

QUANTUM PARAMETER ESTIMATION IN THE LABORATORY

by

Hugo Ferretti

A thesis submitted in conformity with the requirements  
for the degree of Doctor of Philosophy

Graduate Department of Physics  
University of Toronto

© Copyright 2022 by Hugo Ferretti

# Quantum Parameter Estimation in the Laboratory

Hugo Ferretti

Doctor of Philosophy

Graduate Department of Physics

University of Toronto

2022

## Abstract

In this thesis, I report on three projects which I have completed during my PhD. All of these projects are related to quantum parameter estimation.

When simultaneously measuring all the parameters describing a unitary transformation, a decision needs to be made regarding the relative importance of these different parameters. In our first project, we show that the geometry of the group of unitaries in question can guide us toward a natural way of weighting these different parameters. This allows us to transform the matrix Cramér-Rao inequality into a simple parameter-independent scalar inequality. We use this new inequality to calculate the best possible performances for a measurement of the parameters that describe a general  $SU(d)$  unitary and we find an optimal scalable family of quantum states which achieves this performance for  $SU(3)$  unitaries.

In our second project, we experimentally create and characterize the tetrahedron state in the polarization of four photons. This entangled state is the optimal 4-photon state for measuring the three parameters describing a  $SU(2)$  polarization rotation. Interestingly, it possesses the same rotational symmetries as a tetrahedron, giving the state its name. We perform state tomography on our experimental state and measure a quantum state fidelity of  $F = (0.46 \pm 0.02)$  with the ideal state. Our state nevertheless shares many qualitative features with the latter.

Finally, in our last project, we propose a technique named SPLICE which allows us to beat “Rayleigh’s curse”—the inability to measure the separation between two incoherent point sources of light when said separation is below the imaging apparatus’

Rayleigh's criterion. We test this technique in a proof of principle experiment in the laboratory and demonstrate its advantage over traditional methods.

À ma mère et à mon père,

## Acknowledgments

I would first like to thank my supervisor Aephrain Steinberg for his unwavering support and his excellent mentorship. I also give my thanks to two other mentors, postdoctoral fellows Aharon Brodutch and Kent Bonsma-Fisher, for their help and their friendship.

I would further like to thank my colleagues, the graduate students who surrounded me. Edwin, Josiah, David, Aaron, Noah, Arthur, Batuhan, Isabelle, Shreyas, and Ramón, the long hours in the lab would have been quite dreary without your company.

Finally, I would like to thank my girlfriend Chuyun, my parents Vincent and Francine, my sister Camille and my two grandmothers Andrée and Lise. I am greatly indebted to you, and I give you all my love.

# Contents

<b>1</b>	<b>Introduction and Theoretical Background</b>	<b>1</b>
1.1	Introduction . . . . .	1
1.2	Theoretical Background . . . . .	3
1.2.1	Quantum Parameter Estimation . . . . .	3
1.2.2	Classical Fisher Information . . . . .	4
1.2.3	Cramér-Rao Bound . . . . .	5
1.2.4	Fisher Information with Multiple Parameters . . . . .	6
1.2.5	Quantum Fisher Information . . . . .	7
1.2.6	Saturability . . . . .	8
1.2.7	Estimating a Unitary Operation . . . . .	9
<b>2</b>	<b>Intrinsic Sensitivity Limit for Multi-Parameter Quantum Metrology</b>	<b>11</b>
2.1	Motivation . . . . .	11
2.1.1	Estimation of a Magnetic Field with an Atomic Spin . . . . .	12
2.1.2	Estimation of the Action of a Two-Mode Generalized Interferometer with Photons . . . . .	12
2.2	More on the Unitaries' Generators . . . . .	14
2.2.1	Example : Two-Mode Interferometer with Euler Angles . . . . .	16
2.3	The Killing Form and the Cartan Metric . . . . .	17
2.4	SU( $d$ ) on a $d$ -Mode System . . . . .	19
2.5	Optimal States for SU(3) . . . . .	21
2.6	A Note On Mixed States . . . . .	22
2.7	Concluding Remarks . . . . .	22
<b>3</b>	<b>Experimental Generation of the Tetrahedron State</b>	<b>25</b>
3.1	Theory . . . . .	25
3.1.1	Polarization Rotations . . . . .	25
3.1.2	State Visualization . . . . .	27
3.1.3	Choosing the optimal state . . . . .	28

3.1.4	The Tetrahedron State . . . . .	33
3.2	Apparatus . . . . .	35
3.2.1	Concept . . . . .	35
3.2.2	Experimental Scheme . . . . .	37
3.2.3	Hardware . . . . .	37
3.2.4	SPDC . . . . .	39
3.2.5	The N00N State Source . . . . .	43
3.2.6	The Tetrahedron Assembly . . . . .	48
3.2.7	The State Tomography Apparatus . . . . .	50
3.3	Results and Discussion . . . . .	50
3.3.1	Tomography with Hidden Differences . . . . .	50
3.3.2	Results . . . . .	53
3.3.3	Tetrahedral Symmetry . . . . .	54
3.3.4	Reconstructed Eigenstates . . . . .	58
3.3.5	Experimental s-qCRB . . . . .	58
3.4	Background . . . . .	61
3.4.1	Modelling the State Creation . . . . .	62
3.4.2	Modelling the Loss in the PPNRDs . . . . .	64
3.4.3	Calibrating the model . . . . .	67
3.4.4	Model Predictions and Discussion . . . . .	68
3.5	Concluding Remarks . . . . .	69
<b>4</b>	<b>Beating Rayleigh’s Curse Using SPLICE</b>	<b>73</b>
4.1	Motivation . . . . .	73
4.2	Theory . . . . .	74
4.3	Experiment and Results . . . . .	77
4.4	Discussion . . . . .	82
4.5	Concluding Remarks . . . . .	85
<b>5</b>	<b>Conclusion</b>	<b>87</b>
	<b>Bibliography</b>	<b>89</b>
<b>A</b>	<b>Clebsch-Gordan Matrix</b>	<b>97</b>
<b>B</b>	<b>Density Matrix <math>\rho_{exp}</math></b>	<b>99</b>

# Preface

The content of this thesis heavily draws from the following manuscripts :

- A. Z. Goldberg, L. L. Sánchez-Soto, and H. Ferretti, “Intrinsic sensitivity limits for multiparameter quantum metrology”, Phys. Rev. Lett.**127**, 110501 (2021).
- H. Ferretti, K. Bonsma-Fisher, A. Z. Goldberg, N. Lupu-Gladstein, A. O. Pang, A. M. Steinberg, “Creation and Characterization of the Tetrahedron State” (working title), (currently unpublished).
- W.-K. Tham<sup>\*</sup>, H. Ferretti<sup>\*</sup>, and A. M. Steinberg, “Beating Rayleigh’s curse by imaging using phase information”, Phys. Rev. Lett.**118**, 070801 (2017).

<sup>\*</sup>These authors contributed equally to this work.

The co-authors from these publications have contributed to the content of the thesis. This is especially true in Chapter 4, which is in large part borrowed from the last publication in the list above.

During the tenure of my PhD, I also significantly contributed to a few other projects which I won’t discuss in this thesis. They are listed below :

- K. A. Bonsma-Fisher, W.-K. Tham, H. Ferretti, and A. M. Steinberg, “Realistic sub-rayleigh imaging with phase-sensitive measurements”, New Journal of Physics **21**, 09301 (2019).
- W. Tham, H. Ferretti, A. Sadashivan, and A. Steinberg, “Simulating and optimising quantum thermometry using single photons”, Scientific reports **6**, 1–10 (2016).
- W. K. Tham, H. Ferretti, K. Bonsma-Fisher, A. Brodutch, B. C. Sanders, A. M. Steinberg, and S. Jeffery, “Experimental demonstration of quantum fully homomorphic encryption with application in a two-party secure protocol”, Physical Review X **10**, 011038 (2020).

- A. O. Pang, H. Ferretti, N. Lupu-Gladstein, W.-K. Tham, A. Brodutch, K. Bonsma-Fisher, J. Sipe, and A. M. Steinberg, “Experimental comparison of bohm-like theories with different primary ontologies”, *Quantum* **4**, 365 (2020).
- A. Brodutch, N. Lupu-Gladstein, H. Ferretti, W.-K. Tham, A. O. T. Pang, K. Bonsma-Fisher, and A. M. Steinberg, “Do qubits dream of entangled sheep?”, arXiv preprint arXiv:2008.10617(2020).

# List of Tables

- 3.1 Majorana representation of the different eigenstates of  $\rho_{exp}$  with their respective eigenvalues. There are 5 different eigenstates in  $S_2$  and three different eigenstates in  $S_1$ , repeated for each of the 3  $S_1$  sectors. The 2  $S_0$  eigenstates are one dimensional and can therefore not be depicted. 60
- 3.2 List of the different relevant states predicted to be produced incoherently by our experiment. The first row is tetrahedron term followed by background terms due to the loss of a single photon. The first column is an analytical expression for the probability of each of these states. The second column is the numerical evaluation of these probability using the measured value of the parameters of our model. The last column is the density matrix for these states. In order, the origin of loss of a single photon for these 7 background states are: in the tomography apparatus (H photon); in the tomography apparatus (V photon); in the herald mode  $c$  (with another photon present); in the  $b$  mode after the transformation in the tetrahedron assembly stage; in mode  $d$ ; in mode  $e$ ; and in mode  $f$ . . . . . 70



# List of Figures

1.1	Sketch of a typical scenario in Quantum Parameter Estimation. A quantum state $\rho$ is prepared and sent to interact with some environment, which holds information about $\theta$ , the parameters to be estimated. During this process the quantum state $\rho$ is transformed into $\rho_\theta = \Gamma_\theta(\rho)$ . A quantum measurement is then performed on this state, modelled by the POVM $\{M_{\mathbf{x}}\}$ , which results in a specific outcome $\mathbf{x}$ . From this outcome, the parameters $\theta$ can be estimated. . . . .	4
2.1	Scalar Quantum Cramér-Rao Bound $\mathcal{C}$ for the estimation of an SU(d) transformation vs. number of photon used $N$ . Solid lines refer to the performances of generalized N00N states while dashed lines refer to the best possible performances. The N00N states exhibit a shot-noise scaling, with $\mathcal{C} \propto 1/N$ while the optimal bounds satisfy the Heisenberg scaling $\mathcal{C} \propto 1/N^2$ for sufficiently large $N$ . . . . .	21
3.1	Three different visual representations of the spin coherent state $ 4_H, 0_V\rangle$ . (a) Majorana representation of the state. All points are identically placed on the North Pole on the Bloch sphere, aligned with $ H\rangle$ . (b) Wigner distribution on the sphere. The quasi-probability distribution is heavily concentrated on the north pole of the sphere. (c) Equirectangular projection (longitude vs latitude) of the Wigner distribution. . . . .	30
3.2	Three different visual representations of the N00N state $1/\sqrt{2}( 4_H, 0_V\rangle +  0_H, 4_V\rangle)$ . (a) Majorana representation of the state. All points are placed equally spaced on the equator of Bloch sphere. (b) Wigner distribution on the sphere. The quasi-probability distribution is concentrated on the North and South Pole with four fringes along longitude lines of the sphere. (c) Equirectangular projection (longitude vs latitude) of the Wigner distribution. . . . .	31

3.3 Scalar Quantum Cramér-Rao Bound  $\mathcal{C}$  of four different schemes vs Photon Number  $N$  used. The four different colours represent the four different schemes presented in Sec. 3.1.3. While simultaneous parameter estimation with N00N states and sequential parameter estimation with coherent states have shot-noise scalings ( $\mathcal{C} \propto 1/N$ ), sequential parameter estimation with N00N states and simultaneous parameter estimation with Platonic states reach the Heisenberg scaling ( $\mathcal{C} \propto 1/N^2$ ), with the latter beating the former by a factor of 3 for large enough  $N$ . . . . . 34

