R. Osellame and F. Sciarrino. *Optimal photonic indistinguishability tests in multimode networks, preprint at arXiv:1710.03578*, (2017), submitted to Optica [208]

- N. Viggianiello, F. Flamini, L. Innocenti, D. Cozzolino, M. Bentivegna, N. Spagnolo, A. Crespi, D.J. Brod, E. F. Galvao, R. Osellame and F. Sciarrino. *Experimental generalized quantum suppression law in Sylvester interferometers, Preprint at arXiv:1705.08650*, (2017) [193].

The last two works are currently in submission (see "List of Publications" at the end of the manuscript).

5.1 Experimental Scattershot Boson Sampling

According to the extended Church-Turing thesis (ECTT), any physical system can be simulated efficiently, with at most a polynomial overhead in its size, by a classical computer. However, since the advent of quantum computation, the ECTT has been questioned, although the available empirical evidences were not sufficient to dismiss the ECTT as a central dogma of computer science. A more feasible challenge to the ECTT is provided, as introduced in the previous Chapter 2, by Boson Sampling. This task consists on simulating output events of n bosons entering in the n different input ports of a $m > n$ modes interferometer implementing an arbitrary unitary matrix U . Consequently, one draws output events with probability $P(S; T; U)$, through S and T (the input and output state respectively), which corresponds to the permanent of the submatrix of U that contains only the rows and columns matching the occupied input and output modes:

$$P(S, T; U) = |\text{Perm}(U)|^2 \quad M_{p,q} = U_{s_p, t_q} \quad (5.1)$$

It is desirable to choose the matrix U randomly, in order to avoid regularities that could simplify the classical simulation. Unitaries are selected randomly thanks to the Haar measure [194], which defines the uniform distribution analogous on the topological space of the unitary matrices. This particular choice has been made in order to avoid any computational simplification due to alleged symmetries of the unitary operator. Enlarging the number of injected photons n , the permanent calculation becomes intractable with classical algorithms, due to a superpolynomial use of resources. Analysis of the hardness of this problem have been studied in details in Refs. [33, 121]. The believe on the studies and the claims regarding the complexity of this model, makes the Boson Sampling the good candidate to find real evidences of the quantum-over-classical speed up. Moreover, the phenomena at the core of this model is bosonic interference which has several applications in different fields [12]. By employing integrated photonics, Aaronson's model is a good platform to study in detail, the effect of true quantum interference.

To identify signatures of many-body quantum interference we used as testing platform Boson Sampling experiments. We performed such experiments by evolving indistinguishable photons into implementations of Haar-random unitaries matrix on photonics chips with the technique of FLM. From Chapter 3 and 4 we described the experimental tools and the importance to efficiently address such implementations by the search of best architecture. First experimental attempts have been carried out with small scale chips in Refs. [137–140] and showed the first Boson Sampling experiments by paving the way to further studies in large scale implementations. Rather than increasing the number of modes, the hardness of such phenomena relies in increasing the number of photons n which interfere in these device. Thus, it is extremely important to provide an enhancement on the generation rate of n -indistinguishable photon states. Nevertheless, with SPDC sources, this can be really complicated, since the process of photon-pair generation is not-efficient at all. There is, however, a backdoor as illustrated in section 2.5, where we can focus rather than increasing the number of photons but improving the generation of fixed n -photon states. This approach has been called Scattershot Boson Sampling (SBS) and in this section we

will describe the first experimental demonstration that we have carried out in our laboratories. Our work is reported in Ref. [149].

Here we report experimental results on SBS experiments using a 13-mode integrated photonic chip. We use up to six PDC photon sources to obtain data corresponding to 2- and 3-photon interference by heralding all the inputs configurations, thus sampling both from input and output probability distributions. As an additional step we perform the test on a different 9-mode chip where we emulate the same behaviour by injecting into different inputs manually, giving a signature of the robustness of the protocol.

To measure n -boson interference we prepare the state by manipulating polarization and relative delay in order to achieve good indistinguishability with minimum visibility observed for different pairs of 0.87 ± 0.03 . For this reason, 2-photon HOM have been performed for each input combination as a preliminary stage to individuate the precise temporal region by means of delay lines. For the 9-mode chip, the input state was created by a 4-photon SPDC source (crystal C_α in Fig. 5.1b), with one of the photons used as a trigger. Our preliminary experiment involved simulating the statistics of a Scattershot Boson Sampling experiment in the 9-mode chip by manually connecting 20 different sets of input modes to the source, via a fiber array, and uniformly mixing the data corresponding to different input states.

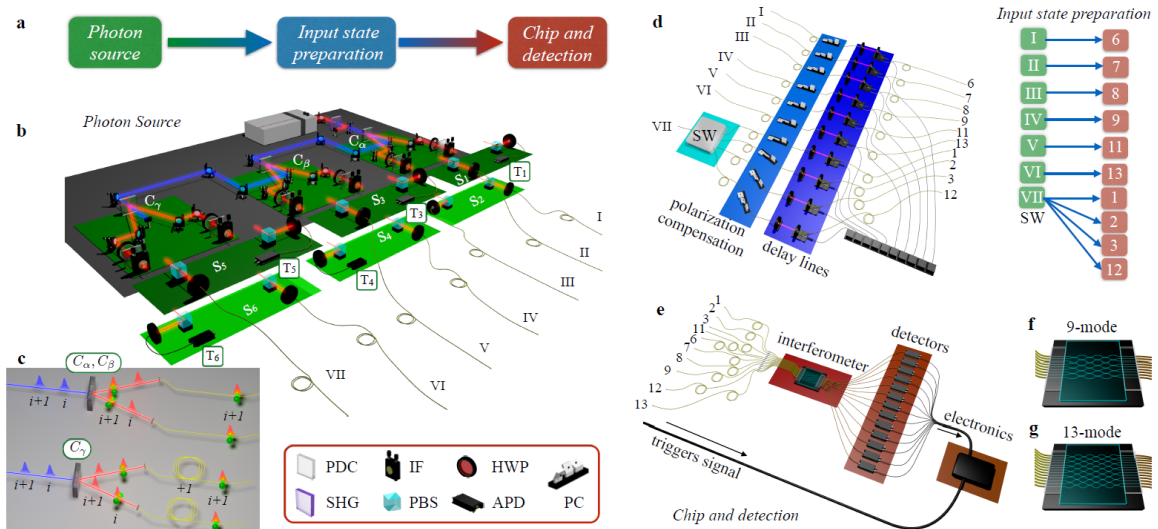


Figure 5.1. Overall conceptual scheme of the experiment. In each of the 3 BBO crystals (C_α , C_β and C_δ), photon pairs are generated via Type II parametric down conversion process. The two possible polarization combinations for the two generated photons, HV and VH, constitute two equal SPDC sources enfolded in the same crystal, each one exciting a different trigger (photon v) and a different input mode (photon H). The only exception is given by source S_2 , whose outputs are both injected in the chip. Sources are also time-multiplexed, since pulses generating photons in crystal C_δ are produced before the ones generated in C_α and C_β . Each input is associated to specific mode of the 13-optical chip. Ref. [149].

We used a total of six PDC sources (S_1 to S_6 in Fig. 5.1b) to perform the SBS experiment. We simplified the implementation by enfolded two equal sources in each crystal, corresponding to the two possible vertical/horizontal polarization combinations for the photon pair generated. Hence, the six sources S_1 - S_6 are created using only the three crystals C_α , C_β and C_γ . One such source (Source S_2) prepares photons I and III, which enter the interferometer respectively in fixed modes 6 and 8. The other five PDC sources produce random, but heralded, single photons, which are coupled to different input ports of the chip. We further increased the input variability by distributing photon VII randomly among four different input ports, via an optical fiber switcher with switching rate comparable to the obtained experimental count rate.

For both chips, the output photons are collected by a multimode fiber array, and multiphoton

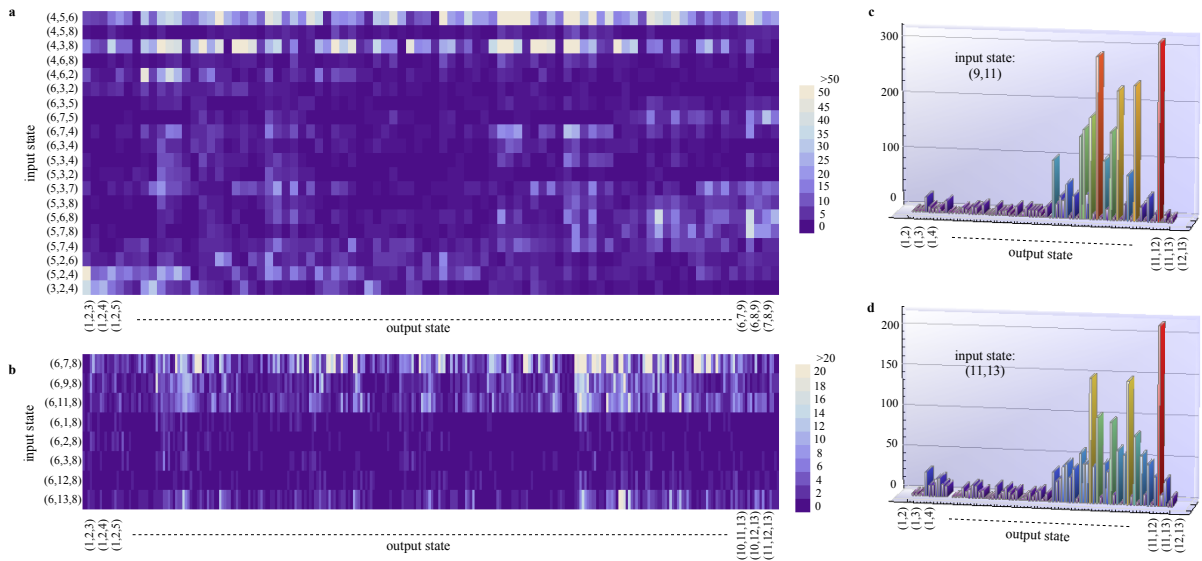


Figure 5.2. Multiple input Boson Sampling in a 9-mode device and Scattershot Boson Sampling in a 13-mode device. **a**, Density plot of the number $n_{i,j}$ of events detected for each of the 1680 input (i) and output (j) combinations used in our Boson Sampling experiments with the 9-mode chip. **b**, Density plot of the number $n_{i,j}$ of events detected for each of the 2288 input (i) and output (j) combinations used in our Scattershot Boson Sampling experiment with the 13-mode chip. **c,d**, Number $n_{i,j}$ of events detected for a two-photon Scattershot experiment with the 13-mode chip for input states (9,11) (**c**) and (11,13) (**d**). Ref. [149]

coincidences are detected by avalanche photodiodes, coordinated by an electronic data acquisition system capable of registering events with an arbitrary number of photons. Synchronization of pairs from the same crystal is achieved by HOM measurements, while synchronization of photons belonging to different crystal is reached exploiting stimulated emission with a pulsed laser seed. The observed number of events corresponding to each input/output combination for the 9- and 13-mode chips are shown respectively in Figs. 5.2a and 5.2b. Note the sparseness of the data set, as only a few events corresponding to each input/output combination are observed (if any). This is an expected feature of more complex Boson Sampling experiments whose number of possible input-output combinations may far exceed the number of observed events. Furthermore, in Figs. 5.2c and 5.2d we show the results for 2-photon experiments, in which each input is a doubly-heralded 2-photon state.

Moreover, in this experiment, we exploited time-multiplexing [195] to achieve quantum interference between photons generated by different pulses. The direct consequence is an increase of the n -photon generation rate using a fixed number of SPDC sources. This was done by introducing appropriate delays, so that photons from Sources S_5 and S_6 are produced by a different pump pulse than those generated by all the other sources (see Fig. 5.1c). It is worth noting how with this approach we can explore higher orders of the Hilbert space by exponentially increase the generation rate of n -fixed photon injected. The same setup has been used also for further studies as we will see in the next chapter. The only difference consists in injecting 2-photon states in our devices instead of 3, thus considerably enhancing the photon counts.

In the next section we will focus on a deeper problem which relies on validating our platform by trying to overcome the problem on identifying signatures of true many-body interference in a reasonable runtime by means of our classical apparatus.

5.2 The problem of certification

Despite several progresses in finding a suitable model to reach the so-called *quantum advantage*, a tricky problem arises which enhance the hardness of the problem even more. As we increase the size of the system in number of modes and photons, predict the output distribution by means of classical calculations become rapidly intractable. What about the correct certification of our data? Supposing that we can reach a situation where in the near-future technology we are able to build a stable and lossless platform to measure true many-body quantum interference and thus perform our quantum protocol. Can we trust on the correctness of our data? In principle we have to validate the output of such devices, but here the same original problem manifests from another point of view. Indeed, under the conditions when classical computers are not able to perform the permanent calculation, we cannot verify if the output distribution we are looking is true many-body interference or not. Verify this feature is hard as the sampling the output distribution, or in another words, it has the same complexity of Boson Sampling itself. Moreover, being a sample problem, since we are dealing with small size platforms, we are able to experimentally to retrieve our data in a reasonable time, while with higher Hilbert spaces, the sampled distributions become considerably flat, thus it gives us not sufficient elements to validate. To sum up, a posteriori verification of the output probability distribution, experimentally obtained, results essential and strictly necessary in order to trust on our data and the functioning of our device.

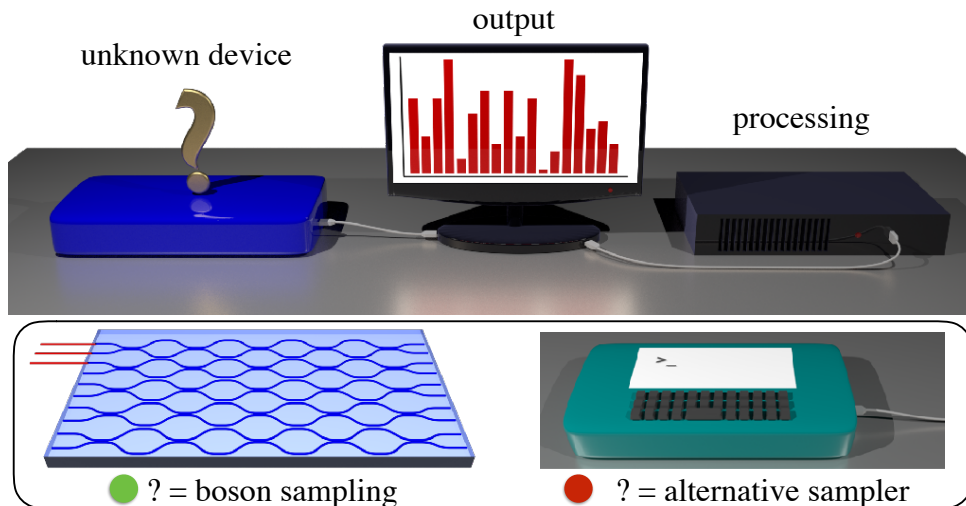


Figure 5.3. Validation of Boson Sampling experiments. An agent has to discriminate whether a finite sample obtained from an unknown device has been generated by a quantum device implementing the Boson Sampling problem or by an alternative sampler.

Even though such experimental conditions have not been achieved yet, their realization in the next future is not so far. Consequently many validation protocols have been proposed, that is, protocols applied to the output probability distribution data and used to certify if such data have been produced by a boson sampler or not. In other words, to certify a boson sampler, one needs to verify predictions following from equation (5.1) and rule out models that yield these predictions without true many-boson interference. A first insight into the validation issue was obtained by Gogolin et al [196], who showed that so-called symmetric algorithms fail to distinguish the distribution of experimental data from the trivial, uniform one. Thus, the first question to answer was how a boson sampling distribution could be distinguished from a uniform sampling one. Aaronson and Arkhipov in 2014 [197] argued that it was unreasonable to restrict the statistical analysis to symmetric algorithms only and, moreover, they proposed a validation protocol that efficiently discern the boson sampling distribution from a uniform sampling distribution.

The Aaronson and Arkhipov test is named Raw Norm Estimator (RNE) and is based on the

estimator P defined by the following expression

$$P = \prod_{i=1}^n \prod_{j=1}^n |A_{i,j}|^2 \quad (5.2)$$

The estimator P has been shown to be correlated with the boson sampler probability associated with that event, but it does not involve any permanent calculation. This feature is extremely important, since makes the test computationally efficient. After having calculated the estimator, the test decides if the outcome is more likely to arise from boson sampling, that is $P > (n/m)^n$, or from uniform sampling, $P < (n/m)^n$. In this sense, at each event, an integer coefficient C is increased or decreased by a unity, depending on which distribution the event belongs.

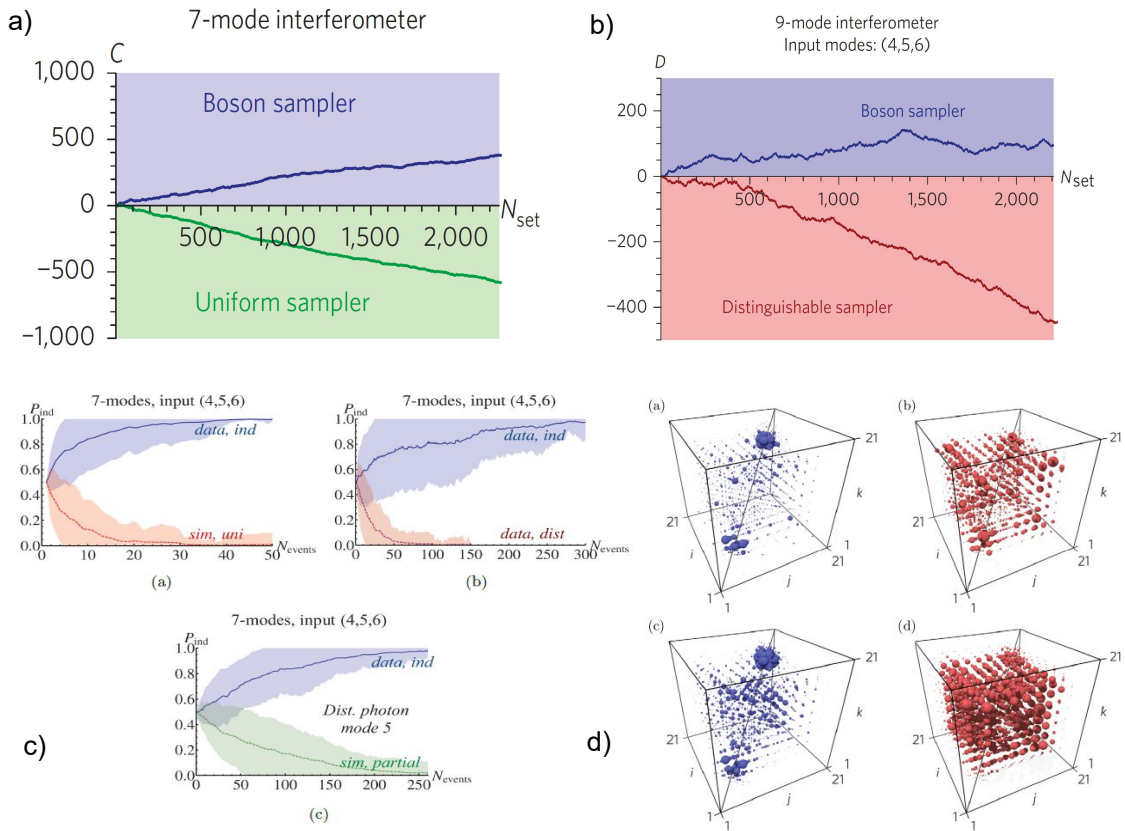


Figure 5.4. List of previously experimentally tested validation protocols. **a)** Validation test with RNE. a 7-Haar random unitary has been experimentally implemented. By collecting data and applying the test of AA they were able to discriminate uniform sampler from Boson sampler. **b)** Likelihood-ratio test performed on 9- mode chip. Brute force approach is not efficient because we have to deal with calculus of the permanent, but we are able to discriminate a more general behaviour for the device, framing on the difference in terms of distinguishability **c)** Bayesian test on Boson sampler data versus uniform and classical samplers. It requires very low samples to discriminate the two behaviour but it is still not efficient in n . **d)** Clouding test which is able to discriminate distinguishable particles in a quantum walk approach (blue), conversely considering a Haar random matrix (red) we cannot distinguish the numbs significantly. Images taken from [198–200].

It is worth noting that in this case, we do not pass through the calculus of the permanent, so it is efficient to calculate scaling n and m . Unfortunately, such single-particle estimator does not contain hints of genuine quantum interference.

The discrimination between distinguishable and indistinguishable photons is the second case we analyse. We aim to validate a boson sampling distribution, obtained with indistinguishable

photons, from the output distribution obtained when distinguishable photons [198] are sent into the same boson sampling input modes and evolve with the same unitary matrix. In this case, the events are associated through a probability p_i (q_i) to come out according to an indistinguishable (distinguishable) output distribution and discriminated with a parameter $R = p_i/q_i$ and subsequently divided according to it. This model is more physical in the sense that we compare a reasonable phenomenon that occurs often during this experiment, but it suffers of a drawback: it must pass through the calculus of the permanent so it does not scale well when we increase the size.

Carolan et al. [199] proposed a method which relies on the natural behaviour of bosons to bunch. The great advantage of this method is that the emergence of bosonic clouds provides away to predict multimode correlations in the transition from the distinguishable to the indistinguishable regime, avoiding the calculation of matrix permanents that describe the relative probabilities for the multimode detection events.

Bentivegna et al. [200, 211] proposed a different road which relies on a bayesian probabilistic approach. Equal *a priori* probabilities, Q and R , are assigned to two different hypothesis, respectively the boson sampling and an alternative one. For each sampled event x , let q_x be the corresponding probability associated to it under the hypothesis Q , and r_x the corresponding probability under hypothesis R . With the Bayes theorem we address more confidence to events through the conditionate probabilities. The main advantage of the Bayesian approach is that it allows to perform validation tests against a large set of alternative hypotheses, since no additional assumption is required for the alternative sampler R . Although it needs only few events to converge, it does not scale well with the number of photons because it has to calculate the permanents in each iteration.

At the end, all of these methods are not sufficient to reach the goal. In figure 5.4 we reported the previous experimental attempts to this protocol. Now, despite problems with the efficiency for large sizes, these models are not able to discriminate a many-body bosons interference respect to semi-classical approximation of such state, namely *mean field state* with a sampling problem [201–203]. Within this description a multimode Fock state is approximated by the macroscopically populated single particle state:

$$|\Psi\rangle = \frac{1}{\sqrt{n}} \sum_{r=1}^n e^{i\theta_r} |\Phi_{j_r}^{in}\rangle \quad (5.3)$$

where the phase θ_r are undefined. The state 5.3 evolves into:

$$\hat{U} |\Psi\rangle = \frac{1}{\sqrt{n}} \left[\sum_{q=1}^m |\Phi_{j_r}^{(out)}\rangle \left(\sum_{r=1}^n e^{i\theta_r} U_{j_r,q} \right) \right] \quad (5.4)$$

The last relation specify that each particle occupies the output mode $1 \leq q \leq m$ with probability given by:

$$P_q^{(mf)} = \left| \langle \phi_q^{(out)} | \hat{U} |\Psi\rangle \right|^2 = \frac{1}{n} \left| \sum_{r=1}^n e^{i\theta_r} U_{j_r,q} \right|^2 \quad (5.5)$$

Hence, the mean-field sampler contains those aspects of many-boson dynamics that survive in the semi-classical limit. The mean-field approximation results as a fundamental test for each certification protocol previously analysed, but we will show in the next chapter the work that we spent during my PhD project in which we will search for an alternative validation test which relies not only to the statistical behaviour of bosonic ensemble but we will exploit several techniques which are based on laws of particular symmetrical matrices belonged to specific classes of symmetry, statistical tools and last but not the least on machine learning techniques.

In the next sections we will see the work that we have carried out experimentally focusing on the techniques and which devices occur to realize the validation protocol implemented in our laboratories. In specific we will discuss and analyse in detail in the approach that we used. For further details at the end of this manuscript there are Appendices that stressed the concepts reported in the following sections, providing other detailed on the techniques used.

5.3 Suppression law of quantum photonic states

As pointed out previously, all protocols seen are not efficient in terms of scaling of photon number since it requires the calculation of $n \times n$ matrix permanents.

The protocol, here described, instead, relies on computable quantities¹. It is based on the *suppression* of specific output configurations in an interferometer implementing an n^p -dimensional Fourier matrix, with $p \in \mathbb{N}$. Experimental test on with this protocol has been done with $n = 2$ and $p = 2, 3$ [182, 201].

The quantum suppression law [201] can be thought as a generalization of the 2-photon/2-modes Hong-Ou-Mandel effect, where only one state were suppressed. It depends on the specific unitary transformation being implemented and on the symmetry of the input state. In particular, let us consider a *cyclic* input, *i.e.* an n -photon Fock state over $m = n^p$ modes, where the occupied modes j_r^s are determined by the relation:

$$j_r^s = s + (r - 1) n^p, \quad (5.6)$$

with $r = 1, \dots, n$ and $s = 1, \dots, n^{p-1}$. The parameter s takes into account the fact that there are n^{p-1} possible n -photon arrangements with periodicity n^{p-1} , which simply differ by a translation of the occupational mode labels. For example, with $n = 2$ and $p = 2$ there are $2^1 = 2$ possible cyclic states, *i.e.* $|1010\rangle$ and $|0101\rangle$.

The evolution determined by the Fourier matrix results in the suppression of all output states for which the following relation does not hold:

$$\text{mod} \left(\sum_{l=1}^n k_l, n \right) = 0, \quad (5.7)$$

where k_l is the output mode of the l -th photon.

Hence, suppression laws are applied to certify the presence of true many-body granular interference during the evolution in the interferometer, ruling out alternative hypotheses which would result in similar output probability distributions. In particular, this interference patterns cannot be simulated by distinguishable photons or any other simulated bosonic state, which, by contrast, can reproduce macroscopic interference effect as the *boson clouding* [199].

The degree of violation \mathcal{D} of the suppression law is quantified by the following expression:

$$\mathcal{D} = \frac{N_{\text{forbidden}}}{N_{\text{events}}} = \sum_{(i,j)_{\text{forbidden}}} P_{i,j}^D (1 - V_{i,j}), \quad (5.8)$$

the ratio between the number of observed events in the forbidden output states divided by the total number of events, where $P_{i,j}^D$ are the classical probabilities and $V_{i,j}$ are the visibilities of HOM dips. If a Fock state is injected in an interferometer implementing a Fourier matrix with no experimental imperfections, a violation $\mathcal{D} = 0$ would be observed. In the case of distinguishable photons there is no suppression law, and the violation would be simply the fraction of suppressed outputs, each one weighted with the number of possible arrangements of the n distinguishable particles (D) in that output combination. In the case of two-photon states, the weighting factor

¹The experimental work has been carried out during my PhD project and the paper is published in Nature Communications in Ref. [182]

is 2 for collision-free outputs and 1 otherwise, and a degree of violation of $1/2$ is expected. On the contrary, in the case of two-photon Mean Field state (MF), bunching effects occur leading to an expected degree of violation of half the weighted fraction of suppressed outputs ($1/4$ for two-photon MF). It has been shown that the fraction of forbidden outputs is always large [201]. Hence, a comparison of the observed value of \mathcal{D} with the expected one represents an efficient way, in terms of necessary experimental runs, to discriminate between Fock states, distinguishable particles states and MFs.

In this work we experimentally tested this validation protocol by exploiting the femtosecond lase micromachining with the Fast scheme (see Fig. 4.3). The observations have been carried out with two-photon Fock states injected into 4-mode and 8-mode qFFT interferometers (see Appendix 8.2 for the detailed implementation). The peculiar behaviour of Fock states compared to other kinds of states is investigated, showing in principle the validity of the certification protocol for the identification of true granular n-particle interference, which is the source of a rich landscape of quantum effects such as the computational complexity of Boson Sampling. The experiment was carried out with the experimental setup described at the end of Chapter 3 with a 4- and 8- mode Fourier matrix. We measured all the output HOM dips and peaks for the cyclic states (where in Fig. 5.5 is shown the measurements for a cyclic input of 8-mode chip) and we calculated all the violations following Eq. (5.8).

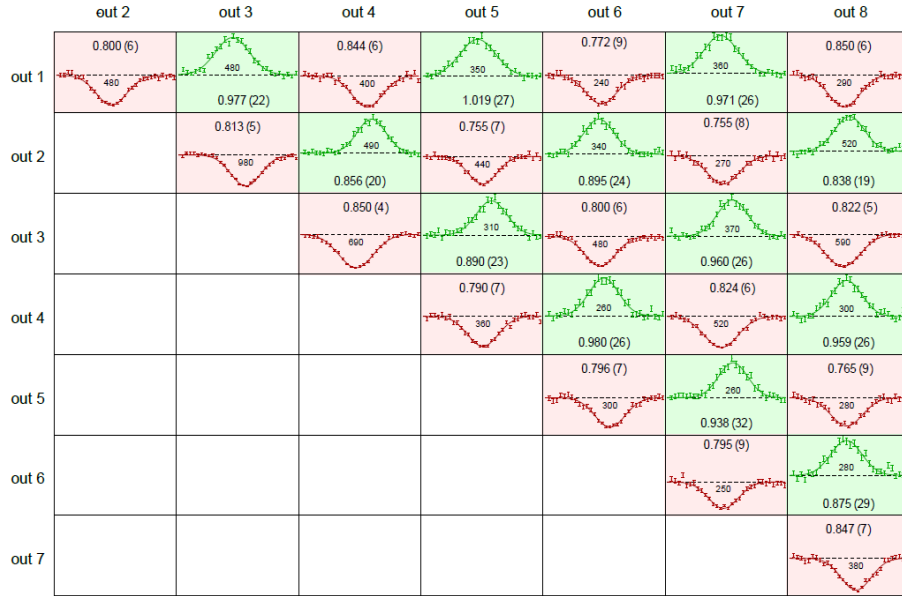
As shown in Fig. 5.6, blue shaded regions in the plots correspond to the cases where the hypothesis of distinguishable particles can be ruled out. Red regions correspond to the cases when both the hypotheses of distinguishable particles and mean field state can be ruled out, and true two-particle interference is present (see Appendices 8.1 and 8.3 for further details on modelling partially distinguishable photons and discrimination from coherent light).

Even if this method is highly promising, it suffers an important disadvantage: conceptually it differs from the boson sampling. Indeed, in the boson sampling model, the unitary matrix has to be chosen randomly, so that the computational complexity is not reduced. Instead, Fourier matrices has well-defined symmetries and moreover we have still to calculate the probabilities for the unsuppressed states for a large set of input-output combinations, thus giving a not efficient protocol to validate quantum interference. This makes, *de facto*, the problem still unsolved.

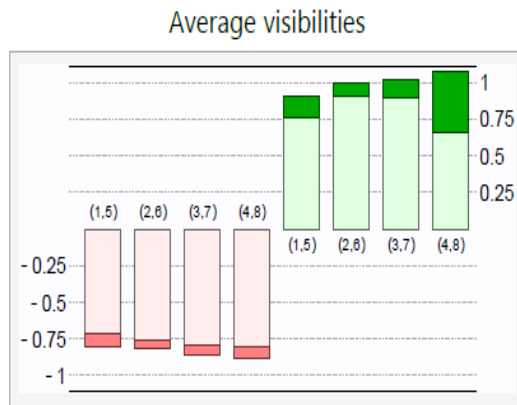
In section 5.5 we will present a new platform which is implemented with the Fast scheme. We will introduce the Sylvester matrix [169] as an optimal platform to ascertain true quantum interference. In the next section, instead, we will provide a further analysis on Fourier matrices using our implementation, focusing on a different perspective which regards the computability.

5.4 Observation of Majorization Principle for quantum algorithms via 3-D integrated photonic circuits

In this section we will discuss of another work carried out during my PhD project where we focused on the computational efficiency of Quantum Fourier Transform (QFT) [45] and validation protocol on Fourier matrices [182, 201]. In fact, quantum computation although holds the promise to greatly improve the capabilities of computational platforms, the unprecedented massive machine resources offered by the parallel processing alone is destined to fail, due to the non-deterministic nature of any measurement process. Follow this road, quantum algorithms have been properly tailored to exploit the power hidden in quantum resources challenging this limitation [204]. In this context, we can yet search for a guiding principle that helps us formulate or discriminate whether a new quantum algorithm is truly efficient. Such criterion may arguably be provided by the Majorization Principle (Majorization Principle (MP)) [205–207]. Intuitively, the Majorization principle says that the notion of quantum parallelism is not a sufficient ingredient to achieve speedup, for which a certain degree of interference is necessary to guarantee that the convergence of the algorithm is reached efficiently. Specifically, MP is believed to provide a



(a)



(b)

Figure 5.5. Suppression law in a 8-mode qFFT integrated chip. (a) Set of 28 measured coincidence patterns (raw experimental data), corresponding to all collision-free output combinations for the input (2,6) of the 8-mode interferometer. For each output combination, the measured coincidence pattern as a function of the time delay is shown (points: experimental data, lines: best-fit curves). Red or green backgrounds correspond to dips and peaks, respectively. For all points, error bars are due to the Poissonian statistics of the events. In each subplot the measured visibility with corresponding error and the sample size are reported. For each visibility, the error is obtained through a Monte Carlo simulation by averaging over 3000 simulated data sets. In each subplot the zero level coincides with the baseline, while a dashed line represents the number of coincidence events in the distinguishable limit. (b) Average visibilities of dips (red bars) and peaks (green bars) observed for the 4 collision-free cyclic inputs [(1,5), (2,6), (3,7), (4,8)]. Darker regions correspond to error bars of ± 1 standard deviation. Image taken from [182].

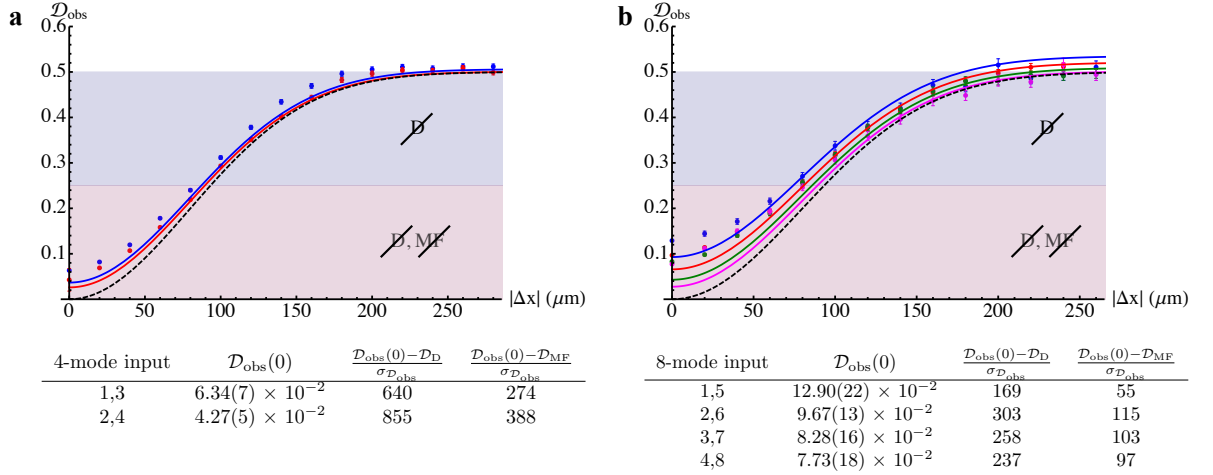


Figure 5.6. Validation with Fourier matrices. Observed violations \mathcal{D}_{obs} as a function of the path difference $|\Delta x| = c|\Delta\tau|$ between the two photons. (a), Data for the 4-mode interferometer. Tables: violations $\mathcal{D}_{\text{obs}}(0)$ at $\Delta x = 0$ and discrepancies (in sigmas) with the expected values for distinguishable particles (\mathcal{D}_{D}) and MFs (\mathcal{D}_{MF}), for the cyclic inputs of the two interferometers. $\mathcal{D}_{\text{obs}}(0)$ are calculated following formula (5.8), while expected values for the other two cases are $\mathcal{D}_{\text{D}} = 0.5$ and $\mathcal{D}_{\text{MF}} = 0.25$. Error bars in all experimental quantities are due to the Poissonian statistics of measured events. All theoretical predictions in solid lines are calculated from the reconstructed unitaries, obtained from different sets of experimental data to ensure statistical independence. Image taken from [182].

necessary condition that must hold to produce an optimal algorithm with a quantum speedup. Moreover, it is considered as part of the general toolbox for the design of quantum algorithms, while the complete set of governing rules remains an open problem.

Here we report on the experimental observation of the Majorization Principle in both a Quantum Fourier Transform (QFT) and a recent quantum validation protocol to certify genuine many-boson interference [201]. In this work it was crucial to exploit femtosecond laser writing technique [24, 157–159] which thanks to its property to design architectures with 3-dimensional topology, it was possible to enable the decomposition in discrete steps [181, 182] in order to observe the protocol.

Majorization Principle

The Majorization Principle lies to the fact that the probability distributions of the outcomes of an algorithm, are updated throughout each step. Given two probability distributions (\vec{x}, \vec{y}) , let $(\vec{x}^\downarrow, \vec{y}^\downarrow)$ be the same vectors with their components sorted in decreasing order. We say that \vec{y}^\downarrow majorizes \vec{x}^\downarrow ($\vec{y}^\downarrow \succ \vec{x}^\downarrow$) if and only if

$$\sum_{i=1}^k x_i^\downarrow \leq \sum_{i=1}^k y_i^\downarrow, \quad \forall k \in \{1, \dots, d\} \quad (5.9)$$

The concept of majorization can be extended in a natural way to quantum algorithms. Let $|\psi^{(s)}\rangle$ represent the state of the register of a quantum computer at a step s of the algorithm. We can associate to $|\psi^{(s)}\rangle$ a vector of probabilities $p^{(s)}$ by writing the register in the computational basis $|j\rangle$, so that $p_j^{(s)} = |\langle j|\psi^{(s)}\rangle|^2$. Consequently, a quantum algorithm is said to undergo a direct (reversed) majorization if and only if $p^{(s)\downarrow} \prec p^{(s+1)\downarrow}$ ($p^{(s)\downarrow} \succ p^{(s+1)\downarrow}$) for all steps s [206]. An intuitive reason for the physical connection between quantum processing and direct (reversed) majorization is that of a neat flux of probability towards (away from) the result of the computation, which makes the distribution steeper (flatter) throughout the algorithm. The principle can now be stated as follows [205]:

Majorization Principle: *In all optimal and efficient quantum algorithms, the set of sorted probabilities associated to the quantum register must obey either a direct or a reverse step-by-step majorization.*

All known quantum algorithms which are both optimal and efficient, *i.e.* with a quantum speedup over the best classical algorithm, have been proven to satisfy the conjectured MP with a direct or reverse majorization [207]. It is worth noting that this validity has already been proved in other algorithms as Grover-like [83] and phase estimation-like [82] as for several ones [207]. The common motivation is that quantum speed-up is always associated to step-by-step majorization, while different story for non-efficient ones.

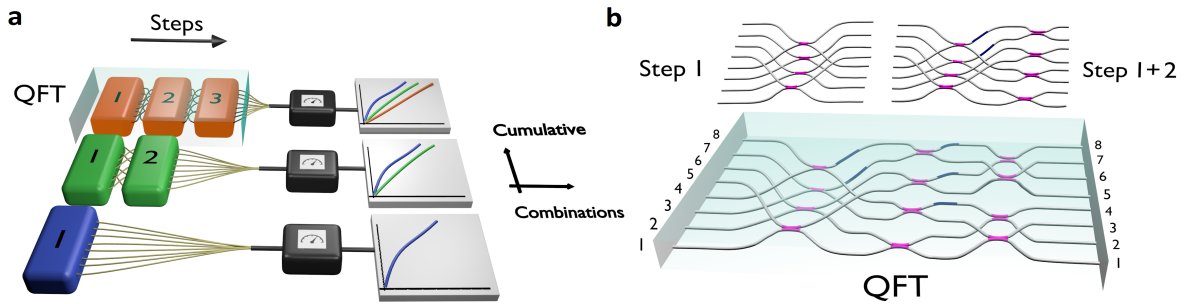


Figure 5.7. Majorization Principle in quantum Fourier transform. (a) Conceptual scheme for the experimental observation of the Majorization Principle in (b) an 8-dimensional quantum Fourier transform: by implementing the QFT with its fast architecture, it is possible to decompose the evolution in three steps, corresponding to the three layers of beam splitters (purple) and phase shifters (blue) between which a step-by-step majorization is observed. Image taken from [170].

Majorization Principle in a fast QFT

Here, to observe the MP we implemented the QFT routine encoding qubits in optical modes of a photonic chip. The transformation has been implemented by adopting the scheme developed by Barak and Ben-Aryeh [181] to reduce the required number of optical elements. As shown in Chapter 3 this scheme represents the quantum analogue of the Fast Fourier Transform (qFFT), the well-known classical algorithm to efficiently calculate the discrete Fourier transform.

We used this scheme not only because it is more efficient than the standard algorithm but also because we are able to observe the reversed majorization step-by-step that naturally emerges from this decomposition, as shown in Fig.5.7. The observation was carried out by injecting single-photon Fock states, generated via a type-II parametric down-conversion source, into three 8-mode integrated interferometers $\{I_1, I_2, I_3\}$ corresponding to partial implementations of the qFFT protocol. The number of fabricated interferometers I_s , each consisting of s layers of beamsplitters and phase shifters, corresponds to the number of layers in the decomposition of an 8-dimensional QFT. The last interferometer I_3 performs the complete 8-mode qFFT, where one photon encodes 3 qubits over the optical modes. The fidelities \mathcal{F}_s between the reconstructed transformations in the I_s and the ideal unitaries obtained with the decomposition are $\mathcal{F}_1 = 0.9954 \pm 0.0002$, $\mathcal{F}_2 = 0.9921 \pm 0.0005$ and $\mathcal{F}_3 = 0.9527 \pm 0.0006$, thus confirming the quality of the fabrication process. For each input state i and each partial transformation I_s , the output probability distributions p_i have been retrieved for the eight output states. The most convenient tool to convey the validity of the Majorization Principle is offered by the Lorenz curves, continuous piecewise linear functions representing the partial cumulatives $C_p(k) = \sum_{i=1}^k p_i^\downarrow$ for the k most probable outcomes. For the MP to be satisfied, the curves C_p at each step of the QFT must not cross, due to the inequality (5.9). As shown in Fig.5.8, a step-by-step reversed majorization is then observed between the output distributions of the three interferometers I_s , *i.e.* by comparing $C_{p(1)}(k)$, $C_{p(2)}(k)$ and $C_{p(3)}(k)$ according to (5.9).

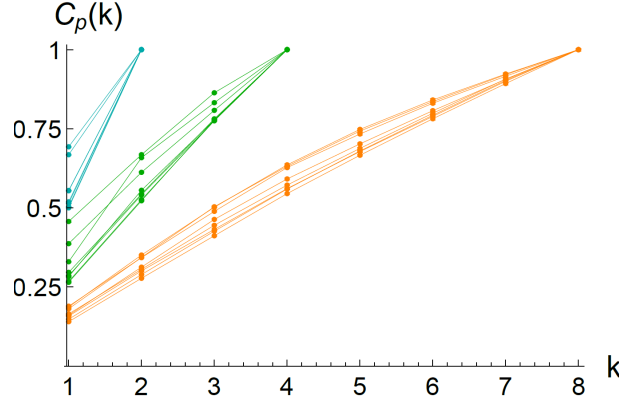


Figure 5.8. Lorenz diagrams for a QFT Majorization experiment, with a single photon encoding three qubit on the first layer I_1 (blue), first and second layer I_2 (green) and complete structure I_3 (orange) of the 8-dimensional Fourier interferometer. For each intermediate I_s , eight diagrams are shown relative to all possible single-photon input states. Each curve represents the partial cumulative probabilities $C_p(k)$ for the k most probable outcomes. Error bars are smaller than the markers. Image taken from [170].

Majorization Principle in a validation protocol

Here we followed the same line to test if the MP is observed also for the validation protocol presented in section 5.3. The efficiency of this validation algorithm are crucial features for the hard-to-simulate regime, hence we expect the MP to be satisfied.

The experimental work was carried out by injecting two-photon states into the three 8-mode integrated interferometers $\{I_1, I_2, I_3\}$ implementing partial instances of the qFFT. While the photonic circuit is still the Fourier interferometer, the computational problem is now fundamentally different from the former. Specifically, the MP in the validation protocol involves a two-photon coalescence mechanism which goes beyond the single-particle evolution of the QFT. Moreover, the very unitary transformation is not the same as the QFT. Now, in fact, the elements of the $\binom{8+2-1}{2} = 36$ -dimensional scattering matrix are the transition amplitudes connecting the two-photon input/output states: the original QFT matrix elements get scrambled by the permanent function and new symmetries appear in this evolution. To observe the MP the whole set of outcomes has to be recorded, which required the measurement of the eight bunching events (i, i) , *i.e.* when photons exit on the same mode. This measurement was carried out by adding, at the end of the fiber array coupled to the output of the interferometer, additional fiber optic splitters to redirect the bunched photons in two separate detectors. A detection system was then able to register all the one-to-one coincidences. For all three I_s , the output distributions $p^{(s)}$ of 4 two-photon input states have been measured and plugged into (5.9) to test the validity of the MP. All 36 patterns for the partial cumulative probabilities $C_p(k)$ can in fact be divided in 4 distinct classes (Fig.5.9). For further details on the analysis performed to these classes and a deeper study on the quality of the protocol for different implementations, refer respectively to Appendices 8.4 and 8.5.

Non-crossing curves are expected when the distribution is the product of two single-photon QFT (Fig.5.9a). We observe non-crossing curves also for non-cyclic input states, which are not employed in the validation protocol (Fig.5.9b-c). This latter observation confirms that majorization does not imply optimality, since the principle does not provide a sufficient condition. Finally, the non-crossing Lorenz curves relative to the cyclic input of the validation protocol (Fig.5.9d), manifesting that $p^{(s)} \succ p^{(s+1)}$ at each stage of the evolution, confirm the operation of the principle along the algorithm.

In the next section we will focus in deeper details in searching for the best platform to test

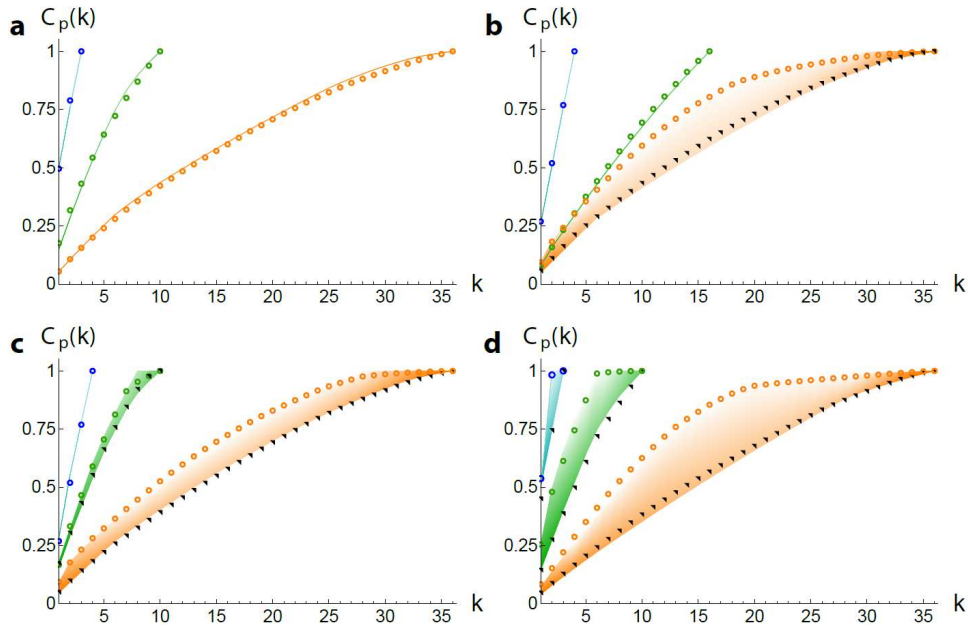


Figure 5.9. Lorenz diagrams for a two-photon Majorization experiment on the first layer I_1 (blue), first and second layer I_2 (green) and complete structure I_3 (orange) of the 8-dimensional Fourier interferometer for the validation algorithm. (a) Two photons in the same input mode (5,5). The distribution is the product of two independent QFT. (b, c) Input modes whose sum is odd (5,6) or even (5,7). The initial crossing between step 2 and 3 in (b) is due to small fabricative imperfections enhanced by the sorting and has no physical consequence for MP in the validation protocol, which works only on the cyclic inputs (d). (d) Cyclic input (2,6) for the validation algorithm. Each curve is obtained by calculating the partial cumulative probabilities $C_p(k)$ for the k most probable outcomes with distinguishable (black triangles) and indistinguishable (circles) photons. Shaded areas are included within the curves corresponding to fully indistinguishable photons (lighter regions) and to fully distinguishable photons (darker regions), as expected from the reconstructed unitary transformations. There is a little offset between shaded regions and experimental points due to the less number of events sorted which imply a larger uncertainty while cumulating. Error bars are smaller than the markers. Image taken from [170].

many-body quantum interference. We will see that we can individuate optimal unitaries which are suitable for this purpose as for example the Sylvester [169] and the Fourier [182] matrices.

5.5 Optimal photonic indistinguishability tests in multimode network

The observation that fundamental particles may be intrinsically indistinguishable is a counter-intuitive feature of quantum mechanics. As a consequence, their collective behaviour is governed by quantum statistics, and results in important phenomena such as Bose-Einstein condensation and the electronic band structure of solids. Moreover, many-body bosonic interference is also at the very heart of post-classical photonic computation [12, 33, 45]. The task of certifying genuine many-boson interference is thus expected to find numerous applications in photonic quantum information, for validating the functioning of Boson Sampling experiments [22, 140, 146, 182, 197, 199–201, 209–213] and, more generally, as a diagnostic tool for quantum optical devices [214, 215]. A well-known approach to experimentally test the degree of indistinguishability between photons involves the use of the Hong-Ou-Mandel effect [14], for which various generalizations have been proposed to account for multiple photon sources [216–221]. A natural problem that arises is then to determine which interferometers allow for an optimal assessment of the indistinguishability of multiphoton states (see Fig. 5.10).

In this section we introduce a new approach for the task of discriminating distinguishable and indistinguishable photon behaviour in multimode photonic networks, and we showed the experimental demonstration with two-photon inputs in laser-written integrated devices with an optimal interferometer design, that we performed in our work². We consider a set-up in which probabilistic single-photon sources are coupled to different sets of input ports of an interferometer, as required for instance in Scattershot Boson Sampling experiments [147, 149, 150]. To assess the optimality of the interferometer we maximize the total variation distance (Total Variation Distance (TVD)) between the output probability distributions corresponding to indistinguishable and distinguishable photons, showing that in certain scenarios the optimal solution is provided by Sylvester interferometers (see Chapter 4 for a full description of this platform). Finally, we perform a Scattershot Boson Sampling experiment in the 4-mode Sylvester device, and verify the capability of the developed approach to validate the collected data with a small number of events.

Sylvester matrices for optimal indistinguishability tests

A linear interferometer is a device that implements a linear map between input and output creation operators, namely: $a_{in_k}^\dagger \rightarrow \sum_j U_{j,k} a_{out_j}^\dagger$. A m -mode interferometer is then completely specified by its corresponding $m \times m$ unitary matrix U , which can be decomposed into two-mode beam splitters and single-mode phase shifters [15]. The evolution of a n -photon state is ruled by the unitary transformation and by the degree of indistinguishability of the input particles. Let $P = \{p_i\}$ ($Q = \{q_i\}$) denote the output probability distribution of an interference experiment when the input photons, all injected in different input modes, are perfectly distinguishable (indistinguishable) and the occupation number is detected at each output port.

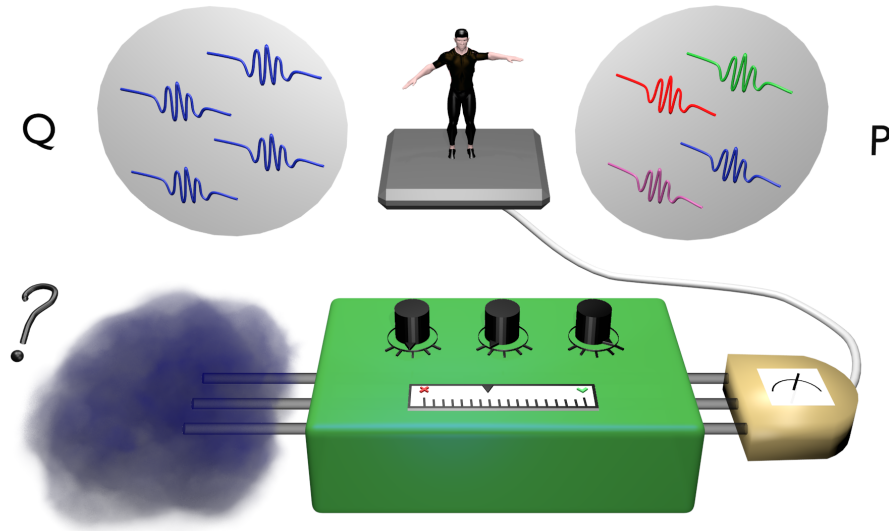


Figure 5.10. Searching for the optimal indistinguishability test. An unknown multiphoton state is injected into a multimode interferometer, designed to discriminate between the two hypotheses: either the photons are indistinguishable (Q) or distinguishable (P). The internal parameters can then be tuned until the most efficient test is found, for which the distributions corresponding to the two hypotheses P and Q are as different as possible, as measured by the total variation distance of the output distributions. Ref. [208]

The total variation distance (TVD) is then defined as

$$T(P, Q) = \frac{1}{2} \sum_i |p_i - q_i|. \quad (5.10)$$

²This work is currently in submission to Optica

The TVD corresponds to the highest possible difference in probabilities that P and Q can assign to the same event. It is also closely related to hypothesis testing, since $1 - T(P, Q)$ is a lower bound on the sum of probabilities of false positive and false negative results [222]. More practically, a larger value of $T(P, Q)$ means that a smaller set of experimental data is sufficient to discriminate between the hypotheses that one is sampling from P or Q up to a given level of confidence.

We consider two scenarios depending on how the input photons enter the interferometer: (i) the input ports are fixed (see Fig. 5.11a) and (ii) the input combination is randomly chosen for each event (see Fig. 5.11b). This second situation is typical of more sophisticated interferometric experiments with multiple probabilistic sources, such as scattershot Boson Sampling experiments [147, 149]. To search for the optimal platform, a natural symmetry we can impose on the interferometer is that the output probability for single photons should be uniform for all input ports. Mathematically, this corresponds to a unitary matrix U with complex entries of modulus $1/d$, known as a Hadamard matrix. A specific subcase is provided by Sylvester transformations [169], described by $m = 2^p$ -dimensional unitary matrices defined by the recursive rule:

$$S(2^p) = \begin{pmatrix} S(2^{p-1}) & S(2^{p-1}) \\ S(2^{p-1}) & -S(2^{p-1}) \end{pmatrix} \quad (5.11)$$

with $S(2^0) = S(1) = (1)$ and p any positive integer. Another notable example of Hadamard matrix is the Fourier one.

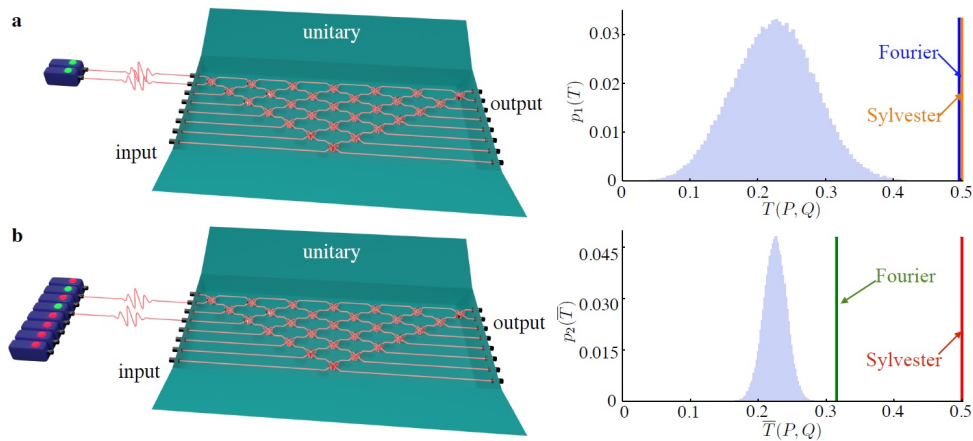


Figure 5.11. Optimal tests for fixed-input and scattershot configurations. **a**, Left figure: fixed input configuration in the $n = 2$, $m = 8$ case, where two fixed input modes are injected with single photons. Right figure: estimate of the distribution $p_1(T)$ of the total variation distance $T(P, Q)$ for uniformly random unitaries in this scenario, obtained numerically by sampling over 10^5 unitaries. For specific choices of input modes, both Fourier (blue vertical bar) and Sylvester (orange vertical bar) interferometers are optimal. **b**, Left figure: scattershot configuration in the $n = 2$, $m = 8$ case, where different sets of two input modes are picked at random. Right figure: estimate of the distribution $p_2(\bar{T})$ of the average total variation distance $\bar{T}(P, Q)$ for uniformly random unitaries in this scenario, obtained numerically by sampling over 10^5 unitaries. In the scattershot case the Sylvester transformation (red vertical bar) is optimal, while the Fourier is not. Ref. [208].

For two-photon experiments, we can apply (5.10) to check that the Sylvester transformation results in a $T_S = 0.5$ independently of which pair of input modes is used, and thus on average (over different pairs of inputs) $\bar{T}_S = 0.5$. As we show later, the Sylvester interferometer is optimal for all possible choices of two-photon input configurations. The Fourier interferometer results instead in a lower average $\bar{T}_F = 0.333$, reaching the highest value $T_F = 0.5$ only for cyclic inputs, i.e. n -photon Fock states in interferometers with n^p modes, where the occupied modes j_r^s are given by $j_r^s = s + (r - 1)n^{p-1}$, with $r = 1, \dots, n$ and $s = 1, \dots, n^{p-1}$. For two-photon experiments in interferometers with $m = 8$ modes, among all known Hadamard matrices the

Sylvester transformation is again the one with the largest average TVD $\bar{T}_S = 0.5$ ($T_S = 0.5$, and thus optimal, for all pairs of inputs). The Fourier interferometer reaches $T_F = 0.5$ for cyclic inputs, while on average we obtain only $\bar{T}_F = 0.3153$ due to inputs with lower TVD. Even if we consider general unitary matrices, for two-photon experiments in $m = 4$ and $m = 8$ modes a random sampling of 10^5 matrices of each size fails to find matrices outperforming the Sylvester matrices, which provides evidence for the optimality of Sylvester interferometers. In the $n = 2$ $m = 4$ case, the optimality can be proved also via a numerical maximization of the TVD over the parameters of general 4-mode interferometers.

Interestingly, we found that for scenarios with different number of photons n and modes m , the optimal interferometers may not even be Hadamard. As an example, for $n = 3$, $m = 4$, the best design we found corresponds to the unitary matrix

$$U_4 = \frac{1}{2} \begin{pmatrix} 1 & 1 & \sqrt{2} & 0 \\ 1 & -1 & 0 & \sqrt{2} \\ 1 & 1 & -\sqrt{2} & 0 \\ 1 & -1 & 0 & -\sqrt{2} \end{pmatrix}, \quad (5.12)$$

with average (over all collision-free inputs) $\bar{T}_{U_4} = 0.53125$, beating for example the Fourier and Sylvester interferometers, which result in $\bar{T} = 0.3125$. In Appendix 4.3 we list the best interferometers we found numerically for a few different scenarios.

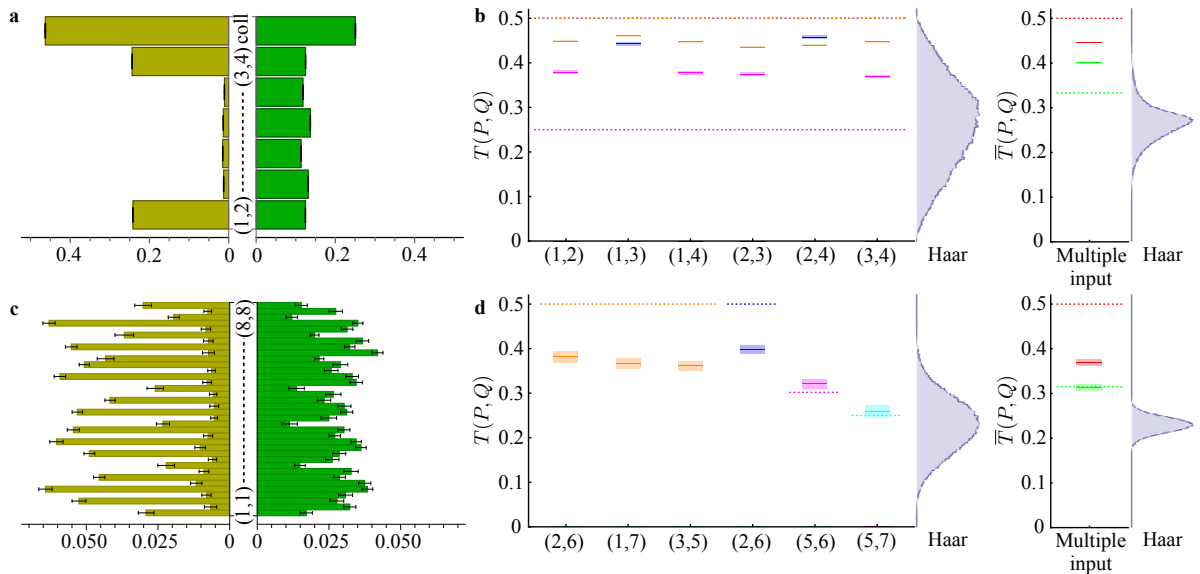


Figure 5.12. Experimental data from Sylvester interferometers and comparison with Fourier ones. Two-photon experimental data from 4-mode and 8-mode Sylvester interferometer. **a,c**, Output probability distributions for **(a)** input (1,2) of the 4-mode Sylvester interferometer and **(c)** input (1,5) of the 8-mode interferometer. Each plot compares the distributions obtained from indistinguishable (dark yellow) and distinguishable (green) input photons. **b,d**, Total variation distances between the measured output distributions for indistinguishable and distinguishable photons in the 4-mode **(b)** and 8-mode **(d)** interferometers, plotted for each fixed input (left) and for the multiple input configuration (right). Orange-Red: Sylvester. Other colour: Fourier. Ref. [208]

Here we report the experimental demonstration of optimal distinguishability tests, which we carried out using two custom-designed integrated photonic circuits fabricated on a glass substrate via the femtosecond laser writing technology [24–26]. A possible approach to the implementation of Fourier and Sylvester matrices on photonic platforms could exploit the decompositions of Refs. [15, 178], which allow to implement arbitrary unitary operations through an in-plane cascade of beam-splitters and phase shifters but are sensitive against losses inside the device. A more effective approach for the implementation of this class of interferometers is found by

adopting an efficient, scalable and reliable three-dimensional interferometer design (as shown in Chapter 4) enabled by the interaction between non-first neighbour modes arranged in a suitable lattice [181] (Fig. 4.4). This architecture presents several advantages with respect to the conventional decomposition in beam splitters and phase shifters [15, 178]. Indeed, this approach is efficient since it allows to significantly reduce the number of optical elements from $O(m^2)$ to $O(m \log m)$. Furthermore, as shown in Appendices (8.7), it is scalable to a larger number of modes. Finally, it is also reliable since the depth of the circuit is small [$O(\log m)$] and the layout is symmetric with respect to the input modes, thus being intrinsically more robust against internal losses than the decomposition of Ref. [15]. These features permit to achieve higher fidelities in the implemented devices.

To probe the device for the reconstruction of its internal operation, we injected the interferometers with one- and two-photon states produced via a spontaneous parametric down-conversion process [13]. The tomographic reconstruction of the processes exploits a priori knowledge of the internal structure, minimizing a suitable χ^2 function with respect to the unknown internal phases and transmissivities [182]. In section 4.3 we provided more details about the design and the tomographic reconstruction. All tests have been carried out also on Fourier interferometers of equivalent dimensions ($m = 4, 8$), to provide a comparison between the two performances and highlight the advantages of the new approach. For more details on the implementation of the Fourier transforms refer to Ref. [182] or Appendix 8.2.

Measurement of the total variation distance

The TVD between distributions obtained from distinguishable and indistinguishable photons has been estimated experimentally for two-photon experiments in 4-mode and 8-mode Sylvester and Fourier interferometers. As previously discussed, a distinctive feature of Sylvester interferometers is that the TVD has the same value for two-photon experiments using any pair of input ports. This value is never outperformed by Fourier interferometers, being matched only for cyclic inputs.

In the case of the 4-mode Sylvester and Fourier interferometers, all $\binom{4}{2} = 6$ possible pairs of different input ports have been injected, measuring for each pair all the collision-free output distributions (that is, events with at most one photon per output port). Additionally, we included an extra bin in the distribution accounting for the probability of collision events (see Fig. 5.12a). For distinguishable photons, this collision probability can be calculated using the measured probabilities for single-photon experiments. For indistinguishable photons, the additional information provided by the measured two-photon Hong-Ou-Mandel visibilities, together with the single-photon data, allows for an estimate of the collision probability. The adoption of this extra bin prevents the appearance of pathological unitaries in our numerical search, with high values of TVD for the collision-free subspace but unacceptably small probability of collision-free events, which strongly reduce the amount of detected signal (see Appendices for more details). Adjusting the distinguishability of the input photons we could then estimate the TVD according to (5.10). The experimental results are reported in Fig. 5.12b. In the fixed input case, the Sylvester interferometer (for all inputs) and the Fourier one (only for cyclic inputs) reaches similar values within the range $0.435 \leq T^{(4)} \leq 0.461$. When averaging over all inputs, the 4-mode Sylvester interferometer results in $\bar{T}_S^{(4)} = 0.4465 \pm 0.0006$, thus clearly outperforming the Fourier interferometer that yields $\bar{T}_F^{(4)} = 0.4003 \pm 0.0016$. In the case of the 8-mode interferometers, all two-photon output events were measured, including those with collisions (see Fig. 5.12c). We probed the interferometers with three different pairs of input ports, corresponding to the three classes of input pairs that, for the Fourier interferometer, result in different TVDs. Again, the TVD for the Fourier interferometer reaches the same value of the Sylvester only for cyclic input pairs. The measured average values were $\bar{T}_F^{(8)} = 0.314 \pm 0.007$ and $\bar{T}_S^{(8)} = 0.370 \pm 0.007$, where the average $\bar{T}_F^{(8)}$ was estimated by weighting the three representative input states by the multiplicity of the corresponding classes with different TVDs (Fig. 5.12d).

Curiously, the measured $\overline{T}_F^{(4)}$ and $\overline{T}_F^{(8)}$ are higher than what was expected theoretically. Indeed, as we show in Appendix 8.7, a particular feature of the Fourier interferometer design is that manufacturing errors result in interferometers which are closer to the Sylvester one, thus increasing the TVD. Our results confirm the theoretical prediction that Sylvester interferometers provide optimal discrimination between distinguishable and indistinguishable photons in all scenarios (fixed and multiple input), outperforming, in particular, Fourier interferometers in the multiple input configuration.

Bayesian hypothesis test

A larger TVD should allow us to discriminate more readily the two hypotheses (distinguishable or indistinguishable photons). To illustrate this feature, in Fig. 5.13 we show the results of a Bayesian analysis with data samples numerically generated from the measured distributions, compared with those expected from ideal interferometers (see Ref. [211], and Appendices 8.9 for more details on the Bayesian validation test). For a given sample size, we use a likelihood ratio test [211] to update a prior which initially assigns equal probabilities to the two hypotheses P and Q , thus assuming no a priori knowledge. To verify the optimality of Sylvester interferometers, the results obtained are compared with a sample of 10^4 Haar-random unitaries. The figure of merit is the confidence probability P_{conf} that a data sample is assigned to the corresponding correct hypothesis.

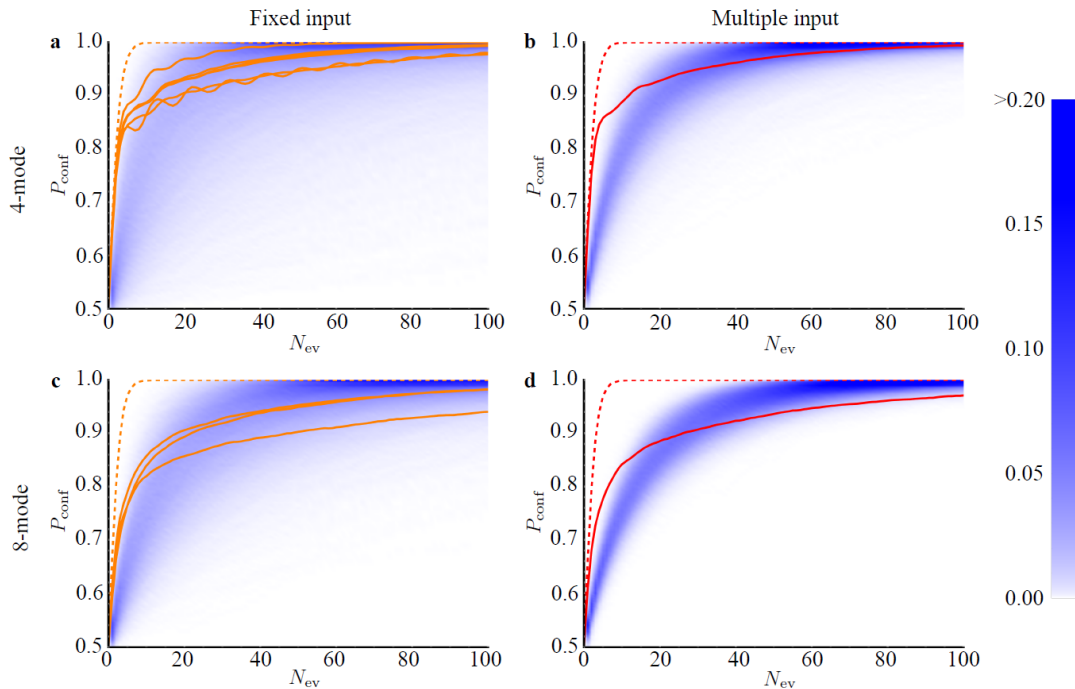


Figure 5.13. Bayesian simulation of hypothesis testing. Probability P_{conf} of correctly identifying the type of data sample (distinguishable or indistinguishable photons) in a Bayesian test, as a function of sample size (number of events), for the Sylvester interferometer. **a**, 4-mode interferometer, fixed input states. **b**, 4-mode interferometer, multiple input configuration. **c**, 8-mode interferometer, fixed input states. **d**, 8-mode interferometer, multiple input configuration. For the 8-mode devices, inputs used were (2,6), (1,7), (3,5). Dashed curves: theoretical predictions. Solid curves: average of 10^4 distinct samples generated numerically from the experimentally measured distributions. Red curves: Sylvester matrix, multiple input configuration. Orange curves: Sylvester matrix, different input states. Blue regions: contour plots obtained from a numerical simulation over 10^4 Haar-random matrices. Simulated data for the Haar-random case include partial photon indistinguishability $x = 0.95$ equal to the experimental data, characterized by performing Hong-Ou-Mandel interference in a 50/50 beam-splitter. Ref. [208].

Note that the dispersion of P_{conf} for Haar matrices in the fixed input case is larger than in the multiple input configuration. The analysis performed on the experimental data shows that, for the tested transformations, very small sample sizes suffice to identify the correct hypothesis with a high probability of success. Numerical simulations performed on the Bayesian test show that in presence of partial photon indistinguishability, quantified by a parameter $x \in [0; 1]$, the photons to pass the test if $x > 0.788$ for the 4-mode interferometer and $x > 0.685$ for the 8-mode case (see Appendices 8.8 and 8.9).

Experimental assessment of Scattershot Boson Sampling

Scattershot Boson Sampling has been shown [149] to enable scalable Boson Sampling experiments using probabilistic single-photon sources, at the cost of using a larger number of single-photon sources at the input. To show our approach works well for this multiple-input configuration, we have performed a Scattershot Boson Sampling experiment by connecting each input mode of our 4-mode Sylvester interferometer to an independent heralded parametric down-conversion source (see Fig. 5.14a). This single setup simultaneously samples all 6 different two-photon input states. We have collected data both for the indistinguishable and distinguishable photon cases, obtained by appropriately adjusting the relative time delay between the input paths. We have then performed a step forward with respect to the binary Bayesian test applied previously (that is, evaluating which is the most likely hypothesis between Q and P). More specifically, we consider a different scenario where the aim is to assess the output data by estimating the value of the mutual indistinguishability of the SPDC sources in the scattershot experiment. The set of hypotheses to be tested is now provided by a convex combination of indistinguishable Q (distinguishable P) photons $H(x) = xQ + (1-x)P$. As previously introduced, $x \in [0; 1]$ is a continuous real parameter quantifying the degree of indistinguishability between the photons generated by different sources. We have then applied Bayesian inference to the measured experimental data to estimate the value of x , starting from a uniform prior $\mathcal{P}(x)$ and subsequently updating the distribution according to the Bayes rule after the observation of a given data sample (see Appendix 8.11 for more details on this approach).

Figs. 5.14b-c show the results of the Bayesian inference performed by using the measured ~ 17000 experimental Scattershot data samples (in this case, no extra bin for collision events has been included). The final estimated value \tilde{x}_{est} of the indistinguishability parameter x , obtained by using the complete set of experimental data, is $\tilde{x}_{\text{est}} = 0.738 \pm 0.004$. Such value is compatible with the one obtained from a characterization with Hong-Ou-Mandel interference in a 50/50 beam-splitter $x_{\text{HOM}} = 0.79 \pm 0.06$, and with the estimated interval obtained with an alternative method based on binary likelihood ratio tests $x_{\text{LR}} \in [0.734; 0.742]$ (see Appendices 8.10 and 8.11).

The performance obtained with the experimental data on Sylvester interferometer are then compared with the one achievable with Haar random transformations, showing the optimality of the Sylvester matrix. Indeed, the value of the estimation error σ_{est} is lower with respect to the Haar ensemble (for the same number of measured events N_{ev}). This confirms that Sylvester interferometers represent a promising platform for the assessment of Scattershot Boson Sampling experiments.

Discussion

In summary, in this Article we propose a method to search for optimal interferometer designs to ascertain photon indistinguishability in multiparticle interference experiments, showing that in certain scenarios Sylvester transformations offer an optimal solution for all input configurations. We then experimentally implement these interferometers for the first time by exploiting a novel 3D-architecture with the femtosecond laser-writing technology. We verify the superior performance of Sylvester interferometers with a Bayesian hypothesis test which correctly identifies

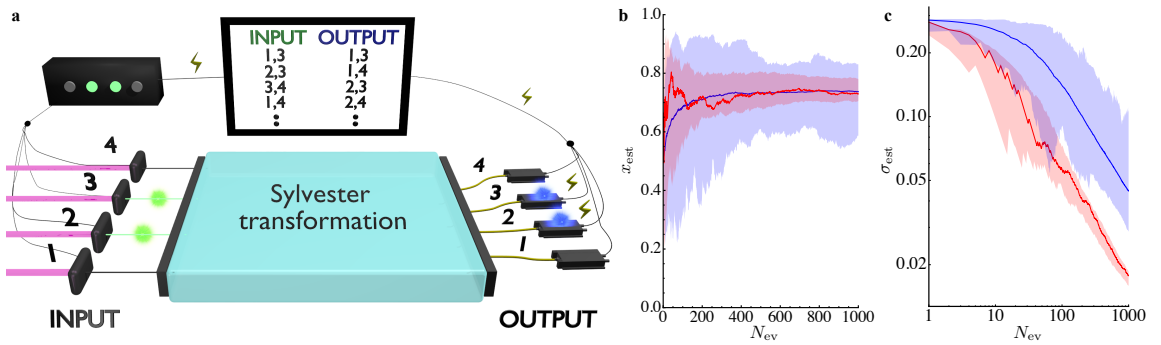


Figure 5.14. Bayesian inference on experimental Scattershot boson sampling. Inference of the photon indistinguishability x as a function of sample size (number of events) from two-photon Scattershot Boson Sampling experimental data in a 4-mode Sylvester chip. **a**, Experimental scheme of the Scattershot approach. We use 4-PDC sources to probabilistically inject a different input state, which is retrieved from heralded 4-fold coincidences (two trigger and two output detectors). **b**, Evolution of the estimated value x_{est} as a function of the number of events and **c**, error σ_{est} inferred on the estimated value x_{est} from the Bayesian posterior distribution with the 4-mode interferometer (experimental data) as a function of the number of events N_{ev} , compared with Haar-random matrices (numerical simulations). Red solid line: single data sequence for the Sylvester interferometer. Red shaded region: interval included within $M = 100$ different data sequences generated by random permutations of the experimentally measured data. Blue solid line: average value obtained by numerically sampled $M = 100$ Haar-random unitary. Blue shaded region: interval obtained by numerically sampling $M = 100$ Haar-random matrices. The value of x adopted for the Haar-random numerical simulations is equal to the final value \tilde{x}_{est} estimated from the full set of experimental data. Ref. [208].

whether photons are distinguishable or not using very small data sets. Furthermore, we perform a complete Scattershot Boson Sampling experiment with the implemented Sylvester device, showing the capability of identifying the collected data in this more complex multiple input configuration. Several perspectives can be envisaged starting from these results.

On the experimental side, we show that the optimality of Sylvester interferometers suggests an immediate application in the assessment of scattershot Boson Sampling, that enables scalable experiments even with probabilistic single-photon sources. Furthermore, this class of transformations can represent a promising platform for near-future investigations on deterministic single photon sources [143, 223], in order to fully characterize multiphoton interference [142, 224], and as a diagnostic tool in reconfigurable photonic devices [22, 23].

On the theoretical side, the identification of optimal interferometer designs for different interferometer sizes and number of photons is a non trivial and interesting open problem. Different scenarios occur in the fixed input and in the Scattershot cases. Based on numerical simulations we can extrapolate that for a single input state, the best unitaries can be obtained by embedding a $n \times n$ unitary matrix in a larger $m \times m$ one, acting as the identity on the unoccupied input modes. Conversely, such operational approach does not hold in the multiple input case, where finding the best unitary requires a global optimization. Hence, while Sylvester matrices are optimal for certain sizes, further investigations are necessary to understand the general structures underlying the optimal unitaries for all (n, m) .

5.6 Suppression law with Sylvester interferometers

The search for the optimal unitary transformation brings us to address the Sylvester as one of the most promising. Following this road, here we introduce³ a generalized law which permits to predict a larger number of suppressed input-output combinations, further providing a sufficient

³This work is online as a pre-print in [193]

condition. Let us write an n -photon state $\mathbf{r} = (r_1, r_2, \dots, r_n)$ evolving in a 2^p -dimensional Sylvester interferometer as an $n \times p$ binary matrix R , whose n rows are the binary representations of $r_1 - 1, r_2 - 1, \dots, r_n - 1$. We then identify with $N^A(\mathbf{r})$ the matrix obtained by negating the columns of R specified by a list of indexes A . The input-output combinations $\mathbf{r} \rightarrow \mathbf{s}$ and $\mathbf{s} \rightarrow \mathbf{r}$ are suppressed if some A exists for which both the following conditions are satisfied:

$$N^A(\mathbf{r}) = \mathbf{r} \quad (5.13)$$

$$\bigoplus_{k=1}^n \bigoplus_{\alpha \in A} S_{k,\alpha} = 1 \quad (5.14)$$

where \oplus denotes the bitwise product, and S is the binary matrix representation of the state \mathbf{s} (for further details of the theoretical model, refer to Appendix 8.17). This new law results in a much higher fraction of efficiently-predictable suppressed input-output combinations with respect to Fourier matrices. More importantly, for fixed m it results in the same large fraction of suppressed output states for every possible two-photon input combination. This feature is of particular interest for Scattershot Boson Sampling experiments, where the input state is injected randomly for each event.

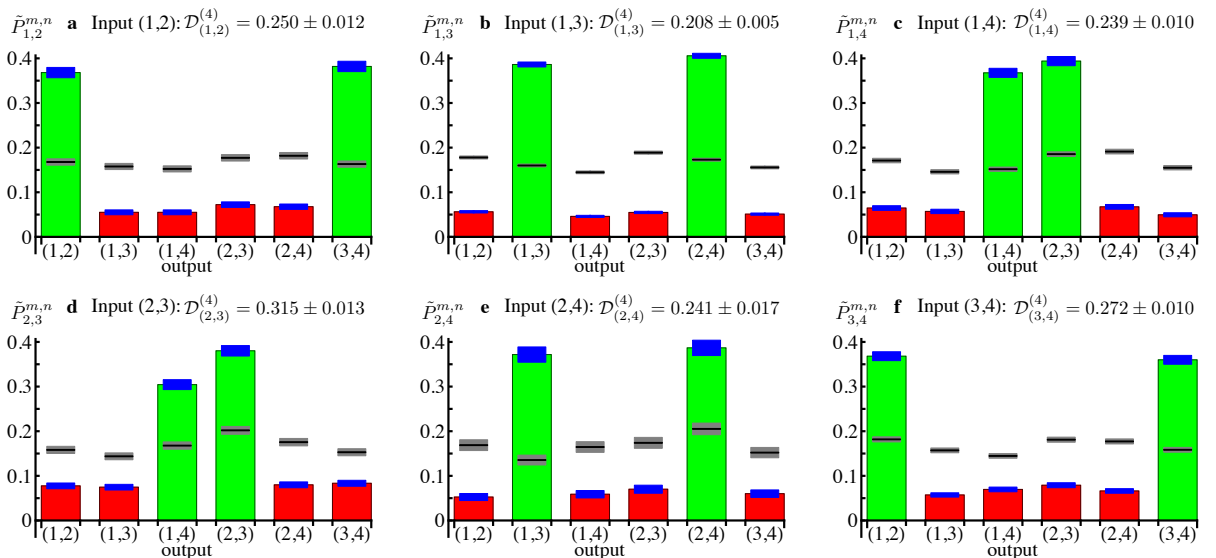


Figure 5.15. Experimental data with 4-mode device. Measured two-photon probability distributions $\tilde{P}_{i,j}^{m,n}$ in the Scattershot regime for all input-output combinations, being (i, j) the input modes and (m, n) the output modes. Contributions which are forbidden by the suppression law are shown in red, while non-suppressed ones are shown in green. Error bars (represented as blue rectangles) are due to Poissonian uncertainties. Black horizontal segments: measured values with distinguishable particles, with corresponding error bars (grey rectangles). On each subfigure the single violations $D_{(i,j)}^{(4)}$ for each input state are reported. Ref. [193].

Consequently, we performed a two-photon Scattershot Boson Sampling experiment [149] to verify the suppression law, by feeding a 4-mode integrated interferometer with one heralded photon from each SPDC pair. The output events corresponding to two-photon injection were then post-selected via four-fold coincidences measurements (two heralding detectors and two detectors at the output of the device) for all the $\binom{4}{2}$ input-output combinations with four SPDC sources. The experimental setup adopted for the Scattershot approach is shown in Fig. 4.4. Further details on the generation and detection are reported in the Methods section.

The observed two-photon output statistics are shown in Fig. 5.15 where it is visible the pattern configuration of peaks and dips described by the Sylvester matrix (U_S^4). An effective figure of merit to estimate if the observed data are compatible with those expected from a fully interfering multiphoton source is the degree of violation [182, 201] $\nu = N_{\text{forbidden}}/N_{\text{events}}$, i.e.

the ratio between the number of observed events in suppressed input-output combinations and the total number of events. Evaluating the observed violation permits to rule out alternative hypotheses, on the nature of the injected state. The fact that suppression laws are only partially

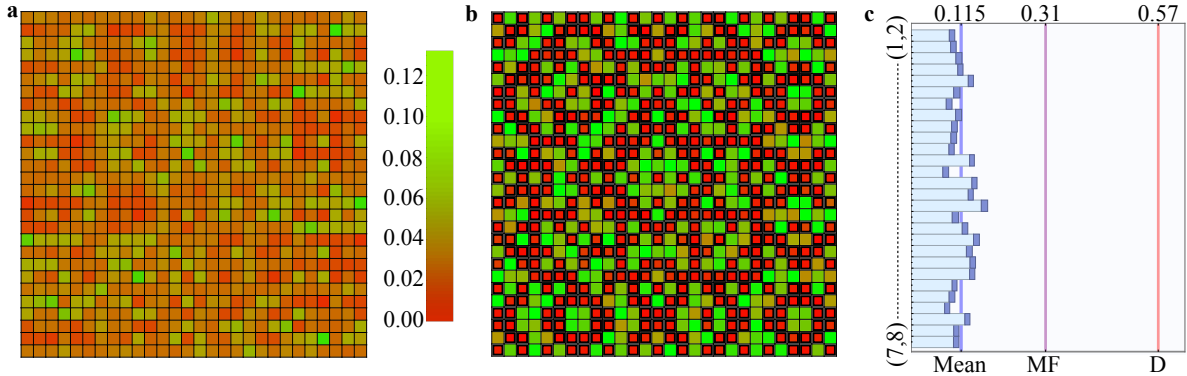


Figure 5.16. Suppression law on an 8-mode Sylvester interferometer. **a**, Two-photon probability distributions in the distinguishable particle regime. **b**, Two-photon probability distributions in the indistinguishable particles regime calculated through the **HOM** visibilities measured shot by shot for all input-output combinations. Combinations forbidden by the suppression law are shown in red, while non-suppressed ones are shown in green. **c**, Measured violations for all the input combinations. Cyan bars are the observed violations, with corresponding error bars in blue. Blue line: mean over all violations. Purple line: Mean Field (MF) threshold. Red line: distinguishable particles (D) threshold. Ref. [193].

fulfilled from statistics coming from Mean Field states, confirms the diagnostic power of these tools. In particular, one can show that the expected degree of violation for distinguishable particles and **MF** states are respectively 0.66 and 0.4 (considering only collision-free two-photon inputs and outputs). The overall experimental violation ν , obtained by summing all the measured events over all possible input-output combination, is found to be $\nu^{(4)} = 0.238 \pm 0.003$. The measured value that are below the two thresholds of respectively 141 (distinguishable) and 54 (**MF**) standard deviations, can thus unambiguously exclude both these hypotheses (Fig. 5.15). Indeed, the experimental values of \mathcal{D} are below the value 0.4 for the **MF** state of at least 6 standard deviations σ for all inputs, and below the threshold 0.66 for distinguishable particles of at least 26σ . The overall violation \mathcal{D} , by summing over all possible input-output combination, is $\mathcal{D} = 0.238 \pm 0.003$.

The suppression law has been experimentally verified also with a Sylvester 8-mode chip, where the full set of $\binom{8}{2} = 28$ two-photon input states has been investigated separately. In Fig. 5.16 we report the full set of 28×28 probabilities, retrieved from the **HOM** visibilities measured for all input-output combinations. The figure shows a good agreement with the expected ideal pattern of suppressions showed in Fig. 5.10. In this case, one can show that the theoretical degree of violations for distinguishable particles and **MF** states are respectively 0.57 and 0.31. Measuring all two-photon coincidences for the Hilbert space with no-bunching contributions (no-collision events) we are able to retrieve the degrees of violation for all input states. The mean value of the violations is $\bar{\nu}^{(8)} = 0.115 \pm 0.002$ which is well below the theoretical bounds predicted for **MF** and distinguishable particles. This provide us a good indication that it is possible to rule out these alternative hypothesis.

Discussion

In this work we introduced and verified experimentally a novel suppression law occurring in Sylvester interferometers with indistinguishable photon inputs. This choice of the unitary transformation ensures a large fraction of suppressed input-output combinations, even for large numbers of photons and modes (see Fig. 5.17). The emergence of forbidden combinations in

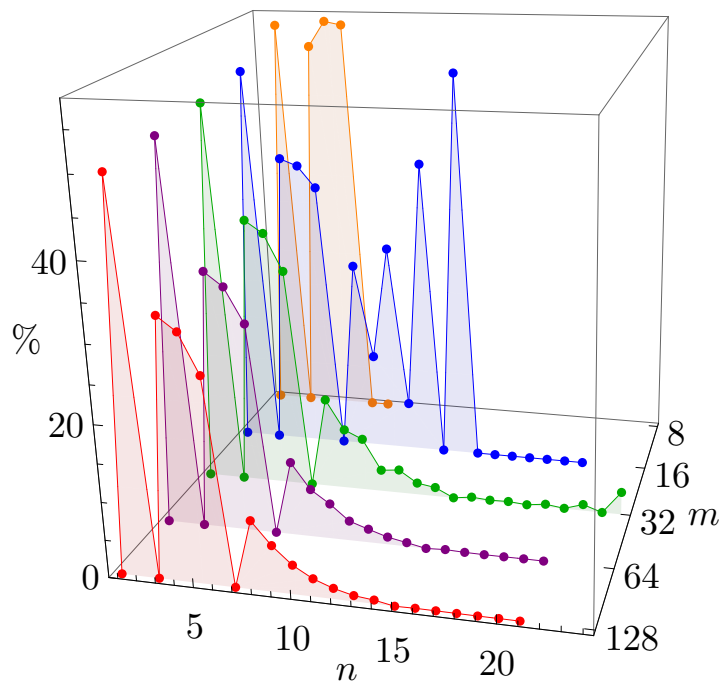


Figure 5.17. Analysis of suppressed input-output combinations. Fraction of suppressed input/output pairs, among the total number of $\binom{m}{n}^2$ pairs of collision-free input/output states for various values of the number of modes m and the number of photons n . The number of forbidden combinations have been evaluated by a numerical approach in order to evaluate the complete landscape.

this scenario can be exploited to certify genuine multiphoton interference in a Scattershot Boson Sampling experiment, where the full set of possible states is employed. Indeed, random input generation has been recognized as a promising approach to greatly enhance the events rate in SPDC-based Boson Sampling experiments. The protocol based on the suppression law is able to rule out efficiently-computable alternative models which show coarse-grained interference effects, thus providing a mid-term large scale validation for Boson Sampling experiments.

Chapter 6

Exploiting machine learning techniques for validation protocols

In this Chapter we introduce a novel approach to the validation problem. Rather than building our tests upon our limited physical intuition of Boson Sampling distributions, we propose using machine learning and statistical tools to teach computers to identify true photon interference. In the following we report the first work based on pattern recognition and clustering techniques and the second one which exploits correlations between output modes and statistical tools to classify the physical behaviours.

These works are reported in the following papers, where I contributed mainly in carrying out the simulations in the first work, while contributing in the experimental realization and in data analysis in the second.

- I. Agresti, N. Viggianiello, F. Flamini, N. Spagnolo, A. Crespi, R. Osellame, N. Wiebe and F. Sciarrino. *Pattern recognition techniques for Boson Sampling validation, preprint at arXiv:1712.06863*, (2017).
- T. Giordani, F. Flamini, M. Pompili, N. Viggianiello, N. Spagnolo, A. Crespi, R. Osellame, N. Wiebe, M. Walschaers, A. Buchleitner and F. Sciarrino. *Experimental statistical signature of many-body quantum interference*, *Nat. Photon.*, **12**, 173–178 (2018).

6.1 Pattern recognition techniques for Boson Sampling validation

In this section we will introduce the first approach that we exploited, which focus on validate true photon interference with pattern recognition techniques. Our approach is based on the theoretical proposal of Wang and Duan [225], and expands the idea behind, exploiting also to machine learning techniques to test the output data. The base model consists in a compatibility test, based on data clustering, performed on experimental sampled data drawn from an untrusted Boson Sampler as well as a trusted device . More specifically, we will tailor the technique to operate in an experimental regime available with current technology ($N \sim 3$, $m \sim 10$ and collision-free case), preserving its robustness also when applied to larger Hilbert spaces. This method relies only on experimental data, not requiring any permanent evaluation, and its goal is to recognize whether two samples drawn from different devices belong the same statistical population or not. Such approach can be applied to validate an unknown device in the scenario when a reliable boson sampler is available. We tested this model using numerical simulations both for small-scale Boson Sampling and Scattershot experiments. We find that, even when tested on Boson Samplers the model had never seen, the model accurately predicts the correct result whereas prior methods did not.

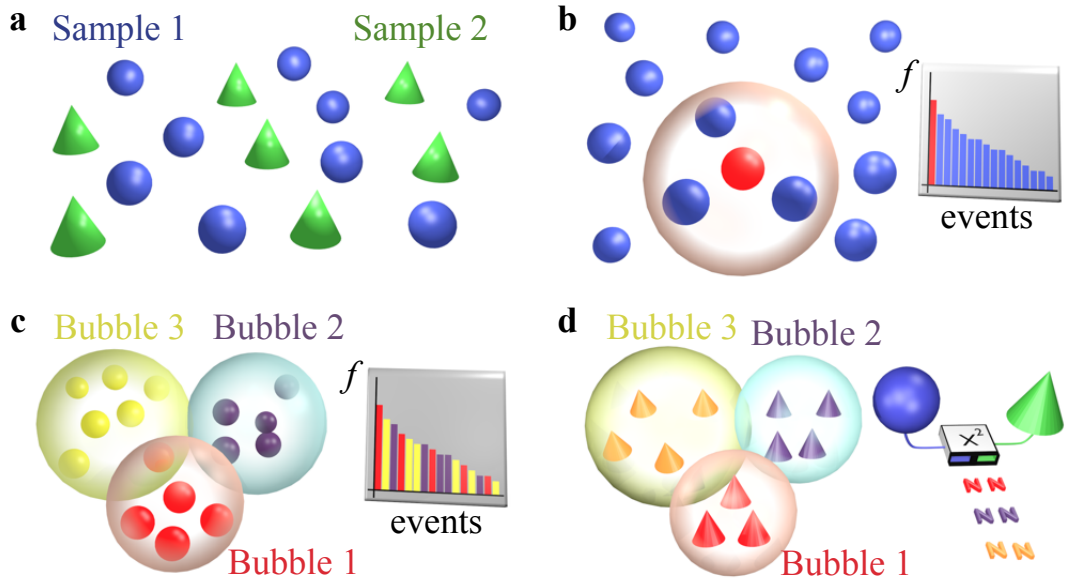


Figure 6.1. Bubble clustering validation scheme. **a**, A sample is drawn from each of the two Boson Samplers to be compared. **b**, The events belonging to one of the two samples are sorted according to their observation frequency. The state with highest frequency is chosen as the center of the first cluster. Those events with distance d from the center smaller than a cut-off radius ρ_1 are included in the cluster. **c**, Starting from the unassigned events, this procedure is iterated until all of the observed events are included in some bubble. At this point, each cluster is characterized by a center and a radius. **d**, The observed events belonging to the second sample are classified by using the structure tailored from the first sample: each event belongs to the cluster with the nearest center. A χ^2 test with $\nu = N_{\text{bubbles}} - 1$ degrees of freedom is performed to compare the number of events belonging to the first and second sample by using the obtained cluster structure. This variable quantifies the compatibility between the samples.

In the regime where a Boson Sampling device is expected to outperform its classical counterpart, the validation problem has inherently to deal with the exponential growth of the number of input/output combinations. A promising approach in this context is provided by the field of machine learning, which studies how a computer can acquire information from input data and learn throughout the process how to make data-driven predictions or decisions [226]. Significant progresses have been achieved in this area over the past few years [227, 228]. One of its main branches is represented by unsupervised machine learning, where dedicated algorithms have to find an inner structure in an unknown data set. The main unsupervised learning approach is clustering, where data are grouped in different classes according to collective properties recognized by the algorithm. Several clustering methods are widely employed [40], such as *K-means* and *Hierarchical clustering*. Being these approaches designed to identify hidden patterns in a large amount of data, they are promising candidates to be applied for the Boson Sampling validation problem.

Let us discuss the general scheme of the proposed validation method based on pattern recognition techniques. Given two samples obtained respectively from a Boson Sampler, that is a trusted device, and a Boson Sampler to be validated, the sequence of operations consists in (i) finding a cluster structure inside the data belonging to the first sample, (ii) once the structure is completed, organizing the data of the second sample by following the same structure of the previous set, and (iii) performing a χ^2 test on the number of events per cluster for the two independent samples. The χ^2 variable is evaluated as $\chi^2 = \sum_{i=1}^N \sum_{j=1}^2 \frac{(n_{ij} - E_{ij})^2}{E_{ij}}$, where index j refers to the samples, index i to the clusters, n_{ij} is the number of events in the i -th cluster belonging to the j -th sample and E_{ij} is the expected value of observed events belonging to the j -th sample in the i -th cluster $E_{ij} = n_i n_j / n$, with $n_i = \sum_{j=1}^2 n_{ij}$, $n_j = \sum_{i=1}^N n_{ij}$ and

$n = \sum_{i=1}^N \sum_{j=1}^2 n_{ij}$. If the null hypothesis of the two samples being drawn from the same probability distribution is correct, the evaluated variable must follow a χ^2 -distribution with $\nu = N_{\text{clusters}} - 1$ degrees of freedom. This scheme can be applied by adopting different metric spaces and different clustering techniques. Concerning the choice of the metric, both 1-norm and 2-norm distances can be employed as distance d between two Fock states Ψ and Φ , namely $d = L_1 = \sum_{i=1}^M |\Psi_i - \Phi_i|$ or $d = L_2 = \sqrt{\sum_{i=1}^M |\psi_i - \phi_i|^2}$, with Ψ_i and Φ_i being respectively the occupation numbers of Ψ and Φ in the i -th mode.

6.1.1 Adopted clustering techniques

Several clustering methods were employed within our validation scheme: (1) a recent proposal by Wang and Duan [225], whose concept is shown in Fig. 6.1, and two unsupervised machine learning techniques, (2) *agglomerative Hierarchical Clustering* and (c) *K-means clustering*. Two variations of the latter approach were also examined, to increase the strenght of our model. A short description of each adopted method follows briefly.