3.4 Three different visual representations of the tetrahedron state  $1/\sqrt{3}(|4_H, 0_V\rangle + \sqrt{2}|1_H, 3_V\rangle)$ . (a) Majorana representation of the state. The points are placed on the vertices of a tetrahedron inscribed in the Bloch sphere (b) Wigner distribution on the sphere. The quasi-probability distribution has fringes in the shape of a tetrahedron. (c) Equirectangular projection (longitude vs latitude) of the Wigner distribution. . . . . 34

3.5 Conceptual illustration of the tetrahedron state's creation process using the Majorana representation. The H polarization of a 3-photon N00N state (three red points on the vertices of an equilateral triangle on the equator) is partially attenuated resulting in a triangle of points with lower latitude (red). Simultaneously, an H polarized heralded single-photon is probabilistically added (green). The result is a 4-photon state represented by four points on the vertices of a tetrahedron. . . . . 36

3.6 Experiment overview. (a) A 808 nm pulsed beam (red) is used to generate a 404 nm beam (blue) through second harmonic generation (SHG) in a 1 mm barium borate (BBO) crystal. Part of this blue beam is sent to generate co-propagating pairs of horizontally-polarized 808 nm photons (red) through type-I spontaneous parametric down-conversion (SPDC) in a 2 mm BBO crystal while the red beam is attenuated and rotated to a vertical polarization. The two beams are recombined on a polarizing beamsplitter (PBS) and polarization rotated to form a three-photon N00N state in the H/V basis. (b) Another blue beam then pumps a second SPDC (2mm BBO) to generate another photon pair. Here, after a 3 nm narrow-band filter, a detection at a single photon counting module (SPCM) labelled “T” heralds the presence of a photon in the upper path. (c) Both the N00N state and heralded single photon enter a displaced polarization Sagnac interferometer. The heralded photon strictly takes the clockwise path while the N00N state is split equally amongst the paths. The half-waveplate (HWP) in the counter-clockwise path is set at  $45^\circ$  to transmit all the light in the upwards output while the HWP in the clockwise path is set at  $22.5^\circ$  to attenuate the V polarization of the N00N state by half and transmit the heralded single-photon half of the time (see Figure 3.5 for context). After being re-combined, all photons pass through a 3 nm narrow-band filter to make sure their spectra are the same. (d) The state is transmitted via single-mode (SM) fibre to a tomography apparatus for projection onto an arbitrary polarization axis. Each output path is coupled to a multi-mode fibre-splitter to allow for partial photon number counting capabilities (a maximum of three detections per polarization). Five-fold detection events between T and any four SPCMs in the tomography setup are recorded. . . . . 38

3.7 (a and c) Normalized coincidence count rate of the colinear SPDC source in the N00N state sources as a function of the collection lens focal length under various settings for the pump lens focal length and the crystal width. (b and d) Collection efficiency as a function of the collection lens focal length the same conditions. (c - d) are taken with the photons from the SPDC going through a 3 nm filter while in (a - b), the filter is removed. . . . . 42

3.8	Sketch of an ideal N00N state source. In (a), the N00N state source is created in the two output mode of the BS. In (b), the N00N state source is created in the two polarization mode of the output beam. . . . .	44
3.9	Normalized single photon spectra of the SPDC photons from the N00N state source (red) and of the coherent state from the N00N state source (blue). The spectra were taken with a grating spectrometer followed by a cooled CCD. . . . .	45
3.10	Coincidence rate after a $\langle D, D  $ polarization projection of the photons from the N00N state source as a function of (a) the translation stage position governing the coarse path length difference between the SPDC photon and the coherent state photons and (b) the voltage in the piezo-electric actuator governing the fine path length difference. . . . .	47
3.11	Triple coincidence rate between the herald detector and the pairs of detectors placed at the output of the tetrahedron assembly stage as a function of the translation stage position controlling the path length of the heralded single photon before the tetrahedron assembly. The peak we see is the result of a reverse HOM effect, where photon bunching is observed in a single output of the PBS in the tetrahedron assembly stage where the heralded single photon meet a coherent state photon from the N00N state source. . . . .	49
3.12	(a) Layout of the full density matrix. $\rho_{sym}$ is the part of the state contained in full-symmetric $S_2$ sector(blue). The $S_1$ (red) and $S_0$ (green) sectors are also shown. (b) The full tomographically reconstructed density matrix, $\rho_{exp}$ . In hidden differences tomography, coherences between spin sectors are neglected and sectors of the same spins are chosen to be identical. (c) Real and imaginary parts of the reconstructed $\rho_{sym}$ . The density matrix of the ideal tetrahedron state is shown for comparison (dashed). Elements of the density matrix are labeled with spin notation $ j, m\rangle$ . (d) Real and imaginary parts of a $S_1$ sector. The population in each $S_0$ sector is less than $10^{-10}$ . . . . .	52
3.13	Illustration of the polarization basis measured during the tomography. All the axes go through the centre of the faces of a small rhombicuboctahedron. Vectors opposite of one another represent the same basis rotated so that the meaning of a transmission and reflection through the PBS are interchanged. For example, the vectors going through the North (South) poles indicate a measurement where the H (V) photons are transmitted through the PBS. . . . .	53

3.14	Wigner distribution of $\rho_{sym}$ , our reconstructed experimentally created tetrahedron state. . . . .	55
3.15	(a - d) Equirectangular projection (Longitude vs. Latitude map) of the Wigner function of $\rho_{sym}$ rotated in 4 different ways to give different perspective on the Wigner function's features. The 4 different rotations are all anti-clockwise $2\pi/3$ rotations around axes defined by each of the vertices of the ideal tetrahedron. All these rotations would leave $ \psi_{tetra}\rangle$ unchanged. The similitude between these four figures is indicative that $\rho_{sym}$ qualitatively possesses tetrahedral symmetry. . . . .	56
3.16	Projection of a $\rho_{exp}$ and $\rho_{ideal} =  \psi_{tetra}\rangle \langle\psi_{tetra} $ onto the the tetrahedron state $\langle\psi_{tetra} $ vs. the rotation angle $\theta$ . The rotation is around three different known axes, $x$ , $y$ and $z$ . Both $\rho_{exp}$ and $\rho_{ideal}$ have three fringes per rotation, a feature of the tetrahedron state. . . . .	57
3.17	Sensitivity of a $\rho_{exp}$ and $\rho_{ideal} =  \psi_{tetra}\rangle \langle\psi_{tetra} $ to the angle of a rotation around a known axis when projected back onto $\langle\psi_{tetra} $ vs. the rotation angle $\theta$ . The rotation is around three different known axes, $x$ , $y$ and $z$ . . . . .	57
3.18	Majorana representation of the eigenstate of $\rho_{exp}$ with the biggest eigenvalue (blue) and Majorana representation of $ \psi_{tetra}\rangle$ (red). . . . .	59
3.19	Scalar quantum Cramér-Rao for the estimation of SU(2) polarization rotation of $\rho_{exp}$ and of its main eigenstate, both with a photon number of 4. These two s-qCRB are plotted with the performances of the multiple measurement strategies discussed in Sec. 3.1.3. We see that the main eigenstate of $\rho_{exp}$ comes close to the optimal strategy while $\rho_{exp}$ is too noisy. . . . .	61
3.20	Schema of the experiment with the added virtual beam splitters modelling the loss channels of the experiments. These introduces the three new spatial modes $d$ , $e$ and $f$ , which are populated by lost photons. . . . .	63
3.21	(a) Schematic description of both the transmitted and the reflected port of the PBS in the state tomography apparatus. The beam is split by optical fibre beam splitter onto three detectors. This allows the detection of a maximum of three photons per port. In this example, the two incoming photons can be both detected if they fall on different detectors or one photon can be lost if they both fall on the same detector. (b) A sketch of a general probabilistic photon number resolving detector, with $m$ detectors. . . . .	64
4.1	Theory plot of Fisher information for IPC, SPLICE and Binary SPADE vs beam separation $\delta$ , normalized to units of $N/4\sigma^2$ and $\sigma$ respectively. . . . .	76

4.2	Shown is the experimental apparatus. In the lower right-hand box is a representation of SPLICE, the measurement scheme tested in this experiment. In the upper right-hand box is a sketch of the spatial profile of the electromagnetic field before the measurement. The rest of the figure depicts the device used to simulate the two light sources, which can be displaced around their centroid by the displacement of the top mirror. . . . .	77
4.3	Raw data plot for SPLICE coarse scans. Dots are experimental photon coincidence counts plotted versus actual beam separation $\delta$ . Solid overlay is a fit to equation 4.6. . . . .	79
4.4	Calibration data plot for SPLICE scans. Dots are experimental photon single counts plotted versus actual beam separation $\delta$ . Solid overlay is a fit to Eq.4.6. From this fit, the value of the beamwaist is taken to be $0.46mm$ . This value is larger than the measured value of the waist of the two beams ( $0.420mm$ and $0.434mm$ ). Furthermore, the fit here overshoots the data for extreme values of $\delta$ . We do not have an adequate explanation for this behavior. However, the important fact is that the fit is in good agreement with the data for central values of $\delta$ , which is the important regime. This allowed us to use this parameter as a calibration to extract a separation from our SPLICE measurement. The other parameters for the calibration were taken from similar data but using coincidence counts. . . . .	80
4.5	Inferred separation vs known actual separation for a) SPLICE (from “lookup” on calibration curve) and b) IPC. Note that the spread in the IPC estimates grows drastically as $\delta \rightarrow 0$ , while the spread for SPLICE remains essentially constant. . . . .	80

4.6 (a) Renormalized Standard Deviation (SD) and (b) Un-normalized Root Mean-Square Error (RMSE) in the estimated separation plotted as functions of actual separation for both IPC and SPLICE. We plot the computed SD multiplied by  $\sqrt{N}$  to compensate for the scaling of the uncertainty with the size of the data set. The solid and dashed curves are the corresponding Monte Carlo simulations. The dotted curve is the CRB for IPC and the dashed horizontal line represents the absolute fundamental limit of  $2\sigma/\sqrt{N}$ . The RMSE (unlike SD) is not similarly rescaled. It allows us to gauge absolute error relative to the known value of the parameter being estimated so that biases are accounted for. Note that two methods were used in the fitting of IPC data to equation 4.7; for small  $\delta$  ( $< 0.65mm$ ), equation 4.7 was expanded to 2nd order and linear regression was performed whereas for large  $\delta$  ( $> 0.4mm$ ), a nonlinear fitting routine built into Mathematica was used. . . . . 81

4.7 Mean estimated  $\delta$  for IPC and SPLICE plotted against known actual  $\delta$ . Two methods were used in the fitting of IPC data to equation 4.7; for small  $\delta$  ( $< 0.65mm$ ), equation 4.7 was expanded to 2nd order and linear regression was performed whereas for large  $\delta$  ( $> 0.4mm$ ), a nonlinear fitting routine built into Mathematica was used. . . . . 83



# Chapter 1

## Introduction and Theoretical Background

### 1.1 Introduction

All the projects presented in this thesis concern Quantum Metrology, a field within which I made my niche. Metrology is a scientific field that consists of the study of measurements. It is a varied practice that has applications in a wide range of human activities. Quantum Metrology, in turn, studies the limitations placed on measurements when quantum effects are taken into account [1–5]. Typically this is done by considering the measurement device to be a quantum system. It is a field of study with much promise, which could lead to the significant improvement of many precision measurements by making use of quantum correlations to outperform their standard counterparts.

One of the classic parameter estimation problems used to demonstrate the advantage of quantum metrology is the measurement of a phase difference  $\theta$  between the two arms  $a$  and  $b$  of a balanced interferometer. For the best measurement scheme using  $N$  separable photons, the error in our estimate of  $\theta$ ,  $\Delta\theta$ , scales as  $\Delta\theta \sim 1/\sqrt{N}$ . This is called the “shot-noise limit” [6–8], and is reflective of the fact that the measurement process is done independently for each photon. On the other hand, the best general measurement strategy for this estimation is to send the photons in a superposition of being all in path  $a$  and of being all in path  $b$ . This highly entangled configuration is called the “N00N state” [9–12] and with it, one can in theory achieve the Heisenberg limit [3, 8, 13] of  $\Delta\theta \sim 1/N$ , which is the best scaling allowed by quantum mechanics. Similar advantages resulting from a proper utilization of quantum correlations can be found in a plethora of different metrological scenarios.