1. The protocol proposed by Wang and Duan [225], and hereafter named *bubble clustering*, determines the inner cluster structure of a sample by (i) sorting in decreasing order the output events according to their frequencies, (ii) choosing the observed state with the highest frequency as the center of the first cluster, (iii) assigning to such cluster all the states belonging to the sample whose distance d from its center is smaller than a cutoff radius ρ_i , and (iv) iterating the procedure with the remaining states until all the observed events are assigned.
2. *Hierarchical clustering*, in its bottom-up version, starts by assigning each observed event to a separate class. Then, the two nearest ones are merged to form a single cluster. This grouping step is iterated, progressively reducing the number of classes. The agglomeration stops when the system reaches a given halting condition pre-determined by the user. In the present case, the algorithm halts when no more than 1% of the observed events is included in some cluster containing less than 5 events. All of these smallest clusters are considered as outliers and removed from the structure when performing the χ^2 test. The distance between two clusters is evaluated as the distance between their centroids. The centroid of a cluster is defined as the point that minimizes the mean distance from all the elements belonging to it.
3. *K-means* is a clustering algorithm where the user has to determine the number of classes (k) [40, 229, 230]. With this method, the starting points for centroid coordinates are chosen randomly. Then, two operations are iterated to obtain the final cluster structure, that are selecting elements and moving centroids. The first one consists in assigning each observed event to the cluster whose centroid has the smallest distance from it. Then, once the k clusters are completed, the centroid of each cluster is moved from the previous position to an updated one, given by the mean of the elements coordinates. These two operations are repeated until the structure is stable. Given a set of k centroids (c_1, \dots, c_k) made of (n_1, \dots, n_k) elements $(e_{11}, \dots, e_{1n_1}, \dots, e_{k1}, \dots, e_{kn_k})$, where $\sum_{i=1}^k n_i = N$, the operations of selecting elements and moving the centroids minimize the objective function $\frac{1}{N} \sum_{j=1}^k \sum_{i=1}^{n_k} d(e_{ij}, c_j)$. Several trials were made to determine the optimal number of clusters, showing that the performance of the test improves for higher values of k and then reaches a constant level. We then chose to balance the two needs of clusters made up of at least 5 elements, since the compatibility test requires a χ^2 evaluation, and of a high efficacy of the validation test.
4. *Variations of K-Means*, different initial conditions can lead to a different final structure. Hence, the algorithm can end up in a local minimum of its objective function. To avoid this

issue, we considered three different strategies: (I-II) replacing the random starting condition with two initialization algorithms, namely (I) *K-means++* and (II) a preliminary run of *Hierarchical clustering* and (III) building on the same data set several cluster structures.

6.1.2 Numerical results

As a first step, we performed a detailed analysis to identify the best algorithm among the mentioned clustering methods. More specifically, we proceeded with the two following steps: (i) a *training stage* and (ii) a *validation stage*. The parameter quantifying the capability of each test to perform correct decisions is the success percentage or, equally, the probability that two samples drawn from the same statistical population are labeled as compatible while two samples drawn from different probability distributions are recognized as incompatible.

(i) In the *training stage*, we applied all the algorithms on a *training set* of numerically generated samples of output states, belonging to the collision free subspace of $N = 3$ photons evolved through a fixed unitary transformation, with $m = 13$ modes. For each considered sample size, the parameters proper of each technique were optimized. To evaluate the success percentages for each configuration of parameters we sample 500 output events from boson sampling and distinguishable output distribution. We observe that the best success percentage is obtained for the *K-means++* method with majority voting and employing the L_2 distance.

(ii) In the validation stage, we validated the algorithm with the highest success percentage according to our tests. We divided this task into two steps, by first (ii.a) validating its functioning for general unitary transformations and then (ii.b) by increasing the dimension of the Hilbert space. Hence, (ii.a) we performed the test with $N = 3$ photons evolving through 20 different Haar-random 13×13 matrices. For each matrix we performed 100 tests between compatible samples and 100 between incompatible ones, by fixing the number of clusters and trials to the values determined in stage (i). We observed that the chosen approach, *K-means++* with majority voting and employing the L_2 distance, permits to achieve better success percentages (Mean $\sim 98\%$). Then (ii.b) we tested the capability of the chosen validation method to successfully operate on Hilbert spaces with larger dimensions. More specifically, we progressively increased the number of photons and of modes, showing that the adopted test permits to validate the data samples even for larger values of N and m (see Tab. 6.1).

Therefore, we collected data samples corresponding to the Boson Sampling distribution, and distinct samples obtained with distinguishable particles. The unitary evolution is implemented by an integrated photonic chip realized exploiting the 3D-geometry capability of femtosecond laser writing [24–26] and performs the same transformation U employed for the numerical results used with the training stage. We then performed the same compatibility tests described previously on experimental data sets with different sizes, by using two methods: *K-means++* with majority voting and *bubble clustering*, both with 2-norm distance (see Appendix 8.13). The results are shown in Fig. 6.2a, for the case of incompatible samples. This implies that the reported percentages represent the capability of the test to recognize Boson Sampler fed with distinguishable photon inputs. Reshuffling of the experimental data was used to have a sufficient number of samples to evaluate the success percentages (see Appendix 8.15). Hence, the tests were performed on samples drawn randomly from the experimental data.

6.1.3 Generalization for Scattershot Boson Sampling

The scattershot version of Boson Sampling [149] is implemented with six independent parametric down-conversion photon pair sources are connected to different input modes of the 13-mode integrated interferometer as for [149]. We considered a generalization of the proposed algorithm to be applied for scattershot Boson Sampling. In this variable-input scenario a Boson Sampler to be validated provides n samples that correspond to n different inputs of the unitary transformation, that is, n Fock states Φ_i with $i \in \{1, n\}$. Hence, our validation algorithm in its standard version

Output Classification							
<i>Photons (N)</i>							
<i>3</i> <i>4</i> <i>5</i>							
Modes (m)	<i>Ind.</i>	<i>Dis.</i>	<i>Ind.</i>	<i>Dis.</i>	<i>Ind.</i>	<i>Dis.</i>	
13	20	0	20	0	20	0	<i>Ind</i>
	0	20	0	20	0	20	<i>Dis.</i>
20	20	0	20	0	20	0	<i>Ind</i>
	0	20	0	20	0	20	<i>Dis.</i>
30	20	0	20	0	19	1	<i>Ind</i>
	0	20	0	20	0	20	<i>Dis.</i>
40	20	0	19	1	19	1	<i>Ind</i>
	5	15	1	19	2	18	<i>Dis.</i>
50	20	0	19	1	20	0	<i>Ind</i>
	3	17	1	19	1	19	<i>Dis.</i>

		<i>Ind.</i>	<i>Dis.</i>	
Overall performance	296	4	<i>Ind.</i>	
	13	287	<i>Dis.</i>	

Table 6.1. Confusion matrix for *K-means++* with majority voting by varying N and m [validation stage, step (ii.b)] Upper table: number of successes for *K-means*, initialized by *K-means++* and with majority voting, obtained for different values of the number of photons N and modes m . We adopted the same parameters obtained from stage. Lower table: overall confusion matrix obtained by summing the results over the number of photons and modes.

needs to perform n separate compatibility tests. Indeed, it would bring n distinct chi-square variables χ_i^2 , where the i -th variable would quantify the agreement between the distribution of the data belonging to the input Φ_i and the distribution of a sample drawn by a trusted Boson Sampler with the same input state. Hence, each input state would be tested separately.

In order to extract only one parameter to tell whether the full data set is validated or not, for all inputs, a new variable can be defined as $\tilde{\chi}^2 = \sum_{i=1}^n \chi_i^2$. This variable is a chi-square one with $\nu = \sum_{i=1}^n \nu_i$ degrees of freedom, provided that the χ_i^2 are independent. We have performed this generalized test on the experimental data by adopting the same clustering technique previously discussed in the single input case. We have collected output samples given by 8 different inputs both with indistinguishable photons and with distinguishable ones. Through the evaluation of the new variable $\tilde{\chi}^2$, the algorithm was able to distinguish between a trustworthy scattershot Boson Sampler and a fake one at the significance level of 5%, using a total number of observed events up to 5000 events (over all inputs), as shown in Fig. 6.2b. The standard version of the test, validating each input separately, would require samples of 2000 events per input to reach a success percentage $\geq 80\%$, that is, an overall amount of 16000 events. Hence, the generalized version of the test permits to significantly reduce the amount of necessary resource to validate scattershot Boson Sampling experiments.

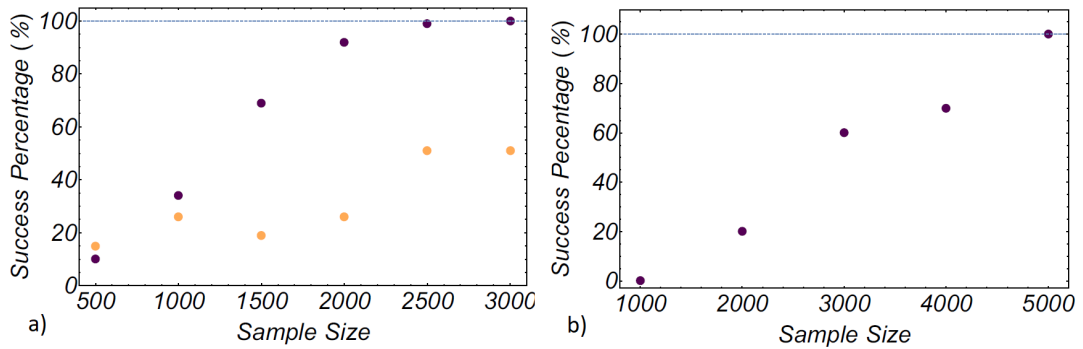


Figure 6.2. Experimental validation of a small-size Boson Sampling and Scattershot experiment with $N=3$ and $m=13$. **a** Success percentages of the compatibility test performed on incompatible experimental data sets with different size (indistinguishable against distinguishable photon inputs), for a significance level of 5%. The darker dots represent the performance of *K-means++* with majority voting clustering algorithm, while the lighter ones were obtained with *bubble clustering*. **b**, Success percentages obtained by applying the generalized version of the compatibility test performed on incompatible experimental data samples of different size, (indistinguishable against distinguishable photon inputs) for a significance level of 5%. Experimental data sets correspond to 8 different inputs.

Discussion

In this section we have shown that pattern recognition techniques can be exploited to identify pathologies in Boson Sampling experiments. The main feature of the devised approach relies in the absence of any permanent evaluation, thus not requiring the calculation of hard-to-compute quantities during the process. Finally, such approach is scalable to larger Hilbert spaces, thus being a promising approach for the validation of mid-term experiments.

Looking forward, it is our hope that when building data-driven (rather than first principles) models for error, cross-validation will be used to report the performance of such algorithms. For example, our method had 100% classification accuracy for the training data but had roughly 95% accuracy in the test data. Had we only reported the performance of the algorithm on the training data it would have provided a misleading picture of the method’s performance for larger Boson Sampling experiments. For this reason it is important that, if we want to use the tools of machine learning to help validate quantum devices, then we should also follow the lessons of machine learning when reporting our results.

6.2 Experimental statistical signature of many-body quantum interference

In this section we want to show a different approach to validate true genuine quantum interference, which exploits statistical tools. Calculate the permanents, as we showed in previous chapters, is not efficient because it belongs to the computational class $\#P$ [42]. The full characterisation of a many-body quantum state requires, unfortunately, the entire set of permanents associated to the scattering matrix which is the pivot of the quantum interference. If we want to investigate higher order moments, we cannot because permanents depends on high-order correlation functions [231].

However, we can see that there is a large set of correlation functions which are suitable and accessible for a standard computation which do not require an exponential amount of time and resources [232]. The important question is whether this correlators can provide a sufficient tool to validate true boson interference among the fraudulent ones [201].

Here, we experimentally tested the protocol proposed in Ref. [210], which is based on the correlator $C_{i,j}$ that exploits the expected value of the number operators:

$$C_{i,j} = \langle \hat{n}_i \hat{n}_j \rangle - \langle \hat{n}_i \rangle \langle \hat{n}_j \rangle. \quad (6.1)$$

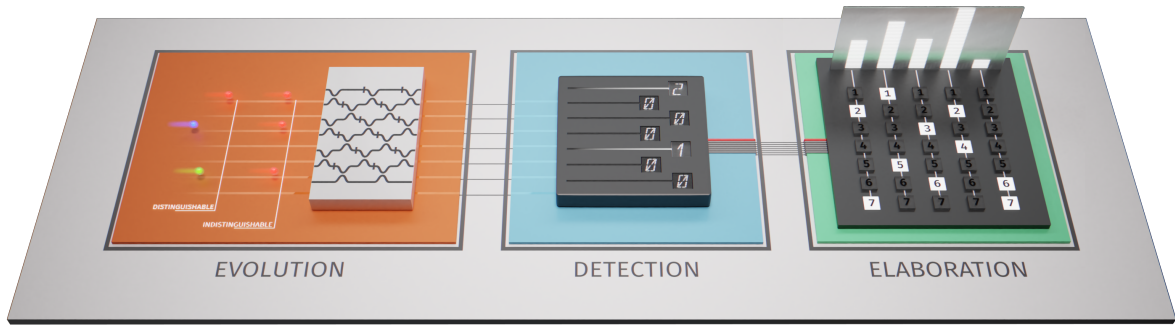


Figure 6.3. Conceptual scheme. The steps of our experimental implementation of the validation protocol: the generation of three-photon state with the possibility to switch between the alternative hypothesis of indistinguishable (\mathcal{I}) and distinguishable (\mathcal{D}) particles, varying the relative time delay and removing the interference filters; the evolution through a network of directional couplers and phase shifters implementing the 7-dimensional random unitary; finally the detection stage to cumulate \mathcal{N}_{klm} events, including those with more than one photon per mode. The setup enables us to compute the whole C-dataset consisting of 21 correlators between the all possible output pairs and to classify the nature of the input state according to its first order moments, in the (NM, CV) plane.

It quantifies how the number of outgoing photons in modes i and j are correlated ($\langle \cdot \rangle$ is the expectation value). The term $\langle \hat{n}_i \hat{n}_j \rangle$ contains all the necessary information of the interaction between particles. As shown in Eq. (6.1) we subtract with the expectation values of the classical contributions on the single output modes, thus focusing only on the interference part. However, a single correlator usually cannot show us with sufficient detail the entire structure. For this reason we have to consider the full set of correlators belonging to the same input, by taking in account all the i, j with $i < j$ to have an hint of interference pattern. In Ref. [210] it is shown that the behaviour of true bosons and their simulated counterparts (which only exhibit bunching, but no multi-particle interference) have a strong resemblance. Consequently, a deeper quantitative understanding is crucial to distinguish the signature of bosonic many-particle interference from the quantum dynamics of other species. We have to take in account instead, the second and third moment of the obtained correlator dataset, which can provide us the desired contrast. To solve the problem, Walschens et al. [210] proposed the *normalised mean (NM)*, *i.e.* the first moment divided by n/m^2 , the *coefficient of variation (CV)*, *i.e.* the standard deviation divided by the mean and the *skewness (S)* as benchmark parameters. For the analytic solutions of this quantities, refer to [210]. It is worth noting that this protocol can be performed efficiently because it requires, if we want to measure the higher moments with sufficient accuracy (independently from the number of modes and photons), at most the calculus of permanent of 4×4 submatrices¹. Nevertheless, to reach a sufficient accuracy we have to do repeated measurements of correlators $C_{i,j}$ for each output combination.

Here we experimentally tested this validation protocol by implementing a Haar random 7×7 unitary matrix injecting a 3-photon state inside the optical chip for different inputs (see Section 3.5 and Fig. 6.3). The full C-dataset of our boson sampler consists of total amount of $\binom{21}{7} = 21$ correlators. In our setup we have collected events in which three particles are detected in the output mode's arrangement (k, l, m) . Let us define \mathcal{N}_{klm} , the number of time in which the (k, l, m) configuration is sampled, N the total sample's size and n_{klm}^i , the number operator's eigenvalue in the mode i of the output state (k, l, m) . The two-mode correlator can be estimated

¹This is due to the upper bound setted by the Variance $Var(C_{i,j})$ which has in the analytic formula, in the worst case a low polynomial power in the creation operators

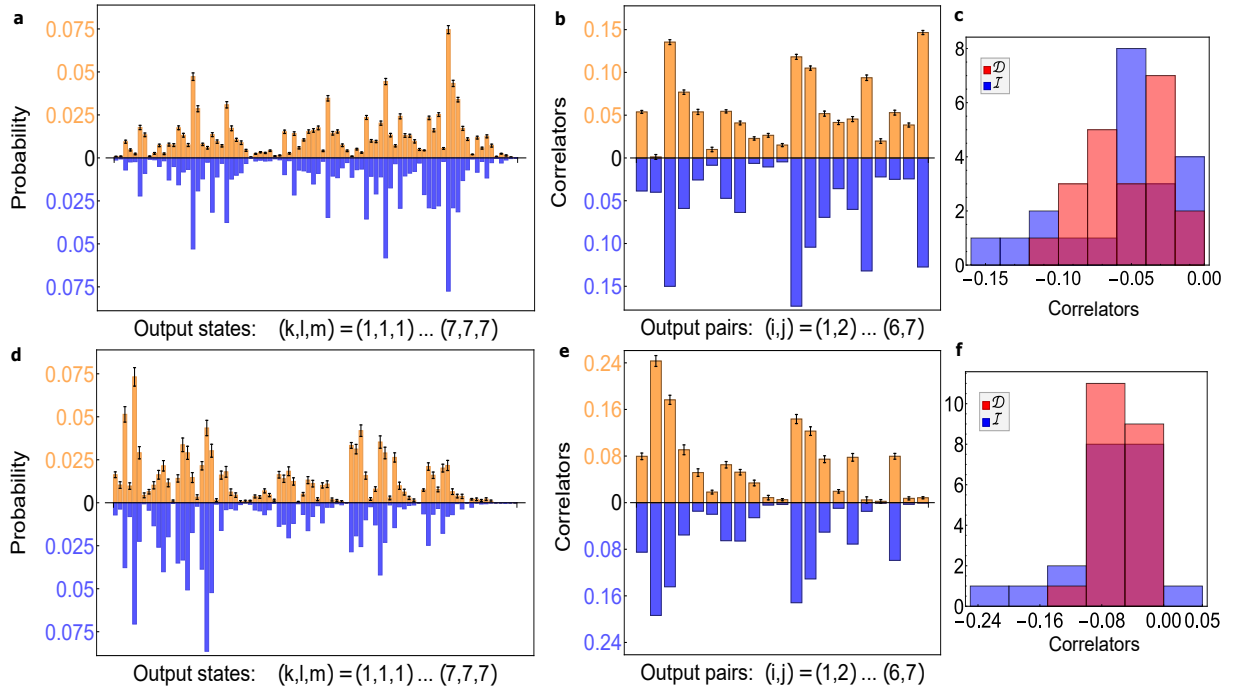


Figure 6.4. Experimental output data samples for indistinguishable particles. For both input states $A=(1,4,5)$ and $B=(1,3,6)$, we measured three-photon data samples (orange) and compared them with the expected distribution from the reconstructed transformation (blue). **a,d**) Output data samples, including all bunching configurations, for input state A with $N_A=10200$ events (**a**) and B with $N_B=1800$ events (**d**). The total variation distances (TVD) between the theoretical distributions and experimental samples are $\text{TVD}^{(A)} = 0.162 \pm 0.004$ (**a**) and $\text{TVD}^{(B)} = 0.205 \pm 0.009$ (**d**). Theoretical distributions take into account partial indistinguishability between the input photons. **b,e**) Experimental C -datasets (orange) corresponding respectively to the input states A and B, compared to the absolute value of the corresponding sets expected from the reconstructed transformation (blue). **c,f**) Histograms of C -datasets for hypothesis \mathcal{D} (distinguishable, red) and \mathcal{I} (indistinguishable, blue) for input A (**c**) and B (**f**). Error bars in the experimental data samples (**a,d**) are due to the Poissonian statistics of photon counting measurements, while error bars in **b,e** are generated via Monte Carlo simulations from the experimental data.

as

$$C_{ij} = \frac{1}{N} \left(\sum_{\substack{klm \\ l \geq k \\ m \geq l}} n_{klm}^i n_{klm}^j \mathcal{N}_{klm} - \sum_{\substack{klm \\ l \geq k \\ m \geq l}} n_{klm}^i \mathcal{N}_{klm} \sum_{\substack{klm \\ l \geq k \\ m \geq l}} n_{klm}^j \mathcal{N}_{klm} \right) \quad (6.2)$$

Note that in Eq. (6.2) output configurations with more than one photon per mode are needed, consequently we took to account this feature. Experimentally, this was done adding cascades of two fiber beam splitters for each output in order to emulate number-resolving detectors [126] and measure all bunching contributions. Data are then corrected and analysed taking into account each loss and splitting ratio of the fiber beam splitters.

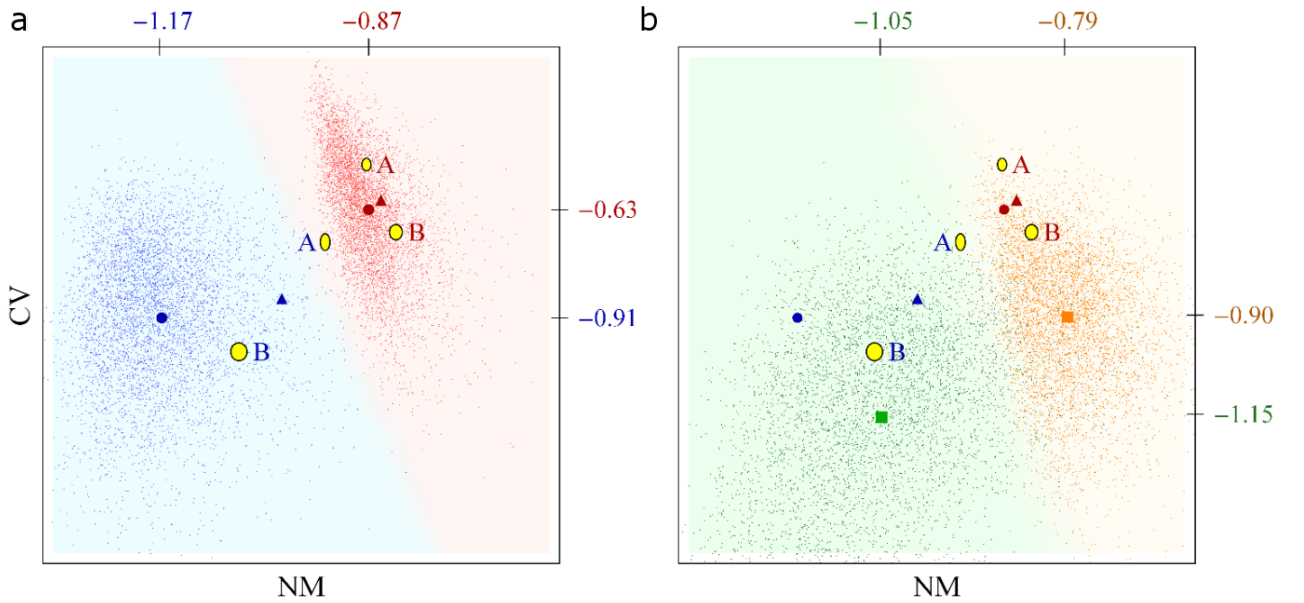


Figure 6.5. Assessment of multi-particle interference. At the final stage of the algorithm we plot our experimental points (yellow disks) in the NM-CV plane, which enhances the separation between data from distinguishable (\mathcal{D}) and indistinguishable (\mathcal{I}) photons. The normalized mean (NM) and the coefficient of variation (CV) are reported for the C -datasets corresponding to the two inputs A and B and for the two alternative hypotheses (\mathcal{I} : blue; \mathcal{D} : red). **a)** Experimental points can be assigned to one of the two clouds numerically generated by 10^4 Haar-random unitary transformations for \mathcal{I} (blue) and \mathcal{D} (red) particles, according to a given classification algorithm. Here, red and blue circles are the centroids predicted by RMT. **b)** Additional information on the adopted implementation is employed. Clouds (\mathcal{I} : green; \mathcal{D} : light orange) are numerically generated exploiting knowledge on the implemented circuit, by sampling 10^4 random unitaries according to a fixed internal structure, equivalent to the one adopted for our integrated circuit, and random parameters for the internal beamsplitters and phase shifters building up the transformation. Furthermore, sampling from different input states is considered. Consequently, centroids of the new clouds (orange and green squares) do not coincide with the RMT predictions (red and blue circles). In both plots, the axes of the yellow disks correspond to 2 standard deviations, as estimated via a Monte Carlo simulation from the experimental data. Triangles, representing the means of the experimental points A and B for distinguishable (red) and indistinguishable (blue), fall well in their respective clouds, allowing for a confident discrimination of the datasets.

We injected two different input states and we calculated for each one the normalized mean and the coefficient of dispersion with the correlators associated to the post-selected events in the output. We changed from indistinguishable to distinguishable particles simply removing interferential filters (gaining a factor 10 in the 3-photon count rate) and delaying of relative time respect to them. In Fig. 6.4 we reported the measured output distributions of the two

indistinguishable input states injected $A = (1, 4, 5)$ and $B = (1, 3, 6)$ with correlators associated to each one calculated with (6.2). In Fig. 6.5 is showed the experimental plot of **NM** versus **CV** where the centroids (black spots) were calculated by means of theoretical prediction given by *random matrix theory* (RMT) [233]. The two numbs represent two collective behaviours of particle sampled with partial distinguishability (dark red) $P = 0.63$ (see Appendices for the model used to describe partial distinguishable states) and fully distinguishable (cyan) where each point is the **NM** and **CV** for any sampled Haar matrix generated. Red and blue points represent the experimental values obtained by injecting the two input states into the integrated device. As it shown, the two behaviour are well confined in the equivalent numbs as predicted by the Haar random simulations.

6.2.1 Generalizing the scheme with random forest classifiers

Inspired by the analysis described in the original proposal [210], we investigate the efficacy of a broader set of estimators to discriminate between datasets with distinguishable and indistinguishable photons. Our approach aims at exploiting summary statistics to identify highly effective signatures of genuine interference. These quantities include common measures of location, dispersion and shape for probability distributions and help quantify key global characteristics of a given dataset in a unique figure of merit. The idea is that data related to particles with different degrees of distinguishability may exhibit a marked separation in these global characteristics, reducing the task of validating multi-particle interference to the assignment of an experimental sample to one of two or more alternative hypotheses. Thus, this task perfectly fits in the scope of classification problems. To this purpose, we choose a set of 10 estimators and study their efficacy in making classification algorithms separate clouds of data associated to the two hypotheses of distinguishable and indistinguishable photons. Following the intuition of Ref. [210], we consider as input for the classification algorithm the two-mode correlators C -dataset given by Eq.(6.2). Indeed, the resources required for its evaluation scale polynomially with the dimension of the problem and they promise to capture hidden correlations highly characteristic for each of the hypotheses under test.

With regard to the classification, we select among others the *random forest classifier* (RFC), a learning method widely adopted for classification tasks for its ability in handling both linear and highly non-linear dependencies [234]. The basic mechanism of a RFC is to build a collection of decision trees over the space of the dataset and point by point output the mode of the classes of the individual trees. Chosen the input data (the C -dataset) and the classifier, we proceed in generating the input data to feed the RFC, for various numbers of photons n and modes m . Specifically, we sampled 10^4 m -dimensional unitary evolutions according to the Haar measure and, for each transformation, we evaluated the C -dataset in the cases of n -photon distinguishable and indistinguishable input photons. We then input the two datasets in the RFC, with hyperparameters properly optimized for this task, and studied the contribution of each estimator to the whole classification procedure as seen by the classifier itself. The basic idea behind this analysis is that not all estimators provide the same amount of information to help the RFC understand the logic on how to assign each point to the various classes. However, thanks to the mechanism proper of decision trees, which is based on iteratively querying each estimator about entropic measures, it is possible to construct a ranking of importance by retracing how successful each estimator has been in performing the assignments (Fig.6.6a). By repeating this analysis for various combinations of (n, m) we identify a clear subset of summary statistics that prove to be effective for discriminating the two hypotheses. Moreover, as we progressively varied the dimensions of the problem, we also retrieve a qualitative trend of their importance, which we summarized in Fig.6.6b.

Our qualitative analysis suggests at least two results: first, we find that random forests indeed select two of the first three moments (**NM** and **CV**) as highly informative features for the classification, retrieving the observations at the core of the validation protocol [210]. Furthermore,

also the qualitative scaling of their importance correctly reproduces the one that was described in the original proposal via direct numerical simulations: **NM** and **CV** become in fact respectively less and more meaningful as the dimension of the problem increases with (n,m) . Similarly, though still less significant for this task in the range of (n,m) reported in Fig.6.6, also the third

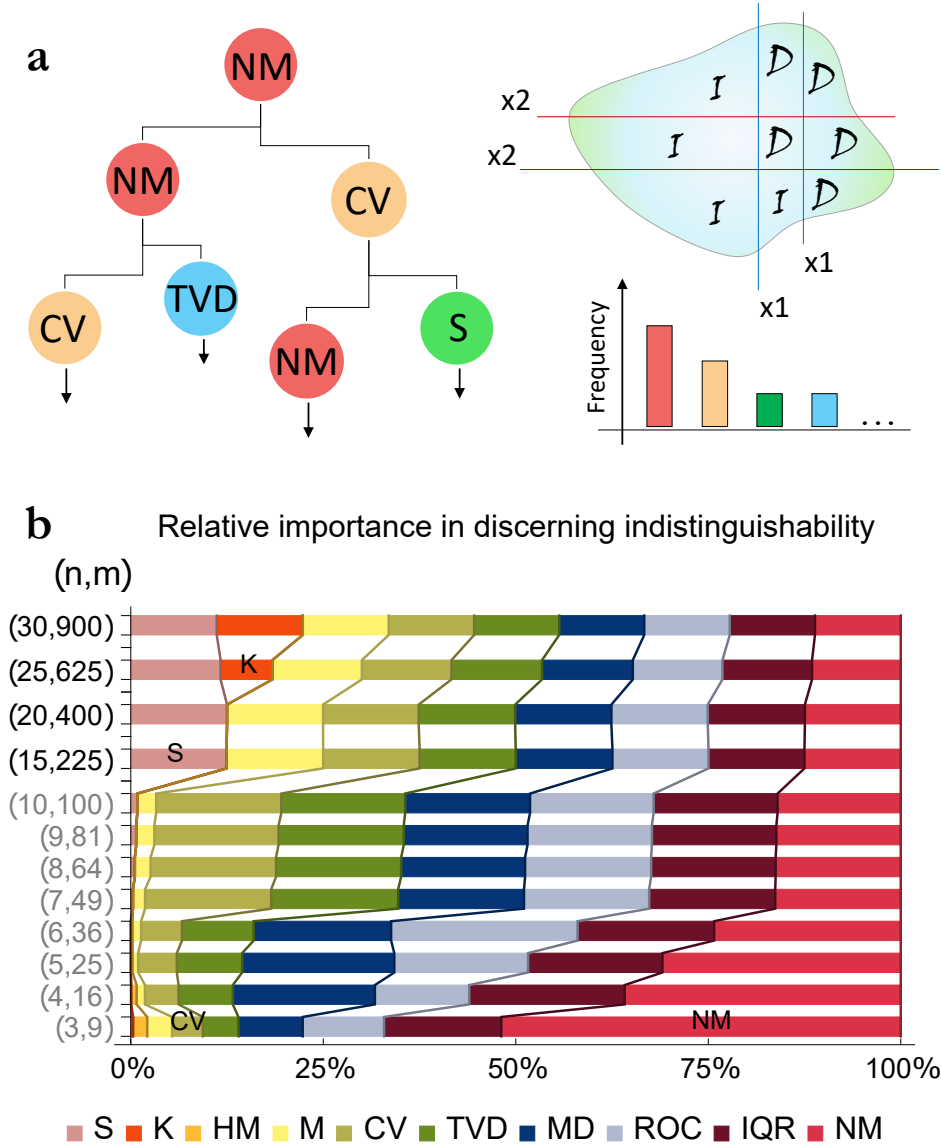


Figure 6.6. Importance of summary statistics for classification. An arbitrarily large set of estimators can be employed to classify datasets from experiments with different levels of distinguishability. To grasp the -unknown- relevance of each estimator, a key tool is provided by feature selection techniques and, in particular, by the *Mean Decrease in Impurity* (MDI) via random forest classifiers (RFC) [234]. MDI estimates the importance of each estimator as the sum over the corresponding number of splits in the RFC, weighted by the number of samples it splits (a). Here we show the relative statistical importance of a specific set of summary statistics according to a RFC, averaged over 200 random extractions of training sets from datasets of 10^4 samples (b). Each sample is constructed by generating a Haar-random unitary transformation and evaluating 10 quantities over the corresponding two-mode set of correlators (C -dataset) from Eq.(6.2). Here, **CV**: coefficient of variation; HM: harmonic mean; IQR: Interquantile range; K: Kurtosis; M: median; MD: median deviation; **NM**: normalized mean; ROC: area under the ROC curve for the normalized C -dataset; S: skewness; **TVD**: total variational distance of the C -dataset, normalized to 1, from the uniform distribution. All bars in the chart are ordered according to the legend below.

moment (S: Skewness) exhibits the same -slowly- increasing trend found in Ref. [210]. Finally, and importantly, our RFC classification scheme allows to identify those quantifiers that are near-to-optimal for a given decision problem in terms of, e.g., size and particle types. These may be quite distinct from, and more efficient than, the lowest order statistical moments of the C -dataset as employed in Ref. [210]. Furthermore, the hierarchical ordering of different quantifiers as achieved by the RFC analysis might be a reflection of specific structural properties of the many-particle interference under scrutiny, and therefore motivates further research. The capability of assessing their importance makes RFCs very useful to gain effective insights, as well as to filter irrelevant figures of merit or to capture unknown connections between them. Moreover, the fact that it does not require detailed knowledge on the system makes this approach flexible and ready for use where a complete theoretical picture is not available.

Discussion

In this work we experimentally tested the protocol proposed by [210]. We showed that by measuring the correlators of all possible output outcomes it is possible to use statistical tools to validate many-bosons interference as *normalised mean* and *coefficient of variation* to distinguish between different particle statistics. We tried to validate also with *skewness* but we think that for the experimental condition that we had, we are not able to use it, because the number of modes and photons are not sufficiently large to discriminate. In principle as pointed out in [210], smaller sizes of the interferometers provide smaller C -datasets that we can use to calculate higher orders statistical functions as the *skewness* and obtain higher statistical significance. Nevertheless, in Fig. 6.5 we showed that our experimental data can successfully discriminate partial distinguishable particles from fully distinguishable particles and the process is efficient because does not require the scaling calculation with n thanks to RMT theory that permits us to predict the centroid of the numbs in a reasonable number of operations. Moreover with RFC we successfully show that NM and CV are important figure of merit in this task even for small sizes. In parallel we confirmed the theoretical predictions about the behaviour of the first moments for large size systems.

From a more general perspective, since these statistical correlators are constructed to individuate many-particle quantum interference, beyond quantum statistical bunching or anti-bunching effects, these results improve our knowledge of the quantum many-body dynamics for distinct particle behaviours. Lastly, all these calculations can be made in an efficient way since the complexity of the protocol do not scale with the number of photons and modes.

For completeness we list and compare in the following table the validation protocols studied so far described in Chapter 5 and 6, by highlighting the strong and the weak points in terms of scalability and efficiency.

Validation protocols	Theoretical proposal	m	n	Efficient?	Multi-Input	Scheme	Scalability	Hypothesis		
								U	D	MF
validation against uniform [149, 198]	Ref. [197]	13	3	Yes	•	Triangular	Large	•		
likelihood ratio [149, 198]	Ref. [198]	13	3	No	•	Triangular	Short		•	
Suppression law with Fourier [170, 182]	Ref. [201]	4,8	2	Yes		Fast	Mid	•	•	•
Suppression law with Sylvester [193]	Ref. [169]	4,8	2	Yes	•	Fast	Mid	•	•	•
Pattern recognition techniques	Ref. [225]	13	3	Yes	•	Triangular	Mid	•	•	•
Benchmarking with statistical tools	Ref. [210]	7	3	Yes	•	Triangular	Large	•	•	•

Table 6.2. Comparison between the different validation protocols experimentally implemented with our photonic platforms. We report the difference in terms of number of modes (m), number of photons (n), if we need to compute the permanent (efficiency), if it is suitable for SBS-like experiments (Multi-Input) the architecture scheme implemented, the discrimination power with the system size (Scalability) and which model can discriminate (Hypothesis).

Chapter 7

Quantum interference with deterministic single photon sources

In this Chapter we will see the work that I carried out in collaboration with the GOSS group led by Pascale Senellart. The main goal of this is to test integrated photonics platforms exploiting on-demand single photon sources. We will address the following two questions: 1) can we scale-up our systems in terms of single photons generations? 2) Can we use our platforms and validation protocols to verify indistinguishability of these sources? In collaboration with IFN-CNR and Politecnico di Milano we built a tunable 3-mode tritter [219] as testing platform. We connected the two platforms in order to verify the genuine quantum interference provided by the merged system. Although this work is still in progress, we were able to answer to question 1, and at the moment we are performing upgrades to the setup in order to answer to the second question.

We will provide a brief description of the quantum dots used in this experiment describing the main properties of these sources. Furthermore we will discuss the properties of micropillar cavities [223] which have been significant for the development of Quantum Dots (QD) in the last years. Finally we will describe the experimental setup and the analysis that we employ for the experiment, showing our results. The work is still in progress and we are focusing on how to validate quantum interference of single photons generated by solid state systems.

The paper of this work is currently in preparation with temporary author list: Carlos Anton, Guillaume Coppola, Juan Loredó, Niko Viggianiello, Fabio Sciarrino, Andrea Crespi, Roberto Osellame and Pascale Senellart.

7.1 Introduction to quantum dots

In the last decades there were significant progresses in the field of quantum technology, especially regarding photon generation and storage. Many systems have been explored, such as: ions [4,5] superconducting qubits, defects in diamond (NV centers) [235], atoms [8,9] and semiconductor quantum dots (QDs) [18,236–238]. In this context, the ideal photon source is a deterministic single-photon source which should shot single photons on-demand. Of course, such properties are very difficult to gather, optical losses inevitably degrade the deterministic nature of the source even if this one is ideal.

In this context, strong efforts have been made by the solid-state community to develop alternative single photon sources based on single quantum emitters. The best results in terms of single photon generation, brightness and photon indistinguishability have been obtained using semiconductor quantum dots. Semiconductor quantum dots are nanometric insertions of a small bandgap material in a larger bandgap matrix. The strong confinement of charges in quantum dots makes the electronic occupancy states discrete, very similarly to the electronic states in a single atom. In a bulk material, single photons are emitted from a QD quasi-isotropically in

high refractive index material, limiting the collection efficiency to a few percent. Controlling the QD emission in photonic structures like microcavities or single mode optical waveguides, it is possible to observe brightness close to 80% as it has been demonstrated in 2010-2013 [239, 240]. The first signal of the good capability of cavities for enhanced emission and the generation of indistinguishable photons has been demonstrated in 2002 by Santori et al. [241].

This chapter describes the work that I have done in collaboration with the group of Pr. Pascale Senellart at the Laboratoire de Photonique et Nanostructures (LPN). The group has developed a long expertise in QD photonic devices, based on cavity quantum electrodynamics (CQED). A great challenge to study CQED with QDs is that QDs are randomly distributed both spatially and spectrally. The LPN team developed in 2008 a technique that allows a full control of the coupling between a single QD and a micro-cavity [242]. This technique, called in-situ lithography technique [250], allowed them to demonstrate on-demand control of spontaneous emission of a QD both in the weak and strong coupling regime. In 2010, they demonstrated the fabrication of an ultrabright source of polarization-entangled photon pairs [243]. Moreover, in 2013, Gazzano et al. fabricated a bright sources of indistinguishable single photons and used it to implement a two-photon quantum logic CNOT gate [244]. All these results opened the road for the development of ultra-bright deterministic sources in the context of linear optics.

7.2 Structure of self-assembled quantum dots

First, we introduce the material and the technique developed to build the structure of the quantum dot for the experiment. A InAs/GaAs quantum dot (QD) is a small cluster of InAs atoms in a GaAs lattice. Since InAs bandgap is lower in energy as GaAs bandgap, it forms a local potential well for the carriers in all the directions of space. Due to the lattice mismatch between GaAs and InAs (the latter having a lattice constant 7% larger), the InAs layer suffers strong mechanical strain. When the InAs reach a critical thick, it becomes energetically favourable for the material to change its configuration, permitting the formation of several clusters of few thousands of atoms. Regarding the surface, the density of quantum dots in the the material can vary with temperature or with indium flow [245]. The typical densities of InAs dots can vary from $10 \mu\text{m}^{-2}$ to $500 \mu\text{m}^{-2}$. The inner structure consists in a series of lenses laying on a 2D InAs layer. The lateral size is typically around 10-20 nm with a height of 3 nm, as shown in Fig. 7.1a and 7.1b. Figure 7.1b shows also the randomness of the QD position as well as the size dispersion. This random distribution and shape diversity make the coupling of a single QD with a cavity a very challenging task as it will be shown in the following sections. The InAs quantum well forming the common base of the QDs is called the wetting layer. In the wetting layer, the confinement of charges is only in the growth direction (1D), the fundamental energy level is higher (1.442 eV, corresponding to a emission wavelength of 860 nm) than in the QD (1.319 eV or 940 nm).

7.3 Occupancy states of a quantum dot

Depending on the laser wavelength, carriers can be generated in the GaAs, in the InAs wetting layer (WL) or even directly in resonance with an energy level of the QD as illustrated in Fig. 7.2. When created non resonantly, the carriers relax to lower energy levels with a typical relaxation time of 30-50 ps [246]. When a quantum dot is occupied by one electron-hole pair, the occupancy state is called *exciton* (denoted \mathbf{X}). Note that contrary to excitons in systems with less confinement, the exciton is formed by the coulombian interaction between the electron and the hole. Here, the exciton in a InAs QD is mainly formed by the 3D confining potential. As a result, it would be more correct to call an exciton an electron hole pair. A negatively or positively charged exciton is named *trion* (denoted X^+ or X^-) [247]. When the dot is occupied by two

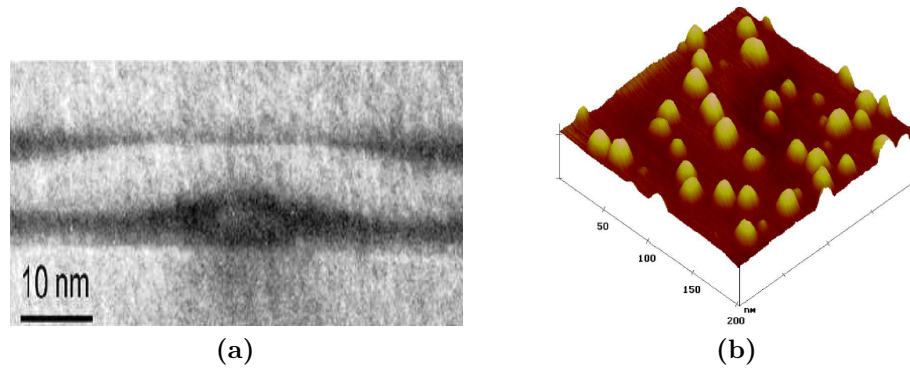


Figure 7.1. (a) Transmission electron microscopy image of a quantum dot. (b) Atomic Force Microscope (AFM) image of InAs quantum dots. Image taken from [256].

excitons, the emission line is called biexciton (**XX**). For all these states, electron/hole pairs can recombine themselves optically, by emitting a photon at a narrow energy band. A biexciton decays to an exciton state by emitting a photon at the wavelength λ_{XX} . The recombination of the second exciton leads to emission at a wavelength λ_X . A negative (positive) trion decays to a single electron (hole) state. This mechanism of successive carriers relaxation is referred as the radiative cascade, depicted in Fig 7.2. Consequently, the recombination from a given occupancy state can be associated to a given wavelength λ_X , λ_{XX} , λ_{X^+} or λ_{X^-} . Because of the Coulomb interaction, these wavelengths are usually different [248, 249].

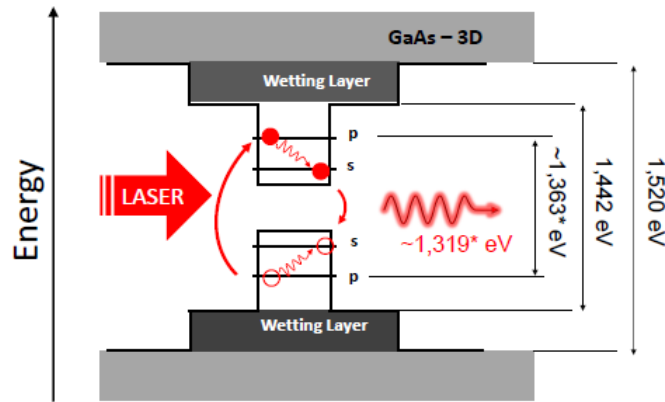


Figure 7.2. Schematics of the energy-level diagram of a QD. Quasi-resonant excitation, where charges are created in the p-shell. After the excitation, charges relax to s-shell before recombining and emitting one photon. Energies marked by a star can change from one sample to another, depending on the growing conditions. Image taken from [256].

We can thus distinguish three main excitations:

- The **non-resonant** (above-band) excitation consists in the creation of electron hole pairs in the continuum associated to the GaAs bulk ($>1.51\text{eV}$ at 4K) or in the continuum of the InAs wetting layer (1.45 eV). Single carriers, electron holes, or excitons scatter in the QD and relax to the lower-energy levels.
- To minimize the delay induced by the capture of charges in the QD, electron-hole pairs can also be directly created in a discrete resonance inside the QD: this is called a **quasi-resonant** excitation. A strong laser power is required than in non resonant excitation since the absorption cross-section of a single QD is very small.

- The **resonant** excitation consists in the optical creation of pairs directly into the considered exciton state.

When the temperature is increased the line shape is modified by the exciton phonon coupling with the broadening of the main emission line (zero phonon line) and the appearance of phonon sidebands [250]. At temperatures lower than 10K the phonon sideband is low populated, therefore, the emission of a phonon is more efficient than the absorption of a phonon. Conversely, increasing temperature, the sidebands become more symmetric and broader. The presence of phonon sidebands degrade indistinguishability of the emitted photons [251]. For these reasons, the QD is at near 4K and filtering out the phonon sidebands with spectral filters.

7.4 Phenomenology of quantum dots in a cavity

In this part, we briefly describe the basic theory of emission of a quantum system as a QD in a cavity. The interaction of an homogeneous system of atoms in a cavity, without considering losses, is described by an Hamiltonian composed by two separate systems and a term of interaction (Jaynes-Cummings model) [252]. The emitter is a two-level system that interacts with one single mode of the electromagnetic field determined by the cavity. The Hamiltonian H_{JC} for the corresponding emitter-radiation system is given by:

$$H_{JC} = \underbrace{\hbar\omega_X |X\rangle\langle X|}_{H_{QD}} + \underbrace{\hbar\omega a^\dagger a}_{H_{CM}} + \underbrace{\frac{\hbar\Omega_1^{(1)}}{2} (|g\rangle\langle X| + |X\rangle\langle g|)}_{H_I} (a^\dagger + a) \quad (7.1)$$

where H_{QD}, H_{CM} and H_I are respectively, the quantum dot two-level system, the photon field and the interaction term. Let us call $\delta = w - w_0$ in order to write respect to the resonance configuration, then the term H_I makes the link between the excited emitter state in a cavity containing n photons $|X, n\rangle$ with the state corresponding to the emitter at its fundamental state in a cavity containing $n + 1$ photons $|g, n + 1\rangle$. If we search for the eigenvectors and eigenvalues of the Hamiltonian and we choose to be in resonance ($\delta = 0$), we obtain the desired state:

$$|\Psi_\pm\rangle = \frac{1}{\sqrt{2}} (|X, 0\rangle \pm |g, 1\rangle) \quad (7.2)$$

Here, we consider the case where the QD is in the excitonic state at $t = 0$. The exciton resonance is assumed to be resonant with the cavity mode ($\delta = 0$) which is empty. The initial state is written $|\Psi(t = 0)\rangle = |X, 0\rangle$, which is not an eigenstate of the Hamiltonian. Expanding the state to the eigenvectors and evolving it through the hamiltonian for a generic time t we obtain that the probability that the QD is at the excitonic state is

$$P_X(t) = |\langle X, 0 | \Psi(t) \rangle|^2 = \cos^2\left(\frac{\Omega_1^{(1)} t}{2}\right) \quad (7.3)$$

The system oscillates with exchange of energy between the QD and the cavity. These oscillations are called vacuum **Rabi oscillations** and at the moment we are considering a lossless system. In an experimental context we have to add losses to the problem, which result is that the photon is stored in the cavity with a finite lifetime and the QD can emit photons in other modes of the electromagnetic field. In the next section we will introduce losses in the model.