At the moment, most of the results of the field consist of theoretical results and

proof-of-principle experiments. Some ideas in the field however have already been implemented in real experiments of considerable importance [14–17]. Furthermore, the current limitation for many ideas from the field, our lack of ability to control quantum systems, is rapidly improving. Although this progress is also driven by an effort to develop practical quantum computers, quantum metrology is well positioned to benefit from this progress and to develop rapidly in the coming years.

In Section 1.2 of this chapter, we introduce a general template of a parameter estimation scenario and we introduce methods typically used to calculate the performance of different estimation schemes. These rely on the Fisher Information (FI), the Cramér-Rao Bound (CRB) and their quantum analogue, the Quantum Fisher Information (QFI) and the quantum Cramér-Rao Bound (qCRB), all of which will be defined later.

In recent years, there has been much discussion regarding multi-parameter metrology, the study of measurement of more than one quantity. It was found that, when making optimal use of quantum correlations, measurement strategies that simultaneously estimate these multiple quantities outperform strategies where the quantities are individually measured [18]. In single-parameter scenarios, the qCRB is a scalar lower bound on the variance of the parameter to be estimated. For a given measurement, it is a property of the quantum state originally used as a probe. Minimizing the qCRB is a commonly used strategy to find the optimal probe quantum state for a given measurement. In multi-parameter scenarios, a similar analysis based on the QFI and the qCRB yields a matrix lower bound and leaves us with no clear scalar figure of merit to find an optimal probe quantum state. Several comprehensive reviews highlight recent progress in this subject area [19–24]. In Chapter 2, we address the case of unitary estimation, where the quantities to be measured are the parameters used as a full description of a unitary operation. Based on a geometric argument, we introduce a scalar form of the qCRB, which allows us to find unambiguous optimal probe quantum states for this task. We explore the desirable properties of this simple bound and apply it to unitaries drawn from different groups.

In Chapter 3, we use our results from Chapter 2 to rediscover the optimal states to characterize unitaries drawn from the  $SU(2)$  group. These states are useful for a wide range of applications, including magnetometry, the measurement of rotations of the orbital angular momentum of light [25], and the measurement of polarization rotations. These states have previously generated interest in the literature due some of their interesting properties, their beautiful symmetries, and their metrological power. In the laboratory, we create the entangled “tetrahedron state” in the polarization of four indistinguishable photons. This is the unique optimal four-photon state for

characterizing polarization rotations. We create it using spontaneous parametric down-conversion (SPDC) and we characterize it by performing the full tomography of the state.

Finally, in Chapter 4, we change gears and take a look at the task of measuring the distance between two incoherent point sources of light. In traditional imaging techniques, light from the point sources is focused in an imaging device onto an image plane. At every position on this image plane, the intensity of the electromagnetic field is recorded. It was found in a 2015 publication [26] that, while traditional imaging techniques provide little information of that separation when it is below the Rayleigh criterion of the imaging apparatus, the information is still present in the traditionally discarded *phase* of the electromagnetic field at the image plane. We designed a simple technique called SPLICE, which allows us to use that information to better estimate the separation of closely separated point sources. This technique differs from previous super-resolution techniques [27, 28] by requiring no control of the illumination of the point-sources, and therefore could play a big role in fields such as astronomy, where such control is impossible. We tested our technique in a proof-of-principle experiment in the laboratory.

## 1.2 Theoretical Background

In this section, we introduce general theoretical concepts that are widely used in quantum metrology. This is aimed to be a quick introduction or a refresher, and will also allow us to set the notation for the rest of the document.

### 1.2.1 Quantum Parameter Estimation

Quantum parameter estimation is the task of estimating numerical quantities encoded in a quantum state. In a typical scenario, these parameters are physical quantities that have been encoded in a quantum probe system through a dynamical process. We depict a sketch of it in Figure 1.1.

In this case, the goal of the measurement is to estimate an array of parameters  $\boldsymbol{\theta}$ . These parameters are simply real numbers, most likely physical quantities of interest. A probe quantum system is prepared in a definite quantum state, represented by the density matrix  $\rho$ , and sent through the measurement. During this process, some information about the parameters  $\boldsymbol{\theta}$  is encoded into the quantum state of the probe through the process map  $\Gamma_{\boldsymbol{\theta}}$ . A subsequent quantum measurement of the probe  $\rho_{\boldsymbol{\theta}} = \Gamma_{\boldsymbol{\theta}}(\rho)$ , modelled by the POVM  $\{\hat{M}_{\mathbf{X}}\}$ , yields some results  $\mathbf{X}$  which reveal these parameters through an estimator function  $\tilde{\boldsymbol{\theta}}(\mathbf{X})$ . For the estimate to be accurate,

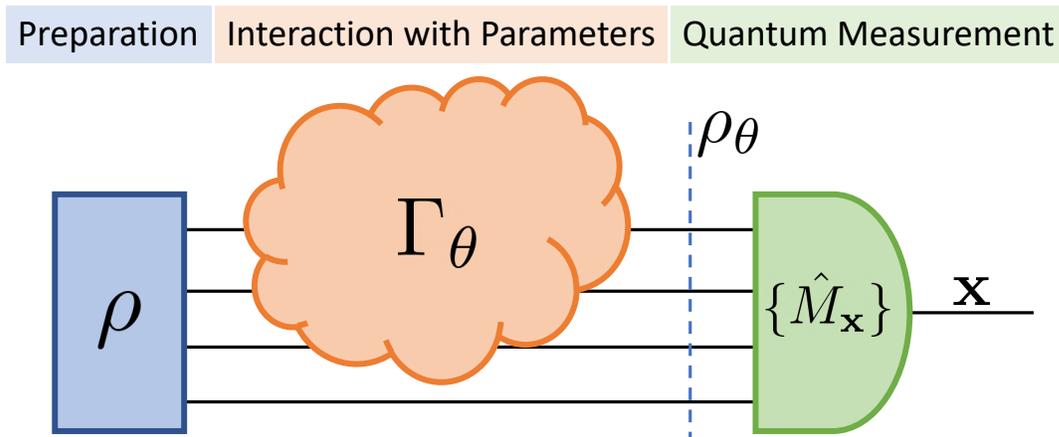


Figure 1.1: Sketch of a typical scenario in Quantum Parameter Estimation. A quantum state  $\rho$  is prepared and sent to interact with some environment, which holds information about  $\theta$ , the parameters to be estimated. During this process the quantum state  $\rho$  is transformed into  $\rho_\theta = \Gamma_\theta(\rho)$ . A quantum measurement is then performed on this state, modelled by the POVM  $\{M_x\}$ , which results in a specific outcome  $x$ . From this outcome, the parameters  $\theta$  can be estimated.

sufficient control of the quantum state needs to be maintained throughout the process. Because of the fundamental probabilistic nature of quantum mechanics, the outcomes of the final measurement are noisy, which puts fundamental limits on the precision with which we can estimate the relevant parameters. In quantum metrology, one aims to optimize the probe's initial state, its interaction with the parameters, and the final quantum measurement to yield a more precise estimate of  $\theta$ .

## 1.2.2 Classical Fisher Information

The Fisher Information (FI) [29–31] or Classical Fisher Information allows us to calculate the amount of information that a set of data contains about previously unknown parameters. When trying to design experiments to measure said physical parameters, optimizing the Fisher Information in the expected data for that parameter is a great way to ensure that the data collected are sensitive to said parameters and can therefore be used to estimate them.

Let's say a physicist wants to measure a parameter  $\theta$  in the laboratory, they will engineer an experiment which yields a vector of data  $\mathbf{X}$ . In the context of Figure 1.1,  $\mathbf{X}$  would correspond to the results of the quantum measurement obtained at the very end of the process. To simplify things for now, we assume only a single parameter is to be estimated. The probability distribution for obtaining a certain data vector given

the value of the parameter  $\theta$  is  $p(\mathbf{X}|\theta)$ . It is often called the likelihood function. In that context however, the parameter  $\theta$  is considered the argument of the function.

For given values of the random variable  $\mathbf{X}$ , we can calculate the score  $s(\theta)$  [29]

$$s(\theta) = \frac{\partial}{\partial\theta} \log(p(\mathbf{X}|\theta)). \quad (1.1)$$

The score is the derivative of the log-likelihood function and setting its value to 0 gives us the maximum likelihood estimator for  $\theta$ . The average value of the score is also 0, which is easy to prove using integration by parts with well-behaved boundary conditions. In general however, its magnitude is an indication of how sensitive the log-likelihood is to a change to the parameter  $\theta$ , which makes the variance of the score an intuitive metric to see if the distribution  $p(\mathbf{X}|\theta)$  is sensitive to the parameter  $\theta$ . The variance of the score is the Fisher Information, which we label  $\mathcal{I}$ . Under the assumption that  $p(\mathbf{X}|\theta)$  is well behaved [30, 31], we can write

$$\mathcal{I} = \text{E} \left[ \left( \frac{\partial}{\partial\theta} \log p(\mathbf{X}|\theta) \right)^2 \right] = \text{E} \left[ -\frac{\partial^2}{\partial\theta^2} \log p(\mathbf{X}|\theta) \right]. \quad (1.2)$$

The second part of this equation makes intuitive sense on its own. It looks like the average of the average of the log-likelihood function's curvature. In cases where the data are very sensitive to  $\theta$ , we expect very sharply peaked likelihood functions and therefore high curvatures close to the peak of the distributions.

### 1.2.3 Cramér-Rao Bound

The Cramér-Rao Bound (CRB) is in part what makes the FI so useful. It gives us a lower bound on the variance of any unbiased estimator of  $\theta$ . In the previous section, we mentioned the maximum-likelihood estimator for  $\theta$ , but in principle, there could be many other estimators. An estimator is simply a function of the data  $\tilde{\theta}(\mathbf{X})$ . The bias of an estimator  $b(\theta)$  is the average difference between the estimator and the true value

$$b(\tilde{\theta}) = \text{E}[\tilde{\theta}(\mathbf{X})] - \theta. \quad (1.3)$$

An unbiased estimator has a bias of 0. For a good measurement of  $\theta$ , we aim for the variance in the unbiased estimator to be as low as possible. The CRB takes a value of the inverse of the FI and bounds this variance from below,

$$\text{Var}[\tilde{\theta}(\mathbf{X})] \geq \frac{1}{\mathcal{I}}. \quad (1.4)$$

The power of this bound is that it gives us an idea of the best possible error we could get for the estimated parameter, regardless of the way we analyze the data. Using the FI to design an experiment can therefore effectively decouple the performance of the experiment itself from the performance of the analysis.

A big caveat however, is that this bound is only valid if the estimator is unbiased. A trivial example of this bound failing for a biased estimator is if we arbitrarily set the estimator function to be  $\tilde{\theta}(\mathbf{X}) = a$ , where the estimator gives the value  $a$  regardless of  $\mathbf{X}$ . Then, the variance in that estimator is trivially 0, beating the CRB. An estimator like this one is not useful. In later sections however, we will discuss realistic cases where biased estimators are to be seriously considered. In these cases, the CRB defined here is not a very useful tool.

#### 1.2.4 Fisher Information with Multiple Parameters

In many experiments, the vector of data  $\mathbf{X}$  is used to estimate more than one parameter. If we aim to estimate the vector of unknown parameters  $\boldsymbol{\theta}$ , the FI takes the form of a square matrix

$$[\mathcal{I}]_{lm} = \text{E} \left[ -\frac{\partial^2}{\partial \theta_l \partial \theta_m} \log p(\mathbf{X} | \boldsymbol{\theta}) \right]. \quad (1.5)$$

Then, given an unbiased estimator  $\tilde{\boldsymbol{\theta}}$ , with

$$\text{E}[\tilde{\boldsymbol{\theta}}] = \boldsymbol{\theta}, \quad (1.6)$$

the CRB takes the form of a bound on the covariance matrix of the estimators

$$\left[ \text{Cov}[\tilde{\boldsymbol{\theta}}(\mathbf{X})] \right]_{l,m} = \text{E} \left[ \left( \tilde{\theta}_l - \theta_l \right) \left( \tilde{\theta}_m - \theta_m \right) \right] \quad (1.7)$$

and the Cramér-Rao inequality becomes a matrix inequality

$$\text{Cov}[\tilde{\boldsymbol{\theta}}(\mathbf{X})] \geq \mathcal{I}^{-1}. \quad (1.8)$$

The right hand side of the inequality here is the matrix inverse of the QFI matrix and the matrix inequality  $A \geq B$  means that  $A - B$  is positive semi-definite.

Mostly, the FI satisfies the same purpose for multi-parameter estimation as it does for single-parameter estimation. One can note however that the FI matrix could be non-invertible, in which case the data are insufficient to separately estimate all the parameters. This kind of problem often arises when, for example,  $p(\mathbf{X} | \theta_1, \theta_2)$  is a function of only a linear combination of  $\theta_1$  and  $\theta_2$ . Then, only multiples of said linear combinations can be estimated but the data are completely insensitive to any other

combinations.

### 1.2.5 Quantum Fisher Information

The Quantum Fisher Information (QFI)[32–34] is similar to its classical analogue. Instead of looking at the amount of information that a classical probability distribution  $p(\mathbf{X} | \boldsymbol{\theta})$  contains about certain parameters, however, we now attempt to quantify the amount of information that a quantum state  $\rho_{\boldsymbol{\theta}}$  contains about said parameters. Once again, the power of the QFI is that it decouples different parts of the experimental design. If we aim to maximize the QFI, we optimize the way the parameters are encoded into the quantum state without having to consider the measurement process yet. The QFI is well liked by theorists since it gives general and powerful results but it should be used with care as it can lead some to ignore the sometimes very relevant less-than-ideal conditions of real experiments. In the context of Figure 1.1, the QFI is a property of  $\rho_{\boldsymbol{\theta}}$ , the quantum state in which the vector of parameter  $\boldsymbol{\theta}$  is encoded.

Again, we aim to estimate  $\boldsymbol{\theta} \in \mathbb{R}^d$  encoded into  $\rho_{\boldsymbol{\theta}}$ . In general, a measurement used to estimate  $\boldsymbol{\theta}$  is a POVM  $\{M_m\}$  with an estimator function  $\tilde{\boldsymbol{\theta}}(m)$ . The observable  $\hat{\boldsymbol{\theta}} = \sum_m M_m \tilde{\boldsymbol{\theta}}(m)$  acts as our estimator. To make use of the quantum CRB (qCRB), which we will define later, we require the estimator to be locally unbiased, a weaker version of unbiasedness. We require that the following expressions evaluated at  $\boldsymbol{\theta}_0$  hold :

$$\begin{aligned} \langle \hat{\boldsymbol{\theta}} \rangle &= \sum_m \text{Tr}[M_m \rho_{\boldsymbol{\theta}}] \tilde{\boldsymbol{\theta}}(m) = \boldsymbol{\theta}_0 \\ \nabla \langle \hat{\boldsymbol{\theta}} \rangle &= \sum_m \text{Tr}[M_m \nabla \rho_{\boldsymbol{\theta}}] \tilde{\boldsymbol{\theta}}(m) = \mathbf{1}, \end{aligned} \tag{1.9}$$

where  $\mathbf{1}$  is the  $d \times d$  identity matrix and the derivatives are with respect to the  $d$  linearly independent parameters. These ensure that the estimator is unbiased in a neighborhood of  $\boldsymbol{\theta}_0$ . The stronger condition of full unbiasedness would be that

$$\langle \hat{\boldsymbol{\theta}} \rangle = \sum_m \text{Tr}[M_m \rho_{\boldsymbol{\theta}}] \tilde{\boldsymbol{\theta}}(m) = \boldsymbol{\theta} \quad \forall \quad \boldsymbol{\theta}. \tag{1.10}$$

The QFI does not depend on the estimator but only on the quantum state. It is often defined in terms of the symmetric logarithmic derivative operators  $\hat{L}$  (SLD). These operators are defined implicitly by the equality

$$\partial_{\theta_l} \hat{\rho}_{\boldsymbol{\theta}} = \frac{1}{2} \left( \hat{\rho}_{\boldsymbol{\theta}} \hat{L}_l + \hat{L}_l \hat{\rho}_{\boldsymbol{\theta}} \right). \tag{1.11}$$

The QFI then takes the form

$$[I_{\rho, \boldsymbol{\theta}}]_{l, m} = \text{Tr} \left[ \hat{\rho}_{\boldsymbol{\theta}} \left\{ \hat{L}_l, \hat{L}_m \right\} \right], \quad (1.12)$$

where  $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$  is the anti-commutator. Again, it is a  $d \times d$  matrix and its inverse is the qCRB with

$$\text{Cov}[\hat{\boldsymbol{\theta}}] = \left\langle (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^\top \right\rangle \geq I_{\boldsymbol{\theta}}^{-1}. \quad (1.13)$$

Note that again, the QFI may not be an invertible matrix, which simply means that the information in the state is insufficient to simultaneously estimate all the parameters. When it comes using the QFI in practice, this inequality is not very useful, as a matrix inequality does not give us a clear metric to quantify the performance of  $\rho_{\boldsymbol{\theta}}$ . In a real life scenario, we would have an idea of the importance of each of the parameters to estimate and we would be able to weight the different element of  $\text{Cov}[\hat{\boldsymbol{\theta}}]$  accordingly. By defining a weight matrix  $W > 0$ , we define those weights and convert the above matrix inequality to a scalar one

$$\text{wMSE}(\hat{\boldsymbol{\theta}}) := \text{Tr}[W \text{Cov}[\hat{\boldsymbol{\theta}}]] \geq \text{Tr}[W I_{\boldsymbol{\theta}}^{-1}] =: \mathcal{C}(\boldsymbol{\theta}). \quad (1.14)$$

Here the  $\text{wMSE}(\hat{\boldsymbol{\theta}})$  refers to the weighted mean squared error, which is simply the weighted sum of the covariances weighted. We also define a scalar version of the qCRB ( $\mathcal{C}$ ), which we will refer to as the scalar qCRB (s-qCRB) for the rest of the document.

### 1.2.6 Saturability

The saturability of the qCRB is far from guaranteed in general. In the case of single parameter, then the saturability is guaranteed for pure states and the optimal POVM simply consists of projections onto the eigenvectors of the SLD [33]. Since the optimal POVM may depend on the underlying value of the parameter being estimated, saturability is only guaranteed in the asymptotic limit of many repetitions [35]. When there are multiple parameters, the saturability is only achieved if the ideal POVM elements (the SLDs) are compatible (i.e. if they can be performed at the same time). And so we require that

$$\langle [L_m, L_l] \rangle = 0 \quad \forall \quad l, m. \quad (1.15)$$

This was proven to be the necessary and sufficient condition for the saturability of the qCRB [35, 36]. The Holévo Cramér-Rao [37] bound is a similar bound to the qCRB but is always saturable. Its main drawback is that it is difficult to calculate. It essentially

involves an optimization over all valid POVM and is often intractable analytically. The good news is that the s-qCRB is always a decent bound as it was proven recently that the Holévo bound is at most a factor of 2 higher than the s-qCRB[38–41].

### 1.2.7 Estimating a Unitary Operation

A particular case of the quantum parameter estimation scheme in Figure 1.1 is the case where the dynamical process during which the parameters  $\boldsymbol{\theta}$  are encoded into  $\rho$  is a unitary acting on the probe state. In quantum mechanics, all time evolutions are represented by unitary operations on some Hilbert space. In this specific case, the unitary time-evolution is to act only on the probe. Typically, the to-be-measured parameters appear either in the Hamiltonian generating this time evolution or as an effective interaction time. Under such evolution, the probe experiences the transformation

$$|\psi\rangle \longrightarrow |\psi_{\boldsymbol{\theta}}\rangle = U_{\boldsymbol{\theta}} |\psi\rangle. \quad (1.16)$$

Here, we assumed that the probe starts in a pure state to simplify the calculations. In the next section, we will give two examples of physically relevant scenarios of this type but first we derive some helpful equations.

For pure states, the Symmetric-Logarithm Derivative operators take the form

$$L_l = 2 \left( (\partial_{\theta_l} U_{\boldsymbol{\theta}}) |\psi\rangle \langle\psi| U_{\boldsymbol{\theta}}^\dagger + U_{\boldsymbol{\theta}} |\psi\rangle \langle\psi| (\partial_{\theta_l} U_{\boldsymbol{\theta}}^\dagger) \right), \quad (1.17)$$

and the Quantum Fisher Information

$$[I_{\psi, \boldsymbol{\theta}}]_{l,m} = \frac{1}{2} \langle\psi| U_{\boldsymbol{\theta}}^\dagger \left\{ \hat{L}_{\theta_l}, \hat{L}_{\theta_m} \right\} U_{\boldsymbol{\theta}} |\psi\rangle. \quad (1.18)$$

Using the identity

$$0 = \partial_{\theta_l} \mathbf{1} = \partial_{\theta_l} (U_{\boldsymbol{\theta}}^\dagger U_{\boldsymbol{\theta}}) = (\partial_{\theta_l} U_{\boldsymbol{\theta}}^\dagger) U_{\boldsymbol{\theta}} + U_{\boldsymbol{\theta}}^\dagger (\partial_{\theta_l} U_{\boldsymbol{\theta}}), \quad (1.19)$$

the above expression can be reduced to

$$\begin{aligned} [I_{\psi, \boldsymbol{\theta}}]_{l,m} &= 2 \langle\psi| (\partial_{\theta_l} U_{\boldsymbol{\theta}}^\dagger) (\partial_{\theta_m} U_{\boldsymbol{\theta}}) + (\partial_{\theta_m} U_{\boldsymbol{\theta}}^\dagger) (\partial_{\theta_l} U_{\boldsymbol{\theta}}) |\psi\rangle \\ &\quad + 4 \langle\psi| U_{\boldsymbol{\theta}}^\dagger (\partial_{\theta_l} U_{\boldsymbol{\theta}}) |\psi\rangle \langle\psi| U_{\boldsymbol{\theta}}^\dagger (\partial_{\theta_m} U_{\boldsymbol{\theta}}) |\psi\rangle. \end{aligned} \quad (1.20)$$

To simplify this expression, we introduce operators

$$\hat{H}_l = i U_{\boldsymbol{\theta}}^\dagger \partial_{\theta_l} U_{\boldsymbol{\theta}}, \quad (1.21)$$

which we call the generators of the unitary. By cleverly inserting the identity  $(U_{\boldsymbol{\theta}} U_{\boldsymbol{\theta}}^\dagger)$

in the first term of the above equation and reusing the identity in Eq. 1.19, we deduce

$$[I_{\psi, \boldsymbol{\theta}}]_{l,m} = 4 C_{\psi}(\hat{H}_l, \hat{H}_m), \quad (1.22)$$

where

$$C_{\psi}(\hat{A}, \hat{B}) = \left\langle \frac{\hat{A}\hat{B} + \hat{B}\hat{A}}{2} \right\rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle \quad (1.23)$$

is a symmetrized version of the operator covariance. We can rewrite this equation

$$I_{\psi, \boldsymbol{\theta}} = 4 \mathbf{C}_{\psi}(\hat{\mathbf{H}}), \quad (1.24)$$

with  $[\mathbf{C}_{\psi}(\hat{\mathbf{H}})]_{l,m} = C_{\psi}(\hat{H}_l, \hat{H}_m)$ . The s-qCRB defined in Equation 1.14 then becomes

$$\text{wMSE}(\hat{\boldsymbol{\theta}}) \geq \frac{1}{4} \text{Tr}[W \mathbf{C}_{\psi}(\hat{\mathbf{H}})^{-1}]. \quad (1.25)$$

This very simple form of the QFI and of the s-qCRB will be very useful in the next section where we tackle real problems.