7.4.1 Considering losses: the Purcell effect

In 1946, Purcell [253] observed that, coupling the system to an electrical resonator, the spontaneous emission rate is increased by a quantity F_P called the Purcell factor. In 2005, the evidence of the Purcell effect with spontaneous emission in a single InAs QD in a 2D photonic crystal has been demonstrated [254].

In a real cavity, photons are stored for only a finite time T . The decay rate τ describes the exponential decrease of the electromagnetic energy in the cavity and corresponds to a quality-factor $Q = w_m/\tau$ (where w_m is the mode linewidth). Evolving the density matrix ρ which takes into account the whole set of electromagnetic modes [255], we can provide a full description of the phenomenon. If we solve the system of equation [110] we can distinguish two regimes, depending on $\Omega_1^{(1)}/\tau$:

- **Strong coupling** ($\Omega_1^{(1)} > \tau/2$): This regime corresponds to the scenario when the cavity damping is weak. As seen before, the losses are small enough as to allow another absorption of the emitted photon. However, the Rabi oscillations are now damped with τ as decay constant, as depicted by Fig. 7.3. Spectrally, the strong regime coupling is characterized by the Rabi splitting around the resonance frequency.
- **Weak coupling** ($\Omega_1^{(1)} < \tau/2$): In this case, the cavity is lossy and the relaxation of the electromagnetic field is efficient. In this regime, the cavity is empty most of the time. Figure 7.3 shows the time signature of the weak coupling regime (bottom left) as well as the spectral signature (top left). The acceleration of the spontaneous emission is spectrally visible with a broadening of the emission linewidth.

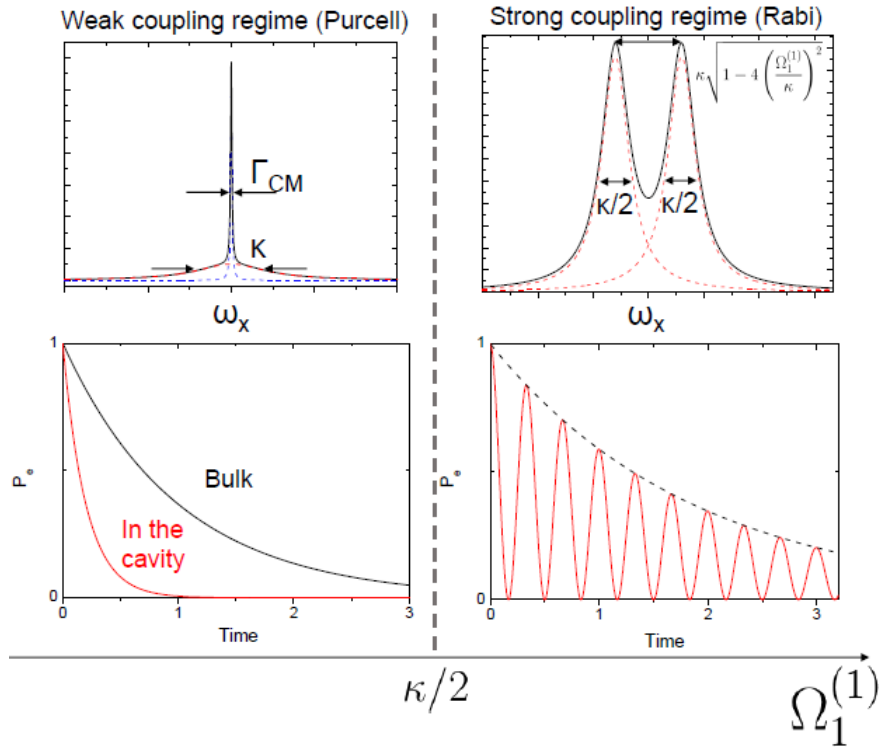


Figure 7.3. Comparison between the two regimes in time and spectrum. Image taken from Ref. [256].

Now in the case of weak coupling, which is of our interest, we can say that if $\Omega_1^{(1)} \ll \tau$, then the decay rate is $\Gamma_{CM} \sim (\Omega_1^{(1)})^2/\tau$. Consequently, the dipole and the mode electric field are along

the same direction, we have

$$\Gamma_{CM} = \frac{2|\mu_x|^2}{\epsilon_0 \hbar} \frac{Q}{V_{eff}} \quad (7.4)$$

The emission rate shown by the eq. (7.4) can be compared to the spontaneous emission of the same emitter in bulk. Consequently, the Purcell factor is defined as the ratio of both rates:

$$F_P = \frac{\Gamma_{CM}}{\Gamma} = \frac{3}{4\pi^2} \frac{Q(\lambda/n)^3}{V_{eff}} \quad (7.5)$$

The results show that the spontaneous emission is largely enhanced with cavities exhibiting small losses (high quality factor Q) and small effective volumes. shows the first acceleration of the spontaneous emission of a QD ensemble. It is possible to see that with resonance into the cavity, the decay rate is extremely faster compared to the bulk and out of resonance.

7.5 Techniques of cavity QD coupling and fabrication

In this section we will describe the devices used in this work. Among different geometries, one can individuate three main structures of semiconductor cavities to control the spontaneous emission of a single emitter: microdisks [257], photonic crystals [258] and micropillars [223, 259]. These three structures, where a single QD is coupled with the optical mode of a cavity, were used to demonstrate the Purcell effect or the strong coupling regime. For simplicity we will focus only on micropillars cavities which are being used in this experiment.

Micropillars: A micropillar is a cylinder with few micrometer diameter is few micrometers and a 10 micrometers height. We can distinguish three main parts: the *bottom mirror*, the *cavity spacer* and the *top mirror*. The mirrors are Bragg reflectors, made of an alternation of layers with thickness equal to $\lambda/4n$ where n is the optical index in the layer and λ the resonant wavelength. In standard structures, the cavity width is λ/n where n is the optical index in the cavity (which is usually GaAs). Using a top-down process [260], micropillars are fabricated from a planar cavity where they are etched. The etching process is a crucial step in the fabrication. The roughness on the edges of the micropillars must be as low as possible in order to maintain a high Q at low volume.

The Purcell enhancement of spontaneous emission requires two main criteria:

- **Spatial matching:** The emitter (or dipole) must be located where the electric field of the mode in the cavity is the strongest. This criterion is one of the most critical: after a Stransky-Krastanov growth, the QDs are randomly distributed in the plane.
- **Spectral matching:** The cavity resonance must be at the frequency of the QD optical transition.

P. Senellart and co-workers developed an original low temperature photolithography technique to deterministically couple a QD with a micropillar cavity. The technique relies on the overlap of three optical paths that are used to excite a target QD (~ 850 nm), to collect their emission (~ 930 nm) and to expose a resist layer (at 550 nm) deposited on a planar cavity sample. The QD position is measured with a ~ 50 nm accuracy by maximizing its emission. When the QD position is determined, the 550 nm excitation line is used to define a disk shape in the resist centred on the QD. This disk will later be used as a direct mask to etch the micropillar. The disk diameter is adjusted so that the fundamental mode pillar matches the QD emission energy. After the in-situ lithography step, the resist is developed leading to disks. The planar cavity is then etched except where the resist was exposed in order to get micropillars that are all centred on QDs. For a detailed description of the in-situ lithography process refer to [260].

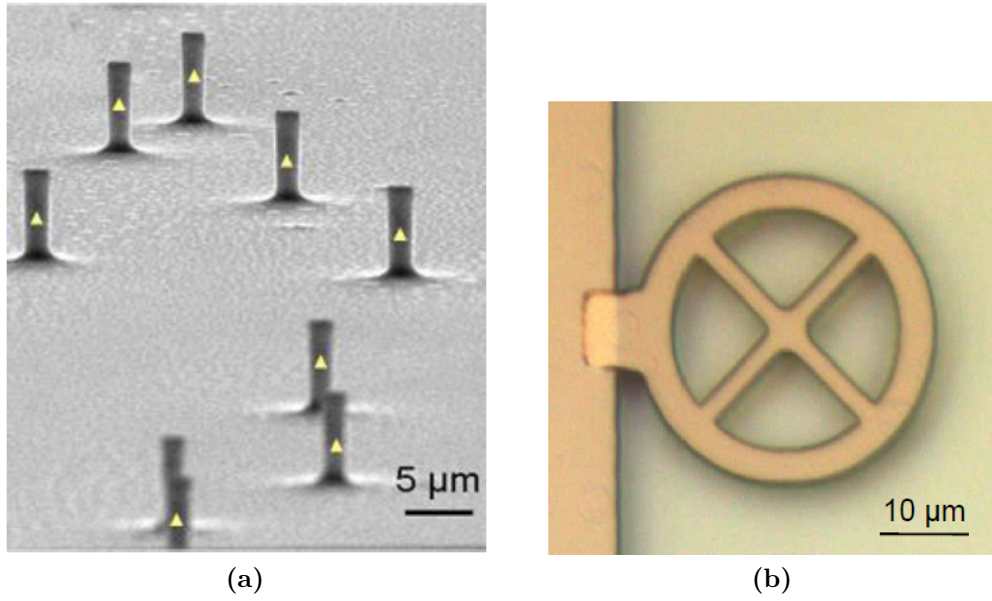


Figure 7.4. a) Image that shows a set of pillar built through *in situ* lithography. b) Top view of a micropillar. Diagonal axis are the H and V collection modes. Image taken from [260].

7.6 Brightness of quantum dots sources

Given a brief description of the physics and the technology of micropillar quantum dots, let us introduce a criteria to quantify the performance of the generation rate of the sources:

Definition 7.1: Brightness

In a pulsed regime, we define the **Brightness** (Br) of a single-photon source as *the probability that one excitation pulse results in the effective emission and collection of one and only one photon in the first lens of the photonic path.*

Hence, the brightness depends both from the extraction efficiency that defines the probability of collecting an emitted photon and also from the probability that the source emits one photon. Contrary to SPDC sources, the single-photon purity of a QD source does not degrade with the pump power. Moreover, the probability that a single QD emits a single photon converges to one at the saturation:

$$P(n = 1) \xrightarrow{P \rightarrow P_{sat}} 1 \quad (7.6)$$

Nevertheless, without any modification of its electromagnetic environment, a QD exhibits a isotropic emission inside a semiconductor material presenting a high optical index. It can be shown that the fraction of emission exiting the sample in the upper half space is $\sim 2\%$. In this case, there is no so much advantage compared to SPDC sources. To overcome these limitations we couple the quantum dot system with the cavity.

In the case of a semiconductor QD in a micropillar, the brightness of the single-photon source is proportional to the fraction of photons emitted in the cavity mode. The QD emission in cavity is similar to that in bulk and the total emission rate is:

$$\Gamma_{tot} = \Gamma + \Gamma_{CM} \quad (7.7)$$

The fraction of photons emitted in the cavity mode is then given by the ratio β :

$$\beta = \frac{\Gamma_{CM}}{\Gamma_{tot}} = \frac{F_P}{F_P + 1} \quad (7.8)$$

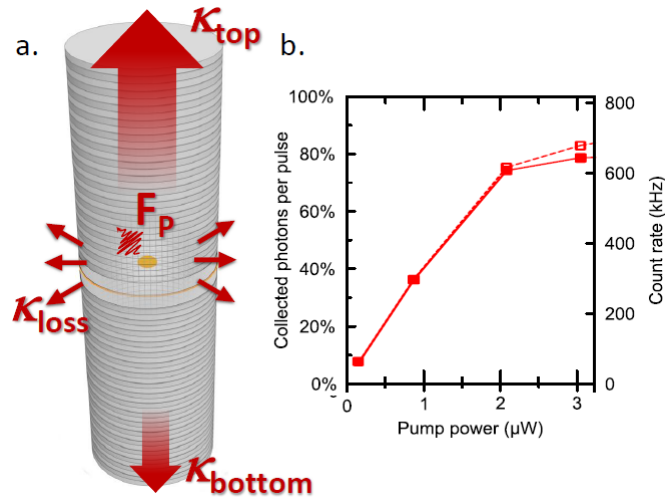


Figure 7.5. a) Schematics of the extraction of the photons from a pillar cavity. $k_1(k_2)$ corresponds to the cavity damping associated to the top (bottom) mirror of the micropillar while k_{loss} corresponds to the leaking rates on the walls of the micropillar. b) Right (Left) axis: Detected count rate (Measured brightness of a micropillar) as a function of the pump power in pulsed regime. Empty symbols and dashed lines correspond to raw data while full symbols and solid line correspond to multiphoton corrected data. Image taken from [261].

This quantity becomes rapidly predominant when F_P increases. For instance, if the QD decay rate is accelerated by a factor 5, the QD emits more than 80% of photons in the mode. Combined to an efficient extraction of the photons from the cavity, this device can be a bright source. Indeed the brightness (Br) of the source is then given by:

$$Br = p_x \beta \eta \quad (7.9)$$

where p_x is the occupation factor of the QD state and η the probability that one photon emitted in the cavity mode is effectively released by the top mirror of the micropillar and collected by the first lens. Here, because the fundamental mode of a micropillar is highly directional:

$$\eta = \frac{k_{top}}{k_{top} + k_{bottom} + k_{loss}} \quad (7.10)$$

where the coefficients k_{top} , k_{bottom} and k_{loss} correspond respectively to the cavity damping rates through the top mirror, the bottom mirror and the wall losses of the pillar as depicted by Fig 7.5. The probabilities p_x and η are respectively inherent to the QD and to the cavity, while $\beta = \frac{F_P}{F_P + 1}$ results from the coupling of the dot with the optical mode of the cavity.

7.7 Experimental properties and details to obtain photoluminescence (PL) emission

The sources subject of our analysis are formed by a set of different QDs built each one in micropillars by the lithographic process. In this section we will briefly describe the main requirements that we need in order to correctly observe single photons signals provided by the dot. First of all, the photoluminescence (PL) signal generated by coupling a laser beam with the cavity is retrieved with a back-scattering geometry. In other words, the output signal from the micropillar is both scattered laser and single photons. Moreover if the laser has a certain polarization, the emission signal is expected to have the same component. Thus if we want to extinguish and separate the laser from dot signal we can not do it by means of waveplates (discriminate with polarization) or gratings (discriminate with wavelength, not possible because it is a resonant excitation $\lambda_{laser} = \lambda_{dot}$).

Luckily, there is a property hidden in the solid structure of the micropillar that solve our problems. Here we describe shortly the qualitative behaviour. We assume that the QD axis are rotated of 45° respect to cavity's axis. When we excite with $|H\rangle$ polarization, we create an exciton state $|H\rangle = (|X\rangle + |Y\rangle)\sqrt{2}$ and we collect a $|V\rangle$ with the minus instead. In presence of fine splitting Δ_{FSS} , the V polarization is an eigenstate of the system, consequently, there is no emission in this polarization. With the absence of fine splitting, the state evolves accordingly to Schrodinger's picture from $|\Psi(t=0)\rangle = (|X\rangle + |Y\rangle)\sqrt{2}$ into

$$|\Psi(t)\rangle = \frac{e^{i\omega_X t}}{\sqrt{2}}(|X\rangle + e^{i\Delta_{FSS}/\hbar}|Y\rangle) \quad (7.11)$$

with $\Delta_{FSS} = \delta(\omega_X - \omega_Y)$. After a short time $t \propto \pi/\Delta_{FSS}$ the state presents a maximal overlap with the V state. Both polarized exciton states can radiatively decay in the cavity H and V modes respectively. The collected PL is thus related to the spontaneous emission of the V exciton state in the V mode of the pillar. The higher the splitting, the faster the evolution of the state is. In practice, the polarization rotation time induced by the fine splitting is of the same order as the accelerated radiative decay time, thus no oscillations of the $|V\rangle$ population are observed. Consequently, we can collect V polarized emitted quantum dots while filtering scattered H polarized excitation laser with linear polarizers and wave plates. For a more detailed theoretical description refer to [256].

Finally, we list three ingredients that we have to take in account in order to extract a bright excitation with good indistinguishability.

1. **Power Rabi's oscillations:** In the presence of an oscillatory driving field, a two-state quantum system follows a cyclic behaviour called Rabi oscillations (15 – 50 ps of duration in a pulse). During the excitation, the system cyclically absorbs photons and emits them by stimulated emission. In the case where the pump is in resonance with the two-level system, the probability of finding the excited state oscillates with a unitary contrast with a pulsation (Ω_1 is the Rabi frequency) is: $P(t) = 1/2 \sin^2(\Omega_1 t/2)$. It is possible to indirectly measure this effect with power measurements.

Since the pulsation of the oscillation depends on the pump power $1/E \propto \sqrt{P}$, the development of the photoluminescence reflects the Rabi oscillations (see Fig. 7.6). During the pulse, the population oscillates due to the interaction with the field. After the pulse, the system relaxes to the ground state by emitting spontaneously a photon. The best situation in terms of indistinguishability (avoiding also scattered laser) and brightness is slight below π -pulse (lowest reflected laser, highest dot emission) which is defined as the lowest pump power where there is the first maximum.

2. **Voltage control:** In order to achieve a full control of the chargers inside the dot, the solid-state system has been doped with Si atoms we changed the structure of the doping in order to improve the charge control, thus stabilizing the Fermi level close to the QD. In Fig. 7.7 we showed the scheme of voltage control, where, pillars after fabricated with the doping, are organized in diodes in order to apply direct voltage on top. The effect of applying voltage is to blue-shift (positive V) or red-shift (negative V) the photoluminescence emission of the QD.

If apply to much positive voltage there is the chance that other carriers interact with the QD disturbing the indistinguishability (different transitions), conversely if apply to much negative voltage, the band is stretched and there is a significative probability to have tunneling effects of the electrons, which gives different line transitions, thus degrading the indistinguishability as well. Typical values of applied voltages on our samples are in range $V_{range} = [-0.4, 0.3] V$ in order to satisfy these conditions.

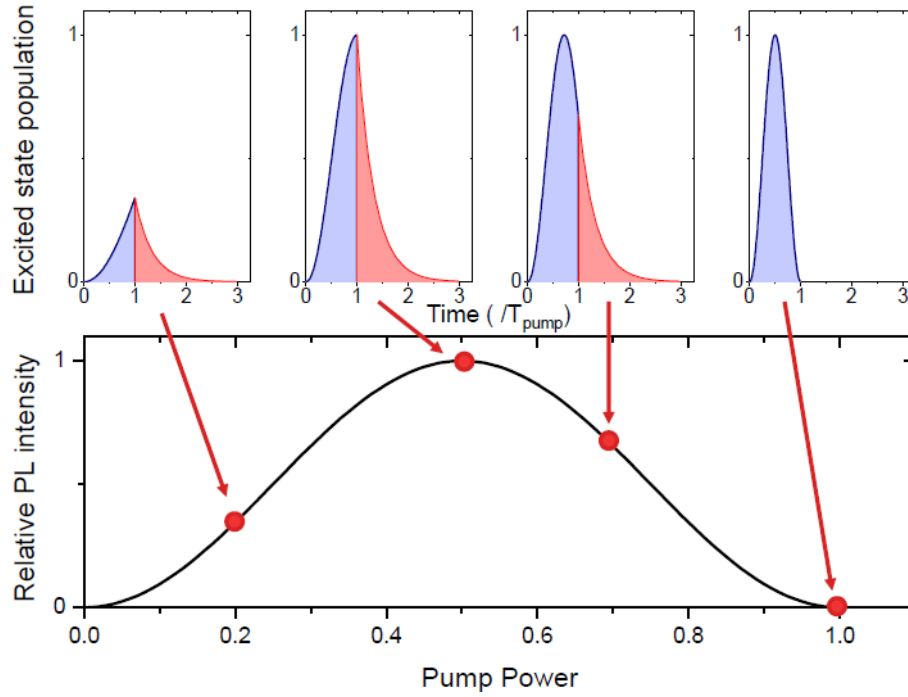


Figure 7.6. Schematics of the observation of the Rabi oscillations observed by changing the excitation power given a fixed pulse duration. The behaviour of the Rabi oscillations are similar of a damped harmonic oscillator in terms of applied power and photoluminescence signal retrieved.

- Filtrage of the sideband:** temperature and optical alignment are crucial to enhance the indistinguishability. We have to confine our excitation in order that only the desired spectral line is excited. In other words we must focus on *resonant excitation* (in our case 925 nm). Moreover, to filter the residual phonon sideband (which grows with temperature, but is partially limited by the cavity, because it forces to emit in a specified mode) we used interferential filters which select only the excitation line at the price of transmittivity which is reduced at least of a factor 2.

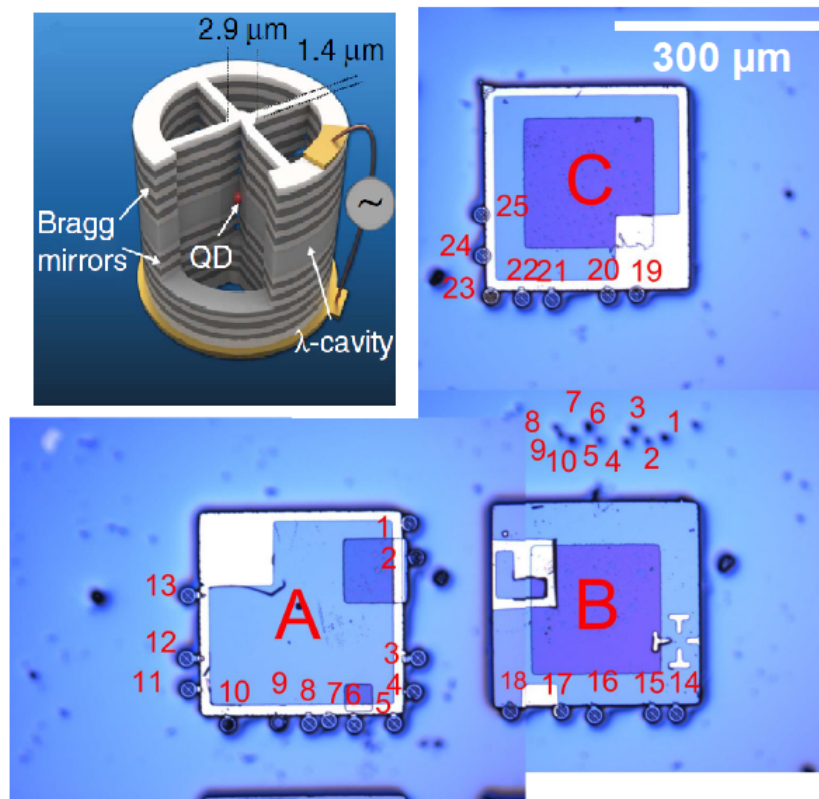
7.8 Merging deterministic single photon sources with integrated devices

In this section we will present the main work by first describing experimental setups and second the analysis of the data retrieved by the merged system of deterministic single photon sources (QD source setup) and integrated photonics (FLM). We used a 3-mode chip which implements a tritter device [219] with a tunable control of an inner phase. The experiment deals with the generation of 3-photon indistinguishable state to inject into the tritter in order to study the enhanced output rate (with respect to SPDC-like experiments) and individuating the quantum interference occurring in, by employing validation protocols techniques.

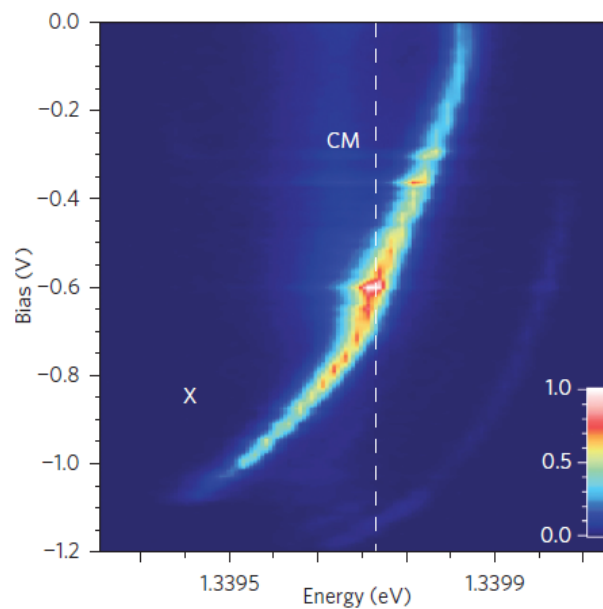
We will subdivide the description in three steps: 1) tritter specifications, 2) experimental setup both for source (S) and tritter (T) with analysis of preliminary characterization of (S) and (T) and 3) analysis of tritter output data and brief discussion on results.

7.8.1 Tritter specifications

The tritter device is built in a borosilicate glass with the technique of FLM, where a golden layer on the top surface (see Fig. 3.6) enables the possibility to apply voltage and tune a phase [161]. We build inside 4 devices, each one with a tunable phase. In this experiment we chose to inject a



(a)



(b)

Figure 7.7. Voltage control. a) Microscope image of the three diodes in our sample (denoted A,B and C). Each diode is connected to several devices made of a surrounding frame contacting the diode and a micropillar embedding a single QD. b) This figure reports the shifting process when we apply the voltage on diodes. Image taken from [223].

3-photon state (one for each input) into the 3-mode chip which implements a U_3 Fourier matrix. If particles are fully indistinguishable, the unitary matrix provides the sequent transformation through the Fock space:

$$|\Psi\rangle_{in} = |1, 1, 1\rangle \xrightarrow{U_3} \frac{1}{\sqrt{3}} |1, 1, 1\rangle + \frac{\sqrt{2}}{3} |3, 0, 0\rangle + \frac{\sqrt{2}}{3} |0, 3, 0\rangle + \frac{\sqrt{2}}{3} |0, 0, 3\rangle \quad (7.12)$$

However, we can observe bunched states only by using number resolving detectors [126] which can discriminate between 1 or n photons on the same detector. For this reason, in this experiment we focus on the performances of the merged system by comparing with SPDC experiment with the same device [219], rather than trying to sample from the full output distribution.

In Fig. 7.8 is shown the circuit implemented in the optical chip with the Reck scheme [15]. In this case we want to implement a Fourier transformation by choosing opportunely the phases into the circuit.

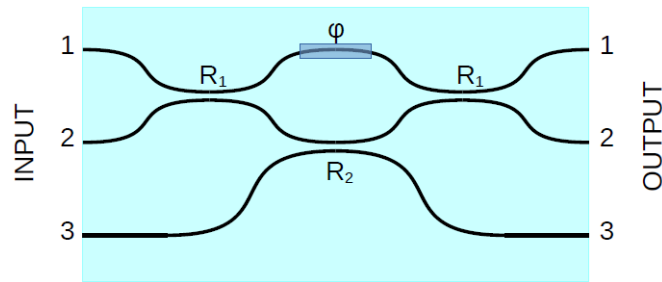


Figure 7.8. Layout of a planar circuit realizing the tritter operation with the Reck scheme [15]. This includes three directional couplers, which should have nominally reflectivities setted as $R1 = 50\%$, $R2 = 33.3\%$, and an electrically tuneable phase-shifter φ .

The phase φ_i of the i -th device is governed by the electrical power P_i dissipated on the related resistor, according to a linear relation $\varphi_i = \alpha_i P_i + \Phi_i$ where Φ_i is constant. Experimental characterizations were performed at the wavelength of 910 nm, however these values do not strongly depend on the wavelength, in fact $\alpha \sim 5.34 \text{ rad/W}$ for both frequencies. The dissipated power follows trivially the Joule's law $P = V^2/R$ where V and R are respectively voltage and resistance. In Fig. 7.9 we report the experimental reflectivities and transmissivities characterized by Milan group, of the four device implemented:

Dev. N°	@ 910 nm		@ 940 nm		INTERP. 925 nm	
	R1	R2	R1	R2	R1	R2
10	58%	47%	47%	34%	53%	40%
11	59%	40%	46%	29%	53%	35%
12	52%	38%	41%	25%	47%	31%
13	49%	33%	38%	20%	43%	26%

Figure 7.9. List of reflectivities the four device implemented. We performed an interpolation between the two wavelength, to extract a value which is near to the experimental condition in which we are working (wavelength of dot emission).

Given these values, we perform a numerical simulation to individuate the phase in which the final implemented transformation is the closest to the ideal U_3 Fourier matrix. We perform a minimization search using the mean (on the three inputs) of the TVDs of single photon

probabilities between the ideal case (I^{Fou}) and the parametric chip (P^{Fou} knowing the inner structure and the reflectivities):

$$f(\varphi) = \left[\frac{1}{3} \sum_{j=1}^m \left[\frac{1}{2} \sum_{i=1}^m |P_{i,j}^{Fou}[\varphi] - I_{i,j}^{Fou}| \right] \right] \implies \varphi^* = \varphi_{Min}[f(\varphi)] \quad (7.13)$$

where φ^* is the phase where the function is minimum. Performing this simulation for all listed device, we chose the implementation which maximizes the fidelity between ideal and simulated case, that is Dev. 12 with $F_{Dev.12} = 0.995$. We extracted the associated voltage to the phase $V^* = \sqrt{\frac{\varphi^* - \phi_0}{\alpha_0}} R_0 = 2.8 V$, where $R_0 = 60 \Omega$, $\phi_0 = 0.81 rad$ and $\alpha_0 = 5.41 rad/W$. In this experiment we will use this Dev. 12 as main platform for our studies.

7.8.2 Setup of quantum dot sources and analysis.

Let us introduce the experimental setup regarding the excitation stage. The conceptual image is reported in Fig. 7.10. We used as a source of excitation, a pulsed tunable Ti-Sapph Tsunami laser with $\Gamma = 82 MHz$ of repetition rate. By employing of monochromators we can increase the pulse duration from $2 - 3 ps$ to $15 - 20 ps$. This is necessary because the cavity modes are larger than the normal one provided by the Tsunami ($\sim 10 ps$). In this way we can excite the sample with both polarizations, depending on the usage that we want to do with the two cavity modes with respect to indistinguishability and brightness. In order to increase both counts rates and indistinguishability, we built a Multiplication of Pulses (MOP) in bulk increasing the number of pulses per second. Without the MOP, we have a pulse each $t_p = 12.3 ns$, while by using cascades of beam splitters with paths delayed properly, we obtained bunches of 3 pulses for each t_p . The relative delays of the pulses in each bunch, if we put the first as a reference are $0, 2.5, 5 ns$.

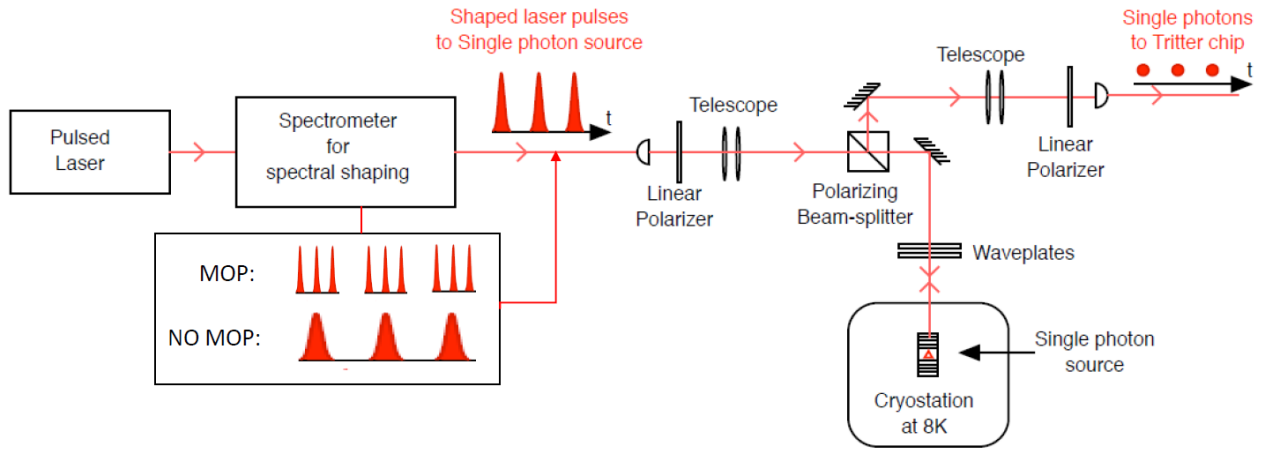


Figure 7.10. Conceptual image of the experimental setup. A Tsunami pulsed laser generates a beam with $\lambda = 925 nm$ and a spectrometer shape the spectra in order to adapt the coupling with the micropillar cavity. We built a multiplication of pulses (MOP) to enhance the generation rate and the indistinguishability between emissions belonging to different pulses. With a linear polarizer we select a polarization and with a telescope we collimate the beam before focusing on the sample with a confocal geometry stage through a $N = 0.75$ objective. The laser beam is aligned with respect the top surface of the cryostation chamber. The QD is conserved at a temperature of $8 K$ in order to avoid sidebands emissions. The photoluminescence signal and the laser are backscattered on the same path. A combination of waveplates and PBS filters the residual laser, while photoluminescence signal according to Eq. (7.11) can pass through the path by extinguishing mostly half of the signal.

For our experiment we chose to use both MOP and standard version to compare the generation rates. By increasing the number of bunches we observed a gain in brightness of a factor 3 with respect to the standard case. After the generation of the bunches we select polarization with

a polarizer (P) and we collimate the beam before the sample. The beam has to be collimated in the far field for two reasons: to prepare it for the focusing objective before the sample and because we want that the photoluminescence signal is maximized if we want to put an etalon (Fabry-Perot interferometer) to increase the indistinguishability. Subsequently using a PBS, we filter the residual polarization and the beam goes to an objective which focus on the sample. The sample in analysis was grown by molecular beam epitaxy (MBE) on a n-doped (100) GaAs substrate. It consists of a GaAs cavity containing an InGaAs quantum dot layer, surrounded by two distributed Bragg reflectors with 30 pairs for the bottom and 20 for the top.

The sample is inserted into a closed-cycle cryostation Montana S50 at a temperature of $T = 8 K$ in order to reduce sideband emissions. Into the cryo chamber we can control the position of the sample with respect to the laser beam by using a 3-axis translational stage with sensibility of $\sim 20 nm$. The pillars are electrically connected with voltage control and combined with a p-i-n doping of the InAs/GaS layers, this geometry allows applying a bias to the structure, a critical tool to finely tune the QD resonance to the cavity mode energy as well as to stabilize the charge environment of the QD [223].

Subsequently, both photoluminescence signal and residual scattered laser are back-reflected towards the PBS. By employing a HWP and QWP, we can minimize the laser component by leaving the intact the PL signal. If we look at eq. (7.11), we know that the emission is in a superposition between H and V cavity modes thanks to the fine splitting Δ_{FSS} and the QD can emit both H and V according to its evolution.

Finally, the photoluminescence signal is sent to the tritter stage after a polarizer and a telescope. All preliminary measures have been made with a high resolution spectrometer.

Before installing the tritter setup, we performed several tests to ascertain the good quality of our sources, such as good brightness and high indistinguishability. All of these measurements have been performed in the condition of resonant excitation. First we analysed the properties of micropillar cavities such as radiative decay time and reflection signal and second we employed different figures of merit to quantify the performances: second-order autocorrelation $g^{(2)}(0)$ measurements for the purity and mean wave-packet overlap M for indistinguishability. Data were extracted as outputs of Hong-Ou-Mandel and Hanbury Brown and Twiss setups. Exploiting APDs and a HydraHarp 400 correlator we are able to retrieve histograms of coincidences events.

Properties of micropillar cavities: we selected the two sources (namely A13 and B18) which are most promising in terms of generation rates and consequently we studied the radiative decay time and we looked the reflectivity at the output to identify the wavelength of the cavity modes:

- *Lifetimes and decay time τ_D :* we measured for A13 the lifetime using a higher resolved in time Fast APD which has a resolution $\Delta t = 4 ps$ which permits to observe a more detailed behaviour. We perform a linear fit on logarithmic scale on data employing this fitting function $f(a, t_0, \tau) = a - (t - t_0)/\tau$, for which we retrieved a lifetime of $\tau_{A13} = 128.9 ps$. We report in Fig. 7.11 the elaborated data. It is worth noting that on the same curve there are two behaviour that we recognize: the effective decay times and the answer of the fastAPD. In order to properly analyse data, we have to focus only on the lifetime region of data, avoiding the second part of the descent (see Fig. 7.11b).
- *Analysis on cavity modes:* the observed reflectivity corresponds to a 3 ps laser coming through the input fiber, going through a HWP (angle choosing the mode you excite and look at), being reflected on the sample, going back through the same HWP (as Identity polar transform), and through the same fiber. We use APC fibers and fibered BS before to be able to collect the back-reflection. From this analysis we retrieved the width and the distance of the H and V modes. In this case for A13 we have $\Delta\lambda_{H,V} = 0.115 \pm 0.005 nm$

while the width for both modes is $\lambda_w = 0.105 \pm 0.008 \text{ nm}$. In Fig. 7.12 we reported the waveplates scan and the peak profile of the H and V mode of our micropillar.

Analysis of purity and indistinguishability: Here we focus on studying the purity and the indistinguishability of the selected sources exploiting the figures of merit $g^{(2)}$ and M . The mean wave-packet overlap M , in the standard configuration (*i.e.* without MOP) is deduced by [241]:

$$M = \frac{1}{1 - \epsilon} \left[2g^{(2)}(0) + \frac{R^2 + T^2}{2RT} - \frac{A_0}{A_{-1} + A_{+1}} \left(2 + g^{(2)}(0) \frac{(R^2 + T^2)}{2RT} \right) \right] \quad (7.14)$$

where $1 - \epsilon$ is the visibility of the interferometer typically around 0.98, $R = 0.48$ and $T = 0.5$ are the reflectivity and transmissivity of the beam splitter in the HOM setup. A_0 is the area of the central peak, while $A_{\pm 1}$ are the previous and subsequent peak with respect to the central one. Areas are corrected taking into account the background noise. If we assume a near 50:50 beam splitter we can rewrite the formula in a more handy version:

$$M_{uncorr} = 1 - \frac{2A_0}{\langle A_{side} \rangle} \quad M_{corr} = 1 - \frac{2A_0}{\langle A_{side} \rangle} - 2g^{(2)} \quad g^{(2)} = \frac{A_0}{\langle A_{side} \rangle} \quad (7.15)$$

where A_0 is the area at the central peak and $\langle A_{side} \rangle$ is the mean of the area for peaks after $A_{\pm 1}$ if we do not use MOP, while if we use it we have to take in account only the nearest peaks. In Fig. 7.13 we show the analysis performed on our source, where all measurements have been done with a time bin of $\Delta t = 126 \text{ ps}$. Retrieved M and $g^{(2)}$ are $M = 65 \pm 5\%$ and $g^{(2)} = 0.032 \pm 0.003$. If we use the etalon for the same analysis we find better purity and indistinguishability values, losing a factor $\times 4$ in the count rates. As an example, in the case of pillar B18 (which was an old pillar not the subject of our experiment) we obtain with etalon a $M_{eta} = 90 \pm 4\%$ and $g_{eta}^{(2)} = 0.012 \pm 0.002$. The value of $g^{(2)}$ depends on multi-photon generation of the source and residual scattered photons of laser which are not extinguished with waveplates. The motivation regarding the low indistinguishability observed in A13 is due to a non-perfect control of cryostation. In fact we suppose that there are leaks in the chamber such that we can not reach the nominal temperature showed by the controller. As a consequence, there are contributions of phonon sideband that spoils the indistinguishability.

7.8.3 Tritter apparatus and analysis of experimental data

Here we describe the second part of the setup which regards the probabilistic generation of a 2- or 3-photon state, the tritter stage and the detection.

After the characterization of our sources, we address the photoluminescence signal in the tritter apparatus. In order to generate a 3-photon state, we must delay subsequent pulse accordingly to the length of each path. We use a cascade of two fiber beam splitter (FBS) and we put two additional fibers (Delay 1 and 2) to synchronize them. Being a passive gate, there is only one configuration of 3 subsequent pulses for which we can work. In this case, being $n = 3$ photon and $C = 3$ different pulses we have $\sigma = n^C = 27$ configurations, thus the probability to observe one specific 3-photon event is $P_{FBS} = 1/27$. HWP and QWP are used to manipulate polarization, while with delay lines we can synchronize each path by adjusting the relative delay with a sensibility of $30 \mu\text{m}$. The tritter is aligned with a single-mode fiber array in input and an multi-mode fiber array in output, where each output is connected with an APD. Finally a HydraHarp 400 correlator elaborates TTL signals from the APD and write in the computer. A post-processed analysis is necessary to retrieve 2- and 3- photon correlations.

Below we list the efficiency in transmission for each step of the propagation:

- efficiency of passive probabilistic demultiplexer $\eta_{PPD} = 0.8$,

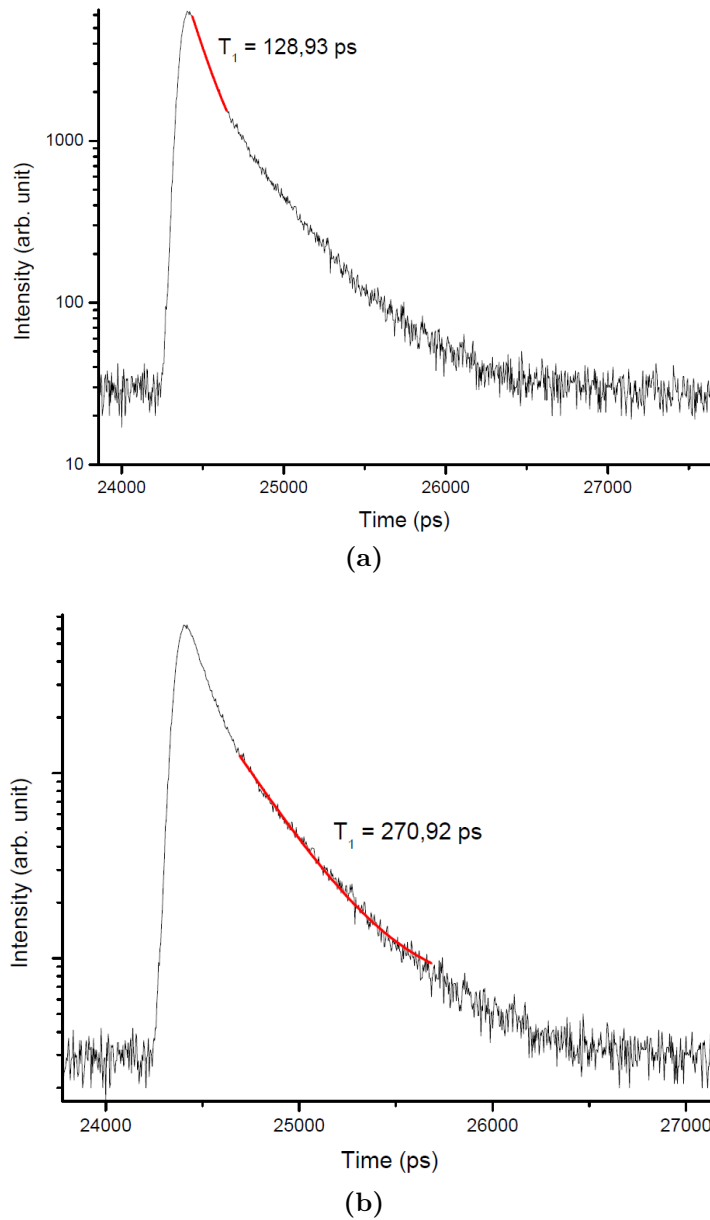


Figure 7.11. Lifetimes of A13 pillar: logarithmic plot the lifetime of A13. It is possible to recognize two curve trend which correspond to real lifetime and fastAPD decay answer. **a)** In log scale, the lifetime of A13 is fitted with linear function. **b)** exponential decrease provided by the fastAPD, in the data analysis we have to take in account also this behaviour and perform the real fit only on a partial set of data.

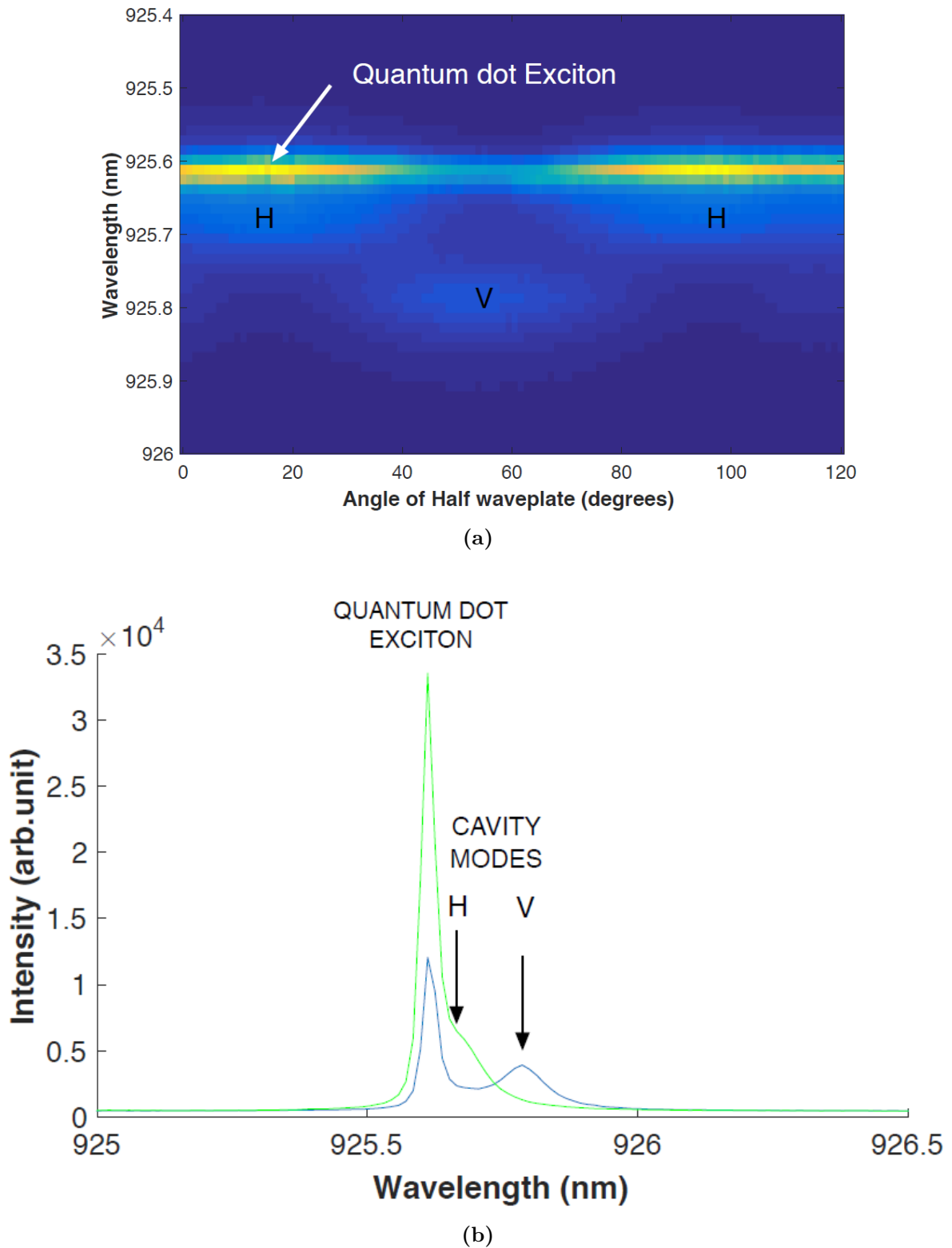
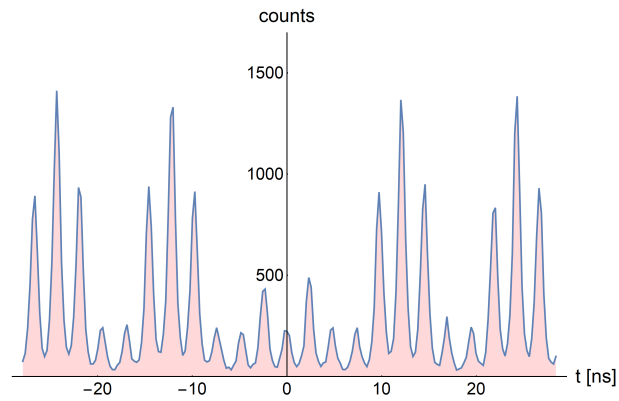
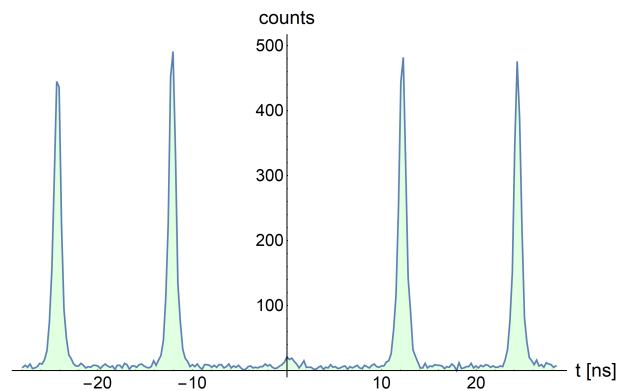
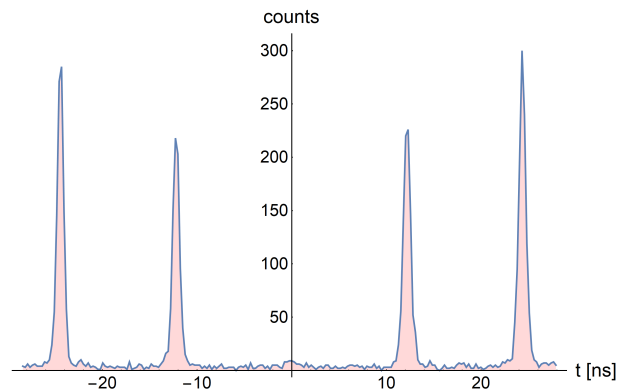


Figure 7.12. Measurements of cavity modes H and V: a) Measurement performed with a scan on the waveplates in order to find the working region (angle of the waveplate) to select H and V mode. b) Analysis of the excitation of the dot to extract informations of the center peaks of the cavity modes and their width. In this way we can choose to perform resonant fluorescence with both cavity modes.