## Chapter 2

# Intrinsic Sensitivity Limit for Multi-Parameter Quantum Metrology

In this chapter, we present theoretical results previously published in [42]. These results stand alone but also greatly motivate the experimental results presented in Chapter 3, as they were derived precisely to further motivate our experiments.

### 2.1 Motivation

The goal of this project is to study how to properly weight the different parameters that characterize a unitary. As we can see from Eq. 1.25, the particular form of the weight matrix  $W$  is important to retrieve a scalar bound to optimize. Different choices of  $W$  can lead to different optimal probe states and optimal experimental designs, making its determination crucial. In many cases, we are interested in measuring physical parameters encoded in the unitary. Then the choice of the weight is specific to the particular needs of the physicists in their given experiment. In other cases, the object to be estimated is the unitary itself and the measured parameters are used simply to mathematically parameterize the unitary. In these cases, the geometry of the unitary group specifies a *natural* way to weight the different parameters. This should yield a result independent of the choice of parametrization. We will try to make a clear distinction between these two objectives using two different examples of estimation tasks.

### 2.1.1 Estimation of a Magnetic Field with an Atomic Spin

In the first example, we want to measure a classical magnetic field  $\mathbf{B}$  in a small region of space. At our disposition, we have one atom with a magnetic moment operator  $\hat{\boldsymbol{\mu}} = \gamma \hat{\mathbf{J}}$ , where the vector  $\hat{\mathbf{J}}$  is the angular momentum operator for the atom and  $\gamma$  is the gyromagnetic ratio. Setting  $\hbar = 1$ , the Hamiltonian describing the interaction of the atom with the field is given by  $\hat{H}_I = -\hat{\boldsymbol{\mu}} \cdot \mathbf{B} = -\gamma \hat{\mathbf{J}} \cdot \mathbf{B}$ . If we interact the atom with the field for a time  $t$ , the unitary time-evolution on the atom is

$$\hat{U} = e^{-i\gamma t \mathbf{B} \cdot \hat{\mathbf{J}}}. \quad (2.1)$$

Unitaries of this form are a representation of the Lie group  $SU(2)$ . The Lie algebra for this group in this case is generated by the angular momentum operators, with their usual commutation relations

$$[\hat{J}_i, \hat{J}_j] = \sum_k i\epsilon_{ijk} \hat{J}_k, \quad (2.2)$$

where  $\epsilon_{ijk}$  is the Levi-Civita symbol. In this case, we truly are interested in the physical parameters  $\mathbf{B}$  and we don't care much about the unitary itself. If for example  $\gamma t |\mathbf{B}| = 2\pi$ , then the unitary would simply be the identity ( $\hat{U} = \hat{\mathbf{1}}$ ) and although we could measure that, we would have no way of knowing the direction of the field  $\mathbf{B}$ . This is an example of a situation where the QFI matrix would be non-invertible and the qCRB would diverge.

The behavior of the QFI is exactly what we desire here. We are interested in the magnetic field and this measurement protocol fails at measuring it. An obvious way to fix this problem would be to change the interaction time  $t$ , which would then in turn change the unitary.

### 2.1.2 Estimation of the Action of a Two-Mode Generalized Interferometer with Photons

In our second example, we aim to measure the action of a two-mode interferometer, a task which may be necessary in photonic quantum computing for example. Two-mode interferometers are used as single-qubit gates and optimizing their frequent re-calibration could be a useful task. A two-mode interferometer is a device that transforms two modes of the electromagnetic field into two others. In the Heisenberg picture, it transforms the annihilation operators of the input modes ( $\hat{a}_1, \hat{a}_2$ ) into new ones for the output modes ( $\hat{a}'_1, \hat{a}'_2$ ). Quantum mechanics enforces the transformation

to be linear

$$\begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix} \longrightarrow \begin{pmatrix} \hat{a}'_1 \\ \hat{a}'_2 \end{pmatrix} = M \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix}, \quad (2.3)$$

where the complex matrix  $M$  represents the linear map, and the commutation relation  $[\hat{a}_i, \hat{a}_j^\dagger] = i\delta_{ij}$  must be valid for both the input and output modes, which constrains  $M$  to be unitary. This makes  $M$  an element of the group  $SU(2)$  up to an irrelevant global phase. A generalized interferometer can make any of the transformations with the restriction above, therefore spanning the full group. In practice these can be realized by Mach-Zender or Michelson interferometer where the beam splitters have variable transmission and reflection coefficients. This is the same transformation group as the group in the previous problem of magnetic field sensing, despite dealing with very different physical systems. To better help describe the unitary, we define the angular momentum operators as

$$\begin{aligned} \hat{J}_x &= \frac{1}{2}(\hat{a}_1^\dagger \hat{a}_2 + \hat{a}_1 \hat{a}_2^\dagger) \\ \hat{J}_y &= \frac{-i}{2}(\hat{a}_1^\dagger \hat{a}_2 - \hat{a}_1 \hat{a}_2^\dagger) \\ \hat{J}_z &= \frac{1}{2}(\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2) = \frac{1}{2}(\hat{n}_1 - \hat{n}_2) \\ \hat{J}_0 &= \frac{1}{2}(\hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2) = \frac{1}{2}(\hat{n}_1 + \hat{n}_2). \end{aligned} \quad (2.4)$$

These have the same commutation relations as Eq. 2.2 and also follow

$$\hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 = \hat{J}_0(\hat{J}_0 + 1). \quad (2.5)$$

This set of operators are generators of the Lie algebra  $\mathfrak{su}(2)$  associated with the Lie group  $SU(2)$ . A quantum state of the two modes can be generally written in the Fock basis

$$|\psi\rangle = \sum_{n_1 n_2} c_{n_1 n_2} |n_1, n_2\rangle = \sum_{n_1 n_2} \frac{c_{n_1 n_2}}{\sqrt{n_1! n_2!}} (a_1^\dagger)^{n_1} (a_2^\dagger)^{n_2} |0\rangle, \quad (2.6)$$

where  $n_1$  and  $n_2$  are the photon numbers in mode 1 and 2 respectively. Here, the choice of parametrization for the unitary is ours to make. No choice of parametrization is preferred since we are interested in characterizing the unitary itself. A popular choice of parametrization is the Euler-Angle parametrization. If we go back to the Schrödinger picture, we can write the evolution of the state through the interferometer as

$$\hat{U} = e^{-i\hat{J}_z\theta_1} e^{-i\hat{J}_y\theta_2} e^{-i\hat{J}_z\theta_3}, \quad (2.7)$$

a rotation around the  $y$ -axis, preceded and followed by two different  $z$ -axis rotations. In this parametrization, we notice that the performance of parameter estimation depends on the coordinates in a different way than in the last example. This is most apparent in the singularities. We notice, for example, that when  $\theta_2 = 0$ , the individual parameters  $\theta_1$  and  $\theta_3$  can no longer be estimated, as the unitary only depends on their sum. These divergences strictly depend on the choice of the coordinate system. In Example 1, if we tried to estimate  $\mathbf{B}$  in the neighbourhood of  $\mathbf{B} = 0$ , where  $\hat{U} = \hat{\mathbf{1}}$ , there were no divergences. Varying any of the three parameters yielded different unitaries. In Example 2, this is no longer the case. The distinction is that the motivation in Example 1 and 2 are not the same. The coordinate divergences are not problematic in Example 2. We have no attachment to the particular value of  $\theta_1$ ,  $\theta_2$  and  $\theta_3$ . Our inability to estimate them precisely does not mean that we can't estimate  $\hat{U}$ . There should be a geometric way to choose the weight matrix  $W$  to remove the coordinate dependent quirks of the s-qCRB.

## 2.2 More on the Unitaries' Generators

In quantum mechanics, unitaries can always be written in the form

$$\hat{U} = e^{i\hat{A}} \quad (2.8)$$

where the operator  $\hat{A}$  is a Hermitian operator. Furthermore, if  $\hat{U}$  is a part of a compact semi-simple Lie group like  $SU(2)$ , then it has an associated Lie algebra,  $\mathfrak{su}(2)$  in this case. Say  $\hat{\mathbf{X}}$  are the generators of such an algebra, then the unitary can be written as

$$\hat{U} = \mathbf{1} + i\nabla\Omega_{(\theta)} \cdot \hat{\mathbf{X}} + \mathcal{O}(\hat{X}^2) = e^{i\Omega_{(\theta)} \cdot \hat{\mathbf{X}}}, \quad (2.9)$$

where the vector function  $\Omega_{\theta}$ , which we assume to be smooth, encodes the parametrization of the unitary. This form is valid for any parametrization and is also independent of the particular Hilbert Space on which these operators act. Using the Wilcox formula[43] for the the derivative of the exponential map

$$\frac{\partial(e^{\hat{A}})}{\partial\lambda} = \int_0^1 e^{(1-\beta)\hat{A}} \frac{\partial\hat{A}}{\partial\lambda} e^{\beta\hat{A}} d\beta, \quad (2.10)$$

we can use Equation 1.21 and Equation 2.9 to find the following expression for the generators of the unitary :

$$\hat{H}_j = -\frac{\partial\Omega_{\theta}}{\partial\theta_j} \cdot \int_0^1 \hat{U}^{-\beta} \hat{\mathbf{X}} \hat{U}^{\beta} d\beta. \quad (2.11)$$

This formula illustrates that, in general, generators  $\hat{H}_j$  are linear combinations of the Lie algebra generators with real coefficients, which makes them elements of the Lie algebra themselves. Knowing that fact, we define the real vectors  $\{\mathbf{h}_i\}$  forming the matrix  $\mathbf{H}$ ,

$$\begin{aligned}\hat{H}_j &= \mathbf{h}_j \cdot \hat{\mathbf{X}} \\ \mathbf{H} &= (\mathbf{h}_1, \dots, \mathbf{h}_j, \dots)^\top.\end{aligned}\tag{2.12}$$

This linear combination is the same for any representation of the group, as neither the Hilbert space, nor its dimensions were used in the calculation. When it comes to calculating what these linear combinations are, our preference is to calculate the derivatives explicitly from a low-dimensional representation of the unitary operators. Then we don't need to worry about putting the unitary in the form of Eq. 2.9, nor do we need to explicitly calculate the integral in Eq. 2.11.

Looking back at Eq. 1.22, 1.23 and 1.24, we see that

$$\begin{aligned}C_\psi(\hat{H}_l, \hat{H}_m) &= \left\langle \frac{(\sum_i \mathbf{H}_{li} \hat{X}_i)(\sum_j \mathbf{H}_{mj} \hat{X}_j) + (\sum_i \mathbf{H}_{mi} \hat{X}_i)(\sum_j \mathbf{H}_{lj} \hat{X}_j)}{2} \right\rangle \\ &\quad - \left\langle \left( \sum_i \mathbf{H}_{li} \hat{X}_i \right) \right\rangle \left\langle \left( \sum_j \mathbf{H}_{mj} \hat{X}_j \right) \right\rangle,\end{aligned}\tag{2.13}$$

becomes

$$C_\psi(\hat{H}_l, \hat{H}_m) = \sum_{ij} \mathbf{H}_{li} \left( \left\langle \frac{\hat{X}_i \hat{X}_j + \hat{X}_j \hat{X}_i}{2} \right\rangle - \langle \hat{X}_i \rangle \langle \hat{X}_j \rangle \right) \mathbf{H}_{jm}^\top,\tag{2.14}$$

after a small change index variable, resulting in a new expression for the QFI

$$I_{\psi, \boldsymbol{\theta}} = 4 \mathbf{C}_\psi(\hat{\mathbf{H}}) = 4 \mathbf{H} \mathbf{C}_\psi(\hat{\mathbf{X}}) \mathbf{H}^\top.\tag{2.15}$$

Then, the s-qCRB becomes

$$\text{wMSE}(\hat{\boldsymbol{\theta}}) \geq \mathcal{C}(\hat{\boldsymbol{\theta}}) = \frac{1}{4} \text{Tr}[(\mathbf{H}^{-1} \mathbf{W} \mathbf{H}^\top)^{-1} \mathbf{C}_\psi(\hat{\mathbf{X}})^{-1}],\tag{2.16}$$

using the cyclic permutability of the trace. This form of the s-qCRB is very elegant as it separates the terms depending on the parameters (in parenthesis in the equation) from the terms which depend on the initial state, which we ultimately aim to optimize. Furthermore, we can see that the all that matters about the state for the purpose of this optimization is the covariance of the generators of the Lie algebra, which can reveal useful information before even calculating  $\mathbf{H}$ . For example, if  $\mathbf{C}_\psi(\hat{\mathbf{X}})$  is non-

invertible, then we already know that the state will be completely unsuitable to the estimation of all the parameters with a diverging wMSE.

### 2.2.1 Example : Two-Mode Interferometer with Euler Angles

We look back at the two-mode interferometer example with the Euler angle parametrization. To calculate the  $\mathbf{H}$  matrix, it is nice to consider the unitary in a very low dimensional Hilbert Space. As we have shown before, the results will be general and the representation of the  $SU(2)$  Lie group won't matter. The representation of  $SU(2)$  with the lowest dimensions describe the transformation on a single photon in this interferometer. The angular momentum operators take the form of the well-known Pauli operators,

$$(\hat{J}_0, \hat{J}_x, \hat{J}_y, \hat{J}_z) = \frac{1}{2}(\hat{\mathbb{1}}, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z) \quad (2.17)$$

We then use the identities

$$\begin{aligned} e^{-ia(\mathbf{n} \cdot \boldsymbol{\sigma})} &= \mathbb{1} \cos a - i(\mathbf{n} \cdot \boldsymbol{\sigma}) \sin a \\ \{\sigma_i, \sigma_j\} &= 2\delta_{ij}\mathbb{1}, \end{aligned} \quad (2.18)$$

after which the derivatives are easy to calculate. We find that

$$\begin{aligned} \hat{H}_1 &= -\sin \theta_2 \cos \theta_3 \hat{J}_x + \sin \theta_2 \sin \theta_3 \hat{J}_y + \cos \theta_2 \hat{J}_z \\ \hat{H}_2 &= \sin \theta_3 \hat{J}_x + \cos \theta_3 \hat{J}_y \\ \hat{H}_3 &= \hat{J}_z. \end{aligned} \quad (2.19)$$

Gathering the coefficient, the  $\mathbf{H}$  matrix is

$$\mathbf{H} = \begin{pmatrix} -\sin \theta_2 \cos \theta_3 & \sin \theta_2 \sin \theta_3 & \cos \theta_2 \\ \sin \theta_3 & \cos \theta_3 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (2.20)$$

After many simplifications, the problem, now represented by the matrix  $\mathbf{H}$ , has a fairly simple form. Given a weight matrix  $W$ , one can now compute  $(\mathbf{H}^{-1}W\mathbf{H}^T)^{-1}$  and find the optimal state covariance matrix. For example, if we are equally interested in these parameters, such that we aim to minimize the sum of the variances on each

parameters, we get

$$\begin{aligned} \text{Var}[\theta_1] + \text{Var}[\theta_2] + \text{Var}[\theta_3] &\geq \frac{1}{4} \text{Tr}[(\mathbf{H}^{-1} \mathbf{1} \mathbf{H}^{-1\top}) \mathbf{C}_\psi(\hat{\mathbf{J}})^{-1}] \\ &= \frac{1}{4} \text{Tr}\left[\begin{pmatrix} \frac{1}{\sin^2 \theta_2} & 0 & -\frac{\cos \theta_2}{\sin^2 \theta_2} \\ 0 & 1 & 0 \\ -\frac{\cos \theta_2}{\sin^2 \theta_2} & 0 & \frac{1}{\sin^2 \theta_2} \end{pmatrix} \mathbf{C}_\psi(\hat{\mathbf{J}})^{-1}\right]. \end{aligned} \quad (2.21)$$

In general, the optimal states and their covariance matrices depend on the value of the parameters. It may seem strange that we require knowledge of the parameters to obtain their measurement. This is a feature of the Fisher Information, which, if we remember the definition, is a very local metric, that simply looks at the difference between one state and a state infinitesimally close to it. As such, we tend to consider the Fisher Information useful when we already have a very good estimate of the parameters. An example of this would be a tracking measurement, where we know that the parameters are in the neighbourhood of certain values because we have recently measured them. Other bounds, typically of Bayesian inspiration[44, 45], don't suffer from this problem, optimizing for a given arbitrary prior distribution. Bayesian bounds are sometimes harder to calculate. In the absence of prior knowledge, it has been shown that adaptive schemes, where better estimates of the parameters are used to better optimize the next iteration of measurement, will asymptotically approach the bounds calculated with the Fisher Information. The Fisher Information and the CRB therefore remain useful performance metrics.

### 2.3 The Killing Form and the Cartan Metric

We have hinted at that there was a *natural* way of defining the weight metric for a given group of unitary. In fact, Lie groups[46] are differential manifolds and have geometrical properties, much like geometrical objects. These formal mathematical properties can help us find a metric *intrinsic* to the Lie group in question, which then imposes itself as the natural weight metric. Importantly, this metric, expressed in a given coordinate system, is a tensor depending purely on the geometry of the group. For the Lie algebra  $\mathfrak{su}(n)$ , which we will restrict ourselves to for this result, the Killing Form  $B(X, Y) = 2n \text{Tr}[X^\dagger Y]$  introduces an inner product on the space.

$$(X, Y) = B(X, Y) \quad (2.22)$$

This inner product can in turn be used to define the Cartan metric

$$ds^2 = (dU, dU) \quad (2.23)$$

This metric is the one that we will use. Using

$$dU = \sum_i \frac{\partial \hat{U}}{\partial \theta_i} d\theta_i = -iU \sum_i \hat{H}_i d\theta_i, \quad (2.24)$$

the metric tensor elements  $g_{ij}$  defined by

$$ds^2 = \sum_{ij} g_{ij} d\theta_i d\theta_j \quad (2.25)$$

become

$$g_{ij} = \text{Tr}[\hat{H}_i \hat{H}_j] \quad (2.26)$$

up to an irrelevant overall constant. This can be further simplified by decomposing  $\hat{H}$  into the orthonormal generators of the Lie algebra  $\hat{\mathbf{X}}$ ,

$$g_{ij} = \text{Tr}[(\mathbf{h}_i \cdot \hat{\mathbf{X}})(\mathbf{h}_j \cdot \hat{\mathbf{X}})] = \mathbf{h}_i \cdot \mathbf{h}_j, \quad (2.27)$$

again up to an insignificant constant. The matrix form of the metric is

$$\mathbf{g} = \mathbf{H}\mathbf{H}^\top. \quad (2.28)$$

We propose that the geometry of the  $\text{SU}(n)$  Lie group is such that the *natural* way to weight the parameters used to describe said unitaries is to use  $W = \mathbf{g}$ . With this choice, we see much simplification to the s-qCRB. Indeed, when substituting this choice of metric in Eq. 2.16, we obtain

$$\text{wMSE}(\hat{\boldsymbol{\theta}}) \geq \mathcal{C}(\hat{\boldsymbol{\theta}}) = \frac{1}{4} \text{Tr}[(\mathbf{H}^{-1}(\mathbf{H}\mathbf{H}^\top)\mathbf{H}^{-1}) \mathbf{C}_\psi(\hat{\mathbf{X}})^{-1}] = \frac{1}{4} \mathbf{C}_\psi(\hat{\mathbf{X}})^{-1}. \quad (2.29)$$

This formula is the main result of our work on this topic. It is a very simple and elegant formula. The s-qCRB now only depends on the state and the covariance matrix of the Lie algebra generators. The parameter dependence has completely dropped off the equation, indicating a certain equivalence between all element of the group, which can be estimated with equal precision. We also note that for any unitary  $\hat{U}$  belonging to the Lie group in question,

$$\mathbf{C}_\psi(\hat{U}\hat{\mathbf{X}}\hat{U}^\dagger) = \mathbf{C}_\psi(\hat{\mathbf{X}}). \quad (2.30)$$

Rotating the state prior to the unitary's action does not affect the performance of a state. Prior to this result, it was well known that when trying to estimate a single parameter encoded in the unitary of the form  $e^{i\hat{A}\theta}$ , the performance of the scheme was constant along the unitary's orbit [35]. Our result generalizes that statement to multi-parameter estimation.

## 2.4 SU(d) on a $d$ -Mode System

With this result in our toolbox, we immediately can tackle much more general problems than before. This is partly due to the fact that we need not worry about specific parametrizations. The calculation of the  $\mathbf{H}$  matrix, which could grow to be very messy in groups with lots of parameters, is no longer necessary. We illustrate this point by tackling the problem of a  $d$ -mode generalized interferometer, corresponding to a SU(d) transformation. A SU(d) transformation can be parametrized by  $d^2 - 1$  parameters. Identically, its Lie algebra,  $\mathfrak{su}(d)$ , has  $d^2 - 1$  generators  $\hat{\mathbf{X}}$ . They can easily be constructed from the generalized Gell-Mann matrices  $\{\lambda_i\}_{d^2-1}$ [47], with

$$\hat{X}_i = \frac{1}{2} \begin{pmatrix} \hat{a}_1^\dagger & \cdots & \hat{a}_d^\dagger \end{pmatrix} (\lambda_i) \begin{pmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_d \end{pmatrix}. \quad (2.31)$$

There is a quadratic Casimir invariant for these representations of the Lie algebra given by

$$\Omega_d(N) = \sum_i \hat{X}_i \hat{X}_i = \frac{d-1}{2d} N(N+d) =: \tilde{\Omega}_d(N) \hat{\mathbf{1}}, \quad (2.32)$$

which can be derived by using the form of the generalized Gell-Mann matrices in Ref. [47].

When trying to find the optimal state to estimate these SU(d) transformations, we use the identity

$$\mathrm{Tr}[\mathbf{1}]^2 = \mathrm{Tr}[M^{1/2}M^{-1/2}]^2 \leq \mathrm{Tr}[M] \mathrm{Tr}[M^{-1}], \quad (2.33)$$

to put a lower bound on the s-qCRB. Then,

$$\mathcal{C}_{\psi,\theta} = \frac{1}{4} \times \mathrm{Tr}[C_\psi(\hat{\mathbf{X}})^{-1}] \geq \frac{(d^2 - 1)^2}{4 \mathrm{Tr}[C_\psi(\hat{\mathbf{X}})]}. \quad (2.34)$$

This identity also guarantees the saturation of the bound when  $M$  or  $C_\psi(\hat{\mathbf{X}})$  is

proportional to  $\mathbb{1}$ . The trace of the covariance matrix is given by

$$\mathrm{Tr}[C_\psi(\hat{\mathbf{X}})] = \sum_i \left( \langle \hat{X}_i^2 \rangle - \langle \hat{X}_i \rangle^2 \right) \leq \left\langle \sum_i \hat{X}_i^2 \right\rangle = \tilde{\Omega}_d(N). \quad (2.35)$$

Therefore, under the conditions that

$$\begin{aligned} \langle \hat{\mathbf{X}} \rangle &= \mathbf{0} \\ C_\psi(\hat{\mathbf{X}}) &= \frac{\tilde{\Omega}_d(N)}{d^2 - 1} \mathbb{1}, \end{aligned} \quad (2.36)$$

the optimal bound of

$$\mathcal{C}_{\psi,\theta} = \frac{(d^2 - 1)^2}{4\tilde{\Omega}_d(N)} = \frac{d^4 + d^3 - d^2 - d}{2N(N + d)} \quad (2.37)$$

is reached. Interestingly, beyond guaranteeing the optimal s-qCRB, these criteria are also sufficient to ensure its saturability. Indeed, it is easy to show that the commutators of the SLDs are linear combinations of  $\{\hat{\mathbf{X}}\}$ , meaning that

$$\langle [L_i, L_j] \rangle = 0 \quad \forall \{i, j\}. \quad (2.38)$$

In Figure 2.1, we plot the optimal  $\mathcal{C}$  compared to the performance of a sub-optimal generalized N00N state

$$|\psi\rangle = \frac{1}{\sqrt{d}}(|N, 0, \dots, 0\rangle + |0, N, \dots, 0\rangle + \dots + |0, 0, \dots, N\rangle). \quad (2.39)$$

We discuss N00N states in Sec. 3.1.3. We show that for different  $d$ , we get the promised Heisenberg scaling of  $\mathcal{O}(1/N^2)$ , as opposed to the N00N state, which only performs as  $\mathcal{O}(1/N)$  in multi-parameter estimation.