(a) M with MOP (A13)

(b) $g^{(2)}$ (A13)

(c) M with etalon (B18, no MOP)

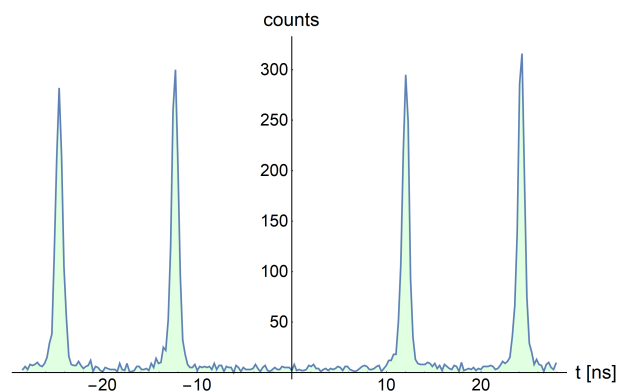
(d) $g^{(2)}$ with etalon (B18)

Figure 7.13. $g^{(2)}$ and M measurements on A13 in resonant fluorescence with and without etalon: **a)** M measurements using MOP. **b)** $g^{(2)}$ measurements without MOP. **c)** and **d)** M (Red) and $g^{(2)}$ (Green) examples measurements on B18 (old micropillar) respectively using an etalon to filter the sideband without employing MOP. All measures have been performed with a standard HBT setup. Quantities have been retrieved from the ratio between central peak area and mean of side peaks areas.

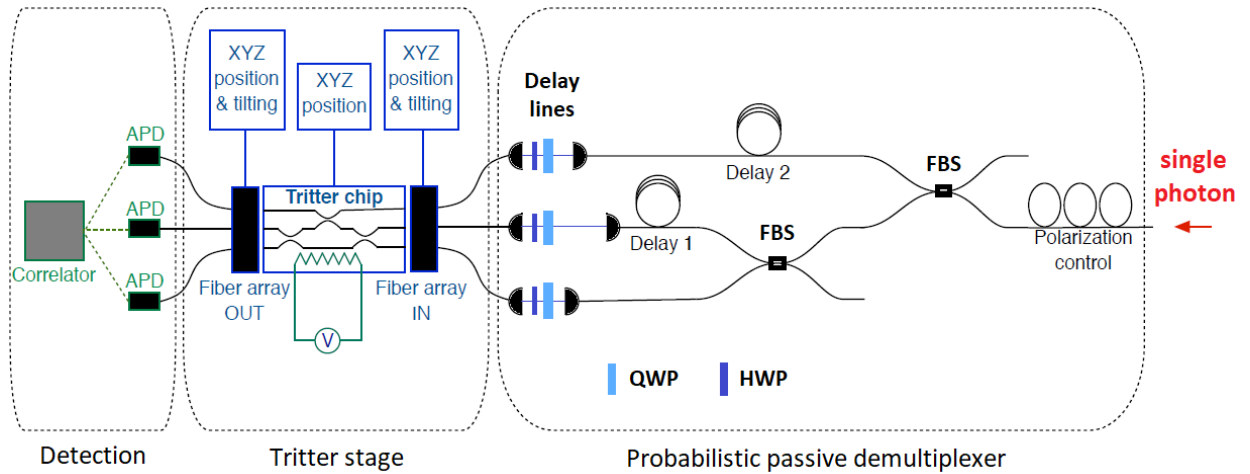


Figure 7.14. Conceptual image of the tritter setup. With a polarization control we are able to change the power of the signal. A 3-photon state is generated with a certain probability by injecting into probabilistic passive demultiplexer (PPD): two fiber beam splitters (FBS) in cascade, delaying with additional fibers (delay 1 and 2) each path (choosing length accordingly to the presence of MOP or not). HWP, QWP and delay lines are used on each path to manipulate polarization and temporal delay with micrometer scale. Fiber arrays single mode (input) and multi mode (output) are aligned with respect to the chosen device in the tritter. With an external voltage control we are able to apply current on the resistance. Single photons data are measured with 3 APDs, each one attached to an output fiber and a correlator (with clock provided by the Tsunami) elaborates the TTL signal writing the result on a computer. Post-processed analysis is performed to retrieve 2,3-fold coincidence measurements.

- efficiency coupling with collimators in delay lines (DL) $\eta_{DL} = 0.9$,
- tritter input/output efficiency $\eta_{trit} = 0.4$,
- detectors efficiency at $\lambda = 925 \text{ nm}$ is $\eta_{APD} = 0.25$.

Analysis of 2-photon correlation of tritter output data: After characterizing losses and properties of our sources, we generate a 3-photon state with the PPD by selecting with delays only 1 of 27 combinations having 3 photons distributing in 3 paths. Referring to Fig. 7.10, if we use MOP we have bunches of three peaks for each pulse of the laser repeated every 12 ns. The relative delays for peaks belonging to each bunch are 0, 2.5, 5 ns. If we consider all possible combinations of clicks between photons distributed to different pulses and paths of the PPD we have: 0, 2.5, 5, 7.5, 10 ns which are labelled as timetags 0,1,2,3,4 respectively. These 3 photons can give correlation peaks. For example, 2 photon arriving at APDs A and B at timetags 1 and 2 will give a count in the peak at +2.5 delay (it is symmetric with respect to negative delays because of different BNC lengths). This configuration makes a complex behaviour at the output pattern of the tritter.

In the following we will briefly describe the model that we employ to analyse our data.

1. Each bunch can give 9 peaks (positive and negative delays), as follows from the description above. Two different bunches can overlap partially with the nearer one.
2. We consider the whole tritter setup (Fig. 7.14) as an implementation of several unitary matrices in cascade. The system is described by 3 waveguides (path 0, 1 and 2). At each gate we apply the beam splitter and phase shifters transformation following how are they

distributed along the path. In other words, for each layer there are matrices

$$U^{BS} = \begin{pmatrix} T & R & 0 \\ R & T & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad U^\phi = \begin{pmatrix} e^{I\phi_1} & 0 & 0 \\ 0 & e^{I\phi_2} & 0 \\ 0 & 0 & e^{I\phi_3} \end{pmatrix} \quad (7.16)$$

where R and T are known and the position of the beam splitter on the matrix is associated with the connection of the paths.

3. We model the degrees of freedom of the three photons, in this way:

$$(a^\dagger)_p^M(t) \quad (7.17)$$

where $p = 0, 1, 2$ is the index of the optical path, $M = i, j, k$ is the *color* index relative to the indistinguishability properties regarding the source (spectra, phase and temporal profile) and $t = 0, 1, 2, 3, 4$ is the timetag. Let's clarify the meaning for M : we say that all three photons are indistinguishable when they have the same index i , we have index j if the second photon is distinguishable respect to the first and index k if the third photon is distinguishable respect to the second. For now, we take into account all possibilities, we will sort out only at very end. If we are interested in characterizing the $g^{(2)}$, we need to introduce, in the same spirit, new colours i_1, j_1, \dots . Then, each contribution is factorized by a $\sqrt{1 - g^{(2)}} + \sqrt{g^{(2)}}$. This means we have a $1 - g^{(2)}$ probability of having the single photon alone, and a $g^{(2)}$ probability of having this single photon mixed with a photon of the laser.

4. We multiply properly each layer described at point (2) according to the whole transformation and we build all the combinations of $C_\sigma = (a^\dagger)_{p1}^{M1}(t1) \cdot (a^\dagger)_{p2}^{M2}(t2) \cdot (a^\dagger)_{p3}^{M3}(t3)$. In this case the first layer is the matrix referred to APDs clicks. At the end we retrieve a 2D matrix X where each element correspond to a particular configuration C_σ of paths, colours and timetags which are distributed according to their collective behaviour (interference or not).
5. From X we select only configurations in which we are interested. We sum over the colors index (because APDs are insensitive to different colors). The simulation sort out combinations by factorizing and simplifying the expressions. We select all the combinations and we assign in specific temporal bins according to the action of APD clicks matrix.
6. This procedure provides the set of peaks for Bunch 1. If we want to see what happens for different bunches, we apply the algorithm on another bunch and we sum peaks at the boundaries of each one in order to take in account also intersections between different bunched contributions. For example Bunch 1 has (2.5, 5, 7.5, 10, 12.5, 15, 17.5, 20 ns) which partially overlap with the Bunch 0.

Performing the simulation for different bunches we found a good agreement between the model that we built and the 2-photon correlation data. The result is reported in Fig. 7.15. We also report a $g^{(3)}$ density plot as a result of 3-photon correlation data between APD A,B and C using coherent laser. In the latter case, the target is to reproduce the experimental conditions and same amount of statistics that we will employ with singles photons. This is just a preliminary step that we did before the measurement with single photons which is still in progress. What we would expect from the single photon measurement is an enhancement of counts at 0 delay for both APDs.

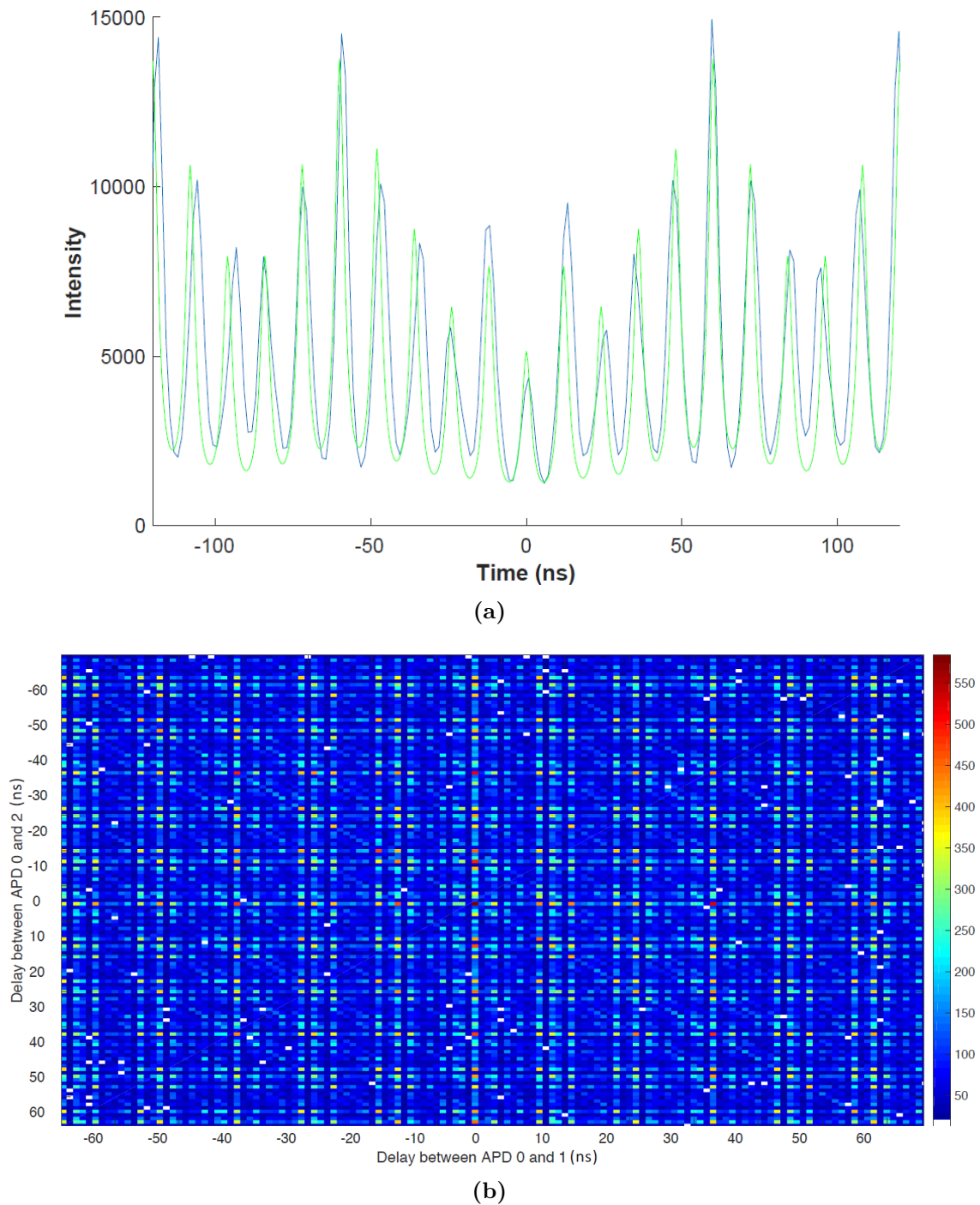


Figure 7.15. 2- photon correlation and preliminary coherent measurements on output data from the tritter device. a) $g^{(2)}$ measurement performed using MOP configuration. Simulation (green) and elaborated data (blue). **b)** $g^{(3)}$ measurement performed using MOP configuration with coherent. In this measure we focused to simulate the final measurement with single photons by setting the same experimental conditions as we will do for single photons. Each line of the 3D plot represent $g^{(2)}$ measurement of APD 1 and 2 with respect APD 0.

7.9 Discussion and future developments

Let us discuss on the comparison between the best SPDC sources and QD systems. Ref. [223] demonstrated how micropillar cavities for quantum dot single photon sources are promising in terms of scalability and indistinguishability. They compared in a brightness versus mean wave-packet overlap graph, the performances of their dot with respect of the state-of-art of SPDC sources, clearly showing an advantage in terms of brightness, given a good indistinguishability. In this experiment we reached $\sim 2 \times 10^6$ Hz only considering the source signal (with MOP we gain a factor 3), while if we consider SPDC-like experiments we obtain an heralded 3-photon state $\sim 100 - 200$ Hz, assuming the same repetition rate and same detector efficiency for both implementations. If we want to build a 3-photon state with the PPD we lose a factor ~ 30 on the rate, due only to the probabilistic gate. If we calculate the ratio between rate of 3-photon state in the case of micropillar systems and SPDC sources, we gain considering QD with respect to SPDC, a factor ~ 300 . The question now is: how does it scale the rate of n -fold photon with the the number of photons? Let us assume the same repetition rate, detector efficiency, the same optical losses in the chip. hence in order to generate n -photon state, with SPDC we have to wait for higher order of pair generation, while with QD system we have to add cascade of FBS for the demultiplexer. Thus, we can distinguish two behaviours in the generation rate R :

$$R_{SPDC} \propto (\eta_{SPDC})^{n/2} \quad R_{QD,PPD} \propto (\eta_{QD,PPD})^n n^{-n} \quad (7.18)$$

where $\eta_{SPDC} \sim 0.015$ is the probability to generate a pair per pulse and $\eta_{QD,PPD} \sim 0.3$ is the probability to generate 1 photon per pulse¹. In Fig. 7.18 is showed the curve trend for the three systems.

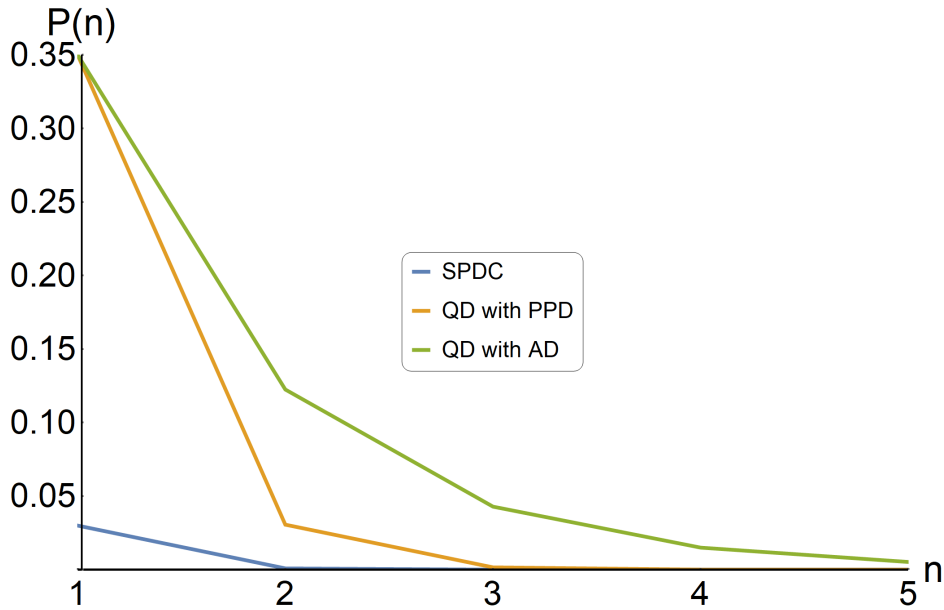


Figure 7.16. Comparison of the performances for large implementations. Probability $P(n)$ to observe a n -photon state at the output of listed sources. Following equations (7.18), SPDC and QD sources with PPD follow the same curve trend with respect to the generation of n -photon states. This is due to the probabilistic generation (for SPDC) and selection (for PPD) that exponentially decrease the $P(n)$. Using an active demultiplexer (AD) we can overlook the exponential factor in PPD approach, thus scaling the generation rate. Simulation takes in account optical losses during the propagation and we assume the same collection efficiency at the output.

¹This efficiency is retrieved by knowing the signal arriving into APD considering reversely the efficiency of detector (0.25), optical (0.9) and coupling losses (0.35) in realistic implementations of micropillar source experiments.

The scaling in the number of photons is linked to the capability of provide a n -photon state. The weakness for scaling the number of photon in QD systems, is referred to the probabilistic gate for which we have only n^{-n} probability to observe the desired synchronized n -photon state, given n paths and n photons. For future tests it should be necessary to upgrade the system changing the probabilistic passive demultiplexer (PPD) to an *Active Demultiplexer* (AD) [262]. In this way, we are able to select actively the n -photon state during the propagation, by tuning reflectivities and switching paths of the fiber beam splitters thus omitting the decreasing exponential probabilistic factor n^{-n} in PPD case.

For future improvements of the setup, we will exploit active demultiplexer and perform further studies on the indistinguishability of our sources trying to solve the problems with cryostation system. At the moment we did not perform other measurements with these condition because we can not correctly select the temperature inside the dot chamber in order to eliminate the contributions of the phonon sideband.

Conclusions and perspectives

The goal to realize quantum computers seem to be nearer than we think. Up to now there are been several improvements not only from the technological point of view (regarding sources [5,8,16,223,259], prototype-calculators [5,6,22,26] and detectors [126]) but also in the development of theories which provides us a better comprehension of the quantum physics associated to this phenomena. In particular it is important to understand the phenomenon of quantum interference which could pave the way for a solid computational scheme. Photonic implementations have demonstrated extreme accuracy and versatility, and the advances in the photonic integration techniques are allowing the realization of more and more complex networks [12]. In this context, Boson Sampling [33] hold the advantage to be the suitable platform to observe this kind of events. In fact if we consider the model proposed by Aaronson in 2011, it has not immediate practical uses, besides a recent proposal using squeezed state to map the problem into another well known in solid state physics which can be used to reconstruct molecular vibronic spectra [263]. Indeed, Boson Sampling is one of the most promising candidate to highlight quantum advantage respect to its classical counterpart with a very simple and feasible platform. In the run of the so-called *quantum advantage* there is still a lot of work to do, in particular by increasing the number of photons and modes. Although several ideas to overcome the problem with different approaches in bulk [147,149] or integrated photonics [22,142,264] have been proposed, however it seems not easy to achieve this target. Moreover, recent studies arose the challenge between classical and quantum, from the boson sampling point view. This because there have been improvements in classical computation [134] and in theory [135] that extend the countdown even further. However, the main interest in Boson Sampling remains in the proof-of-principle of quantum computational advantage, which would have deep implications in many fields.

However, what I would like to transmit with this thesis, is not the importance of Boson Sampling itself, but the very phenomenon at the core of this model. It relies on many-body indistinguishable bosonic interference which is important because we can use it as an optimal platform to find new interesting topics in quantum computation. The task of certifying genuine many-boson interference is thus expected to find numerous applications in photonic quantum information [12], for validating the functioning of Boson Sampling experiments, as a parallelism for quantum simulation [265] and more generally, as a diagnostic tool for quantum optical devices.

In this work we begin with two question: (1) "how to scale our systems" and (2) "how to validate genuine quantum interference efficiently". Regarding point (1) we exploit an advanced setup for random-input generation SPDC sources in parallel and subsequently we focused on analyse and compare the performances of deterministic single photon sources with respect to probabilistic ones. Regarding point (2) we have seen different approaches to certificate true quantum interference reporting our experimental verifications of validation protocol proposing different approaches to the problem exploiting specific unitaries and machine learning techniques. We investigated on the search of the optimal unitary useful for this task and we found out that the best choice for this task among all possible unitaries extractable from the Haar measure are the Fourier and Sylvester matrices. We experimentally observed the validity of the suppression laws, discovering that Sylvester matrices as suitable platforms for Scattershot-like experiments. Furthermore, we have been usage also of machine learning techniques and statistical tools to highlight patterns provided by these classification protocols. We believe our work contributed to

enrich the analysis of how to identify true quantum interference by employing several validation techniques which are suitable for medium-large size systems. Moreover, these results could pave the way for pushing optical community in the search and the development of tools and techniques which focus on how to scale our system and how to certificate interference.

To perform our experiments we exploited femtosecond laser micromachining² as a promising solution for experiments in small size optical chips, as shown in chapter 3, where we fully investigated and described the capabilities and the limitations in terms of implementations. Indeed, the presence of chip losses, summed to not efficient generation and non unity detector efficiency can slow down anyway the scalability of our systems. There are several perspectives to list, but the most important one is to reach the situation in which we are able to merge generation, photonic platforms and detection all together³. For this reason we worked on realize a merged system between quantum dot sources and integrated device observing a significant increase in the count rate. A challenge for the future will be to find a way to develop this approach to a larger number of particles by employing active demultiplexers. Another important requirement, is to develop detectors with efficiency close to the unity, because, as we mentioned in previous chapters, even if we have bright sources and high transmissive optical chips, we still have to deal with the exponential decrease of the rate of simultaneous n-boson state measured with post-selection.

²Devices realized in collaboration with IFN-CNR and Politecnico of Milan.

³This is the main target of the research project QUCHIP in which partially I collaborate: <http://www.quchip.eu/>

Chapter 8

Appendices

8.1 Appendix Note 1: Multiphoton emission terms [5.3]

The main sources of experimental imperfections in the state preparation are due to multi-photon emission terms occurring in **SPDC** processes and to non-unitary indistinguishability of the two photons. Here we perform a numerical analysis on how these two effects lead to the presence of added noise in the suppressed contributions, making harder to violate the validation criteria.

The role of non-unitary indistinguishability in multiphoton interference, and in particular in Boson Sampling, has been discussed in several works [220, 266, 267]. In our works, we model this effect by considering an input state of the form:

$$\rho = \alpha|1, 1\rangle\langle 1, 1| + (1 - \alpha)|1_a, 1_b\rangle\langle 1_a, 1_b|, \quad (\text{A1})$$

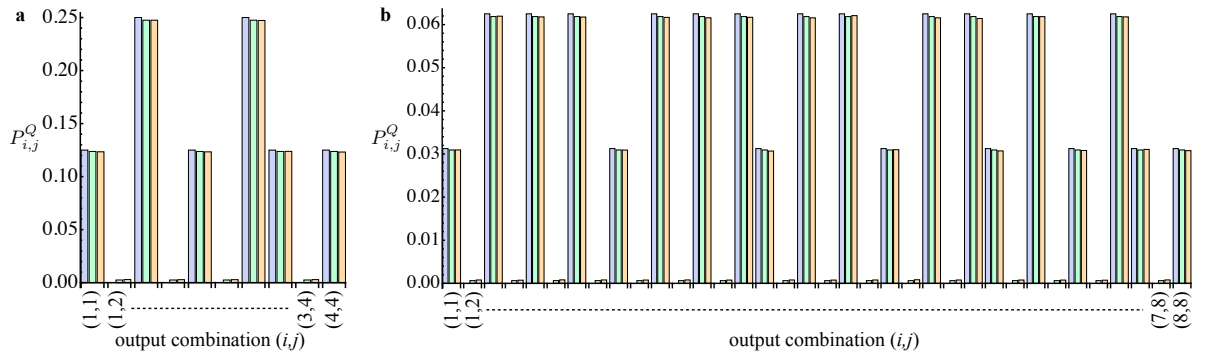
where α represents the degree of overlap of the two-photon state, and $|1_a, 1_b\rangle$ stands for two distinguishable photons. Hence, the output probability distribution is the weighted mixture of the distributions $P_{i,j}^Q$ (indistinguishable photons) and $P_{i,j}^D$ (distinguishable photons).

Multiphoton emission is due to the probabilistic nature of the photon-pair generation in **SPDC** sources. Being g the nonlinear gain of the process, the output of an **SPDC** source in the low gain limit can be written as

$$|\psi\rangle \sim |0, 0\rangle + g|1, 1\rangle + g^2|2, 2\rangle. \quad (\text{A2})$$

Multiphoton emission then leads to the probabilistic injection of input states different from the expected one. Since losses are present in the apparatus and non photon-number resolving detectors are employed, these multiphoton events generated by multiphoton emission cannot be discriminated from the events generated by the correct input state. Furthermore photons belonging to different pairs can present a lower value of distinguishability than photons belonging to the same pair. Note that in the discussed configuration the lower is the distinguishability between different pairs, the higher is the noise contribution due to multiphoton emission. Indeed, for full indistinguishability between different photon pairs the noisy state injected in the cyclic inputs presents the same symmetry of the correct input $|1, 1\rangle$, and thus some suppressed events are present reducing the amount of additional noise.

In Appendix Figure A1 we show the results reported in our work [182] for numerical calculations for the two noise effects for values of α (non-unitary overlap), g (non linear gain), and losses corresponding to our experimental implementation. In our case $\alpha \sim 0.98$ (estimated from the visibility of an Hong-Ou-Mandel experiment in a 50/50 beam-splitter), while $g \sim 0.12$ (estimated from single-photon counts and two-photon coincidences of the source). For the sake of simplicity, losses have been grouped in two main contributions (input and output). The overall input η_{in} and output η_{out} transmissivities, including also coupling efficiency of the source and detection



Appendix Figure A1. Numerical analysis of noise contributions due to imperfect state preparations. Comparison of the output probability distributions for a cyclic input with an ideal two-photon state (blue bars), a two-photon state with non-unit overlap (green bars), and an input state with multiphoton emission terms (orange bars). The unitary adopted is the Fourier matrix. The parameters used in the different cases are described in the text. The distributions for the multiphoton emission terms are obtained from a numerical simulation of the process with 10^6 events. (a) 4-mode interferometer and (b) 8-mode interferometer.

efficiency, were in our case ($\eta_{in} \sim 0.1, \eta_{out} \sim 0.28$ for 4-mode device) and ($\eta_{in} \sim 0.085, \eta_{out} \sim 0.23$ for 8 mode device). Propagation losses in the device have been equally distributed in the two parameters (η_{in} and η_{out}). We observe that the two effects of multiphoton emission and non-unit overlap lead to similar noise contributions, that is, the observation of events in suppressed outputs. Both imperfections in our experiment can thus be modeled with the input state of Eq. (A1), by considering an effective overlap parameter $\alpha_{eff} \sim 0.95$.

8.2 Appendix Note 2: Interferometer design in the 2^p -modes case [5.3]

In the following, we will schematically discuss the procedure to design an interferometer by implementing the quantum Fast Fourier Transform (qFFT) in the general case of $m = 2^p$ modes, for some integer p .

- Each optical mode (waveguide) $k \in [1, m]$ can be numbered according to its binary representation and can be associated to the set

$$(b_1, b_2, \dots, b_p)_k$$

where $b_i = -1$ or $+1$ if the i -th bit of the binary representation of k is 0 or 1 respectively. The ordering of the b_i s is intended to be from the most significant bit b_1 to the least significant bit b_p .

- The p -element vector (b_1, b_2, \dots, b_p) represents the coordinates of the vertices of a p -dimensional hypercube in \mathbb{R}^p .
- The optical qFFT algorithm [1] consists of p steps. The j -th step connects, making them interact, all the couples of modes that differ only for the j -th bit. (Proper phase terms are also added and a final relabelling of the outputs is needed, but this is not relevant here).
- The position of each waveguide in the cross-section plane can be defined by suitably projecting in two dimensions the vertices of the hypercube. Each step of the qFFT corresponds to connecting, by directional couplers, the couples of modes that have to interact in that step. In particular, such connections corresponds to edges of the hypercube with a given direction, all parallel to each other (note that the edges of a p -dimensional hypercube are placed along p possible directions).

- It is worth noting that with this layout, in each step, modes that have to be connected by couplers are all at the same relative distance and the projection of the couplers on the plane are all parallel lines. This means that such couplers can be all identical (except for deformations that may be needed for introducing phase terms) and waveguides never cross. This avoids parasitic coupling between modes that should stay separated and unwanted differences in the optical paths of the mode connections.

8.3 Appendix Note 3: Discrimination from classical light [5.3]

The measured two-photon data can be discriminated from classical light by using a test which involves both collision-free and bunching contributions [266, 267]. We performed this test to discriminate classical light in our work [182]. It is based on the two-mode correlation matrix, which is defined as $\Gamma_{i,j} = \langle a_i^\dagger a_j^\dagger a_j a_i \rangle$. For two-photon input states, this reads $\Gamma_{i,j} = (1 + \delta_{i,j})P_{i,j}^Q$, where $P_{i,j}^Q$ is the output probability. The correlation matrix for classical light must obey the following inequality for all (i, j) :

$$W_{i,j} = \frac{1}{3} \sqrt{\Gamma_{i,i}^{(\text{cl})} \Gamma_{j,j}^{(\text{cl})}} - \Gamma_{i,j}^{(\text{cl})} < 0 \quad (\text{A3})$$

A violation of this inequality indicates that the collected data present correlations which are stronger than those achievable with classical light.

We then applied this test to our experimental data, corresponding to the two cyclic inputs injected in the 4-mode interferometer. The experimental correlation matrix $\tilde{\Gamma}_{i,j}$, shown in Appendix Figure A3a-b, has been retrieved from single-photon measurements and two-photon visibilities. The violation parameters $\tilde{W}_{i,j}$ and the corresponding errors $\sigma_{i,j}$ can then be obtained from $\tilde{\Gamma}_{i,j}$. We observe a violation of inequality (A3) of more than 68 standard deviations, for all suppressed contributions in both cyclic inputs (see Appendix Figure A3c). This shows that the observed correlations cannot be reproduced by classical light.

8.4 Appendix Note 4: Classes for 2-photon output probabilities with Fourier interferometers [5.4]

Due to the symmetries of a Fourier transform it is possible to classify the input states in groups with the same 2-photon output distribution $p_{i,j}^{k,l}$, being (i, j) and (k, l) the input and output 2-photon states respectively. For pairs of indistinguishable photons injected into an interferometer implementing $(U)_{i,j} = u_{i,j}$ we have

$$p_{i,j}^{k,l} \propto |\text{per } c_{i,j}^{k,l}|^2 = |u_{k,i}u_{l,j} + u_{k,j}u_{l,i}|^2 \quad (\text{8.1})$$

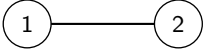
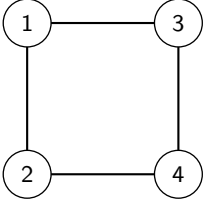
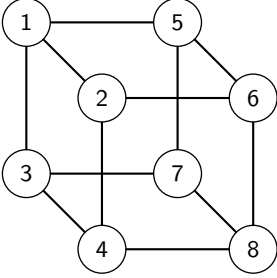
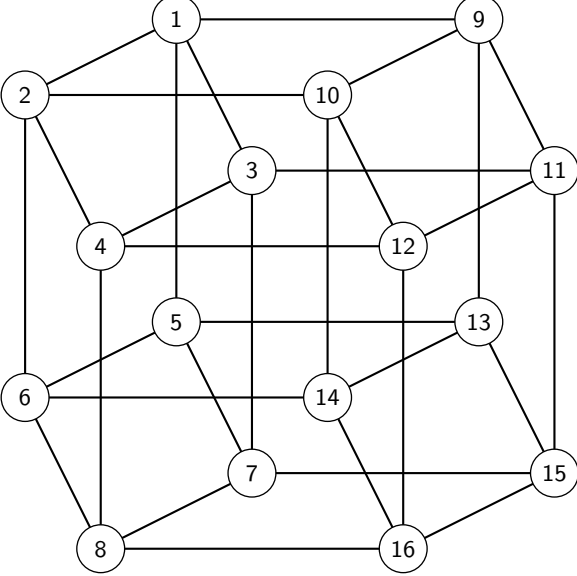
which in the case of a Fourier transformation becomes proportional to the quantity $|e^{2\pi\beta\frac{ki}{d}} e^{2\pi\beta\frac{lj}{d}} + e^{2\pi\beta\frac{kj}{d}} e^{2\pi\beta\frac{li}{d}}|^2$. From here, observing that $|e^{\beta\alpha} + e^{\beta\beta}|^2 \propto \cos^2(\frac{\beta-\alpha}{2})$, we directly find

$$p_{i,j}^{k,l} \propto 1 + \cos\left(\frac{2\pi}{d} (l-k)(j-i)\right) \quad (\text{8.2})$$

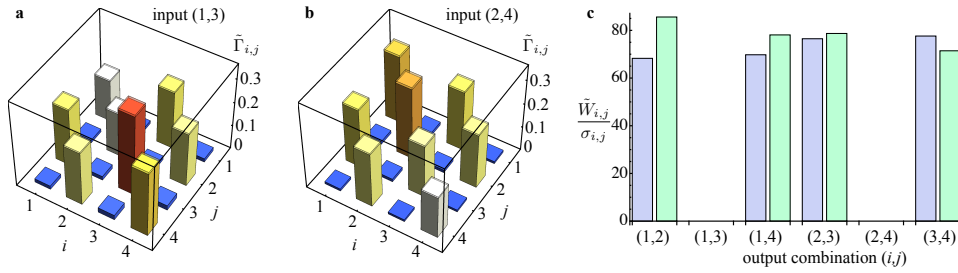
Let us now define the quantities $\Delta_{kl} = |k-l|$ and $\Delta_{ij} = |i-j|$, whence

$$p_{i,j}^{k,l} \propto 1 + \cos\left(\frac{2\pi}{d} \Delta_{kl} \Delta_{ij}\right) \quad (\text{8.3})$$

Suppose that the input modes (i, j) belong to the same class $C_{a,b}$ of the input modes (a, b) , over a set of classes whose cardinality is yet to be determined. Let us also write $\Delta_{ij} = \Delta_{ab} \pm r$, with

Dimension	Planar projection of the hypercube
$p = 1$	
$p = 2$	
$p = 3$	
$p = 4$	

Appendix Figure A2. Planar projections of hypercubes of dimension p up to 4. The numbered vertices correspond to the possible position of the waveguides in the cross-section of a 3D interferometer which implements the quantum Fast Fourier Transform. Each connection between the vertices corresponds to a directional coupler in a given section of the interferometer.



Appendix Figure A3. Discrimination of two-photon data with respect to classical light.

Two-photon correlations $\tilde{\Gamma}_{i,j}$ obtained from experimental data for the two cyclic inputs in the 4-mode interferometer: (a) input modes (1,3) and (b) input modes (2,4). (c) Experimental results for the discrimination test against classical light with the 4-mode interferometers and two-photon input: number of standard deviation for which the parameter $\tilde{W}_{i,j}$ violates Eq. (A3) for each output combination. Blue bars: input modes (1,3). Green bars: input modes (2,4).

$r < d$. We then require, for the two pairs of inputs (i, j) and (a, b) to belong to the same class, that $p_{i,j}^{k,l} = p_{a,b}^{k,l}$, i.e.

$$\cos\left(\frac{2\pi}{d} \Delta_{kl} \Delta_{ij}\right) = \cos\left[\frac{2\pi}{d} \Delta_{kl} (\Delta_{ij} \pm r)\right] \quad (8.4)$$

from which we need to satisfy $\text{mod}(\Delta_{kl} r, d) = 0$. Thus, for each Δ_{kl} there are $\frac{d}{2} + 1$ different classes, corresponding to the values $r = 0, 1, 2, \dots, \frac{d}{2}$. \square

8.5 Appendix Note 5: Majorization and fidelity of the operation [5.4]

We have investigated the quality of the qFFT and the validation protocol for instances with a different strength of majorization. To this aim, we have introduced a new figure of merit for majorization, which captures most of the relevant aspects of a Lorenz diagram:

$$\mathcal{M} = \frac{1}{2(m-1)(n-1)} \sum_{s=1}^{n-1} \sum_{k=1}^{m-1} \delta_k^{(s)} \left(1 + \text{sign}(\delta_k^{(s)})\right) \quad (8.5)$$

where m is the number of states in the quantum register, n is the number of steps in which the algorithm is decomposed and the $\delta_k^{(s)}$ represent the differences between the Lorenz curves at steps s and $s+1$ for a fixed level of imperfections σ . The estimator \mathcal{M} can be easily understood as the arithmetic mean over all the steps and numbers of components involved in the cumulatives. Using the new estimator, we show in Fig. 8.1 how the likeness to the ideal QFT (here judged with the Fidelity F) and the probability of success for the validation protocol (judged by the ratio of successes R_s of correctly witnessing the injection of Fock states) change for increasing values of \mathcal{M}_σ , for as many levels of fabrication imperfections σ . The results in Fig. 8.1 show that, for increasing values of \mathcal{M}_σ , the values of F and R_s increase monotonically, confirm once more the significance to the Fidelity of the two operations for majorization. Therefore, MP is a valuable tool in realistic experimental scenarios for testing whether the efficiency of a quantum algorithm is maintained throughout the whole implementation process.

We have also carried out numerical investigations about the occurrence of majorization in a broader scenario, in the scope of realistic devices involving fabricative noise and more unstructured architectures. As a first case, we checked if the Majorization emerges when a different decomposition of the unitary transformation is employed. More specifically, we considered the Reck decomposition [15] which can be exploited to decomposed a generic Haar random unitary in a sequence of beam-splitters and phases. By picking a large sample of random

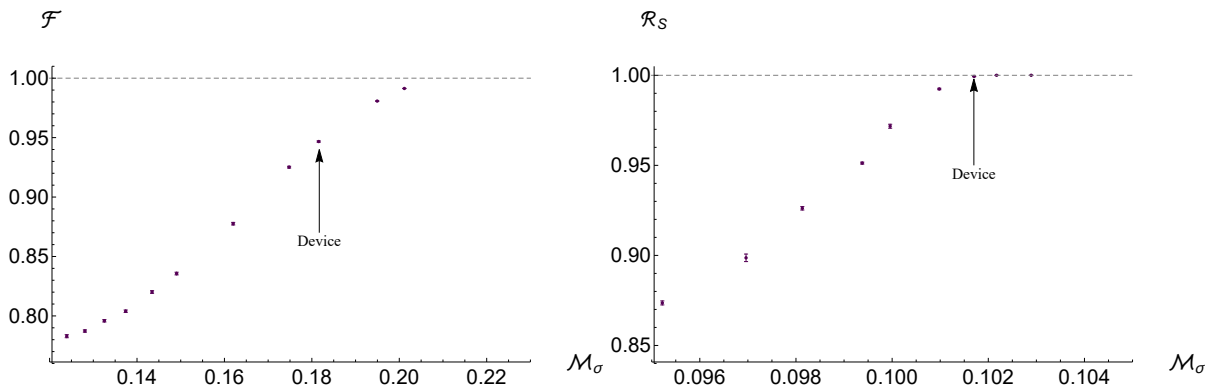


Figure 8.1. Significance to the fidelity of an operation for majorization. The strength of the majorization, estimated by the figure of merit \mathcal{M}_σ , is evaluated at different levels of noise and is put in correlation with the Fidelity \mathcal{F} of the QFT (left) and with the probability of success \mathcal{R}_S for the quantum validation protocol (right). Our devices have an error tolerance of $\sigma=0.05$, which corresponds to $\mathcal{M}_{0.05} = 0.18$ and $\mathcal{F}_{0.05} = 0.95$, which is compatible with the Fidelity measured between our reconstructed transformation and the ideal QFT.

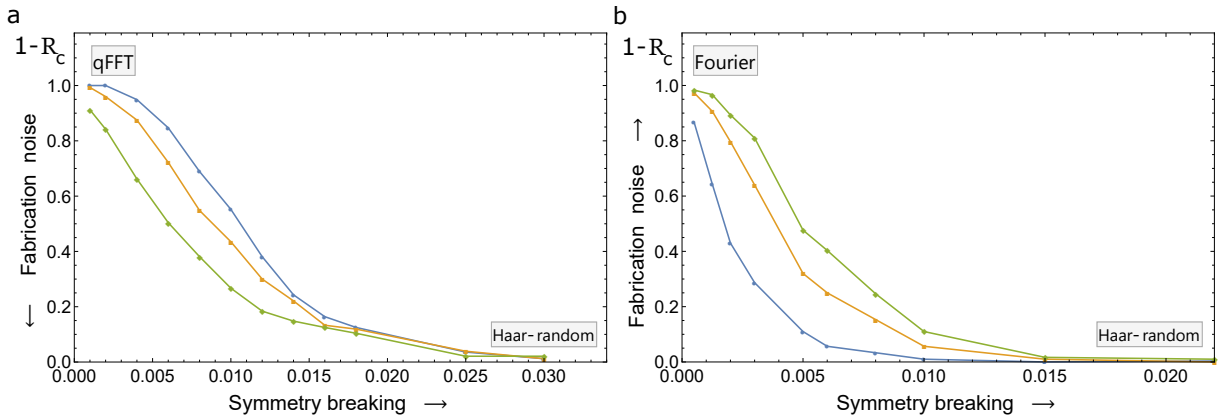


Figure 8.2. Occurrence of majorization in realistic devices. The probability P of having a majorization between the steps of the qFFT (a) and of the Fourier validation protocol (b) in the Barak and Ben-Aryeh scheme ($P = 1 - R_c$) is studied numerically for increasing levels of symmetry of the architecture and of noise in the fabrication process (blue, corresponding to our device and behaving like the ideal qFFT for this plot: $\sigma = 0.05$; orange: $\sigma = 0.1$; green: $\sigma = 0.15$). The X-axis represents the gaussian noise σ admitted on the decomposition of the virtual quantum walk, modeled in the same way as for the fabrication noise. For each curve, corresponding to a different level of fabrication noise, each point is obtained by averaging over 10^3 noisy transformations. This noise leads to a different majorization in the two algorithms, since imperfections can make the probability distributions more peaked (a) or flat (b). For a sufficiently strong degree of decoherence, modelled by the quantum walks introduced between each step, the qFFT behaves more similarly to a Haar-random unitary and no majorization is asymptotically observed in this case.

unitaries from the Haar measure, and by decomposing them according to the Reck decomposition, the Majorization has never been observed for different ways of defining the steps.

To further investigate this aspect, in Fig.8.2 we study the probability of a crossing between any of the partial transformations with a transition from the ideal efficient qFFT to the Reck-decomposed Haar-random unitary, for different levels of fabrication error tolerances (qFFT breaking) and decoherence strength between the steps (symmetry breaking). Here, the fabricative imperfections are modelled with a gaussian noise centered on the ideal value of the internal parameters, while for the progressive disintegration of the high level of symmetry present in the BB scheme we have adopted the following model. We have inserted a weak quantum walk between each step, coupling each mode with all the others so as to introduce a degree of decoherence in the structured architecture. To model the quantum walk we take the identity transformation between each step, decompose it with the Reck scheme and introduce some gaussian noise in the parameters of its beam splitters and phase shifters. Putting back together the decomposed

noisy identity we are able to simulate a quantum walk of tunable intensity. As shown in Fig.8.2, any undesired noise allowed in the architecture breaks the conditions for the observation of majorization constraints (and thus the Majorization Principle). This analysis also confirms the quality of the implementation and the efficacy of the novel 3D architecture adopted for the decomposition.

8.6 Appendix Note 6: Optimal distinguishability for different n and m [5.5]

In this Appendix Note we report results on the optimal interferometer designs we found for the task of testing the hypotheses of distinguishable vs. indistinguishable photons, as we increase the number of photons n and modes m . As in the main text, the figure of merit is the total variation distance between the two distributions. In Appendix Table A1, we report the TVD for combinations of two possible choices. The first is whether we only consider a single input, in which case we choose that input which displays the greater TVD (in order to eliminate ambiguities related to relabeling of the modes), or the average TVD over all possible inputs, which might be more useful for a Scattershot implementation, where one wants to test several different probabilistic sources at once. The second choice is whether we allow collision outcomes (i.e. where multiple photons exit in the same mode) or not. In the latter case, one might wish to compute the TVD between the normalized distributions restricted to coincidence outcomes, but that poses a problem: in some cases (e.g. in the Hong-Ou-Mandel effect) there are **no** collision-free outcomes, and this means that the renormalized distributions, and thus the TVD, are not well-defined. This has the undesired side-effect of creating pathologies in the numerical search. In order to deal with this, for the collision-free situation we actually bin all collision events as a single “no-detection” event in the probability distribution before computing the TVD.

For each combination of these choices and of values of n and m , we also sampled 10000 Haar-random matrices (except $\{m, n\}$ equal to $\{3, 8\}$ and $\{2, 16\}$, where we sampled 2500, and $\{4, 8\}$, where we sampled 100 matrices, due to computational constraints) in order to search for non-Hadamard interferometers that performed better than Sylvester or Fourier matrices. Often the best Haar-random matrix suggested an analytical optimal which we obtained by inspection. These are represented by U_i in the table and discussed below.

Let us now discuss some interesting features of Appendix Table A1:

- (i) If we restrict ourselves to a fixed set of inputs, then for a given number of photons n , adding extra modes does not seem to help. In fact, the best TVD in this case for $m > n$ seems to be obtained by embedding an $n \times n$ matrix in the $m \times m$ one that acts as the identity in the remaining modes. As an example, consider the case $n = 2$ - the best TVD is the same as obtained by the HOM effect, and the corresponding unitary is simply a beam splitter between two of the modes.
- (ii) The same as above naturally does not hold when we average over all inputs. In this case, considering a smaller matrix embedded in the larger one is clearly only optimal for a small fraction of the inputs. A notable example is for the case of 3 photons in 4 modes. In that case, embedding a 3-mode Fourier transform is optimal for the single-input scenario (case U_5), but the optimal interferometer for the average case is U_4 , which consists of a cascade of one 50:50 beam splitter between two modes (say, pair $\{1, 2\}$) followed by two parallel beam splitters (between pairs $\{1, 3\}$ and $\{2, 3\}$).
- (iii) Comparing only the Fourier (F_m) and Sylvester (S_m) matrices, for the $n = 2$ case, we see that if we only look at a single input they are always tied. However, when looking at the average case, the value for the Sylvester is unchanged, but the Fourier decreases. This is a consequence of the fact that, while the Sylvester matrix is symmetric over all inputs, the

m	n	U	Col, Max	No-col, Max	Col, Avg	No-col, Avg
2	2	F_2	0.5	0.5	0.5	0.5
3	2	F_3	0.3333	0.3333	0.3333	0.3333
		U_1	0.5	0.5	0.3333	0.3333
		U_2	0.3951	0.3951	0.3951	0.3951
3	3	F_3	0.6667	0.1111	0.6667	0.1111
		U_3	0.5	0.5	0.5	0.5
4	2	S_4	0.5	0.5	0.5	0.5
		F_4	0.5	0.5	0.3333	0.3333
	3	S_4	0.3125	0.125	0.3125	0.125
		F_4	0.3125	0.125	0.3125	0.125
		U_4	0.5625	0.5	0.5313	0.375
		U_5	0.6667	0.3333	0.4167	0.2778
		U_6	0.5	0.5	0.5	0.5
		S_4	0.75	0.1563	0.75	0.1563
	4	F_4	0.75	0.0938	0.75	0.0938
		U_7	0.5	0.5	0.5	0.5
8	2	S_8	0.5	0.5	0.5	0.5
		F_8	0.5	0.5	0.3153	0.3153
	3	S_8	0.3125	0.2188	0.3125	0.2188
		F_8	0.4112	0.2813	0.3407	0.2545
		U_8	0.6667	0.3333	0.1012	0.0912
		U_9	0.5	0.5	0.0536	0.0536
		U^a	0.4893	0.4311	0.3658	0.3087
		S_8	0.75	0.2813	0.5375	0.275
	4	F_8	0.75	0.3047	0.4337	0.2501
		U_{10}	0.75	0.5	0.2536	0.2013
16	2	S_{16}	0.5	0.5	0.5	0.5
		F_{16}	0.5	0.5	0.3147	0.3147

Appendix Table A1. TVD between distinguishable and indistinguishable photon distributions. Columns indicates the setting: Max (highest TVD over inputs) or Avg (TVD averaged over inputs), and Col (collision outcomes are considered) and No-col (all collision outcomes are gathered into a single no-click event). S_n and F_n are the Sylvester and Fourier matrices of size n , resp., and U_i are other noteworthy matrices as described in the text. Bold numbers correspond to the highest values in that setting and $\{n, m\}$ pair.