More precisely, the optimal scaling is  $\mathcal{O}(d^4/N^2)$ , with the number of parameters scaling as  $d^2$ . Some schemes known in the literature as “sequential estimation scheme” or “individual estimation schemes” [48] require splitting photons into  $\mathcal{O}(d^2)$  groups. In each of these groups, they use the best state to measure a single parameter. These states only have a Heisenberg scaling for that single parameter and are either insensitive, or have a bad scalings to estimate the other parameters. The generalized N00N states are examples of such states. When the results of the different groups are used together to estimate the unitary, the wMSE also leads to a Heisenberg scaling. However, splitting the physical resources among the  $\mathcal{O}(d^2)$  groups remains inefficient. With a variance of  $\mathcal{O}(d^4/N)$  for each individual parameter, their sum goes as  $\mathcal{O}(d^6/N)$ , a factor of  $d^2$  more than the optimal scaling. This differing factor

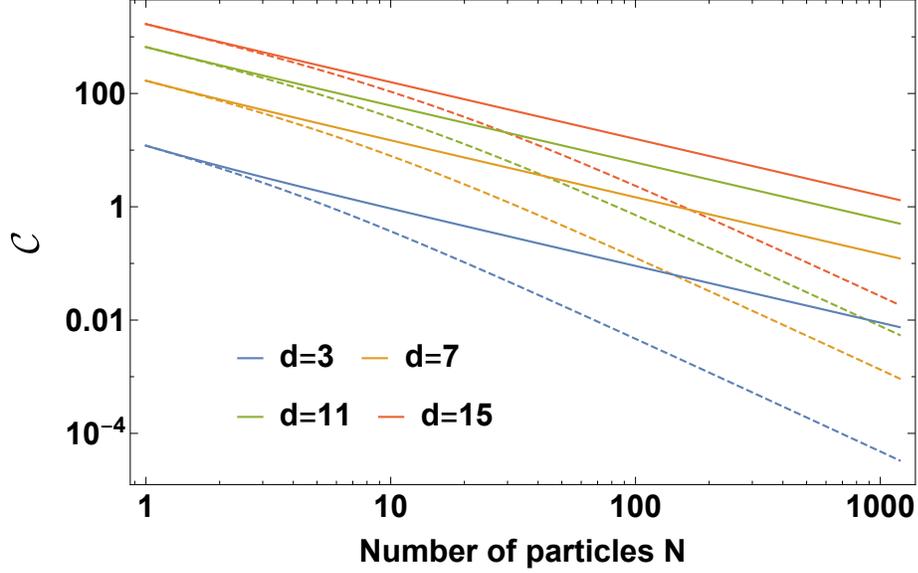


Figure 2.1: Scalar Quantum Cramér-Rao Bound  $\mathcal{C}$  for the estimation of an  $SU(d)$  transformation vs. number of photon used  $N$ . Solid lines refer to the performances of generalized N00N states while dashed lines refer to the best possible performances. The N00N states exhibit a shot-noise scaling, with  $\mathcal{C} \propto 1/N$  while the optimal bounds satisfy the Heisenberg scaling  $\mathcal{C} \propto 1/N^2$  for sufficiently large  $N$ .

of the number of parameters  $d^2$  between the best sequential estimation schemes and the best simultaneous estimation schemes seems general. Similar famous results have been found in the case of multiple phase estimation[48, 49].

## 2.5 Optimal States for SU(3)

Although we have presented criteria for the optimality of the s-qCRB, we have not shown that states which satisfy them exist.  $SU(2)$  is a well studied group and even though the form of the s-qCRB presented in this paper has only been recently published, families of states have been found which satisfy our criteria in Eq. 2.36. We shall discuss them more in Section 3.1.3. Here, we tackle the problem of  $SU(3)$  operations on 3 modes of the field. We were quickly able to find a family of optimal state which provides the ideal scaling. In the second-quantized formalism, states which obey

$$|\psi\rangle = \frac{1}{\sqrt{3}}(|k-l, k, k+l\rangle + |k, k+l, k-l\rangle + |k+l, k-l, k\rangle) \quad (2.40)$$

for non-zero integers  $k$  and  $l$  such that  $k = N/3$  and  $4l^2 = 3k(k+1)$ , satisfy all the required properties. The set of values of  $k$  with a corresponding  $l$  is sparse but infinite, meaning that a scheme relying on these state would be scalable. If one wanted to tackle the problem in earnest it would be fruitful to spend more time looking for an

alternate family containing an optimal state for more values of photon number  $N$ . We think that the ease with which we were able to provide a meaningful analysis of the metrological landscape of  $SU(3)$  transformations demonstrates the power of the bound we derived. A parametrization-dependent analysis of  $SU(3)$  would inevitably be very difficult with a minimum of 8 parameters to deal with and many  $8 \times 8$  matrices to calculate and optimize.

## 2.6 A Note On Mixed States

Up until now, we have restricted ourselves to pure states. Fortunately, this was not for any fundamental reason, but rather to simplify the mathematics. For the mixed state formulation of this bound, we need to start with the equivalent Lyapunov form of the QFI [35]

$$[I_{\rho\theta}]_{lm} = 2 \int_0^\infty ds \text{Tr}[\partial_{\theta_l}(\hat{U}_\theta \rho \hat{U}_\theta^\dagger) e^{-(\hat{U}_\theta \rho \hat{U}_\theta^\dagger)s} \partial_{\theta_m}(\hat{U}_\theta \rho \hat{U}_\theta^\dagger) e^{-(\hat{U}_\theta \rho \hat{U}_\theta^\dagger)s}]. \quad (2.41)$$

If we substitute in the generators  $\hat{H}$ , we get that

$$\partial_{\theta_l}(\hat{U}_\theta \rho \hat{U}_\theta^\dagger) = (\partial_{\theta_l} \hat{U}_\theta) \rho \hat{U}_\theta^\dagger + \hat{U}_\theta \rho (\partial_{\theta_l} \hat{U}_\theta^\dagger) = i \hat{U}_\theta (\rho \hat{H}_j - \hat{H}_j \rho) \hat{U}_\theta^\dagger, \quad (2.42)$$

and so

$$I_{\rho\theta} = 2 \int_0^\infty ds \text{Tr}[(\rho \hat{H} - \hat{H} \rho) e^{-\rho s} (\hat{H}^\top \rho - \rho \hat{H}^\top) e^{-\rho s}]. \quad (2.43)$$

In fact, a quick inspection reveals that simply substituting the covariances in the expressions for pure state with

$$\mathbf{C}_\psi(\hat{X}) \mapsto \frac{1}{2} \int_0^\infty ds \text{Tr}[(\rho \hat{X} - \hat{X} \rho) e^{-\rho s} (\hat{X}^\top \rho - \rho \hat{X}^\top) e^{-\rho s}], \quad (2.44)$$

yields correct expressions for the s-qCRB. With this slightly modified formula, the main feature of our result, namely the parameter independence of the s-qCRB when using the *intrinsic* metric of the corresponding Lie group, remains, allowing the for the possibility of using this technique to analyze scenarios with noise.

## 2.7 Concluding Remarks

In conclusion, we found a way to weight the different variances of different parameters in multi-parameter estimation that seems to be natural given the geometry of the group of unitaries in which the parameters are encoded. Using this weighting gives a very simple form for the scalar quantum Cramér-Rao Bound, which could lead to an

easier optimization. Before this result, analyzing the performances of different states in estimating unitaries was often derailed by discussion of the group itself and the specific choice of its parametrization. Because this result completely removes these distractions, we hope it finds great use in the field. That said, it must be stressed that all the caveats regarding the saturability of the CRB discussed in Sec. 1.2.6 are applicable to our recent results. In the published paper[42], we demonstrated the power of the method using the group  $SU(2)$  and  $SU(3)$  as example. The former was already well studied[50, 51] but our simple result quickly led us to progress on the latter. We shall revisit the  $SU(2)$  Lie group in Chapter 3, in which we will further discuss unitary estimation from an experimental perspective.



## Chapter 3

# Experimental Generation of the Tetrahedron State

In this chapter, we present the results of an experiment soon to be submitted for publication. The goal of this experiment is to create a tetrahedron state in the laboratory, a 4-photon polarization-entangled state optimal for characterizing polarization rotations.

### 3.1 Theory

We start by continuing the theory presented in Chapter 2. First, we introduce the formalism necessary to discuss polarization rotations. Second, we introduce different ways to visualize quantum states. Finally, we discuss different families of quantum states and quantify their performance in estimating polarization rotations.

#### 3.1.1 Polarization Rotations

Polarization is a degree of freedom of light living in a two dimensional vector space. In quantum mechanics, the polarization of a single photon constitutes a qubit, or a Spin-1/2 system( $S_{1/2}$ ). An orthonormal basis for this qubit is the H/V basis,

$$\begin{aligned} |H\rangle &= \hat{a}_H^\dagger |0\rangle = |1_H, 0_V\rangle \\ |V\rangle &= \hat{a}_V^\dagger |0\rangle = |0_H, 1_V\rangle. \end{aligned} \tag{3.1}$$

Unitaries acting on qubits generally form the  $SU(2)$  group. In this case, the generators of the  $\mathfrak{su}(2)$  algebra take the form of Pauli operators, a special case of their representation stated in Equation 2.4. If we consider  $N$  distinguishable such photons, their polarization state is contained in the Hilbert Space  $(S_{1/2})^{\otimes N}$ . Global polariza-

tion rotations, which rotate the polarization states of each of the photons identically, are still elements of  $SU(2)$ . They can be described by the operator  $(\hat{R}_{(1/2)})^{\otimes N}$ , where  $\hat{R}_{(1/2)}$  is the operator for a rotation acting on  $S_{1/2}$ . It is known that

$$(S_{1/2})^{\otimes N} = \bigoplus_{j=[0,1/2]}^{N/2} \underbrace{S_j \oplus \cdots \oplus S_j}_{d_j \text{ times}}, \quad (3.2)$$

where  $S_j$  is the Hilbert Space of a spin with total angular momentum  $j$ . The lower bound in the tensor sum is 0 if  $N$  is even and  $1/2$  if  $N$  is odd. The integers  $d_j$ , representing the multiplicity of each spin sectors, vary depending on the photon number  $N$  but  $d_{N/2}$  is always 1. A given state can be expressed in bases naturally following the two equivalent Hilbert Space decompositions above. The transfer from one basis to the other can be done using a Clebsch-Gordan transformation matrix. In Appendix A, we give the example of a Clebsch-Gordan transform matrix for  $N = 4$ . The spin Hilbert spaces  $S_j$  are commonly studied and so are the spin operators  $\hat{\mathbf{J}}^{(j)}$  acting on them. These are the generators of rotations on these spaces. In the Spin tensor sum decomposition of the Hilbert Space, the generators of the global polarization rotation are

$$\hat{\mathbf{J}} = \bigoplus_{j=[0,1/2]}^{N/2} \underbrace{\hat{\mathbf{J}}^{(j)} \oplus \cdots \oplus \hat{\mathbf{J}}^{(j)}}_{d_j \text{ times}}. \quad (3.3)$$

An example of parametrization of a general global rotation is given by the Cartesian parametrization

$$\hat{R} = e^{i\theta \cdot \hat{\mathbf{J}}}, \quad (3.4)$$

with the  $\hat{\mathbf{J}}$  operators defined above. For each of the spin sectors in Eq. 3.2, there is an associated symmetry related to particle exchanges. For example, when  $N = 2$  and  $(S_{1/2})^{\otimes 2} = S_1 \oplus S_0$ , the spin-1 sector contains the triplet states  $|HH\rangle$ ,  $\frac{1}{\sqrt{2}}(|HV\rangle + |VH\rangle)$  and  $|VV\rangle$ , all fully symmetric under the exchange of the two particles. The spin-0 sector on the other hand contains the singlet state  $\frac{1}{\sqrt{2}}(|HV\rangle - |VH\rangle)$ , anti-symmetric under that same exchange. More can be learned of these symmetries by analyzing Young's diagrams[52]. For the purpose of this work, all we need to know is that the unique largest spin sector,  $j = N/2$ , is fully symmetric under the exchange of any particle. We will sometimes refer to it as the full-symmetric sector.