^a For these sizes, it was not always possible to identify the best matrices by inspection, so each column might represent a different Haar-random matrix.

Fourier matrix has different TVDs for different sets of inputs, as observed in Fig. 4 in the main text.

The notable matrices in Appendix Table A1 are (up to permutation of the modes) as follows:

U_1 : A 50:50 beam splitter between $\{1, 2\}$ followed by another between $\{2, 3\}$.

U_2 : It is the matrix

$$U_2 = \begin{pmatrix} 1/3 & 2/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \\ 2/3 & 1/3 & -2/3 \end{pmatrix} \quad (8.6)$$

U_3 : A single 50:50 beam splitter, with identity on the remaining mode.

U_4 : A 50:50 beam splitter between $\{1, 2\}$ followed by two parallel ones, between $\{1, 3\}$ and $\{2, 4\}$.

U_5 : A 3-mode Fourier transform, with identity on the remaining mode.

U_6 : Two parallel 50:50 beam splitters, between $\{1, 3\}$ and $\{2, 4\}$.

U_7 : A single 50:50 beam splitter, with identity on the remaining modes.

U_8 : A 3-mode Fourier transform, with identity on the remaining modes.

U_9 : A single 50:50 beam splitter, with identity on the remaining modes.

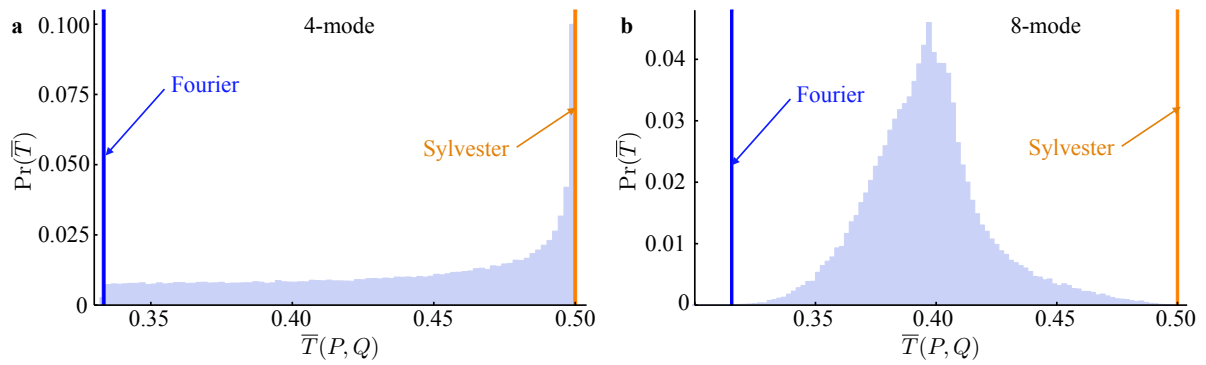
U_{10} : A 4-mode Fourier transform, with identity on the remaining modes.

8.7 Appendix Note 8: Total variation distance for 3D interferometers with fast architecture

Here we discuss the TVD between the distributions with indistinguishable and distinguishable particles when the interferometer is implemented with fast architecture. The layout for the 4-mode and 8-mode devices are shown in Appendix Figure 4.5. We restrict our attention to the case where all directional couplers are symmetric (and thus with transmittivities $\tau_{s,k_s}^{(n)} = 2^{-1/2}$). Sylvester or Fourier interferometers can be recovered by inserting appropriate values for the fabrication phases $\phi_k^{(n)}$.

We performed a numerical simulation of the TVD by generating $N = 10^5$ different unitary transformations according to the Barak and Ben Aryeh architecture, with transmittivities $\tau_{s,k_s}^{(n)} = 2^{-1/2}$ and random, uniformly distributed, phases $\phi_k^{(n)}$ for $n = 2$ photons and $m = 4, 8$ modes. This permits to evaluate the effects of fabrication noise in the output distributions of the implemented interferometers. In the fixed input case, we observe that cyclic inputs always lead to the optimal value $T(P, Q) = 0.5$, independently of the values of phases $\phi_k^{(n)}$. Fabrication errors in the optical phases within the interferometers will thus not affect the obtained results in this scenario. In the multiple input configuration, on the other hand, we observe a different scenario. In the 4-mode case, the maximum and the minimum values of the TVD are obtained respectively for the Sylvester and the Fourier transformation (see Appendix Figure A4a). In the 8-mode case, the maximum is still obtained for the Sylvester interferometer, while a (very) small subset of unitaries presents a lower value of the TVD with respect the Fourier transformation (see Appendix Figure A4b). For the $N = 10^5$ set of randomly generated unitaries, the obtained minimum is $\bar{T}(P, Q) \sim 0.298$, while the value for the Fourier is $\bar{T}(P, Q) \sim 0.315$.

Our simulations thus indicate that fabrication noise in the implementation of Fourier interferometers will tend to result in an increased value of the TVD, with respect to the ideal Fourier transformation. As we have seen in the main text, this was indeed what was observed experimentally in the two devices we characterized.



Appendix Figure A4. Total variation distance in the multiple input configuration for 3D interferometers with fast architecture, with transmittivities $\tau_{s,k_s}^{(n)} = 2^{-1/2}$ and random, uniformly distributed, phases $\phi_k^{(n)}$. Histograms of the TVD for $N = 10^5$ unitaries for **a** $n = 2, m = 4$ and **b** $n = 2, m = 8$. Orange vertical bar: Sylvester interferometer. Blue vertical bar: Fourier interferometer.

8.8 Appendix Note 9: Bayesian validation test

The aim is to discriminate between two hypotheses Q , corresponding to a Boson Sampling device and P , corresponding to the evolution obtained with distinguishable particles. To this end, it is possible to employ a Bayesian test as shown in [200]. At the initial stage, no a priori knowledge is assumed on the system. This translates into uniform prior probabilities $\Pr(Q) = \Pr(P) = 0.5$. After the measurement of N_{events} events, the ratio between the conditional probabilities associated to the two hypotheses is updated according to Bayes' rule:

$$R = \frac{\Pr(Q|N_{\text{events}})}{\Pr(P|N_{\text{events}})} = \prod_{i=1}^{N_{\text{events}}} \left(\frac{q_i}{p_i} \right), \quad (8.7)$$

where x is the index of each sampled event, q_x and p_x are the corresponding probabilities associated to the two hypotheses Q and P respectively. Thus, the confidence probability for hypothesis Q of indistinguishable particles after N_{events} reads:

$$\Pr(Q|N_{\text{events}}) = \frac{1}{\mathcal{N}} \prod_{i=1}^{N_{\text{events}}} \left(\frac{q_i}{p_i} \right), \quad (8.8)$$

where \mathcal{N} is a normalization constant.

This test can be applied in two different scenarios, namely when the incoming data are generated by a Boson Sampling device (ind) or by distinguishable particles (dis). When the incoming data are generated by indistinguishable particles, $\Pr_{\text{ind}}(Q|N_{\text{events}})$ represents the conditional probability associated to hypothesis Q (that is, the correct guess), while $\Pr_{\text{ind}}(P|N_{\text{events}})$ represents the probability associated to hypothesis P (that is, the wrong guess). Analogous definitions are obtained when the data are generated by distinguishable particles, where $\Pr_{\text{dis}}(Q|N_{\text{events}})$ corresponds to the wrong guess and $\Pr_{\text{ind}}(Q|N_{\text{events}})$ to the correct guess.

The confidence probability of the test can be then defined as:

$$P_{\text{conf}} = \frac{1}{2} [\Pr_{\text{ind}}(Q|N_{\text{events}}) + \Pr_{\text{dis}}(P|N_{\text{events}})]. \quad (8.9)$$

Conversely, $P_{\text{err}} = 1 - P_{\text{conf}}$ represents the average error probability of the test.

8.9 Appendix Note 10: Bayesian test on the experimental data

Here we provide more details on the application of the Bayesian validation test on the experimental data. For each input state, $M_{\text{sample}} = 10^4$ distinct samples were generated numerically

from the experimentally measured two-photon distributions with indistinguishable photons and distinguishable particles, the latter obtained by applying a suitable time delay $\Delta\tau$ between the photons larger than their coherence time. For each data sample, Gaussian noise was introduced in the distribution according to the experimental errors. Then, for each data sample the Bayesian test was applied using event data sets of increasing size N_{events} . In the 4-mode case, the full set of input states have been measured in the no-collision subspace, and the cumulative contribution of collision events has been included in an extra bin (see Section “Measurement of the TVD” of main text). The confidence probability P_{conf} is thus averaged over all input configurations. In the 8-mode case, three input states have been measured by detecting also the collision events. For the Sylvester interferometer, all input states are equivalent, and thus an average with uniform weights has been performed.

8.10 Appendix Note 11: Partial photon distinguishability in the binary Bayesian test

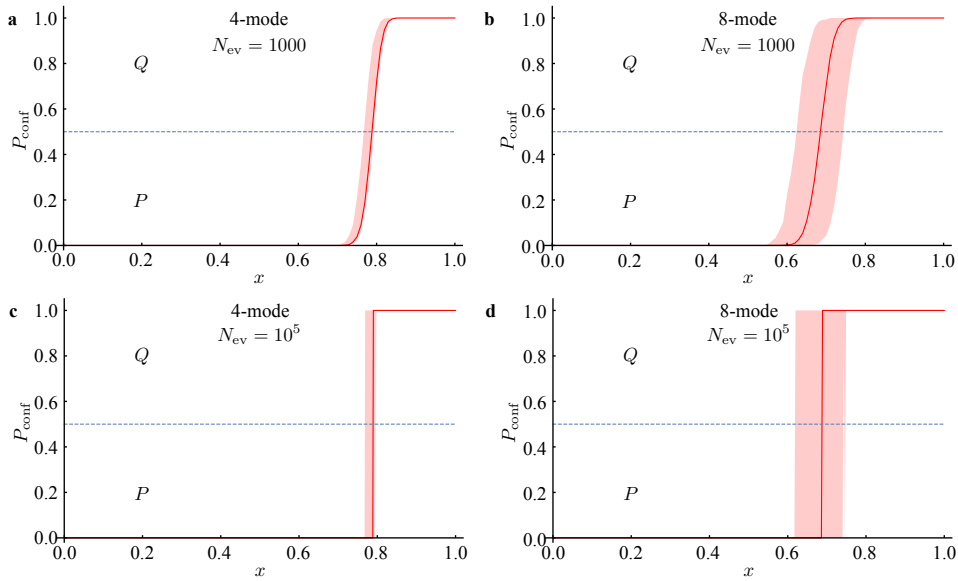
Here we discuss the role of partial photon distinguishability in the Bayesian test when discriminating between hypotheses P and Q . More specifically, we analyzed the case when the experimental data are generated from photon sources with partial photon distinguishability. In the two-photon case, this means that the actual two-photon distribution can be approximated as a convex combination $h_i(x) = xq_i + (1-x)p_i$, where $x \in [0, 1]$ is a real parameter quantifying the degree of indistinguishability between the photons ($x = 1$ corresponds to the perfect indistinguishable case). We keep the Q (P) hypotheses to be the distributions with perfectly indistinguishable (distinguishable) photons, while we allow partial distinguishability on the data to determine the threshold on the indistinguishability x below which the Bayesian test inverts its outcome. To this end, we performed some numerical simulations shown in Appendix Figure A5 for the 4-mode and 8-mode implemented Sylvester interferometers by applying the test to numerically generated data samples calculated from the reconstructed unitaries $\tilde{U}_{\text{fs}}^{(4)}$ and $\tilde{U}_{\text{fs}}^{(8)}$. We then evaluated the asymptotic confidence probability P_{conf} after collecting a large number of events ($N_{\text{ev}} \sim 1000$), as a function of the parameter x . When $P_{\text{conf}} > 0.5$ ($P_{\text{conf}} < 0.5$), the full experiment is attributed to hypothesis Q (P). We observe that, in the 4-mode case, data with $x > 0.788$ are still assigned to indistinguishable particles even in presence of partial distinguishability ($x > 0.685$ in the 8-mode case).

8.11 Appendix Note 12: Bayesian test on scattershot Boson Sampling experimental data

In this Appendix Note we describe the Bayesian inference method applied in the main text on the scattershot Boson Sampling data, and describe a related alternative approach based on binary likelihood ratio tests.

Bayesian inference to ascertain photon indistinguishability

Bayesian inference has been applied in the main text on the scattershot Boson Sampling data. More specifically, the binary set of hypotheses Q and P is replaced by the convex combination $h_i(x) = xq_i + (1-x)p_i$ described in the sequent section. Thus, experimental data are assumed to be described by the combination h_i for a given value x_{true} of the indistinguishability parameter x . Then, the value of x for the collected data is retrieved from a given data sample by performing single-parameter Bayesian estimation. Namely, a uniform prior $\mathcal{P}(x)$ is assumed for x , quantifying the total a-priori ignorance on the parameter. The conditional probability of x after N measured



Appendix Figure A5. Confidence probability P_{conf} as a function of the photon distinguishability x , averaged over 10^4 distinct samples generated from the distribution calculated from the reconstructed unitary transformations $\tilde{U}_{\text{fs}}^{(4)}$ and $\tilde{U}_{\text{fs}}^{(8)}$. After $N_{\text{ev}} \sim 1000$ events: **a**, 4-mode Sylvester interferometer and **b**, 8-mode Sylvester interferometer. After $N_{\text{ev}} \sim 10^5$ events: **c**, 4-mode Sylvester interferometer and **d**, 8-mode Sylvester interferometer. Shaded regions: fixed input configuration, interval comprising the curves for each individual state. Solid Lines: multiple input configuration.

events is obtained from the Bayes rule:

$$\mathcal{P}(x|i_1, \dots, i_N) = \frac{\mathcal{P}(i_1, \dots, i_N|x)\mathcal{P}(x)}{\mathcal{N}}, \quad (8.10)$$

where (i_1, \dots, i_N) is the recorded data sample, $\mathcal{P}(i_1, \dots, i_N|x)$ is the conditional probability of obtaining the data sample for a given value of x , and \mathcal{N} is a normalization constant. The estimated value of the parameter (and its associated error) after N measured events are retrieved from the distribution $\mathcal{P}(x|i_1, \dots, i_N)$ as:

$$x_{\text{est}} = \int_0^1 x \mathcal{P}(x|i_1, \dots, i_N) dx, \quad (8.11)$$

and

$$\sigma_{\text{est}} = \left[\int_0^1 (x - x_{\text{est}})^2 \mathcal{P}(x|i_1, \dots, i_N) dx \right]^{1/2}. \quad (8.12)$$

The method has been applied on scattershot Boson Sampling experimental data, and the results are shown in Fig. 5.13 of the main text. More specifically, starting from the overall collected data sample of ~ 17000 events, we generated numerically 100 different data sequences, in order to evaluate the variability of the test when the data sequence is changed. By using the overall data set, the final estimated value of x is $\tilde{x}_{\text{est}} = 0.738 \pm 0.004$, which is compatible with the one retrieved from Hong-Ou-Mandel interference with two of the SPDC sources in a 50/50 beam-splitter: $x_{\text{HOM}} = 0.79 \pm 0.06$.

Convex hypothesis test with likelihood ratios

A different approach based on binary likelihood ratio tests can be employed to identify the value of x that better describes the experimental data. As in the previous case, the hypothesis is that data are described by the convex combination $h_i(x) = xq_i + (1-x)p_i$ for some value x_{true} of x .

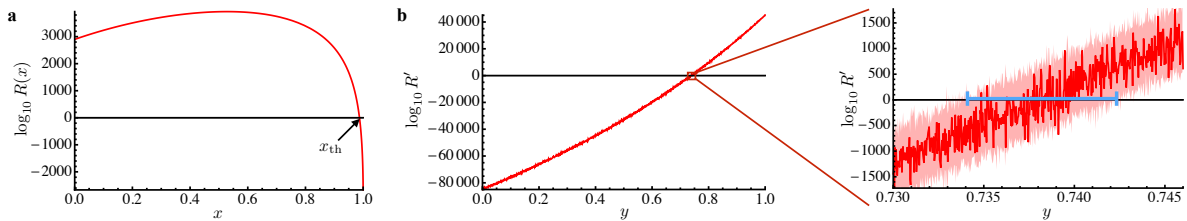
This approach relies on the likelihood ratio test described in Appendix Note 4, and is divided in two stages.

Stage A. The first stage of the test exploits experimental data to identify a threshold value x_{th} where the binary decision test becomes ambiguous. More specifically, let us consider the case where the experimental data pass the binary test for the hypothesis Q . Then, the test is performed with the same data by fixing the hypothesis Q and by using as alternative hypothesis $H(x)$. The test is repeated by tuning the value of x until the ratio $R(x)$ between the hypotheses Q and $H(x)$ lead to an ambiguous results:

$$R(x) = \prod_{i=1}^{N_{\text{events}}} \frac{q_i}{xq_i + (1-x)p_i}. \quad (8.13)$$

The value of x_{th} then corresponds $R(x_{\text{th}}) = 1$. Analogously, if experimental data pass the test for the hypothesis P the test is performed between P and $H(x)$ thus evaluating the ratio:

$$R(x) = \prod_{i=1}^{N_{\text{events}}} \frac{xq_i + (1-x)p_i}{p_i}. \quad (8.14)$$



Appendix Figure A6. Evaluation of the indistinguishability x through the convex hypothesis test based on the likelihood ratio approach. **a**, First stage of the test with the experimental data to determine the value of x_{th} . **b**, Second stage of the test with numerically simulated data. Left figure: plot of $\log_{10} R'$ as a function of y (indistinguishability of the numerically generated) data for the full interval $y \in [0; 1]$. Right figure: highlight in an interval close to $R' = 1$. Solid line: variability of R' as a function of y when each point is evaluated from a single-data set. Shaded region: interval for R' obtained by generating 100 distinct data samples for each value of y . The interval included within the blue ruler identifies the interval for the estimate y' .

Stage B. The second stage employs simulated data from the unitary transformation. The test is now performed by keeping fixed the two hypothesis to Q and $H(x_{\text{th}})$ (or P and $H(x_{\text{th}})$ if the data pass the initial test for P). Numerically simulated data sample of size N_{sim} are generated from the distribution $H(y)$ for different value of y , and the binary likelihood ratio test is performed by evaluating:

$$R' = \prod_{i=1}^{N_{\text{sim}}} \frac{q_i}{x_{\text{th}}q_i + (1-x_{\text{th}})p_i} \text{ or } R' = \prod_{i=1}^{N_{\text{sim}}} \frac{x_{\text{th}}q_i + (1-x_{\text{th}})p_i}{p_i} \quad (8.15)$$

This procedure is repeated by tuning y until a value y' is obtained for which the test is ambiguous ($R' = 1$). This value of y' is an estimate of the parameter y . Intuitively, this second stage determines the value of the photon indistinguishability that leads to an ambiguous result for the same value of x_{th} obtained from experimental data.

We have then applied the test to the experimentally measured scattershot Boson Sampling data. In Stage A, we obtained $x_{\text{th}} \simeq 0.9872$ by using the full data sample (see Appendix Figure A6a). In Stage B, we performed the likelihood test between hypotheses P and $H(x_{\text{th}})$ by generating samples of $N_{\text{sim}} = 10^5$ data from the distributions $H(y)$ for different values of y . The distributions p_i and q_i for the convex combination are calculated from the reconstructed unitary

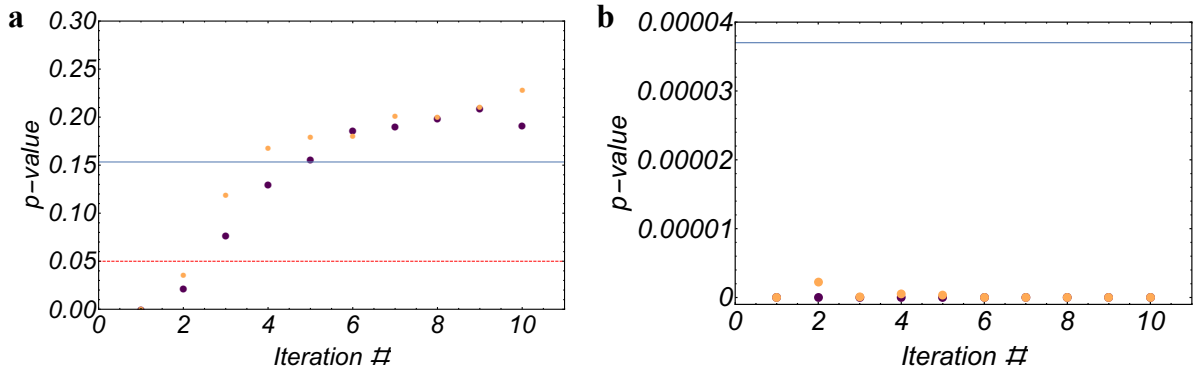
transformation $\tilde{U}_{\text{IS}}^{(4)}$. To give an estimate of the error associated to y' and take into account the effect of using a finite number of samples, in Stage B we generated 100 distinct data samples of size N_{sim} for each value of y . This allows to obtain an interval for the estimate y' . By naming $y' = x_{\text{LR}}$, we obtain $x_{\text{LR}} \in [0.734; 0.742]$ (see Appendix Figure A6b). This interval is compatible with the value \tilde{x}_{est} obtained via Bayesian inference.

8.12 Appendix Note 13: Efficiency of the different pattern recognition techniques [6.1]

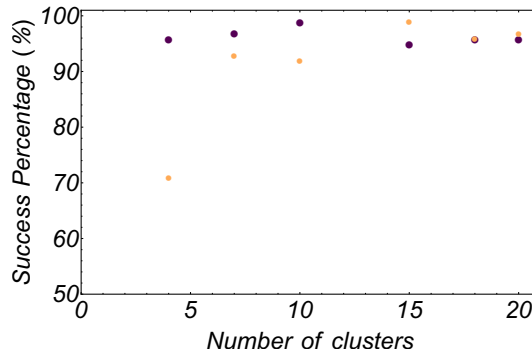
From Appendix Note 13 to 16 we will discuss the pattern recognition techniques used in our work to validate Boson Sampling data. Specifically, the main strength of *K-means clustering* [189] lies in the ability of the algorithm to learn, through a certain number of iterations, a locally optimal cluster structure for the experimental data to be analyzed. This feature is clearly highlighted when we compare the performances of the compatibility test, that is the resulting p-values, when we adopt *K-means clustering* and *bubble clustering*. In Appendix Fig. A7 a we compare two incompatible samples, i.e. one drawn from a trustworthy Boson Sampler and one from a Boson Sampler fed with a distinguishable three photon input. In Appendix Fig. A7 b we instead compare two compatible samples, both drawn from a trustworthy Boson Sampler. *Bubble clustering* is not iterative and the resulting cluster structure is deterministic for a given data set. Indeed, the algorithm chooses the cluster centers accordingly to the observation frequency of the output states and assigns the elements only once. On the other hand, *K-means* is not deterministic, since the final cluster structure depends on the centroids initialization, and, through the selection of the centroids and the corresponding assigned elements, it is guaranteed to locally optimize its objective function, $\frac{1}{N} \sum_{j=1}^k \sum_{i=1}^{n_k} d(e_{ij}, c_j)$ [190]. Having considered this feature of *K-means*, we performed 200 tests both on the two compatible samples and on the incompatible ones, using 100 different uniformly drawn random initializations, and 100 *K-means++* initializations. The pattern recognition techniques based on *K-means*, given its objective function convergence to local minima property, improve the test's result after each iteration, being able to recognize whether the samples are compatible or not with better results than *bubble clustering*, that is, leading to a smaller p-value, with incompatible samples and a greater p-value with compatible samples. The p-value quantifies the significance of the χ^2 test result, since it gives the probability of obtaining a greater value of the χ^2 variable, if the null hypothesis of compatible samples is true.

8.13 Appendix Note 14: Finding the optimal number of clusters for K-Means [6.1]

All clustering techniques investigated in the present study are characterized by several parameters that have to be tuned to ensure the right operation for the algorithm. *K-means clustering*, in particular, requires the user to set the number of clusters forming the structure. The approach we used to set this parameter in the training stage consisted in multiple trials of the compatibility test on sets of the same size and varying the number of clusters. The success percentage of the test increases significantly as the number of cluster is raised, especially in the case of incompatible samples, as shown in Appendix Fig. A8. The increase in the number of clusters was halted by requiring that clusters are composed of at least 5 elements, to ensure the correct operation of the compatibility test (that requires the evaluation of a χ^2 variable).



Appendix Figure A7. *K-means clustering vs bubble clustering.* **a)**, P-values obtained by the application of our compatibility test on two compatible samples of 2000 events, both drawn from a trustworthy Boson sampler, with $N=3$ and $m=13$. The blue horizontal line indicates the result obtained by using the cluster structure given by *bubble clustering*. The darker and lighter dots represent the p-values obtained respectively from the application of the test on the cluster structures obtained by *K-means*, initialized by uniformly drawn random centroids and by *K-means++*, at each iteration. Since the cluster structure obtained by v is not deterministic, we performed the mean on the p-values corresponding to 100 different inicializations. The red dashed line indicates the p-value above which the two samples are recognized as compatible. **b)**, P-values obtained by the application of our compatibility test on two incompatible samples of 2000 events, one drawn from a trustworthy Boson sampler and the other from a Boson Sampler fed with a distinguishable photon input. The blue horizontal line indicates the result obtained via *bubble clustering*. The darker and lighter dots represent, respectively, the p-values obtained by *K-means*, initialized by uniformly drawn random centroids and via *K-means++* inicialization. Since the cluster structure obtained by *K-means* is not deterministic, we performed the mean on the p-values corresponding to 100 different inicializations.



Appendix Figure A8. Success percentages vs number of clusters (k). Success Percentages of the compatibility test, with *K-means clustering*, obtained by comparing numerically generated samples of 1000 events with $N=3$ and $m=13$ drawn from the Boson Sampling distribution and from the distribution given by distinguishable photon inputs, as a function of the number of cluster. The darker and lighter dots represent the percentages corresponding respectively to compatible and incompatible samples. The percentages here indicated are evaluated considering 5% as significance level for the χ^2 test.

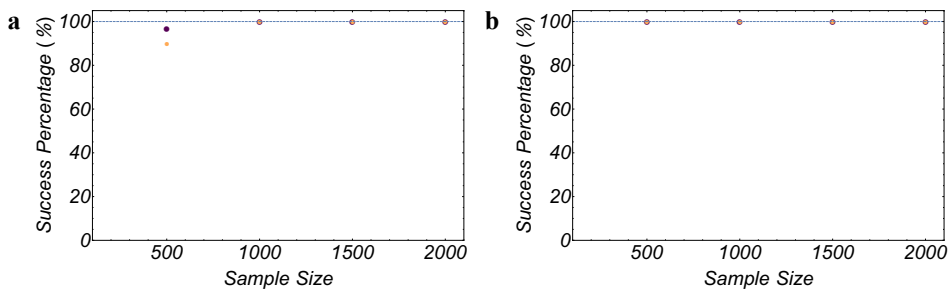
8.14 Appendix Note 15: Halting condition for hierarchical clustering [6.1]

The choice of the halting condition for *Hierarchical clustering* attempts to balance two conflicting requisites: the number of clusters and their minimum size. Firstly, there is the need to have clusters with at least 5 elements, to make the χ^2 test meaningful. However, the minimum size is not the optimal choice as halting condition, since there is no control over the number of clusters and there is the risk that the algorithm stops with too few clusters remaining. An excessively low

number of clusters, less than 3, can compromise the correct operation of the test. We therefore chose to remove outlier elements in the data, neglecting them in the final cluster structure. This is done by requiring that the algorithm stops when a predetermined fraction of the elements belongs to clusters with less than 5 elements. The fraction was then tuned to maximize the efficiency of the validation test, while monitoring the amount of data taken away.

8.15 Appendix Note 16: Reshuffling of the experimental data [6.1]

Our evaluation of the experimental success percentages of the validation test uses 100 sets of three data samples, two compatible and one incompatible. Having such a number of data allows to perform 100 compatibility tests between compatible samples and 100 between incompatible samples. The required sample size to ensure more than 0.8 probability to recognize two incompatible samples is 500 events. This sample size implies that, for each input mode occupation list, we would need at least 5×10^4 events drawn from the distribution with distinguishable photon input and 10×10^5 events drawn from the corresponding Boson sampling distribution. Since the amount of experimental data available was not sufficient, we adopted a reshuffling of the experimental data. To this purpose, starting from a set of N_{exp} output events sampled experimentally, we picked randomly the number of events corresponding to the required sample size N_e among the available data, as long as $N_e < N_{exp}$. This approach may bring a small bias in the results. To quantify this bias, we performed a numerical simulation, comparing the trends of the success percentage as a function of the sample size, with and without reshuffling. The trend is the same, but the growth is slower. Specifically, as reported in Appendix Figure A9, numerical simulations show that, for the investigated sample sizes, reshuffling of data does not affect the result of the validation algorithm adopting *K-means clustering*, initialized with *K-means++*. This analysis implies that the success percentages obtained experimentally and showed in Fig. 3 b and Fig. 4 of the main text are reliable.



Appendix Figure A9. Reshuffling of the data. **a** Success Percentages of the compatibility test performed on incompatible samples, with $N=3$ and $m=13$, with *K-means clustering* initialized with *K-means++*. The lighter and darker dots represent the percentages corresponding respectively to reshuffled and to not reshuffled data. **b** Success Percentages of the compatibility test on incompatible samples, with $N=3$ and $m=13$, adopting *K-means clustering* initialized with *K-means++*, with majority voting. The lighter and darker dots represent the percentages corresponding respectively to reshuffled and to not reshuffled data. The percentages here indicated are evaluated considering 5% as significance level for the χ^2 test.

8.16 Appendix Note 17: Assigning new data with machine learning techniques [6.2]

Machine learning techniques can be employed to assign new data to one of the two classes of the statistical test, namely indistinguishable photons (class *B*) or distinguishable particles (class *D*).

We have here verified the performances obtained with three different techniques in the $NM-CV$ plane, namely (i) *nearest centroid*, (ii) *k-nearest neighbor* and (iii) *support vector machines* with a linear classifier. (i) The nearest centroid method assigns the new datum to the class with lowest distance d between the datum and the class centroid. The latter are calculated (efficiently) by employing the random matrix theory results reported in Ref. [210]. Note that the nearest centroid method requires the calculation of only two distances for each new datum to be classified, and thus the method is computationally efficient. (ii) The *k-nearest neighbor* method exploits the shape of the points distribution for the two classes. More specifically, given a training set of N_{training} points for each of the two classes, the distance d is calculated between the new datum and all points of the training set for B and D and sorted in increasing order. Majority voting is applied to the k lowest distance values, where k is a parameter that can be chosen by the user. At variance with the previous technique (i), $2N_{\text{training}}$ distances have to be evaluated for each new datum to be assigned. Furthermore, the parameter k can be optimized to minimize the error probability. (iii) A support vector machine is constructed to find the optimal way to divide the parameter space in two regions corresponding to the two classes. As for technique (ii), a training set of N_{training} points for both class D and B . In this case, we employed a linear classifier which searches for the optimal hyperplane separating the two regions. The new datum is then assigned according the determined separation. In all cases (i)-(iii), the figure of merit is the error probability P_{err} , defined as $P_{\text{err}} = (P_{B \rightarrow D} + P_{D \rightarrow B})/2$. Here, $P_{B \rightarrow D}$ is the probability to wrongly assign a datum corresponding to indistinguishable photons to the class D , and conversely for $P_{D \rightarrow B}$.

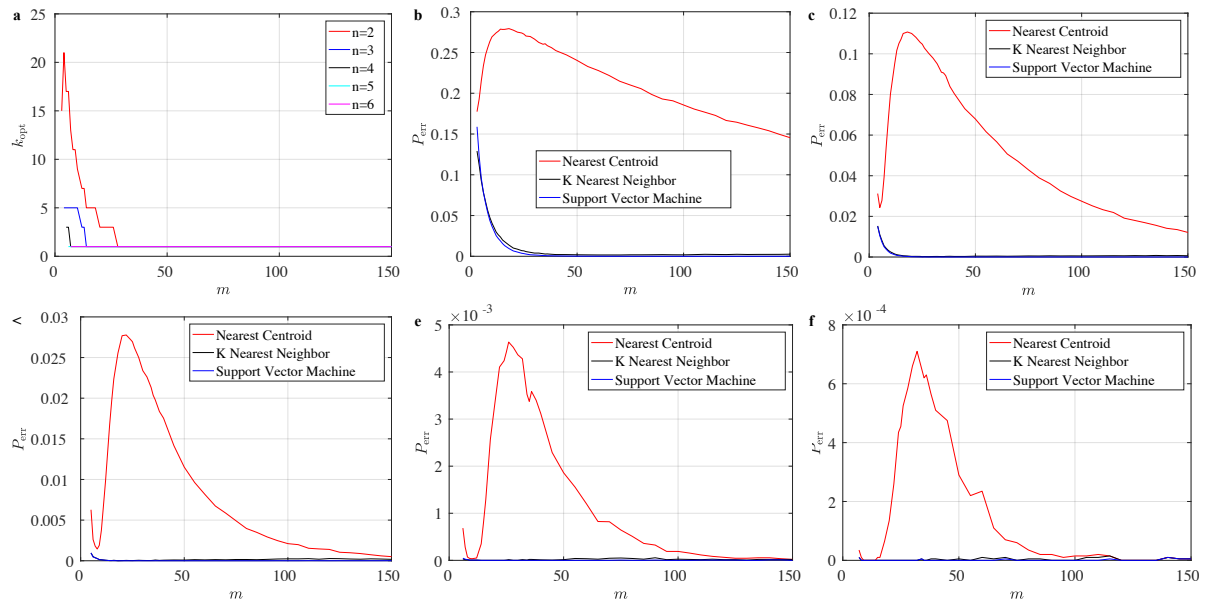


Figure 8.3. **a** Optimal k value obtained from numerical simulations for the *k-nearest neighbour* method. **b-f** Error probability obtained for the three methods for **b**, $n = 2$, **c**, $n = 3$, **d**, $n = 4$, **e**, $n = 5$, **f**, $n = 6$. Red points: *nearest centroid* approach. Black points: *k-nearest neighbour*. Blue points: *support vector machine* with linear classifier.

We then performed numerical simulations for each of the three techniques (i)-(iii). For the *nearest centroid* method, we numerically generated $N_s = 100000$ Haar random matrix, evaluated the statistical quantities (NM, CV) from the two-mode correlators $C_{i,j}$ for both indistinguishable photons and distinguishable particles, and then estimated the average error probability P_{err} . For the *k-nearest neighbour* and the *support vector machine*, the simulations have been performed by generating $N_r = 100$ different training sets of size $N_{\text{training}} = 1000$, and for each training set the error probability has been estimated from $N_s = 1000$ Haar random matrices. In the *k-nearest neighbour case*, the analysis has been performed by finding numerically the optimal value for k

(see Fig. 8.3a). In Figs. 8.3b-f we report the error probabilities for the three techniques (i)-(iii), for $2 \leq n \leq 6$. We observe that the *k-nearest neighbour* and the *support vector machine* methods present comparable error probabilities, greatly outperforming the *nearest centroid* one.

8.17 Appendix Note 18: Suppression laws for Sylvester interferometers [5.6]

In this Section we discuss in detail the suppression law for Sylvester interferometers adopted in section 5.6.

Preliminary definitions

Let us begin by defining some notation used throughout this Section. We are interested in the transition amplitudes between states of n photons in m modes, and restrict ourselves to the case $n \leq m$. We denote such states as $|r\rangle := |r_1, \dots, r_m\rangle$, where r_k is the number of photons in the k th mode and $\sum_{k=1}^m r_k = n$. The set of all such states is denoted by $G_{n,m}$. We refer to the particular states where all r_i s take only values 0 or 1 as collision-free, and the set of all collision-free states is denoted by $Q_{n,m}$. It is easy to see that $|G_{n,m}| = \binom{m+n-1}{n}$ and $|Q_{n,m}| = \binom{m}{n}$. Let us now define two useful alternative representations for such states.

Definition 8.1

(MAL and BM representations). Let $|r\rangle = |r_1, \dots, r_m\rangle \in G_{n,m}$. We can define the following two representations of this state.

- Mode Assignment List (MAL): an n -dimensional vector $\tilde{r} := (\tilde{r}_1, \dots, \tilde{r}_n)$ listing the mode occupied by each particle.
- Binary matrix (BM): States with $m = 2^p$, for some positive integer p , can also be represented as an $n \times p$ binary matrix R as follows. For a given MAL $\tilde{r} := (\tilde{r}_1, \dots, \tilde{r}_n)$ describing the state, the i th row of R corresponds to the binary representation of $\tilde{r}_i - 1$ (padded with zeros on the left so it has length p).

Since photons are indistinguishable, the ordering of elements in the MAL is meaningless, and we conventionally choose the list to be in increasing order. Correspondingly, the ordering of the rows of two binary matrices is irrelevant, i.e. if R and R' are related only by a permutation of the rows, they represent the same physical state, and this we denote as $R \sim R'$.

Example 1. For $m = 8$ and $n = 4$, consider the state $|r\rangle = |1, 1, 1, 0, 0, 0, 0, 1\rangle$. It has the corresponding MAL $\tilde{r} = (1, 2, 3, 8)$, and its BM representation is the 4×3 matrix

$$R = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}. \quad (8.16)$$

The following definitions are also convenient shorthands used throughout this Section.

Definition 8.2

Let R be the $n \times p$ BM representation of some state $|r\rangle$. Let A be some subset of the columns of R . We denote by $\mathcal{N}^A(R)$ the matrix obtained by flipping each bit in the columns of R specified in A . If \tilde{r} is the MAL associated with R , we denote by $\mathcal{N}^A(\tilde{r})$ the MAL associated with $\mathcal{N}^A(R)$.

Example 2. For $m = 8$ and $n = 4$, consider the state $|r\rangle = |1, 1, 1, 1, 0, 0, 0, 0\rangle$, corresponding MAL $\tilde{r} = (1, 2, 3, 4)$ and BM representation

$$R = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}. \quad (8.17)$$

Then we have the following possibilities when A consists of a single element

$$\mathcal{N}^{\{1\}}(R) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad \mathcal{N}^{\{2\}}(R) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{N}^{\{3\}}(R) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \quad (8.18)$$

Clearly $\mathcal{N}^{\{2\}}(R) \sim \mathcal{N}^{\{3\}}(R) \sim R$, since they are related to each other by permutations of their rows, whereas $\mathcal{N}^{\{1\}}(R)$ represents a different state. The set A can also contain more than a single element. For example:

$$\mathcal{N}^{\{1,3\}}(R) = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad \mathcal{N}^{\{1,2,3\}}(R) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad (8.19)$$

and so on.

Definition 8.3

(Sylvester matrices) Let U^S be an m -dimensional unitary matrix of the form $U^S(m) \equiv H(m)/\sqrt{m}$, with $H(m)$ defined recursively as

$$H(2^p) := \begin{pmatrix} H(2^{p-1}) & H(2^{p-1}) \\ H(2^{p-1}) & -H(2^{p-1}) \end{pmatrix}, \quad (8.20)$$

for each positive integer p , and with $H(1) := 1$. We refer to $U^S(m)$ as *normalized Sylvester matrix* and to $H(m)$ as *Sylvester matrix*. The m dependence of U^S and H will be omitted when clear from the context.

An analytic expression for the (i, j) element of a Sylvester matrix can be given in the form:

$$[H(2^p)]_{i,j} = (-1)^{i_B \odot j_B}, \quad (8.21)$$

where i_B and j_B are the binary representations of $i - 1$ and $j - 1$ respectively, and \odot is the bitwise dot product.

Suppression law for Sylvester matrices

In this Section we present a test which generalizes that of [169], predicting a higher fraction of suppressed pairs. This test can be assessed with a computational cost increasing only polynomially in m and n .

Let us begin with the following two straightforward Lemmas, the proof of which we leave as an exercise for the reader.

Lemma 8.1

Let S_n be the set of permutations of $\{1, \dots, n\}$, and let $\tau \in S_n$ be a permutation different from the identity such that $\tau^2 = \mathbb{K}$. Then we can uniquely associate to each $\sigma \in S_n$ another (different) permutation $\sigma_\tau \equiv \tau \circ \sigma$, where \circ denotes the composition of permutations.

Lemma 8.2

Let \tilde{r} , R and \mathcal{N}^A be as in Definitions 8.1 and 8.2. Then $\mathcal{N}^A(\tilde{r}) = \tilde{r}$ if and only if $\mathcal{N}^A(R) \sim R$. This, in turn, happens if and only if there is a permutation $\tau \in S_n$ such that $\mathcal{N}^A(R) = R^\tau$, where R^τ is obtained from R by applying the permutation τ to the rows of R . Now, if τ is such a permutation, we have

1. $\tau \neq \mathbb{K}$,
2. $\tau^2 = \mathbb{K}$,
3. for each $\sigma \in S_n$, $\mathcal{N}^A(R^\sigma) = R^{\tau \circ \sigma}$
4. all columns of R in A have an equal number of 1s and 0s,
5. all columns of R not in A have an even number of 1s and an even number of 0s.

We are now ready to state the main result of the Section:

Theorem 8.1

Let $|r\rangle$ and $|s\rangle$ be two states of n particles in $m = 2^p$ modes, with corresponding MAL representations \tilde{r} and \tilde{s} , and BM representations R and S . If there is a subset A of the columns of R such that

$$\begin{cases} \mathcal{N}^A(R) \sim R, & (8.22a) \\ \bigoplus_{k=1}^n \bigoplus_{\alpha \in A} S_{k\alpha} = 1, & (8.22b) \end{cases}$$

then the transition from $|r\rangle$ to $|s\rangle$ (and consequently also that from $|s\rangle$ to $|r\rangle$), when transversing the Sylvester interferometer, is suppressed.

Proof. The transition amplitude from $|r\rangle$ to $|s\rangle$ can be written as

$$\mathcal{A} := \frac{1}{\sqrt{r_1! \dots r_m! s_1! \dots s_m!}} \text{Per}(U_{r,s}), \quad (8.23)$$

where $\text{Per}(U)$ denotes the permanent of U . In the equation above, $U_{r,s}$ is the matrix defined element-wise as

$$[U_{r,s}]_{i,j} := \frac{1}{\sqrt{m}} [H(2^p)]_{\tilde{r}_i, \tilde{s}_j}, \quad (8.24)$$

where, recall, $H(2^p)$ is the Sylvester matrix (cf. Definition 8.3).

Using Eqs. (8.21), (8.23), and (8.24), and denoting \mathbf{M}_i the i th row of matrix M , we obtain

$$\mathcal{A} = D \sum_{\sigma \in S_n} \prod_{k=1}^n (-1)^{\mathbf{R}_{\sigma(k)} \odot \mathbf{S}_k} = D \sum_{\sigma \in S_n} (-1)^{\mathcal{E}(\sigma)}, \quad (8.25)$$

where $\mathbf{B} \odot \mathbf{C}$ denotes the *bitwise dot product* between vectors \mathbf{B} and \mathbf{C} , defined as $\mathbf{B} \odot \mathbf{C} := \bigoplus_{\alpha=1}^p B_\alpha C_\alpha$, D is a constant factor, and we defined

$$\mathcal{E}(\sigma) \equiv \bigoplus_{k=1}^n \mathbf{R}_{\sigma(k)} \odot \mathbf{S}_k = \bigoplus_{k=1}^n \bigoplus_{\alpha=1}^p R_{\sigma(k),\alpha} S_{k,\alpha}. \quad (8.26)$$

The actual value of $\mathcal{E}(\sigma)$ is unimportant here, only its parity matters. Since we are interested in whether $\mathcal{A} = 0$, we will ignore the constant factor D in Eq. (8.25).

Clearly, for \mathcal{A} to vanish, we need exactly half of the permutations to be such that $(-1)^{\mathcal{E}(\sigma)} = 1$. A necessary and sufficient condition for this to hold is if, for each permutation σ , we can uniquely assign another permutation σ' such that $\mathcal{E}(\sigma') = 1 \oplus \mathcal{E}(\sigma)$. From Lemmas 8.1 and 8.2, we know that if condition (8.22a) holds we can uniquely associate to each σ another permutation $\sigma_\tau \equiv \tau \circ \sigma$, where τ is the permutation such that $\mathcal{N}^A(R) = R^\tau$. Using σ_τ in Eq. (8.26) we have

$$\mathcal{E}(\sigma_\tau) = \bigoplus_{k=1}^n \bigoplus_{\alpha=1}^p R_{\tau(\sigma(k)),\alpha} S_{k,\alpha} = \bigoplus_{k=1}^n \left[\left(\bigoplus_{\alpha \in A} R_{\tau(\sigma(k)),\alpha} S_{k,\alpha} \right) \oplus \left(\bigoplus_{\alpha \notin A} R_{\tau(\sigma(k)),\alpha} S_{k,\alpha} \right) \right]. \quad (8.27)$$

Using now the explicit expression for $\mathcal{N}^A(R)$ and Lemma 8.2, we have

$$\mathcal{N}^A(R^\sigma) = R^{\tau \circ \sigma} \iff \begin{cases} 1 \oplus R_{\sigma(k),\alpha} = R_{\tau(\sigma(k)),\alpha}, & \alpha \in A, \\ R_{\sigma(k),\alpha} = R_{\tau(\sigma(k)),\alpha}, & \alpha \notin A. \end{cases} \quad (8.28)$$

Inserting these into Eq. (8.27), we obtain

$$\mathcal{E}(\sigma_\tau) = \mathcal{E}(\sigma) \oplus \left[\bigoplus_{k=1}^n \bigoplus_{\alpha \in A} S_{k,\alpha} \right] = \mathcal{E}(\sigma) \oplus 1, \quad (8.29)$$

where in the last step we used Eq. (8.22b). Using this last result into Eq. (8.25) we conclude that

$$\mathcal{A} = C \sum_{\sigma \in S_n} (-1)^{\mathcal{E}_{R,S}(\sigma)} = C \sum_{\sigma \in S_n: \mathcal{E}(\sigma) \text{ even}} \left[(-1)^{\mathcal{E}(\sigma)} + (-1)^{\mathcal{E}(\sigma_\tau)} \right] = 0, \quad (8.30)$$

which proves that the input-output pair $(|r\rangle, |s\rangle)$ is suppressed. \square

Example 4. Consider the state $|r\rangle = |1, 1, 1, 1, 0, 0, 0, 0\rangle$ from Example 2. We have $\mathcal{N}^A(\tilde{r}) = \tilde{r}$ for $A = \{2\}$, $A = \{3\}$, and $A = \{2, 3\}$. Theorem 8.1 predicts suppression of all output states whose BM representation has an odd number of 1s in either the second column, the third column, or in the second and third columns combined. For example, the states $\tilde{s} = (3, 6, 7, 8)$, $\tilde{s} = (2, 6, 7, 8)$, and $\tilde{s} = (4, 6, 7, 8)$, having BM representations

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad (8.31)$$

respectively, are all suppressed.

Remark. (Efficiency) To check if Theorem 8.1 applies to a given input-output pair, one has to verify condition 8.22 for each one of the $2^p - 1 = m - 1$ possible (non-empty) subsets of the p columns of R and S , which requires only a polynomial (in n and m) number of elementary operations. Hence, the proposed suppression law is efficiently verifiable.

While Eq. (8.22b) gives a sufficient condition for an input-output pair to be suppressed, it is not necessary. For most input states, not *all* suppressed outputs satisfy Theorem 8.1. In the next Section we give estimates of the which fraction of all states our test identifies as suppressed.

Estimates on fraction of suppressed states

In this Section, we give estimates on the fraction of suppressed input-output pairs identified by the conditions of Theorem 8.1, particularly focusing on upper bounds and asymptotic limits (in the number of modes m and photons n). Throughout this Section, we are still restricted to $m = 2^p$ for some integer p , which means all binary matrices are $n \times p$, and to $n \leq m$. The fractions we obtain consider mainly the set of all possible states, $G_{n,m}$, which has $|G_{n,m}| = \binom{m+n-1}{n}$ elements. At the end of this Section we discuss the applicability of our results to the restriction of no-collision states.

We begin by restating the two conditions of Theorem 8.1 informally, as they would be used in a test. For simplicity, we refer to the states $|s\rangle$ and $|t\rangle$ of the Theorem as input and output states, respectively, although any conclusions can be extended to the case where the roles of input and output are reversed. With this in mind, we note that Condition 8.22a concerns only inputs, and we restate it as:

Condition I: For a given input with BM representation R , check whether R has any subsets of columns A such that negating those columns of R results in R up to a permutation of the rows (i.e. $\mathcal{N}^A(R) \sim R$).

Let us call \mathcal{A} the set of all such subsets of columns A . For the state in Example 2, for instance, $\mathcal{A} = \{\{2\}, \{3\}, \{2, 3\}\}$. If Condition I finds no such A , the test fails to identify any suppressed transitions. Also, as a consequence of Lemma 8.2, the test only works if n is even. Assuming that Condition I yielded some non-empty \mathcal{A} , we can test for Condition 8.22b, which is a test only on the outputs and which we restate as:

Condition II: Consider any output with BM representation S and any $A \in \mathcal{A}$. If the columns A of S contain an odd number of 1s, the transition from R to S is suppressed.

For simplicity, we begin by estimating how many output states are suppressed *given that* Condition I identified a non-empty set \mathcal{A} for some input. Before that, we need one final definition. We say that the elements of \mathcal{A} are independent if none can be replaced by a sequence of the others. That is, if there is no $A \in \mathcal{A}$ and $A_1, A_2, \dots, A_k \in \mathcal{A} \setminus \{A\}$ such that $\mathcal{N}^{A_1}(\mathcal{N}^{A_2} \dots (\mathcal{N}^{A_k}(X))) = \mathcal{N}^A(X)$ for all binary matrices X . To illustrate this, consider Example 2. There, $\mathcal{A} = \{\{2\}, \{3\}, \{2, 3\}\}$. Clearly, negating columns $\{2, 3\}$ of a binary matrix is the same as negating column $\{2\}$ followed by column $\{3\}$. Furthermore, Condition II can only be satisfied for $\{2, 3\}$ if it is satisfied by either $\{2\}$ or $\{3\}$. Thus, including $\{2, 3\}$ in \mathcal{A} does not give any new suppressions beyond those identified by $\{2\}$ and $\{3\}$, so we can safely drop it (we could have dropped either $\{2\}$ or $\{3\}$ instead, to the same effect). Since the binary matrices have p columns, there can be at most p independent elements in \mathcal{A} . We are now ready to state the following.

Corollary 8.1

Let \mathcal{A} be the set identified by Condition I for some input state, and suppose it contains q

independent elements. Then the fraction of outputs in $G_{n,m}$ (i.e. including collision states) that Condition II identifies as suppressed is equal to $1 - \frac{1}{2^q} + O\left(\frac{\log m}{n}\right)$.

Proof. Suppose initially that there is a single element $A \in \mathcal{A}$, say $A = \{1\}$. The corresponding suppressed outputs are those whose BM representation contains an odd number of 1s in the first column. These consist of approximately half of all possible states, which can be seen as follows. Let us write the binary matrix S of some such output as

$$S = \begin{pmatrix} \mathbf{S}_1 & S' \end{pmatrix}, \quad (8.32)$$

where \mathbf{S}_1 is its first column and S' a matrix of the remaining $p - 1$ columns. Since all matrices that are equivalent up to a permutation of the rows correspond to the same state, we can assume without loss of generality that the 1s in \mathbf{S}_1 occupy the first slots. This means there are only $n + 1$ possibilities for \mathbf{S}_1 , and it is easy to see that $n/2$ of them satisfy Condition II. Since this holds irrespective of the choice of S' , we conclude that $(n/2)/(n + 1) = 1/2 + O(1/n)$ out of *all* states are suppressed. The argument follows through almost unchanged for any A , even if it spans several columns.

Suppose now \mathcal{A} has q independent elements. By the previous paragraph, the first element of \mathcal{A} , let us call it A_1 , leads to a suppression of approximately (i.e. up to $O(1/n)$) half of all outputs. The second element, A_2 , also leads to a suppression of approximately half of all outputs—but now there is an overlap with those identified by A_1 . Since the two are independent, approximately half of the elements identified by A_2 have already been identified by A_1 (e.g., approximately half of the matrices with an odd number of 1s in the first column also have an odd number of 1s in the second column). Thus the new suppressions identified by A_2 correspond only to $1/4 + O(1/n)$ of all states. Each subsequent independent element of \mathcal{A} further divides the remaining set of unsuppressed states by half, and we conclude that Condition II in fact identifies $1 - 1/2^q + O(q/n)$ of all states as suppressed. Since $q \leq p = \log(m)$, this gives the claimed asymptotic behavior. \square

Corollary 8.1 shows that any input which satisfies Condition I has at least $1/2$ of all outputs suppressed, and this fraction can, in principle, be as high as $1 - 1/m$ if Condition I identifies p independent elements in \mathcal{A} . It is thus essential to identify how many inputs effectively satisfy Condition I in order to determine the overall fraction of suppressed pairs. Indeed, Theorem 8.1 treats input and outputs asymmetrically, and as a consequence Condition I is much more stringent than Condition II.

Corollary 8.2

Only an exponentially-vanishing subset of the inputs is detected by the test of Theorem 8.1.

Proof. Let us begin by counting how many inputs have a particular element A in their set \mathcal{A} . Consider initially that $A = \{1\}$. For Condition I to hold in this case, we need half of the elements of the first column of R to be 1 (cf. Lemma 8.2). Since we are free to rearrange the rows as desired, we can assume that R is ordered as follows

$$R = \begin{pmatrix} \mathbf{1}_{n/2} & R_1 \\ \mathbf{0}_{n/2} & R_2 \end{pmatrix}, \quad (8.33)$$

where $\mathbf{1}_{n/2}$ and $\mathbf{0}_{n/2}$ are vectors of $n/2$ 1s and 0s respectively, and R_1 and R_2 are $(n/2) \times (p - 1)$ binary matrices. For R to satisfy $\mathcal{N}^{\{1\}}(R) \sim R$, we must have $R_1 \sim R_2$, and in fact we can further reorder the rows of R have such that $R_1 = R_2$. We can now count how many binary matrices satisfy these constraints. Clearly we have no choice over the first column and over R_2 , so we only need count all possibilities for R_1 . Given the form chosen for R above, it is clear that

any two choices for R_1 that are equal upon permutation of the rows represent the same state, and should not be counted twice. This leads to a cumbersome combinatorial problem, since we need to tally all possibilities for R_1 up to permutations, but the number of permutations changes depending on whether R_1 has repeated rows. A shortcut to this calculation is to realize that we can formally map R_1 back to a MAL representation, as in Section 8.17 and, subsequently to a quantum state of $n/2$ photons in $2^{p-1} = m/2$ modes. Thus, the number of possibilities for R_1 is $\binom{(n+m)/2-1}{n/2}$.

Consider next another possibility, that $A = \{1, 2\}$. The counting in this case is similar to before, but a little trickier, because we need to keep track of how many choices we have for the first two columns such that $\mathcal{N}^{\{1,2\}}(R) \sim R$. Let us once more order the matrix as follows:

$$R = \begin{pmatrix} \mathbf{1}_{n/2} & \mathbf{a}_1 & R_1 \\ \mathbf{0}_{n/2} & \mathbf{a}_2 & R_2 \end{pmatrix}, \quad (8.34)$$

where $\mathbf{1}_{n/2}$ and $\mathbf{0}_{n/2}$ are defined as before, \mathbf{a}_1 and \mathbf{a}_2 are two binary vectors of length $n/2$ and R_1 and R_2 are $(n/2) \times (p-2)$ binary matrices. We want R to be ordered in such a way that $\mathcal{N}^{\{1,2\}}(R) \sim R$ implies $R_1 = R_2$, to reuse part of the previous argument. This immediately implies that \mathbf{a}_2 is obtained from \mathbf{a}_1 by negating its elements. We know that \mathbf{a}_1 and \mathbf{a}_2 , together, must contain $n/2$ 1s, due to Lemma 8.2, but that is already guaranteed by the fact that they are negations of each other¹. From the freedom of reordering rows, we can assume that the 1s in \mathbf{a}_1 occupy the first positions. This leaves $(n/2 + 1)$ possibilities for \mathbf{a}_1 . Combining that with the $\binom{(2n+m)/4-1}{n/2}$ possibilities for R_1 as in the previous paragraph, we get a total of $(n/2 + 1)\binom{(2n+m)/4-1}{n/2}$ possibilities. How does this compare with the case $A = \{1\}$? We can use the following asymptotic expression for the binomial coefficient $\binom{n}{k}$

$$\binom{n}{k} \approx \frac{(n-k/2)^k}{k^k e^{-k} \sqrt{2\pi k}}, \quad (8.35)$$

which holds when n is both large and much larger than k . By using this expression it is easy to see that, in the limit of both m and n large with $n \leq m$, $(n/2 + 1)\binom{(2n+m)/4-1}{n/2}$ grows exponentially slower than $\binom{(n+m)/2-1}{n/2}$, and so we can use the latter as an upper bound. In fact, as we consider sets A comprising of more columns, the constraints tend to become more restrictive and the number of matrices that satisfy them decreases. So we will use $\binom{(n+m)/2-1}{n/2}$ as an upper bound on the number of states that satisfy Condition I for *any* A .

We are now ready to give an upper bound on the number of states that satisfy Condition I for some nonempty \mathcal{A} . Since there are p columns in the binary matrices, There are $2^p - 1 = m - 1$ possible A s that can appear in \mathcal{A} . By the inclusion-exclusion principle, we would need to sum the number of states that satisfy Condition I for each possible A , then subtract those that have been counted multiple times because they satisfy it for more than one A . It is simpler, however, to use the union bound, which in this case says that the number of states is upper-bounded by $(m-1)\binom{(n+m)/2-1}{n/2}$. Recall now that the total number of states is $\binom{n+m-1}{n}$. Using an asymptotic formula for the binomial coefficient, it is clear that the fraction of states detected by the test is exponentially small in the limit of large n and m , as claimed. \square

So far, we have considered only the full set of states (i.e, including collision states) in the estimates of suppressed fractions, but the restriction to no-collision states is often more useful. For example, no-collision outputs are the only detected outcomes when the experiment is performed using bucket detectors (i.e. that do not distinguish one photon from many). More importantly,

¹Note that, whenever \mathbf{a}_1 has $1/4$ of the 1s, we are double-counting some matrices that were already included in the case $A = \{1\}$. But in the asymptotic limit they form a negligible fraction of the cases, so we do not worry about this correction.

experimental implementations typically consider inputs with no more than a single photon per mode. This is also relevant for boson sampling applications, being input state with at most one photon per mode the appropriate choice for its computational hardness. Thus, it would be interesting to obtain versions of Corollaries 8.1 and 8.2 where both the suppressed pairs and the set of all states included this restriction. Unfortunately, some pathological instances arise when we try to specialize the previous results in that way. To see that, consider the case where $n = m$. There, we have a single no-collision state, and it satisfies Condition I. Thus, we conclude that 100% of inputs in that case have suppressed outputs! As we now argue, it is still possible to show a weaker version of Corollary 8.2 for no-collision inputs.

Consider the regime where $m = O(n^2)$. Experiments are often done in this limit, especially since it seems to be a requirement for the computational hardness of the boson sampling model [33]. It is easy to show that the set of no-collision states is not a negligibly subset of all states in this regime, due to the so-called birthday paradox. To illustrate this suppose $m = n^2$ holds exactly, in which case the fraction of no-collision states among all states is $\binom{n^2}{n} / \binom{n^2+n-1}{n}$. Using Stirling's approximation, one obtains that this tends to $1/e$ in the limit of large n . Since the set of no-collision states is only polynomially small in the set of all states, a no-collision version of Corollary 8.2 must still hold—even if *all* inputs that satisfy Condition I were concentrated in the no-collision subset, they would still be an exponentially small fraction of it. This argument shows that the conclusion of Corollary 8.2 can be extended to the no-collision case in the limit $m = O(n^2)$, and we leave it as an open question whether it holds in general.

Corollary 8.2 also has consequences for the application of Theorem 8.1 as a test for validating boson sampling experiments. As argued in [201], suppressed events in Hadamard matrices (such as the Sylvester or Fourier matrices) could be useful as a way to witness partial photonic indistinguishability.

Corollary 8.2 shows that in Scattershot BosonSampling experiments [147–149], where inputs are chosen uniformly at random from all no-collision states, the number of suppressed events detected by Theorem 1 vanishes exponentially. On the other side, when specific input states that satisfy Condition I for many different A s are employed, Theorem 1 might provide a favorable scaling. Indeed, we showed that as many as $1 - O(1/m)$ out of all outcomes can be suppressed, but we leave a formal description of such a test for future work, as well as the question of whether the Sylvester matrix is optimal for this task.

Bibliography

- [1] Feynman, R. Simulating physics with computers. *Int. J Theor. Phys.* **21**, 476 (1982). <https://people.eecs.berkeley.edu/~christos/classics/Feynman.pdf> [Cited on pages 7 and 22.]
- [2] Deutsch, D. and Jozsa, R. Rapid solutions of problems by quantum computation. *Proceedings of the Royal Society of London A* **439**, 553 (1992). <http://rspa.royalsocietypublishing.org/content/439/1907/553>
- [3] Boixo, S., Ronnow, T., Isakov, S., Wang, Z., Wecker, D., Lidar, D., Martinis, J. and Troyer, M. Evidence for quantum annealing with more than one hundred qubits. *Nat. Phys.* **10**, 218–224 (2014). <http://www.nature.com/nphys/journal/v10/n3/full/nphys2900.html> [Cited on pages 7, 22, and 29.]
- [4] Schindler, P., Nigg, D., Monz, T., Barreiro, J. T., Martinez, E., Wang, S.X, Quint, S., Brandl, M.F., Nebendahl, V., Roos, C. F., Chwalla, M., Hennrich M. and Rainer Blatt A quantum information processor with trapped ions. *New J. Phys* **15**, 123012 (2013). <http://iopscience.iop.org/article/10.1088/1367-2630/15/12/123012/meta> [Cited on page 7.]
- [5] Barros, H. G., Stute, A., Northup, T.E., Russo, C. Schmidt, P. O. and Blatt, R. Deterministic single-photon source from a single ion. *New J. Phys.* **11** (10), 103004 (2009). <http://iopscience.iop.org/article/10.1088/1367-2630/11/10/103004/meta> [Cited on pages 7, 22, and 118.]
- [6] Martinis, J., Nam, S., Aumentado, J. and Urbina, C. Rabi Oscillations in a Large Josephson-Junction Qubit. *Phys. Rev. Lett.* **89** (11), 117901 (2002). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.89.117901> [Cited on pages 7, 118, and 141.]
- [7] Chiorescu, I., Nakamura, Y, Harmans, C.J and Mooij, J.E Coherent Quantum Dynamics of a Superconducting Flux Qubit. *Science* **299** (5614), 1869–71 (2003). <http://science.sciencemag.org/content/299/5614/1869> [Cited on pages 7, 22, and 141.]
- [8] Duan, L. M., Lukin, M. D., Cirac, J. I. and Zoller, P. Long-distance quantum communication with atomic ensembles and linear optics. *Nature* **414** (6862), 413-418 (2001). <http://www.nature.com/nature/journal/v414/n6862/abs/414413a0.html> [Cited on pages 7 and 22.]
- [9] Castelli, F., Plez, B., Lallement, R., Ferlet, R., Kalas, P.G., Walker, H. J. and Holland, W. S. Deterministic Generation of Single Photons from One Atom. *Science* **80**, 1992-1995 (2004). <http://science.sciencemag.org/content/303/5666/1992> [Cited on pages 7, 118, and 141.]
[Cited on pages 7 and 118.]
- [10] Rudolph, T. Why I am optimistic about the silicon-photonics route to quantum computing. *APL Photonics* **2**, 030901 (2017). <http://aip.scitation.org/doi/abs/10.1063/1.4976737> [Cited on pages 7 and 31.]

- [11] Raussendorf, R. and Briegel, H. J. A One-Way Quantum Computer. *Phys. Rev. Lett.* **86**, 5188–5191 (2001). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.86.5188> [Cited on pages 7 and 31.]
- [12] Knill, E., Laflamme, R. and Milburn, G. J. A scheme for efficient quantum computation with linear optics. *Nature* **409**, 46–52 (2001). <http://www.nature.com/nature/journal/v409/n6816/abs/409046a0.html?foxtrotcallback=true> [Cited on pages 7, 8, 31, 32, 82, 94, and 141.]
- [13] Burnham, D. C. and Weinberg, D. L. Observation of simultaneity in parametric production of optical photon pairs. *Phys. Rev. Lett.* **25**, 84–87 (1970). <https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.25.84> [Cited on pages 7, 57, and 98.]
- [14] Hong, C. K., Ou, Z. Y. and Mandel, L. Measurement of subpicosecond time intervals between two photons by interference. *Phys. Rev. Lett.*, **59** (18), 2044–2046 (1987). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.59.2044> [Cited on pages 7, 37, 39, and 94.]
- [15] Reck, M., Zeilinger, A., Bernstein, H. J. and P. Bertani. Experimental realization of any discrete unitary operator. *Phys. Rev. Lett.* **73** (1), 5861 (1994). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.73.58> [Cited on pages 7, 8, 50, 51, 60, 61, 62, 63, 72, 79, 95, 97, 98, 129, and 148.]
- [16] Carman, R. L., Chiao, R. Y. and Kelley, P. L. Observation of degenerate stimulated four-photon interaction and four-wave parametric amplification. *Phys. Rev. Lett.* **17** (26), 1281 (1966). <https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.17.1281> [Cited on pages 8, 59, and 141.]
- [17] Stolen, R. H. Phase-matched-stimulated four-photon mixing in silica-fiber waveguides. *IEEE J. Quantum Electron.* **11** (3), 100 (1975). <http://ieeexplore.ieee.org/document/1068571/?reload=true> [Cited on pages 8 and 59.]
- [18] Lodahl, P., Mahmoodian, S. and Stobbe, S. Interfacing single photons and single quantum dots with photonic nanostructures. *Rev. Mod. Phys.* **87**(2), 347–400 (2015). <https://journals.aps.org/rmp/issues/87/2> [Cited on pages 8 and 118.]
- [19] Gisin, N. and Thew, R. Quantum communication. *Nat. Photonics* **1**, 165–171 (2007). <http://www.nature.com/nphoton/journal/v1/n3/abs/nphoton.2007.22.html> [Cited on page 8.]
- [20] Gates, B.D., Xu, Q., Love, J.C., Wolfe, D.B. and Whitesides, G.M. Unconventional nanofabrication. *Ann. Rev. Mater. Res.* **15**, 339–372 (2004). <http://www.annualreviews.org/doi/abs/10.1146/annurev.matsci.34.052803.091100> [Cited on pages 8 and 53.]
- [21] Gates, B.D., Xu, Q., Stewart, M., Ryan, D., Willson, C.G. and Whitesides, G.M. New approaches to nanofabrication: Molding, printing, and other techniques. *Chem. Rev.* **105**, 1171–1196 (2005). <http://pubs.acs.org/doi/pdf/10.1021/cr030076o> [Cited on pages 8 and 53.]
- [22] Carolan, J., Harrold, C., Sparrow, C., Martín-López, E., Russell, N.J., Silverstone, J.W., Shadbolt, P.J., Matsuda, N., Oguma, M., Itoh, M., Marshall, G.D., Thompson, M., Matthews, J., Hashimoto, T., O’Brien, J. and Laing, A. Universal linear optics. *Science* **349**, 711–716 (2015). <http://science.sciencemag.org/content/349/6249/711> [Cited on pages 8, 47, 58, 72, 74, 76, 94, 101, and 141.]

- [23] Harris, N. C., Steinbrecher, G. R., Mower, J., Lahini, Y., Prabhu, M., Baehr-Jones, T., Hochberg, M., Lloyd, S. and Englund, D. Quantum transport simulations in a programmable nanophotonic processor. *Nat. Photon.***11**, 447-452 (2017). <http://www.nature.com/nphoton/journal/v11/n7/full/nphoton.2017.95.html> [Cited on pages 8, 58, 74, 77, and 101.]
- [24] Sugioka, K., Xu, J., Wu, D., Hanada, Y., Wang, Z., Cheng, Y. and Midorikawa, K. Femtosecond laser 3d micromachining: a powerful tool for the fabrication of microfluidic, optofluidic, and electrofluidic devices based on glass. *Lab on a Chip* **14** (18), 3447-3458 (2014). <http://pubs.rsc.org/en/content/articlelanding/2014/lc/c4lc00548a> [Cited on pages 8, 54, 91, 97, and 108.]
- [25] Malinauskas, M., Farsari, M., Piskarskas, A. and Juodkazis, S. Ultrafast laser nanostructuring of photopolymers: A decade of advances. *Physics Reports* **533** (1), 1-31 (2013). <http://www.sciencedirect.com/science/article/pii/S0370157313002718> [Cited on pages 8, 54, 97, and 108.]
- [26] Della Valle, G., Osellame, R. and Laporta, P. Micromachining of photonic devices by femtosecond laser pulses. *Journal of Optics A: Pure and Applied Optics* **11** (1), 013001 (2008). <http://iopscience.iop.org/article/10.1088/1464-4258/11/1/013001/meta> [Cited on pages 8, 54, 97, 108, and 141.]
- [27] Sansoni, L., Sciarrino, F., Vallone, G., Mataloni, P., Crespi, A., Ramponi, R. and Osellame, R. Two-Particle Bosonic-Fermionic QuantumWalk via Integrated Photonics. *Phys. Rev. Lett.* **108**, 010502 (2012). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.108.010502> [Cited on pages 8 and 36.]
- [28] Sansoni, L., Sciarrino, F., Vallone, G., Mataloni, P., Crespi, A., Ramponi, R. and Osellame, R. Polarization entangled states measurement on a chip. *Phys. Rev. Lett.* **105** (20), 200503 (2010). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.105.200503> [Cited on pages 8 and 54.]
- [29] DiVincenzo, D. P. The Physical Implementation of Quantum Computation. *Fortschritte der Physik* **48**, 771 (2000) *Preprint at arXiv:quant-ph/0002077*. <https://arxiv.org/pdf/quant-ph/0002077.pdf> [Cited on pages 8 and 26.]
- [30] Aharonov, Y., Davidovich, L. and Zagury, N. Quantum random walks. *Phys. Rev. A* **48**, 1687 (1993). <https://journals.aps.org/prl/abstract/10.1103/PhysRevA.48.1687> [Cited on pages 8, 26, and 30.]
- [31] Farhi, E. and Gutmann, S. Quantum computation and decision trees. *Phys. Rev. A* **58**, 915 (1998). <https://journals.aps.org/prl/pdf/10.1103/PhysRevA.58.915> [Cited on pages 8 and 30.]
- [32] Kaplan, A.E., Marzoli, I., Lamb, Jr., W.E. and Schleich, W.P. Multimode interference: Highly regular pattern formation in quantum wave-packet evolution. *Phys. Rev. A* **61**, 032101 (2000). <https://journals.aps.org/prl/abstract/10.1103/PhysRevA.61.032101> [Cited on pages 8 and 31.]
- [33] Aaronson, S. and Arkhipov, A. The computational complexity of linear optics. *Proceedings of the 43rd annual ACM symposium on Theory of Computing*. Ed. by ACM Press, 333-342 (2011). <https://www.scottaaronson.com/papers/optics.pdf> [Cited on pages 8, 40, 41, 43, 44, 45, 46, 82, 94, 141, and 168.]
- [34] Church, A. An unsolvable problem of elementary number theory. *Am. J. Math* **58**, 345 (1936). <https://www.ics.uci.edu/~lopes/teaching/inf212W12/readings/church.pdf> [Cited on pages 12 and 14.]

- [35] Turing, A.M. On computable numbers, with an application to the Entscheidungsproblem. *Proc. Lond. Math. Soc.* **2**, 42-230 (1936). https://www.cs.virginia.edu/~robins/Turing_Paper_1936.pdf [Cited on pages 12 and 14.]
- [36] Gödel, K. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I. *Monatshefte für Mathematik und Physik* **38**, 173-98 (1931). <http://www.w-k-essler.de/pdfs/goedel.pdf> [Cited on page 12.]
- [37] Fernández, M. Models of Computation: An Introduction to Computability Theory. *Undergraduate Topics in Computer Science. Springer* (2009). <http://www.springer.com/gp/book/9781848824331> [Cited on page 13.]
- [38] Solovay, R. M. and Strassen, V. A fast Monte-Carlo test for primality. *SIAM Journal on Computing* **6** (1), 84-85 (1977). <http://epubs.siam.org/doi/abs/10.1137/0206006> [Cited on page 15.]
- [39] Mitchell, T. Machine Learning. *McGraw Hill. p. 2* (1997). [Cited on pages 16 and 17.]
- [40] Wittek, P. Quantum Machine Learning - What Quantum Computing Means to Data Mining. *Elsevier AP, Academic Press* (2014). <https://www.elsevier.com/books/quantum-machine-learning/wittek/978-0-12-800953-6> [Cited on pages 16, 17, 18, 106, and 107.]
- [41] Allender, E., Loui, M.C. and Regan, K.W. Complexity Classes. *Chapter 27 of the forthcoming CRC Handbook on Algorithms and Theory of Computation.* (2010). <http://dl.acm.org/citation.cfm?id=1882779> [Cited on page 20.]
- [42] Valiant, L. G. The complexity of computing the permanent. *Theoretical Comput. Sci.* **8** (2), 189-201 (1979). <http://www.sciencedirect.com/science/article/pii/0304397579900446> [Cited on pages 20, 37, 44, and 110.]
- [43] Sipser, M. A complexity theoretic approach to randomness. *Proceedings of the 15th annual ACM Symposium on Theory of Computing, ACM Press*, 330-335 (1997). <http://dl.acm.org/citation.cfm?id=808762> [Cited on pages 20, 21, and 44.]
- [44] Toda, S. PP is as Hard as the Polynomial-Time Hierarchy. *SIAM Journal on Computing* **20**, 865-877 (1991). <http://epubs.siam.org/doi/abs/10.1137/0220053> [Cited on pages 20 and 21.]
- [45] Nielsen, M. and Chuang, I. Quantum Computation and Quantum Information. *Cambridge University Press*, (2000). [Cited on pages 21, 27, 28, 29, 30, 31, 33, 61, 89, and 94.]
- [46] Deutsch, D. Quantum theory, the Church-Turing principle and the universal quantum computer. *Proc. R. Soc. Lon. A* **400**, 97 (1985). <http://rspa.royalsocietypublishing.org/content/400/1818/97> [Cited on page 21.]
- [47] Bernstein E. and Vazirani U. Quantum Complexity Theory. *Proceedings of the 25th annual ACM Symposium on Theory of Computing*, 11-20 (1993). *SIAM J. Computing*, **26**, 1411-1473 (1997). <http://epubs.siam.org/doi/abs/10.1137/S0097539796300921> [Cited on pages 21, 26, and 44.]
- [48] Benioff, P. The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines. *Journal of statistical physics* **22** (5), 563-591 (1980). <https://link.springer.com/article/10.1007/BF01011339> [Cited on page 22.]

- [49] Manin, Yu. I. Vychislimoe i nevyshislimoe [Computable and Noncomputable] (in Russian). *Proceedings of the Royal Society of London A*. 13-15 (1980). [Cited on page 22.]
- [50] Jozsa, R. and Entanglement and Quantum Computation. *Geometric Issues in the Foundations of Science, Oxford University Press*(1997). <https://arxiv.org/pdf/quant-ph/9707034.pdf> [Cited on page 22.]
- [51] Jozsa, R. and Linden, N. On the role of entanglement in quantum-computational speed-up. *Proc. R. Soc. Lond. A* **459**, 2011-2032 (2003). <http://rspa.royalsocietypublishing.org/content/royprsa/459/2036/2011.full.pdf> [Cited on page 22.]
- [52] Vidal, G. Efficient Classical Simulation of Slightly Entangled Quantum Computations. *Phys. Rev. Lett.* **91**, 147902 (2003). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.91.147902> [Cited on page 22.]
- [53] Biamonte, E., Brassard, G., Kenigsberg, D. and Mor, T. Quantum computing without entanglement Theoretical Computer. *Science* **320**, 15-33 (2004). <http://dl.acm.org/citation.cfm?id=1016401> [Cited on page 22.]
- [54] O'Brien, J.L., Furusawa, A. and Vuckovic, J. Photonic quantum technologies. *Nat. Photonics* **3** (12), 687-695 (2009). <http://www.nature.com/nphoton/journal/v3/n12/abs/nphoton.2009.229.html> [Cited on page 22.]
- [55] Northup, T.E. and Blatt, R. Quantum information transfer using photons. *Nat. Photonics* **8** (5), 356-363 (2014). <http://www.nature.com/nphoton/journal/v8/n5/full/nphoton.2014.53.html> [Cited on page 22.]
- [56] Jones, J.A. Quantum Computing with NMR. *Prog. NMR Spectrosc.* **59**, 91-120 (2011). <https://www.ncbi.nlm.nih.gov/pubmed/21742157> [Cited on page 22.]
- [57] Aharonovich, I., Greentree, A.D. and Prawer, S. Diamond photonics. *Nat. Photonics* **5** (7), 397-405 (2011). <http://www.nature.com/nphoton/journal/v5/n7/full/nphoton.2011.54.html> [Cited on page 22.]
- [58] Haroche, S. Controlling photons in a box and exploring the quantum to classical boundary (Nobel Lecture). *Rev. Mod. Phys.* **52** (39), 10158-10178 (2013). <https://journals.aps.org/rmp/pdf/10.1103/RevModPhys.85.1083> [Cited on page 22.]
- [59] Bloch, I. Review Article Quantum coherence and entanglement with ultracold atoms in optical lattices. *Nature* **453**, 1016-1022 (2008). <https://www.nature.com/nature/journal/v453/n7198/full/nature07126.html> [Cited on page 22.]
- [60] Duty, T., Gunnarsson, D., Bladh, K. and Delsing, P. Coherent dynamics of a Josephson charge qubit. *Phys. Rev. B* **69** (14), 140503 (2004). <https://journals.aps.org/prb/abstract/10.1103/PhysRevB.69.140503> [Cited on page 22.]
- [61] Einstein, A., Podolsky, B. and Rosen, N. Can quantum-mechanical description of physical reality be considered complete?. *Phys. Rev.* **47**, 777 (1935). <https://journals.aps.org/pr/abstract/10.1103/PhysRev.47.777> [Cited on page 26.]
- [62] Bell, J. On the Einstein Podolsky Rosen paradox. *Physics* **1**, 195 (1964). https://cds.cern.ch/record/111654/files/vol1p195-200_001.pdf [Cited on page 26.]
- [63] Aspect, A., Grangier, P. and Roger, G. Experimental realization of Einstein-Podolsky-Rosen-Bohm gedanken experiment: a new violation of Bell's inequalities. *Phys. Rev. Lett.* **49**, 91 (1982). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.49.91> [Cited on page 26.]

- [64] Hensen, B., Bernien, H., Dréau, A. E., Reiserer, A., Kalb, N., Blok, M. S., Ruitenbergh, J., Vermeulen, R. F. L., Schouten, R. N., Abellán, C., Amaya, W., Pruneri, V., Mitchell, M. W., Markham, M., Twitchen, D. J., Elkouss, D., Wehner, S., Taminiau T. H. and Hanson R. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **526**, 682–686 (2015). <https://www.nature.com/nature/journal/v526/n7575/full/nature15759.html> [Cited on page 26.]
- [65] Giustina, M., Versteegh, M.A., Wengerowsky, S., Handsteiner, J., Hochrainer, A., Phelan, K., Steinlechner, F., Kofler, J., Larsson, J.-A., Abellán, C., Amaya, W., Pruneri, V., Mitchell, M.W., Beyer, J., Gerrits, T., Lita, A., Shalm, L.K., Nam, S.W., Scheidl, T., Ursin, R., Wittmann, B. and Zeilinger, A. Significant-Loophole-Free Test of Bell’s Theorem with Entangled Photons. *Phys. Rev. Lett.* **115**, 250402 (2015). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.115.250401> [Cited on page 26.]
- [66] Shalm, L. K. et al. Strong Loophole-Free Test of Local Realism. *Phys. Rev. Lett.* **115**, 250402 (2015). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.115.250402> [Cited on page 26.]
- [67] Raussendorf, R. and Briegel, H.J. A one-way quantum computer. *Phys. Rev. Lett.* **86**, 5188 (2001). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.86.5188> [Cited on page 26.]
- [68] Farhi, E., Goldstone, J., Gutmann, S., Lapan, J., Lundgren, A. and Preda, D. A. Quantum adiabatic evolution algorithm applied to random instances of an np-complete problem. *Science* **292**, 472 (2001). <http://science.sciencemag.org/content/292/5516/472> [Cited on page 26.]
- [69] Jordan, S. P. Permutational quantum computing. *Quantum Info. Comput.* **10**, 470 (2010). <http://dl.acm.org/citation.cfm?id=2011369> [Cited on page 26.]
- [70] Moussa, O., Ryan, C., Cory, D., and Laflamme, R. Testing contextuality on quantum ensembles with one clean qubit. *Phys. Rev. Lett.* **104**, 160501 (2010). <https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.104.160501> [Cited on page 26.]
- [71] Finilla, A. B., Gomez, M. A., Sebenik, C. and Dol, D. J. Quantum annealing: A new method for minimizing multidimensional functions. *Phys. Lett. Chem.* **219**, 343 (1994). [Cited on page 26.]
- [72] Santoro, G. E. and Tosatti, E. Optimization using quantum mechanics: quantum annealing through adiabatic evolution. *J. Phys. A* **39**, R393 (2006). <http://iopscience.iop.org/article/10.1088/0305-4470/39/36/R01/meta> [Cited on page 26.]
- [73] Das, A. and Chakrabarti, K.B. Colloquium: Quantum annealing and analog quantum computation. *Rev. Mod. Phys.* **80**, 1061 (2008). <https://journals.aps.org/rmp/pdf/10.1103/RevModPhys.80.1061> [Cited on page 26.]
- [74] Devoret, M. H. and Schoelkopf, R. J. Superconducting Circuits for Quantum Information: An Outlook. *Science* **339**, 1169 (2013). <http://science.sciencemag.org/content/339/6124/1169> [Cited on page 26.]
- [75] Barends, R. et al. Digitized adiabatic quantum computing with a superconducting circuit. *Nature* **534**, 222–226 (2016). <https://www.nature.com/nature/journal/v534/n7606/full/nature17658.html> [Cited on page 26.]
- [76] Seung Woo, S., Smith, G., Smolin J.A. and Vazirani, S. How ‘Quantum’ is the D-Wave Machine?. *Preprint at arXiv:1401.7087* (2014). <https://arxiv.org/pdf/1401.7087.pdf> [Cited on page 26.]

- [77] Benenti, G. Principles of quantum computation and information: Basic tools and special topics, vol 1. *World Scientific Publishing Company* (2007). [Cited on pages 27, 28, and 61.]
- [78] Benenti, G. Principles of quantum computation and information: Basic tools and special topics, vol 2. *World Scientific Publishing Company* (2007) [Cited on pages 27 and 28.]
- [79] Kitaev, A. Y. Quantum computations: algorithms and error corrections. *Russ. Math. Surv.* **52** (6), 1191-1249 (1997). <http://iopscience.iop.org/article/10.1070/RM1997v052n06ABEH002155/meta> [Cited on page 28.]
- [80] Barenco, A., Bennett, C., Cleve, R., DiVincenzo, D., Margolus, R., Shor, P., Sleator, T., Smolin, J. and Weinfurter, H. Elementary gates for quantum computation. *Phys. Rev. A* **52**, 3457 (1995). <https://journals.aps.org/prb/abstract/10.1103/PhysRevA.52.3457> [Cited on page 29.]
- [81] Vazirani, U. A Survey of Quantum Complexity Theory. *Proceedings of Symposia in Applied Mathematics* (2000). <http://athena.nitc.ac.in/~kmurali/Quantum/ams/chapter2.4.pdf> [Cited on page 29.]
- [82] Shor, P.W. Algorithms for quantum computation: Discrete logarithms and factoring. *Proc. 35nd Annual Symposium on Foundations of Computer Science, (Shafi Goldwasser, ed.), IEEE Computer Society Press* 124-134 (1994). <http://ieeexplore.ieee.org/document/365700/> [Cited on pages 29, 78, and 92.]
- [83] Grover, L.K. A fast quantum mechanical algorithm for database search. *Proceeding STOC '96 Proceedings of the twenty-eighth annual ACM symposium on Theory of computing* 212-219 (1996). <http://dl.acm.org/citation.cfm?doid=237814.237866> [Cited on pages 29, 30, and 92.]
- [84] Hales, L. and Hallgren, S. An improved quantum Fourier transform algorithm and applications. *Proceedings of the 41st Annual Symposium on Foundations of Computer Science* 515 (2000). <http://ieeexplore.ieee.org/document/892139/> [Cited on page 29.]
- [85] Simon, D.R. On the power of quantum computation. *SIAM J. Comput.*, **26** (5) 1474–1483 (1997). <http://epubs.siam.org/doi/abs/10.1137/S0097539796298637> [Cited on page 29.]
- [86] Rivest, R., Shamir, A. and Adleman, L. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM* **21**, 120-126 (1978). <https://people.csail.mit.edu/rivest/Rsapaper.pdf> [Cited on page 30.]
- [87] Pearson, K. The Problem of the Random Walk. *Nature* **72**, 294 (1905). <http://www.nature.com/nature/journal/v72/n1865/abs/072294b0.html> [Cited on page 30.]
- [88] Berry, M., Marzoli, I. and Schleich, W.P. Quantum carpets, carpets of light. *Physics World* **14**, 39 (2001). <http://iopscience.iop.org/article/10.1088/2058-7058/14/6/30> [Cited on page 31.]
- [89] Robinett, R. Quantum wavepacket revivals. *Physics Reports* **392**, 1 (2004). http://www.personal.psu.edu/rq9/Robinett/Physics_Reports_Quantum_Wave_Packet_Revivals_Robinett.pdf [Cited on page 31.]
- [90] Longhi, S. Periodic wave packet reconstruction in truncated tight-binding lattices. *Phys. Rev. B* **82**, 041106 (2010). <https://journals.aps.org/prb/abstract/10.1103/PhysRevB.82.041106> [Cited on page 31.]

- [91] Mohseni, M., Rebentrost, P., Lloyd, S. and Aspuru-Guzik, A. Environment-assisted quantum walks in photosynthetic energy transfer. *J. Chem. Phys.* **129** (17), 174106 (2008). <http://aip.scitation.org/doi/full/10.1063/1.3002335> [Cited on page 31.]
- [92] Cirac, J., Zoller, P., Kimble, H. and Mabuchi, H. Quantum State Transfer and Entanglement Distribution among Distant Nodes in a Quantum Network. *Phys. Rev. Lett.* **78**, 3221 (1997). <http://authors.library.caltech.edu/2635/1/CIRpr197.pdf> [Cited on page 31.]
- [93] Childs, A., Gosset, D. and Webb, Z. Universal computation by multi-particle quantum walk. *Science* **339**, 791-794 (2013). <http://science.sciencemag.org/content/339/6121/791.full> [Cited on page 31.]
- [94] Dyer, M., Frieze, A. and Kannan, R. A random polynomial-time algorithm for approximating the volume of convex bodies. *Journal of the ACM* **38** (1), 1-17 (1991). <http://dl.acm.org/citation.cfm?doid=102782.102783> [Cited on page 31.]
- [95] Schoning, U. A probabilistic algorithm for k-SAT and constraint satisfaction problems. In *40th Annual Symposium on Foundations of Computer Science*, 17-19 (1999). <http://homepages.cwi.nl/~rdewolf/schoning99.pdf> [Cited on page 31.]
- [96] Jerrum, Sinclair, M. A. and Vigoda, E. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *Journal of the ACM* **51**, 673 (2004). <https://www.cc.gatech.edu/~vigoda/Permanent.pdf> [Cited on pages 31 and 41.]
- [97] Raussendorf, R., Harrington, J. and Goyal, K. Topological fault-tolerance in cluster state quantum computation. *New J. Phys.* **9** (6), 199 (2007). <http://iopscience.iop.org/article/10.1088/1367-2630/9/6/199/meta> [Cited on page 31.]
- [98] Kieling, K., Rudolph, T. and Eisert, J. Percolation, Renormalization, and Quantum Computing with Non-deterministic Gates. *Phys. Rev. Lett.* **99**, 130501 (2007). <http://iopscience.iop.org/article/10.1088/1367-2630/9/6/199/meta> [Cited on page 31.]
- [99] Kok, P., Munro, W.J., Nemoto, K., Ralph, T.C., Dowling, J.P. and Milburn, G.J. Linear optical quantum computing with photonic qubits. *Rev. Mod. Phys.* **79**, 135–174 (2007). <https://journals.aps.org/rmp/abstract/10.1103/RevModPhys.79.135> [Cited on pages 31 and 32.]
- [100] Briegel, H.J. and Raussendorf, R. Persistent Entanglement in arrays of Interacting Particles. *Phys. Rev. Lett.* **86** (5), 910–3 (2001). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.86.910> [Cited on page 32.]
- [101] Rahimi-Keshari, S., Lund, A. P. and Ralph, T. C. What Can Quantum Optics Say about Computational Complexity Theory?. *Phys. Rev. Lett.* **114**, 060501 (2015). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.114.060501> [Cited on pages 32 and 44.]
- [102] Bartlett, S. D., Sanders, B. C., Braunstein, S. L. and Nemoto, K. Efficient classical simulation of continuous variable quantum information processes. *Phys. Rev. Lett.* **88**, (2002). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.88.097904> [Cited on page 32.]
- [103] Gurvits, L. On the complexity of mixed discriminants and related problems. *Mathematical Foundations of Computer Science* **88**, 447-458 (2005). https://link.springer.com/chapter/10.1007/11549345_39 [Cited on page 32.]
- [104] Loudon, R. The quantum theory of light. *Oxford Science Publications, Third edition*, (1971). [Cited on page 32.]

- [105] Crespi, A., Ramponi, R., Osellame, R., Sansoni, L., Bongioanni, I., Sciarrino, F., Vallone, G. and Mataloni, P. Integrated photonic quantum gates for polarization qubits. *Nat. Comm.* **2**, (2011). <https://www.nature.com/articles/ncomms1570> [Cited on page 33.]
- [106] O' Brien, J. et al. Demonstration of an all-optical quantum controlled-NOT gate. *Nature* **426**, 264–267 (2003). <http://www.nature.com/nature/journal/v426/n6964/abs/nature02054.html> [Cited on page 33.]
- [107] O'Brien, J. L., Pryde, G. J., Gilchrist, A., James, D. F. V., Langford, N. K., Ralph, T. C. and White, A. G. Quantum process tomography of a controlled-NOT gate. *Phys. Rev. Lett.* **93**, 080502 (2004). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.93.080502> [Cited on page 33.]
- [108] Born, M. and Wolf, E. Principles of optics. *Pergamon Press Ltd., Fifth edition*, (1980). [Cited on page 33.]
- [109] Boyd, R. Nonlinear Optics. *Academic Press, Third edition*, (2007). [Cited on page 33.]
- [110] Saleh, B.E.A. and Teich, M.C. Fundamentals of photonics. *Wiley, Second edition*, (2007). [Cited on pages 33, 34, 52, 59, and 122.]
- [111] Rullière, C. Femtosecond Laser Pulses, Principles and Experiments. *Springer, Second edition*, (2003). [Cited on page 33.]
- [112] Sakurai, J.J. and Napolitano, J. Modern quantum mechanics. *Addison-Wesley*, (2011). [Cited on pages 35 and 36.]
- [113] Pauli, W. The connection between spin and statistics. *Phys. Rev.* **58**, 716-722 (1940). http://hermes.ffn.ub.es/luisnavarro/nuevo_maletin/Pauli_spin_statistics_1940.pdf [Cited on page 36.]
- [114] Leinaas, J. M. and Myrheim, J. On the theory of identical particles. *Il Nuovo Cimento* **37 b (1)**, 1-23 (1977). <http://www.ifi.unicamp.br/~cabrera/teaching/referencia.pdf> [Cited on page 36.]
- [115] Wilczek, F. Magnetic Flux, Angular Momentum, and Statistics. *Phys. Rev. Lett.* **48 (17)**, 1144-1146 (1982). <https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.48.1144> [Cited on page 36.]
- [116] Lahtinen, V. T. and Pachos, J. K. A short introduction to topological quantum computation. *SciPost Phys.* **3**, 021 (2017). <https://scipost.org/SciPostPhys.3.3.021/pdf> [Cited on page 36.]
- [117] Fetter, A.L. and Walecka, J.D. Quantum theory of many-particle systems. *Courier Corporation*, (2003). [Cited on page 36.]
- [118] Zeilinger, A. General properties of lossless beam splitters in interferometry. *Amer. J. Phys.*, **49**, 882 (1981). <http://aapt.scitation.org/doi/10.1119/1.12387> [Cited on page 38.]
- [119] Gard, B. T. and Motes, K. R. and Olson, J. P. and Rohde, P. P. and Dowling, J. P. An introduction to boson-sampling. *From Atomic to Mesoscale: The Role of Quantum Coherence in Systems of Various Complexities*, **1**,167-192 (2014). <https://arxiv.org/abs/1406.6767> [Cited on pages 41 and 44.]
- [120] Arkhipov, A. and Kuperberg, G. The bosonic birthday paradox. *Geometry & Topology Monographs*, **18**,1-7 (2012). <http://msp.org/gtm/2012/18/gtm-2012-18-001s.pdf> [Cited on page 41.]

- [121] Scheel, S. Permanents in linear optical networks. *preprint at arXiv quant-ph/0406127*, (2004). <https://arxiv.org/abs/quant-ph/0406127> [Cited on pages 42 and 82.]
- [122] Troyansky, L. and Tishby, N. Permanent uncertainty: On the quantum evaluation of the determinant and the permanent of a matrix. *Proc. of Phys. Comp.*, (1996). <https://www.cs.huji.ac.il/labs/learning/Papers/perm.pdf> [Cited on pages 43 and 44.]
- [123] Stockmeyer, L. J. The complexity of approximate counting. *Proc. ACM STOC.*,118–126, (1983). <http://dl.acm.org/citation.cfm?id=808740> [Cited on page 43.]
- [124] Haar, A. Der Massbegriff in der Theorie der kontinuierlichen Gruppen. *In Annals of Mathematics*, **34** (1),147-169 (1933). [Cited on page 44.]
- [125] Ryser, H. J. Combinatorial Mathematics. *Carus Mathematical Monograph*, **14**, (1963). <https://www.cambridge.org/core/books/combinatorial-mathematics/8AB6985C13895FAA27999FC5EDABA7AD> [Cited on page 45.]
- [126] Pomarico, E., Sanguinetti, B., Thew, R. and Zbinden, H. Room temperature photon number resolving detector for infrared wavelengths. *Opt. Express*, **18**,10750-10759 (2010). <https://www.osapublishing.org/oe/abstract.cfm?uri=oe-18-10-10750> [Cited on pages 46, 60, 113, 129, and 141.]
- [127] Aaronson, S. and Brod, D. J. BosonSampling with Lost Photons. *Phys. Rev. A*, **93**,012335 (2015). <https://journals.aps.org/prapdf/10.1103/PhysRevA.93.012335> [Cited on page 46.]
- [128] Leverrier, A. and Garcia-Patron, R. Analysis of circuit imperfections in BosonSampling. *Quantum Information and Computation*, **15**,0489–0512 (2015). <https://arxiv.org/abs/1309.4687> [Cited on page 46.]
- [129] Arkhipov, A. BosonSampling is robust against small errors in the network matrix. *Phys. Rev. A*, **92**,062326 (2015). <https://journals.aps.org/prapdf/10.1103/PhysRevA.92.062326> [Cited on page 46.]
- [130] Brod, D.J. Complexity of simulating constant-depth BosonSampling. *Phys. Rev. A*, **91**,042316 (2015). <https://journals.aps.org/prapdf/10.1103/PhysRevA.91.042316> [Cited on page 46.]
- [131] Rohde, P. P. and Ralph, T. C. Error tolerance of the boson-sampling model for linear optics quantum computing. *Phys. Rev. A*, **85**,022332 (2012). <https://journals.aps.org/prapdf/10.1103/PhysRevA.85.022332> [Cited on page 46.]
- [132] Renema, J.J., Menssen, A., Clements, w., Triginer, G., Kolthammer, S. and Walmsley, I. Efficient algorithm for boson sampling with partially distinguishable photons. *Preprint at arXiv:1707.02793*, (2017). <https://arxiv.org/abs/1707.02793> [Cited on page 46.]
- [133] Motes, K.R., Dowling, J.P., Gilchrist, A. and Rohde, P. P. Implementing BosonSampling with time-bin encoding: Analysis of loss, mode mismatch, and time jitter. *Phys. Rev. A*, **92**,052319 (2015). <https://journals.aps.org/prapdf/10.1103/PhysRevA.92.052319> [Cited on page 46.]
- [134] Wu, J., Liu, Y., Zhang, B., Jin, X., Wang, Y., Wang, H. and Yang, X. Computing Permanents for Boson Sampling on Tianhe-2 Supercomputer. *Preprint at arXiv:1606.05836*, (2016). <https://arxiv.org/abs/1606.05836> [Cited on pages 46 and 141.]

- [135] Neville, A., Sparrow, C., Clifford, R., Johnston, E., Birchall, P. M., Montanaro, A. and Laing, A. Classical boson sampling algorithms with superior performance to near-term experiments. *Nat. Phys.*, doi:10.1038/nphys4270 (2017). <http://www.nature.com/nphys/journal/vaop/ncurrent/full/nphys4270.html> [Cited on pages 46 and 141.]
- [136] Clifford, P. and Clifford, R. The Classical Complexity of Boson Sampling. *preprint in arxiv.org/abs/1706.01260*, (2017). <https://arxiv.org/pdf/1706.01260.pdf> [Cited on page 46.]
- [137] Broome, M. A., Fedrizzi, A., Rahimi-Keshari, S., Dove, J., Aaronson, S., Ralph, T. C. and White, A. G. Photonic boson sampling in a tunable circuit. *Science*, **339.6121**,794–798 (2013). <http://science.sciencemag.org/content/339/6121/794> [Cited on pages 46, 47, and 82.]
- [138] Spring, J. B., Metcalf, B. J., Humphreys, P.C., Kolthammer, W. S., Jin, X. M., Barbieri, M., Datta, A., Thomas-Peter, N. Langford, N. K., Kundys, D., Gates, J. C., Smith, B. J. and Walmsley I. A. Boson Sampling on a photonic chip. *Science*, **339. 6121**,798–801 (2013). <http://science.sciencemag.org/content/339/6121/798> [Cited on pages 46, 47, 56, and 82.]
- [139] Tillmann, M., Dakic, B., Heilmann, R., Nolte, S., Szameit, A. and Walther, P. Experimental Boson Sampling. *Nature Photon.*, **7**,540–544 (2013). <https://www.nature.com/nphoton/journal/v7/n7/full/nphoton.2013.102.html> [Cited on pages 46, 47, and 82.]
- [140] Crespi, A., Osellame, R., Ramponi, R., Brod, D. J. Galvão, E. F., Spagnolo, N., Vitelli, C., Maiorino, E., Mataloni, P. and Sciarrino, F. Integrated multimode interferometers with arbitrary designs for photonic boson sampling. *Nature Photon.*, **7**,545–549 (2013). <http://www.nature.com/nphoton/journal/v7/n7/full/nphoton.2013.112.html> [Cited on pages 46, 47, 82, and 94.]
- [141] Goldstein, S., Korenblit, S., Bendor, Y., You, H., Geller, M.R. and Katz, N. Decoherence and interferometric sensitivity of boson sampling in superconducting resonator networks. *Phys. Rev. B*, **95**,020502 (2017). <https://journals.aps.org/prb/abstract/10.1103/PhysRevB.95.020502> [Cited on page 47.]
- [142] Wang, H., He, Y., Li, Y-H., Su, Z-E., Li, B., Huang, H-L., Ding, X., Chen, M-C., Liu, C., Qin, J., Li, J-P., He, Y-M., Schneider, C., Kamp, M., Peng, C-Z., Hoffing, S., Lu, C-Y and Pan, J.W. High-efficiency multiphoton boson sampling. *Nat. Photon.*, **11**, 361-365 (2017). <https://www.nature.com/nphoton/journal/v11/n6/full/nphoton.2017.63.html> [Cited on pages 47, 101, and 141.]
- [143] Loredo, J.C., Broome, M.A., Hilaire, P., Gazzano, O., Sagnes, I., Lamaitre, A., Almeida, M.P., Senellart, P. and White, A.G. Boson Sampling with Single-Photon Fock States from a Bright Solid-State Source. *Phys. Rev. Lett.*, **118**, 130503 (2017). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.118.130503> [Cited on pages 47 and 101.]
- [144] Motes, K.R., Dowling, J. P. and Rohde, P. P. Spontaneous parametric down-conversion photon sources are scalable in the asymptotic limit for boson sampling. *Phys. Rev. A*, **88**,063822 (2013). <https://journals.aps.org/pra/pdf/10.1103/PhysRevA.88.063822> [Cited on page 47.]
- [145] Y. He, X. Ding, Z-E. Su, H-L. Huang, J. Qin, C. Wang, S. Unsleber, C. Chen, H. Wang, Y-M. He, X-L. Wang, W-J. Zhang, S-J. Chen, C. Schneider, M. Kamp, L-X. You, Z. Wang, S. Höfling, C-Y. Lu and J.W. Pan, "Time-Bin-Encoded Boson Sampling with a Single-Photon Device", *Phys. Rev. Lett.*, **118 (19)**, 190501, 2017. <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.118.190501> [Cited on page 47.]

- [146] Spagnolo, N., Vitelli, C., Sansoni, L., Maiorino, E., Mataloni, P., Sciarrino, F., Brod, D. J., Galvao, E. F., Crespi, A., Ramponi, R. and Osellame, R. General rules for bosonic bunching in multimode interferometers. *Phys. Rev. Lett.* **111**, 130503 (2013). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.111.130503> [Cited on pages 47 and 94.]
- [147] Lund, A. P., Laing, A., Rahimi-Keshari, S., Rudolph, T., O'Brien, J. L. and Ralph, T. C. Boson Sampling from a Gaussian state. *Phys. Rev. Lett.*, **113**,100502 (2014). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.113.100502> [Cited on pages 48, 95, 96, 141, and 168.]
- [148] <http://www.scottaaronson.com/blog/?p=1579> [Cited on pages 48 and 168.]
- [149] Bentivegna, M., Spagnolo, N., Vitelli, C., Flamini, F., Viggianiello, N., Latmiral, L., Mataloni, P., Brod D.J., Galvao, E. F., Crespi, A., Ramponi, R., Osellame, R. and Sciarrino, F. Experimental scattershot boson sampling. *Science Advances*, **1**,e1400255 (2015). <http://advances.sciencemag.org/content/1/3/e1400255/tab-pdf> [Cited on pages 47, 48, 49, 81, 83, 84, 95, 96, 100, 102, 108, 117, 141, and 168.]
- [150] Latmiral, L., Spagnolo, N., and Sciarrino, F. Towards quantum supremacy with lossy scattershot boson sampling. *New J. Phys.*, **18**,113008 (2016). <http://iopscience.iop.org/article/10.1088/1367-2630/18/11/113008> [Cited on pages 49 and 95.]
- [151] Danakas, S. and Aravind P. K. Analogies between two optical systems (photon beam splitters and laser beams) and two quantum systems (the two-dimensional oscillator and the two-dimensional hydrogen atom). *Phys. Rev. A* **45** (3), 1973 (1992). <https://journals.aps.org/prl/abstract/10.1103/PhysRevA.45.1973> [Cited on page 51.]
- [152] Politi, A., Cryan, M. J., Rarity, J. G., Yu, S. and O'Brien, J. L. Silica-on-silicon waveguide quantum circuits. *Science* **320**, 646 (2008). <http://science.sciencemag.org/content/320/5876/646> [Cited on page 52.]
- [153] Bogaerts, W., Baets, R., Dumon, P., Wiaux, V., Beckx, S., Taillaert, D., Luyssaert, B., Van Campenhout, J., Bienstman, P. and Van Thourhout, D. Nanophotonic waveguides in silicon-on-insulator fabricated with CMOS technology. *J. Lightwave Technol.* **23** (1), 401-412 (2005). <https://www.osapublishing.org/jlt/abstract.cfm?uri=jlt-23-1-401> [Cited on page 53.]
- [154] Davanço, M., Ong, J., Shehata, A., Tosi, A., Agha, I., Assefa, S., Xia, F., Green, W., Mookherjee, S. and Srinivasan, K. Telecommunications-band heralded single photons from a silicon nanophotonic chip. *Appl. Phys. Lett.* **100**, 261104 (2012). <http://aip.scitation.org/doi/full/10.1063/1.4711253> [Cited on page 53.]
- [155] Spring, J, Salter, P, Metcalf, B, Humphreys, P, Moore, M., Thomas-Peter, N., Barbieri, M., Jin, X., Langford, N., Kolthammer, S., Booth M. and Walmsley I. On-chip low loss heralded source of pure single photons. *Opt. Express.* **21**, 13522-32 (2013). <https://www.osapublishing.org/oe/abstract.cfm?uri=oe-21-11-13522> [Cited on page 53.]
- [156] Bogaerts, W., De Heyn, P., Van Vaerenbergh, T., DeVos, K., Selvaraja, S. K., Claes, T., Dumon, P., Bienstman, P., Van Thourhout, D. and Baets, R. Silicon Microring Resonators. *Laser Photonics Rev.* **6** (1), 47-73 (2012). <http://onlinelibrary.wiley.com/doi/10.1002/lpor.201100017/pdf> [Cited on page 53.]
- [157] Du, D., Liu, X., Korn, G., Squier, J. and Mourou, G. Laser-induced breakdown by impact ionization in SiO₂ with pulse widths from 7 ns to 150 fs. *Applied physics letters* **64** (23), 3071–3073 (1994). <http://aip.scitation.org/doi/abs/10.1063/1.111350> [Cited on pages 54 and 91.]

- [158] Pronko, P.P., Dutta, S.K., Squier, J., Rudd, J.V., Du, D. and Mourou, G. Machining of sub-micron holes using a femtosecond laser at 800 nm. *Opt. Comm.* **114** (1-2), 106–110 (1995). <http://www.sciencedirect.com/science/article/pii/0030401894005851> [Cited on pages 54 and 91.]
- [159] Davis, K. M., Miura, K., Sugimoto, N., and Hirao, K. Writing waveguides in glass with a femtosecond laser. *Opt. Lett.* **21** (21), 1729–1731 (1996). <https://www.osapublishing.org/ol/abstract.cfm?uri=ol-21-21-1729> [Cited on pages 54 and 91.]
- [160] Li, G., Winick, K., Said, A., Dugan, M. and Bado, P. Waveguide electro-optic modulator in fused silica fabricated by femtosecond laser direct writing and thermal poling. *Opt. Lett.* **31** (6), 739-741 (2006). <https://www.osapublishing.org/ol/abstract.cfm?uri=ol-31-6-739> [Cited on pages 56, 60, and 74.]
- [161] Flamini, F., Magrini, L., Rab, A.S., Spagnolo, N., D’ambrosio, V., Mataloni, P., Sciarrino, F., Zandrini, T., Crespi, A., Ramponi, R. and Osellame, R. Thermally-reconfigurable quantum photonic circuits at telecom wavelength by femtosecond laser micromachining. *Light: Science and Applications* **4**, e354 (2015). <http://www.nature.com/lsa/journal/v4/n11/full/lsa2015127a.html> [Cited on pages 56, 57, 58, 74, and 127.]
- [162] B. J. Smith, D. Kundys, N. Thomas-Peter, P. G. R. Smith, and I. A. Walmsley. Phase-controlled integrated photonic quantum circuits. *Optica* **17** (16), 13516–13525 (2009). [doi:10.1364/OE.17.013516](https://doi.org/10.1364/OE.17.013516) [Cited on page 58.]
- [163] J. C. F. Matthews, A. Politi, A. Stefanov, and J. L. O’Brien. Manipulation of multiphoton entanglement in waveguide quantum circuits. *Nat. Photon.* **3**, 346–350 (2009). [doi:10.1038/nphoton.2009.93](https://doi.org/10.1038/nphoton.2009.93) [Cited on page 58.]
- [164] P. J. Shadbolt, M. R. Verde, A. Peruzzo, A. Politi, A. Laing, M. Lobino, J. C. F. Matthews, M. G. Thompson, and J. L. O’Brien. Generating, manipulating and measuring entanglement and mixture with a reconfigurable photonic circuit. *Nat. Photon.* **6**, 45–49 (2012). [doi:10.1038/nphoton.2011.283](https://doi.org/10.1038/nphoton.2011.283) [Cited on page 58.]
- [165] D. Bonneau, M. Lobino, P. Jiang, C. M. Natarajan, M. G. Tanner, R. H. Hadfield, S. N. Dorenbos, V. Zwiller, M. G. Thompson, and J. L. O’Brien. Fast path and polarization manipulation of telecom wavelength single photons in lithium niobate waveguide devices. *Phys. Rev. Lett.* **108**, 053601 (2012). [doi:10.1103/PhysRevLett.108.053601](https://doi.org/10.1103/PhysRevLett.108.053601) [Cited on page 58.]
- [166] D. Bonneau, E. Engin, K. Ohira, N. Suzuki, H. Yoshida, N. Iizuka, M. Ezaki, C. M. Natarajan, M. G. Tanner, R. H. Hadfield, S. N. Dorenbos, V. Zwiller, J. L. O’Brien, and M. G. Thompson. Quantum interference and manipulation of entanglement in silicon wire waveguide quantum circuits. *New J. Phys.* **14**, 045003 (2012). [doi:10.1088/1367-2630/14/4/045003](https://doi.org/10.1088/1367-2630/14/4/045003) [Cited on page 58.]
- [167] Jin H. et al. On-chip generation and manipulation of entangled photons based on reconfigurable Lithium-Niobate waveguide circuits. *Phys. Rev. Lett.* **113**, 053601 (2014). [doi:10.1103/PhysRevLett.113.053601](https://doi.org/10.1103/PhysRevLett.113.053601) [Cited on page 58.]
- [168] B. J. Metcalf, J. B. Spring, P.C. Humphreys, N. Thomas-Peter, M. Barbieri, W. S. Kolthammer, X. Jin, N. K. Langford, D. Kundys, J. C. Gates, B. J. Smith, P. G. R. Smith, and Walmsley I. A. Quantum teleportation on a photonic chip. *Nat. Photon.* **8**, 770–774 (2014). [doi:10.1038/nphoton.2014.217](https://doi.org/10.1038/nphoton.2014.217) [Cited on page 58.]
- [169] Crespi, A. Suppression laws for multiparticle interference in Sylvester interferometers. *Phys. Rev. A* **91**, 013811 (2015). <https://journals.aps.org/pra/pdf/10.1103/PhysRevA.91.013811> [Cited on pages 72, 78, 89, 94, 96, 117, and 163.]

- [170] Flamini, F., Viggianiello, N., Giordani, T., Bentivegna, M., Spagnolo, N., Crespi, A., Corrielli, G., Osellame, R., Martin-Delgado, M. A. and Sciarrino, F. Observation of Majorization Principle for quantum algorithms via 3-D integrated photonic circuits. *Preprint at arXiv:1608.01141* (2016). <https://arxiv.org/abs/1608.01141> [Cited on pages 61, 81, 92, 93, 94, and 117.]
- [171] Kwiat, P., Mattle, K., Weinfurter, H. and Zeilinger, A. New high-intensity source of polarization-entangled photon pairs. *Phys. Rev. Lett.***75**, 4337–4341 (1995). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.75.4337> [Cited on page 57.]
- [172] Mandel, L. and Wolf, E. Optical Coherence and Quantum Optics. *Cambridge University Press, New York* (1995). [Cited on page 57.]
- [173] Walls, D. Squeezed states of light. *Nature* **306**, 141 (1983). <https://www.nature.com/nature/journal/v306/n5939/abs/306141a0.html> [Cited on page 59.]
- [174] Crespi, A., Bentivegna, M., Pitsios, I., Rusca, D., Poderini, D., Carvacho, G., D’Ambrosio, V., Cabello, A. Sciarrino, F. and Osellame, R. Single-photon quantum contextuality on a chip. *Preprint at arXiv:1705.06237* (2017). <https://arxiv.org/abs/1705.06237> [Cited on pages 60 and 74.]
- [175] Pitsios, I., Banchi, L., Rab, A.S., Bentivegna, M., Caprara, D., Crespi, A., Spagnolo, N., Bose, S., Mataloni, P., Osellame, R. and Sciarrino, F. Photonic Simulation of Entanglement Growth After a Spin Chain Quench. *Preprint at arXiv:1603.02669* (2017). <https://arxiv.org/abs/1603.02669> [Cited on pages 58, 60, and 74.]
- [176] Flamini, F., Spagnolo, N., Viggianiello, N., Crespi, A., Osellame, R. and Sciarrino, F. Benchmarking integrated linear-optical architectures for quantum information processing. *Sci. Rep.* **7**, 15133(2017). <https://www.nature.com/articles/s41598-017-15174-2> [Cited on pages 61, 71, 72, 74, and 79.]
- [177] Flamini, F., Viggianiello, N., Bentivegna, M., Spagnolo, N., Mataloni, P., Crespi, A., Ramponi, R., Osellame, R. and Sciarrino, F. Generalized Quantum Fast Transformations via Femtosecond Laser Writing Technique. *Interdisciplinary Information Sciences***23**, 115-118 (2017). https://www.jstage.jst.go.jp/article/iis/23/1/23_2017.A.14/_article [Cited on pages 61, 66, and 78.]
- [178] Clements, W. R., Humphreys, P. C., Metcalf, B. J., Kolthammer, W. S. and Walmsley, I. A. An optimal design for universal multiport interferometers. *Optica***3**, 1460-1465 (2016). <https://www.osapublishing.org/optica/abstract.cfm?uri=optica-3-12-1460> [Cited on pages 61, 62, 63, 72, 74, 75, 76, 79, 97, and 98.]
- [179] Cooley, J. W. and Tukey, W. An Algorithm for the Machine Calculation of Complex Fourier Series. *Math. Comput.***19**, 297-301 (1965). <https://www.jstor.org/stable/2003354> [Cited on pages 61, 63, and 79.]
- [180] Torma, P., Jex, I. and Stenholm, S. Beam splitter realizations of totally symmetrical mode couplers. *J. Mod. Opt.***43**, 245-251 (1996). <http://www.tandfonline.com/doi/abs/10.1080/09500349608232738> [Cited on pages 61, 63, and 66.]
- [181] Barak, R. and Ben-Aryeh, Y. Quantum fast fourier transform and quantum computation by linear optics. *J. Opt. Soc. Am. B***24**(2), 231–240 (2007). <https://www.osapublishing.org/josab/abstract.cfm?uri=josab-24-2-231> [Cited on pages 61, 63, 64, 66, 67, 72, 78, 79, 91, 92, and 98.]

- [182] Crespi, A., Osellame, R., Ramponi, R., Bentivegna, M., Flamini, F., Spagnolo, N., Viggianiello, N., Innocenti, L., Mataloni, P. and Sciarrino, F. Suppression law of quantum states in a 3D photonic fast Fourier transform chip. *Nat. Commun.* **7**, 10469 (2016). <https://www.nature.com/articles/ncomms10469> [Cited on pages 47, 61, 65, 66, 78, 81, 88, 89, 90, 91, 94, 98, 102, 117, 144, and 146.]
- [183] Peruzzo, A. et al. Multimode quantum interference of photons in multiport integrated devices. *Nat. Comm.* **2**, 224 (2011). <https://www.nature.com/articles/ncomms1228> [Cited on pages 67 and 68.]
- [184] Laing, A. et al. Super-stable tomography of any linear optical device. *preprint at arXiv:1208.2868* (2012). <https://arxiv.org/abs/1208.2868> [Cited on page 67.]
- [185] Lobino, M. et al. Complete characterization of quantum optical processes. *Science* **322**, 563-566 (2008). <http://science.sciencemag.org/content/322/5901/563> [Cited on page 67.]
- [186] Rahimi-Keshari, S. et al. Quantum process tomography with coherent states. *N. Jour. Phys.* **13**, 013006 (2011). <http://iopscience.iop.org/article/10.1088/1367-2630/13/1/013006/meta> [Cited on page 67.]
- [187] Tillmann M. et al. On unitary reconstruction of linear optical networks. *Journal of Optics* **18** (11), 114002 (2016). <http://iopscience.iop.org/article/10.1088/2040-8978/18/11/114002/meta> [Cited on page 67.]
- [188] Rahimi-Keshari, S. et al. Direct characterization of linear-optical networks. *Opt. Expr.* **21**, 13450-13458 (2013). <https://www.osapublishing.org/oe/abstract.cfm?uri=oe-21-11-13450> [Cited on page 67.]
- [189] MacQueen, J. Some methods for classification and analysis of multivariate observations. *In of Calif. Press, U. (ed.) Proc. Fifth Berkeley Symp. on Math. Statist. and Prob.* **1**, 281-297 (1967). <https://projecteuclid.org/euclid.bsmmsp/1200512992> [Cited on page 157.]
- [190] Bottou, L. and Bengio, Y. Convergence properties of the k-means algorithms. *In Advances in Neural Information Processing Systems* **7**, ((MIT press, 1995)). <https://pdfs.semanticscholar.org/2352/d9105de31032538900dfb2ce7c95f6402963.pdf> [Cited on page 157.]
- [191] Burgwal, R., Clements, W.R., Smith, D.H., Gates, J.C., Kolthammer, W.S., Renema, J.J. and Walmsley, I. Implementing random unitaries in an imperfect photonic network. *Preprint at arXiv:1704.01945* (2017). <https://arxiv.org/abs/1704.01945> (Not cited.)
- [192] D. A. Miller, Perfect optics for imperfect implementations. *Optica* **2**, 747-750 (2015). [Cited on page 75.]
- [193] Viggianiello, N., Flamini, F., Innocenti, L., Cozzolino, D., Bentivegna, M., Spagnolo, N., Crespi, A., Brod, D.J., Galvao, E. F., Osellame, R. and Sciarrino, F. Experimental generalized quantum suppression law in Sylvester interferometers. *preprint at arXiv:1705.08650* (2013). <https://arxiv.org/abs/1705.08650> [Cited on pages 82, 101, 102, 103, and 117.]
- [194] Mezzadri, F. How to generate random matrices from the classical compact groups. *Notice of the AMS* **54** (5), 592-604 (1983) (2007). <http://www.ams.org/notices/200705/fea-mezzadri-web.pdf> [Cited on page 82.]
- [195] X. Ma, S. Zotter, J. Kofler, T. Jennewein, and A. Zeilinger. Experimental generation of single photons via active multiplexing. *Phys. Rev. A*, 83:043814, 2011. [Cited on page 84.]

- [196] Gogolin, C., Kliesch, M., Aolita, L. and Eisert, J. Boson-sampling in the light of sample complexity. *preprint at arXiv:1306.3995* (2013). <https://arxiv.org/abs/1306.3995> [Cited on page 85.]
- [197] Aaronson, S. and Arkhipov, A. Bosonsampling is far from uniform. *Quantum Information & Computation* **14** (15-16), 1383–1423 (2014). <https://www.scottaaronson.com/papers/response.pdf> [Cited on pages 85, 94, and 117.]
- [198] Spagnolo, N., Vitelli, C., Bentivegna, M., Brod, D. J., Crespi, A., Flamini, F., Giacomini, S., Milani, G., Ramponi, R., Mataloni, P., Osellame, R., Galvão, E. F. and Sciarrino, F. Experimental validation of photonic boson sampling. *Nat. Photon.* **8** (8), 615–620 (2014). <http://www.nature.com/nphoton/journal/v8/n8/full/nphoton.2014.135.html> [Cited on pages 47, 86, 87, and 117.]
- [199] Carolan, J., Meinecke, J. D. A., Shadbolt, P. J., Russell, N. J., Ismail, N., Wörhoff, K., Rudolph, T., Thompson, M. G., O’Brien, J. L., Matthews, J. C. F. and Laing, A. On the experimental verification of quantum complexity in linear optics. *Nat. Photon.* **8** (8), 621–626 (2014). <http://www.nature.com/nphoton/journal/v8/n8/full/nphoton.2014.152.html> [Cited on pages 47, 86, 87, 88, and 94.]
- [200] Bentivegna, M., Spagnolo, N., Vitelli, C., Brod, D. J., Crespi, A., Flamini, F., Ramponi, R., Mataloni, P., Osellame, R., Galvão, E. F. and Sciarrino, F. Bayesian approach to boson sampling validation. *International Journal of Quantum Information* **12** (07n08), 1560028 (2014). <http://www.worldscientific.com/doi/abs/10.1142/S021974991560028X> [Cited on pages 86, 87, 94, and 153.]
- [201] Tichy, M. C., Mayer, K., Buchleitner, A. and Mølmer, K. Stringent and efficient assessment of boson-sampling devices. *Phys. Rev. Lett.* **113** (2), 020502 (2014). <https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.113.020502> [Cited on pages 78, 87, 88, 89, 91, 94, 102, 110, 117, and 168.]
- [202] Hadzibabic, Z., Stock, S., Battelier, B., Bretin, V. and Dalibard, J. Interference of an array of independent bose-einstein condensates. *Phys. Rev. Lett.* **93**, 180403 (2004). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.93.180403> [Cited on page 87.]
- [203] Cennini, G., Geckler, R., Ritt, G. and Weitz, M. Interference of a variable number of coherent atomic sources. *Phys. Rev.* **A72**, 051601 (2005). <https://journals.aps.org/prl/pdf/10.1103/PhysRevA.72.051601> [Cited on page 87.]
- [204] Cleve, R., Ekert, A., Macchiavello, C. & Mosca, M. Quantum algorithms revisited. In *Proceedings of the Royal Society*, 339–354 (1998). <http://rspa.royalsocietypublishing.org/content/454/1969/339> [Cited on page 89.]
- [205] Latorre, J. I. & Martin-Delgado, M. A. Majorization arrow in quantum-algorithm design. *Physical Review A* **66**, 022–305 (2002). <https://journals.aps.org/prl/abstract/10.1103/PhysRevA.66.022305> [Cited on pages 89 and 91.]
- [206] Orus, R., Latorre, J. I. & Martin-Delgado, M. A. Natural majorization of the quantum fourier transformation in phase-estimation algorithms. *Quantum Information Processing* **1** (4), 283–302 (2002). [Cited on pages 89 and 91.]
- [207] Orus, R., Latorre, J. I. & Martin-Delgado, M. A. Systematic analysis of majorization in quantum algorithms. *The European Physical Journal D (EPJ D)* **29**, 119–132 (2004). [Cited on pages 89 and 92.]

- [208] Viggianiello, N., Flamini, F., Bentivegna, M., Spagnolo, N., Crespi, A., Brod, D. J., Galvao, E. F., Osellame, R. and Sciarrino, F. Optimal photonic indistinguishability tests in multimode networks. *preprint at arXiv:1710.03578* (2017). <https://arxiv.org/abs/1710.03578> [Cited on pages 82, 95, 96, 97, 99, and 101.]
- [209] Liu, K., Lund, A. P., Gu, Y. J. and Ralph, T. C. A certification scheme for the boson sampler. *J. Opt. Soc. Am. B***33**, 1835–1841 (2016). <https://www.osapublishing.org/josab/abstract.cfm?uri=josab-33-9-1835> [Cited on page 94.]
- [210] Walschaers, M., Kuipers, J., Urbina, J.-D., Mayer, K., Tichy, M.C., Richter, K. and Buchleitner, A. Statistical benchmark for bosonsampling. *New J. Phys.***18**, 032001 (2016). <http://iopscience.iop.org/article/10.1088/1367-2630/18/3/032001/meta> [Cited on pages 94, 110, 111, 114, 116, 117, and 160.]
- [211] Bentivegna, M., Spagnolo, N. and Sciarrino, F. Is my boson sampler working?. *New J. Phys.***18**, 041001 (2016). <http://iopscience.iop.org/article/10.1088/1367-2630/18/4/041001> [Cited on pages 87, 94, and 99.]
- [212] Shchesnovich, V. S. Universality of generalized bunching and efficient assessment of boson sampling. *Phys. Rev. Lett.***116**, 123601 (2016). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.116.123601> [Cited on page 94.]
- [213] Wang, J., Bonneau, D., Villa, M., Silverstone, J.W., Santagati, R., Miki, S., Yamashita, T., Fujiwara, M., Sasaki, M., Terai, H., Tanner, M.G., Natarajan, C.M., Hadfield, R., O’Brien, J., and Thompson, M. Chip-to-chip quantum photonic interconnect by path-polarization interconversion. *Optica***3**, 407–413 (2016). <https://www.osapublishing.org/optica/abstract.cfm?uri=optica-3-4-407> [Cited on page 94.]
- [214] Aolita, L., Gogolin, C., Klieasch, M. and Eisert, J. Reliable quantum certification of photonic state preparations. *Nat. Comm.***6**, 8498 (2015). <https://www.nature.com/articles/ncomms9498> [Cited on page 94.]
- [215] Spring, J.B., Mennea, P., Metcalf, B., Humphreys, P., Gates, J., Rogers, H., Soller, C., Smith, B., Kolthammer, S., Smith, P. and Walmsley, I. Chip-based array of near-identical, pure, heralded single-photon sources. *Nat. Comm.***4** (1), 90-96 (2017). <https://www.osapublishing.org/optica/abstract.cfm?uri=optica-4-1-90> [Cited on page 94.]
- [216] Ou, Z. Y. and Mandel, L. Violation of Bell’s inequality and classical probability in a two-photon correlation experiment. *Phys. Rev. Lett.***61**, 50 (1988). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.61.50> [Cited on page 94.]
- [217] Walborn, S. P., Oliveira, A. N., Padua, S. and Monken, C. H. Multimode Hong-Ou-Mandel interference. *Phys. Rev. Lett.***90**, 143601 (2003). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.90.143601> [Cited on page 94.]
- [218] Sagioro, M. A., Olindo, C., Monken, C. H. and Padua, S. Time control of two-photon interference. *Phys. Rev. A***69**, 053817 (2004). <https://journals.aps.org/pra/pdf/10.1103/PhysRevA.69.053817> [Cited on page 94.]
- [219] Spagnolo, N., Vitelli, C., Aparo, L., Mataloni, P., Sciarrino, F., Crespi, A., Ramponi, R. and Osellame, R. Three-photon bosonic coalescence in an integrated tritter. *Nature Commun.***4**, 1606 (2013). <https://www.nature.com/articles/ncomms2616> [Cited on pages 94, 118, 127, and 129.]
- [220] Tillmann, M., Tan, S.H., Stoeckl, S.E., Sanders, B.C., De Guise, H., Heilmann, R., Nolte, S., Szameit, A. and Walther, P. Generalized multiphoton quantum interference. *Phys. Rev.*

- X5, 041015 (2015). <https://journals.aps.org/prx/references/10.1103/PhysRevX.5.041015> [Cited on pages 47, 94, and 144.]
- [221] Menssen, A.J., Jones, A.E., Metcalf, B.J., Tichy, M.C., Barz, S., Kolthammer, S. and Walmsley, I. Distinguishability and many-particle interference. *Phys. Rev. Lett.***118**, 153603 (2017). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.118.153603> [Cited on page 94.]
- [222] Lehmann, E. L. and Romano, J. P. Testing Statistical Hypotheses. *Springer Texts in Statistics* (2005). [Cited on pages 77 and 96.]
- [223] Somaschi, N., Giesz, V., De Santis, L., Loredo, J. C., Almeida, M. P., Hornecker, G., Portalupi, S. L., Grange, T., Antón, C., Demory, J., Gómez, C., Sagnes, I., Lanzillotti-Kimura, N. D., Lemaître, A., Auffeves, A., White, A. G., Lanco, L. and Senellart, P. Near-optimal single-photon sources in the solid state. *Nature Photon.***10**, 340345 (2016). <http://www.nature.com/nphoton/journal/v10/n5/full/nphoton.2016.23.html> [Cited on pages 101, 118, 123, 128, 131, 139, and 141.]
- [224] Loredo, J.C., Zakaria, N., Somaschi, N., Anton, C., De Santis, L., Giesz, V., Grange, T., Broome, M., Gazzano, O., Coppola, G., Sagnes, I., Lemaitre, A., Auffeves, A., Senellart, P., Almeida, M.P. and White, A. Scalable performance in solid-state single-photon sources. *Optica***3**, 433 (2016). <https://www.osapublishing.org/optica/abstract.cfm?uri=optica-3-4-433> [Cited on page 101.]
- [225] Wang, S. and Duan, L. *Certification of Boson Sampling Devices with Coarse-Grained Measurements Preprint at arXiv:1601.02627*(2016). <https://arxiv.org/abs/1601.02627> [Cited on pages 105, 107, and 117.]
- [226] Simon, P. *Too Big to Ignore: The Business Case for Big Data* (Wiley, 2013). [Cited on page 106.]
- [227] Bishop, C.M. *Pattern Recognition and Machine Learning* (Springer, 2006). [Cited on page 106.]
- [228] Murphy, K.P. *Machine Learning: A Probabilistic Perspective* (MIT press, 2012). [Cited on page 106.]
- [229] Biometrics, E. F. Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics* **21**, 768-769 (1965). [Cited on page 107.]
- [230] Lloyd, S. P. Least squares quantization in pcm. *IEEE Transactions on Information Theory* **28**, 129-137 (1982). <http://ieeexplore.ieee.org/document/1056489/> [Cited on page 107.]
- [231] Urbina, J.D., Kuipers, J., Matsumoto, S., Hummel, Q. and Richter, K. Multiparticle Correlations in Mesoscopic Scattering: Boson Sampling, Birthday Paradox, and Hong-Ou-Mandel Profiles. *Phys. Rev. Lett.* **116**, 100401 (2016). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.116.100401> [Cited on page 110.]
- [232] Peruzzo, A., Lobino, M., Matthews, J., Matsuda, N., Politi, A., Poulios, K., Zhou, X-Q., Lahini, Y., Ismail, N., Worhoff, K., Bromberg, Y., Silberberg, Y., Thompson, M. and O'Brien, J. Quantum walks of correlated photons. *Science* **329**, 1500-3 (2010). <http://science.sciencemag.org/content/329/5998/1500> [Cited on page 110.]
- [233] Mehta, M.L. *Random Matrices*. (Amsterdam: Elsevier) (2004). [Cited on page 114.]

- [234] T. K., Ho Random decision forests. *Proceedings of the 3rd International Conference on Document Analysis and Recognition*, 278–282 (1995). <https://dl.acm.org/citation.cfm?id=844681> [Cited on pages 114 and 115.]
- [235] Kurtsiefer, C., Mayer, S., Zarda, P. and Weinfurter, H. Stable solid-state source of single photons. *Phys. Rev. Lett.* **85** (2), 290-293 (2000). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.85.290> [Cited on page 118.]
- [236] Volz, T., Reinhard, A., Winger, M., Badolato, A., Hennessy, K.J., Hu, E. L. and Imamoglu, A. Ultrafast all-optical switching by single photons. *Nat. Photonics* **6**(9), 607-611 (2012). <http://www.nature.com/nphoton/journal/v6/n9/full/nphoton.2012.181.html> [Cited on page 118.]
- [237] Gao, W.B., Fallahi, P., Togan, E., Miguel-Sanchez, J. and Imamoglu, A. Observation of entanglement between a quantum dot spin and a single photon. *Nature* **491**(7), 424-426 (2012). <https://www.nature.com/nature/journal/v491/n7424/full/nature11573.html> [Cited on page 118.]
- [238] Gao, W. B., Imamoglu, A., Bernien, H. and Hanson, R. Coherent manipulation, measurement and entanglement of individual solid-state spins using optical fields. *Nat. Photonics* **9**(6), 363-373 (2015). <https://www.nature.com/nphoton/journal/v9/n6/full/nphoton.2015.58.html> [Cited on page 118.]
- [239] Gazzano, O., Michaelis de Vasconcellos, S., Arnold, C., Nowak, A., Galopin, E., Sagnes, I., Lanco, L., Lemaître, A. and Senellart, P. Bright solid-state sources of indistinguishable single photons. *Nat. Commun.* **4**, 1425 (2013). <https://www.nature.com/articles/ncomms2434> [Cited on page 119.]
- [240] Claudon, J., Bleuse, J., Malik, N.S., Bazin, M., Ja Rennou, P., Gregersen, N., Sauvan, C., Lalanne, P. and Gerard, J.M. A highly efficient single-photon source based on a quantum dot in a photonic nanowire. *Nat. Photon.* **4** (3), 174-177 (2010). <http://www.nature.com/nphoton/journal/v4/n3/full/nphoton.2009.287.html> [Cited on page 119.]
- [241] Santori, C., Fattal, D., Vuckovic, J., Solomon, G. and Yamamoto, Y. Indistinguishable photons from a single-photon device. *Nature* **419** (6907), 594-597 (2002). <http://www.nature.com/nature/journal/v419/n6907/abs/nature01086.html> [Cited on pages 119 and 132.]
- [242] Dousse, A., Lanco, L., Suczynski, J., Semenova, E., Miard, A., Lemaître, A., Sagnes, I., Roblin, C., Bloch, J. and Senellart, P. Controlled light-matter coupling for a single quantum dot embedded in a pillar microcavity using far-field optical lithography. *Phys. Rev. Lett.* **101** (26), 30-33 (2008). <https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.101.267404> [Cited on page 119.]
- [243] Dousse, A., Suczynski, J., Beveratos, A., Krebs, O., Lemaître, A., Sagnes, I., Bloch, J., Voisin, P. and Senellart, P. Ultrabright source of entangled photon pairs. *Nature* **466**, 17-20 (2010). <https://www.nature.com/nature/journal/v466/n7303/full/nature09148.html> [Cited on page 119.]
- [244] Gazzano, O., Almeida, M. P., Nowak, A.K., Portalupi, S. L., Lemaître, A., Sagnes, I., White, A.G. and Senellart, P. Entangling Quantum-Logic Gate Operated with an Ultrabright Semiconductor Single-Photon Source. *Phys. Rev. Lett.* **110** (25), 250501 (2013). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.110.250501> [Cited on page 119.]

- [245] Joyce, P.B., Krzyzewski, T. J., Bell, G. R., Jones, T. S., Malik, S., Childs, D. and Murray, R. Effect of growth rate on the size, composition, and optical properties of InAs/GaAs quantum dots grown by molecular-beam epitaxy. *Journal of Crystal Growth* **227-228**, 1000-1004 (2001). <http://www.sciencedirect.com/science/article/pii/S0022024801009678> [Cited on page 119.]
- [246] Toda, Y., Moriwaki, O., Nishioka, M. and Arakawa, Y. Efficient Carrier Relaxation Mechanism in InGaAs/GaAs Self-Assembled Quantum Dots Based on the Existence of Continuum States. *Phys. Rev. Lett.* **82 (20)**, 4114-4117 (1999). <https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.82.4114> [Cited on page 119.]
- [247] Suczynski, J., Kazimierczuk, T., Goryca, M., Piechal, B., Trajnerowicz, A., Kowalik, K., Kossacki, P., Golnik, A., Korona, K. P., Nawrocki, M., Gaj, J. A. and Karczewski, G. Excitation mechanisms of individual Cd Te Zn Te quantum dots studied by photon correlation spectroscopy. *Phys. Rev. B* **74 (8)**, 085319 (2006). <https://journals.aps.org/prb/pdf/10.1103/PhysRevB.74.085319> [Cited on page 119.]
- [248] Banin, U., Cao, Y., Katz, D. and Millo, O. Identification of atomic-like electronic states in indium arsenide nanocrystal quantum dots. *Nature* **400 (6744)**, 542-544 (1999). <http://www.nature.com/nature/journal/v400/n6744/abs/400542a0.html> [Cited on page 120.]
- [249] Medeiros-Ribeiro, G., Pikus, F. G., Petro, P. M. and Efros, A. L. Single-Electron Charging and Coulomb Interaction in InAs Self-Assembled Quantum Dot Arrays. *Phys. Rev. B* **55 (3)**, 1-7 (1996). <https://journals.aps.org/prb/pdf/10.1103/PhysRevB.55.1568> [Cited on page 120.]
- [250] Moreau, E., Robert, I., Gerard, J. M., Abram, I., Manin, L. and Thierry-Mieg, V. Single-mode solid-state single photon source based on isolated quantum dots in pillar microcavities. *Appl. Phys. Lett.* **79 (18)**, 2865 (2001). <http://aip.scitation.org/doi/abs/10.1063/1.1415346> [Cited on pages 119 and 121.]
- [251] Gammon, D., Snow, E. S., Shanabrook, B. V., Katzer, D. S. and Park, D. Homogeneous Linewidths in the Optical Spectrum of a Single Gallium Arsenide Quantum Dot. *Science* **273 (5271)**, 87-90 (1996). <http://science.sciencemag.org/content/273/5271/87> [Cited on page 121.]
- [252] Jaynes, E.T. and Cummings, F.W. Comparison of quantum and semi-classical radiation theories with application to the beam maser. *Proc. IEEE* **51 (1)**, 89-109 (1963). <http://ieeexplore.ieee.org/document/1443594/> [Cited on page 121.]
- [253] Purcell, E.M. Spontaneous emission probabilities at radio frequencies. *Phys. Rev.* **69**, (1946). <https://journals.aps.org/pr/abstract/10.1103/PhysRev.69.674.2> [Cited on page 122.]
- [254] Englund, D., Fattal, D., Waks, E., Solomon, G., Zhang, B., Nakaoka, T., Arakawa, Y., Yamamoto, Y. and Vulckovic, J. Controlling the Spontaneous Emission Rate of Single Quantum Dots in a Two-Dimensional Photonic Crystal. *Phys. Rev. Lett.* **95 (1)**, 013904 (2005). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.95.013904> [Cited on page 122.]
- [255] Haroche, S. and Raimond, J.M., Exploring the quantum : atoms, cavities, photons. *Oxford Graduate texts*, (2012). [Cited on page 122.]
- [256] Giesz, V. Cavity-enhanced Photon-Photon Interactions With Bright Quantum Dot Sources. *Quantum Physics [quant-ph]* (2015). <https://tel.archives-ouvertes.fr/tel-01272948/document> [Cited on pages 120, 122, and 126.]

- [257] Srinivasan, K., Stintz, A., Krishna, S. and Painter, O. Photoluminescence measurements of quantum-dot-containing semiconductor microdisk resonators using optical fiber taper waveguides. *Phys. Rev. B* **72** (20), 205318 (2005). <https://core.ac.uk/download/pdf/4871116.pdf> [Cited on page 123.]
- [258] Lodahl, P., Van Driel, F., Nikolaev, I.S., Irman, A., Overgaag, K., Vanmaekelbergh, D. and Vos, W.L. Controlling the dynamics of spontaneous emission from quantum dots by photonic crystals. *Nature* **430** (7000), 654-657 (2004). <http://www.nature.com/nature/journal/v430/n7000/abs/nature02772.html> [Cited on page 123.]
- [259] Ding, X., He, Y., Duan, Z.C. Gregersen, N., Chen, M.C., Unsleber, S., Maier, S., Schneider, C., Kamp, M., Hofling, S., Lu, C-Y. and Pan, J-W. On-Demand Single Photons with High Extraction Efficiency and Near-Unity Indistinguishability from a Resonantly Driven Quantum Dot in a Micropillar. *Phys. Rev. Lett.* **116**, 020401 (2016). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.116.020401> [Cited on pages 123 and 141.]
- [260] Dousse, A., Lanco, L., Suczynski, J., Semenova, E., Miard, A., Lemaitre, A., Sagnes, I., Roblin, C., Bloch, J. and Senellart, P. Controlled light-matter coupling for a single quantum dot embedded in a pillar microcavity using far-field optical lithography. *Phys. Rev. Lett.* **101** (26), 30-33 (2008). <https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.101.267404> [Cited on pages 123 and 124.]
- [261] Gazzano, O., Almeida, M. P., Nowak, A. K., Portalupi, S. L., Lemaitre, A., Sagnes, I., White, A. G. and Senellart, P. Entangling Quantum-Logic Gate Operated with an Ultrabright Semiconductor Single-Photon Source. *Phys. Rev. Lett.* **110** (25), 250501 (2013). <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.110.250501> [Cited on page 125.]
- [262] Lenzini, L., Haylock, B., Loredano, J.C., Abrahao, R., Zakaria, N.A., Kasture, S., Sagnes, I., Lemaitre, A., Phan, H-P., Dao, D.V., Senellart, P., Almeida, M.P., White, A.G. and Lobino, M. Active demultiplexing of single-photons from a solid-state source. *Laser Photonics Rev.* **11** (3), (2017). <http://onlinelibrary.wiley.com/doi/10.1002/lpor.201770034/abstract> [Cited on page 140.]
- [263] Huh, J., Guerreschi, G.G., Peropadre, B., McClean, J. and Aspuru-Guzik, A. Boson sampling for molecular vibronic spectra. *Nat. Photon.* **9**, 615–620 (2015). <http://www.nature.com/nphoton/journal/v9/n9/full/nphoton.2015.153.html> [Cited on page 141.]
- [264] Meany, T., Ngah, L., Collins, M., Clark, A., Williams, R., Eggleton, B., Steel, M.J., Withford, M., Alibart, O. and Tanzilli, S. Hybrid photonic circuit for multiplexed heralded single photons. *Laser and photonics reviews* **8** (3), L42–L46 (2014). <http://onlinelibrary.wiley.com/doi/10.1002/lpor.201400027/abstract> [Cited on page 141.]
- [265] Crespi, A., Osellame, R., Ramponi, R., Giovannetti, V., Fazio, R., Sansoni, L., De Nicola, F., Sciarrino, F. and Mataloni, P. Anderson localization of entangled photons in an integrated quantum walk. *Nat. Photon.* **7**, 322–328 (2013). <http://www.nature.com/nphoton/journal/v7/n4/full/nphoton.2013.26.html> [Cited on pages 74, 76, and 141.]
- [266] Shchesnovich V. S. Partial indistinguishability theory for multiphoton experiments in multiport devices. *Phys. Rev. A* **91**, 013844 (2015). <https://journals.aps.org/prl/pdf/10.1103/PhysRevA.91.013844> [Cited on pages 144 and 146.]

- [267] Tichy M.C. Sampling of partially distinguishable bosons and the relation to the multi-dimensional permanent. *Phys. Rev. A* **91**, 022316 (2015). <https://journals.aps.org/prapdf/10.1103/PhysRevA.91.022316> [Cited on pages 144 and 146.]

List of publications

- M. Bentivegna, N. Spagnolo, C. Vitelli, F. Flamini, N. Viggianiello, L. Latmiral, P. Mataloni, D. J. Brod, E. F. Galvao, A. Crespi, R. Ramponi, R. Osellame and F. Sciarrino. *Experimental scattershot boson sampling*. *Sci. Adv.*, **1**, e1400255 (2015). <http://advances.sciencemag.org/content/1/3/e1400255/tab-pdf>
- A. Crespi, R. Osellame, R. Ramponi, M. Bentivegna, F. Flamini, N. Spagnolo, N. Viggianiello, L. Innocenti, P. Mataloni and F. Sciarrino. *Suppression law of quantum states in a 3D photonic fast Fourier transform chip*, *Nat. Comm.*, **7**, 10469 (2016). <https://www.nature.com/articles/ncomms10469>
- F. Flamini, N. Viggianiello, M. Bentivegna, N. Spagnolo, P. Mataloni, A. Crespi, R. Ramponi, R. Osellame, and F. Sciarrino. *Generalized Quantum Fast Transformations via Femtosecond Laser Writing Technique*, *Interdisciplinary Information Sciences*, **23**, 115-118 (2017). https://www.jstage.jst.go.jp/article/iis/23/1/23_2017.A.14/_article
- F. Flamini, N. Spagnolo, N. Viggianiello, A. Crespi, R. Osellame and F. Sciarrino. *Benchmarking integrated linear-optical architectures for quantum information processing*, *Sci. Rep.*, **7**, 15133 (2017). <https://www.nature.com/articles/s41598-017-15174-2>
- T. Giordani, F. Flamini, M. Pompili, N. Viggianiello, N. Spagnolo, A. Crespi, R. Osellame, N. Wiebe, M. Walschaers, A. Buchleitner and F. Sciarrino. *Experimental statistical signature of many-body quantum interference*, *Nat. Photon.*, **12**, 173–178 (2018). <https://www.nature.com/articles/s41566-018-0097-4>
- N. Viggianiello, F. Flamini, L. Innocenti, D. Cozzolino, M. Bentivegna, N. Spagnolo, A. Crespi, D.J. Brod, E. F. Galvao, R. Osellame and F. Sciarrino. *Experimental generalized quantum suppression law in Sylvester interferometers*, *New J. Phys.* doi.org/10.1088/1367-2630/aaad92 (2018). <https://arxiv.org/abs/1705.08650>
- F. Flamini, N. Viggianiello, T. Giordani, M. Bentivegna, N. Spagnolo, A. Crespi, G. Corrielli, R. Osellame, M. A. Martin-Delgado and F. Sciarrino, *Observation of Majorization Principle for quantum algorithms via 3-D integrated photonic circuits*, preprint at arXiv:1608.01141, (2016). <https://arxiv.org/abs/1608.01141>
- N. Viggianiello, F. Flamini, M. Bentivegna, N. Spagnolo, A. Crespi, D. J. Brod, E. F. Galvao, R. Osellame and F. Sciarrino. *Optimal photonic indistinguishability tests in multimode networks*, preprint at arXiv:1710.03578, (2017). <https://arxiv.org/abs/1710.03578>.
- I. Agresti, N. Viggianiello, F. Flamini, N. Spagnolo, A. Crespi, R. Osellame, N. Wiebe and F. Sciarrino. *Pattern recognition techniques for Boson Sampling validation*, preprint at arXiv:1712.06863, (2017) <https://arxiv.org/abs/1712.06863>