Photons are bosons, which means that in the absence of distinguishability, their wave-functions are full-symmetric. If we look at a system where the photons are indistinguishable, it can be convenient to describe it with a second-quantized formalism. Here, the system consists of two modes, one for each polarization, and is occupied by

$N$  photons. We have analyzed this scenario in Sec. 2.1.2. This system, with its Lie algebra as we previously defined it, is equivalent to the  $S_{N/2}$  sector as expected. For this reason, we will also use notations from the second-quantized picture to denote the different states. For example, with  $N = 3$ , the following notations for the state with  $j = 3/2$  and  $m = 1/2$  will be considered interchangeable:

$$|3/2, 1/2\rangle = |2_H, 1_V\rangle = \frac{1}{\sqrt{3}}(|HHV\rangle + |H VH\rangle + |VHH\rangle). \quad (3.5)$$

### 3.1.2 State Visualization

We now present two ways to visualize quantum states in the full-symmetric quadrant of the Hilbert Space.

#### Majorana Representation

The Majorana Representation[53] is a way to visualize full-symmetric pure states, with  $N$  points spread on the Bloch sphere. Given the full-symmetric nature of such states, there is always an ensemble of  $N$  single photon pure states  $\{\psi_i\}$  such that

$$|\psi\rangle \propto \sum_{\alpha \in S_N} |\psi_{\alpha_1}\rangle \otimes \cdots \otimes |\psi_{\alpha_N}\rangle, \quad (3.6)$$

where, in this instance,  $S_N$  is the set of permutations of  $\{1, \dots, N\}$ . To find this family of single photon states  $\{\psi_k\}$ , we write the state in the second-quantized formulation of quantum mechanics as a polynomial of creation operators acting on vacuum,

$$|\psi\rangle \propto \sum_{k=0}^N c_k |k_H, (N-k)_V\rangle = \left( \sum_{k=0}^N c_k \frac{(a_H^\dagger)^k (a_V^\dagger)^{N-k}}{\sqrt{k!} \sqrt{(N-k)!}} \right) |0\rangle. \quad (3.7)$$

The latter expression is a polynomial of two variables  $a_H^\dagger$  and  $a_V^\dagger$ . It can always be factored into product of single particle creation operators by finding the  $N$  roots of the polynomial. Then,

$$|\psi\rangle \propto \left( \prod_{i=1}^N (\alpha_i a_H^\dagger + \beta_i a_V^\dagger) \right) |0\rangle \quad (3.8)$$

and we can conclude that

$$|\psi_k\rangle = (\alpha_k a_H^\dagger + \beta_k a_V^\dagger) |0\rangle = \alpha_k |H\rangle + \beta_k |V\rangle. \quad (3.9)$$

In order to represent these states  $\{\psi_k\}$  as points on the Bloch sphere, we discard the global phase and express them as

$$|\psi_k\rangle = \cos \frac{\theta_k}{2} |H\rangle + e^{i\phi_k} \sin \frac{\theta_k}{2} |V\rangle. \quad (3.10)$$

We then use  $(\theta_k, \phi_k)$  as the spherical coordinates for each of the  $N$  points.

### Wigner Distribution

The Wigner distribution offers another visualization of states contained in the full-symmetric sector. In an analogue to the Wigner distributions of infinite dimensional quantum states, the Wigner distribution of states in  $S_{N/2}$  is a quasi-probability distribution. In this case, its domain is not the complex plane but the Bloch sphere, the natural phase-space for a spin system. Much like on the complex plane, many such distributions can be created, the more well-known ones are the Glauber, the Husimi and the Wigner distribution. These distributions can be plotted as colour maps on the sphere, or on 2D projections of the sphere, to quickly visualize the phase-space features of the states that they represent. Furthermore, as opposed to the Majorana representation, these distributions are well defined for mixed states, which is invaluable in an experimental scenario. We use the definition of the Wigner Function in [54–56]. We first introduce the state multipole operators

$$\hat{T}_{KQ} = \sum_{m=-N/2}^{N/2} \sum_{m'=-N/2}^{N/2} (-1)^{N/2-m} \sqrt{2K+1} \begin{pmatrix} N/2 & K & N/2 \\ -m & Q & m' \end{pmatrix} |N/2, m\rangle \langle N/2, m'|, \quad (3.11)$$

where the notation for the kets is the  $|j, m\rangle$  spin notation and the object with six inputs in parenthesis is the Wigner-3j coefficient. The Wigner representation is then

$$W_{\hat{\rho}}(\theta, \phi) = \sum_{K=0}^N \sum_{Q=-K}^K \text{Tr}[\hat{\rho} \hat{T}_{KQ}] Y_{KQ}^*(\theta, \phi), \quad (3.12)$$

where  $Y_{KQ}(\theta, \phi)$  are the spherical harmonic functions and  $(\theta, \phi)$  are the spherical coordinates. Note that certain sources have an insignificant discrepancy of a constant factor  $\sqrt{\frac{4\pi}{N+1}}$ .

#### 3.1.3 Choosing the optimal state

When trying to find the optimal state to estimate  $SU(2)$  rotations, we can use our results from Sec. 2.4 to find the best scalar quantum Cramér-Rao bound. In that section, we were concerned about  $SU(d)$  operation on  $d$ -modes of light. This system

is the particular case with  $d = 2$ , representing the two polarization modes. The full-symmetric nature of the states however are no longer being assumed. We can immediately see that the optimal state in our current case lies in the full-symmetric subspace and that the results are therefore identical.

Following the proof in Sec 2.4, we see that the conditions

$$\begin{aligned} \langle \hat{\mathbf{J}} \rangle &= \mathbf{0} \\ C_\psi(\hat{\mathbf{J}}) &\propto \mathbb{1} \end{aligned} \quad (3.13)$$

should still be respected to reach the optimum performance. The only difference is that we now have multiple spin sectors, each with their own quadratic Casimir invariant

$$\hat{J}_x^{(j)^2} + \hat{J}_y^{(j)^2} + \hat{J}_z^{(j)^2} = j(j+1)\hat{\mathbf{1}}. \quad (3.14)$$

Therefore, the trace of the spin covariance matrix is

$$\text{Tr}[C_\psi(\hat{\mathbf{J}})] = \sum_{i=1}^3 \langle \hat{J}_i^2 \rangle = \sum_{j=[0,1/2]}^{N/2} P_j(j(j+1)) \leq \frac{N}{2} \left( \frac{N}{2} + 1 \right) \quad (3.15)$$

with  $P_j$  being the probability that the state is in a  $S_j$  quadrant. Only when the state is fully in the  $j = N/2$  sector is the inequality saturated. Then, the s-qCRB takes an optimal value of

$$\mathcal{C}_{\psi, \boldsymbol{\theta}} = \frac{9}{4 \left( \frac{N}{2} \left( \frac{N}{2} + 1 \right) \right)} = \frac{9}{N(N+2)}. \quad (3.16)$$

As we can see, this is the same bound found in Sec 2.4. To gather more intuition about the state space, we next look at three families of states of physical significance. These states are all contained in the full-symmetric sector. The last one, which contains what we call second-order unpolarized states, is optimal.

If someone was unconvinced by the importance this form of the s-qCRB and was questioning our choice of the weight metric  $W$ , it should be noted that the same bound can be found for a different problem. In the Cartesian parametrization of the rotation in Eq. 3.4, if we look at the neighbourhood around  $\boldsymbol{\theta} = 0$  ( $\hat{U} = \hat{\mathbf{1}}$ ), we find that  $\mathbf{H} = \hat{\mathbf{1}}$ . If we give an equal weighting to the three parameters ( $W = \mathbb{1}$ ), we recover the same form of the s-qCRB, meaning that in the optimal case,

$$(\Delta\theta_1)^2 + (\Delta\theta_2)^2 + (\Delta\theta_3)^2 \geq \mathcal{C}_{\psi, \boldsymbol{\theta}} = \frac{9}{N(N+2)}. \quad (3.17)$$

This result was first published in [50]. It was then argued that if one wants to make sure that two Cartesian reference frames stay aligned, sending states with this optimal

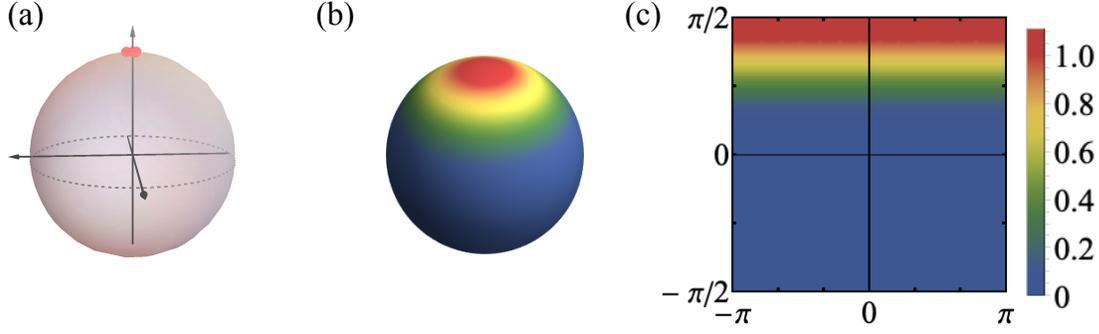


Figure 3.1: Three different visual representations of the spin coherent state  $|4_H, 0_V\rangle$ . (a) Majorana representation of the state. All points are identically placed on the North Pole on the Bloch sphere, aligned with  $|H\rangle$ . (b) Wigner distribution on the sphere. The quasi-probability distribution is heavily concentrated on the north pole of the sphere. (c) Equirectangular projection (longitude vs latitude) of the Wigner distribution.

bound between the two frames and continuously measuring  $\theta$  would be the optimal strategy.

### Spin coherent states

Spin coherent states[57], depicted in Fig. 3.1 (a), are separable states in the symmetric subspace of  $N$  photons. They consist of all  $N$  photons sharing the same polarization properties. One of those states is

$$|\psi_{coh}\rangle = |N_H, 0_V\rangle, \quad (3.18)$$

where all the photons are horizontally polarized. The rest of the spin coherent states can be obtained by rotating this state. The spin covariance matrix for this state is given by

$$C_{\psi_{coh}}(\hat{\mathbf{J}}) = \begin{pmatrix} N/4 & 0 & 0 \\ 0 & N/4 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (3.19)$$

and is not invertible, which leads to a divergence in the s-qCRB and an “infinitely bad” performance at fully characterizing  $SU(2)$  rotations. Intuitively, this is due to the state being completely undisturbed by any rotation around the polarization axis of the separable particles. It is therefore incapable of estimating one of the three parameters necessary to characterize the rotation, which leads to a divergence in the weighted sum of the variance of those parameters. In the example above, all the polarizations of all photons were aligned with H, which leads to a variance of 0 for  $\hat{J}_z$ , the spin operator aligned with that axis.

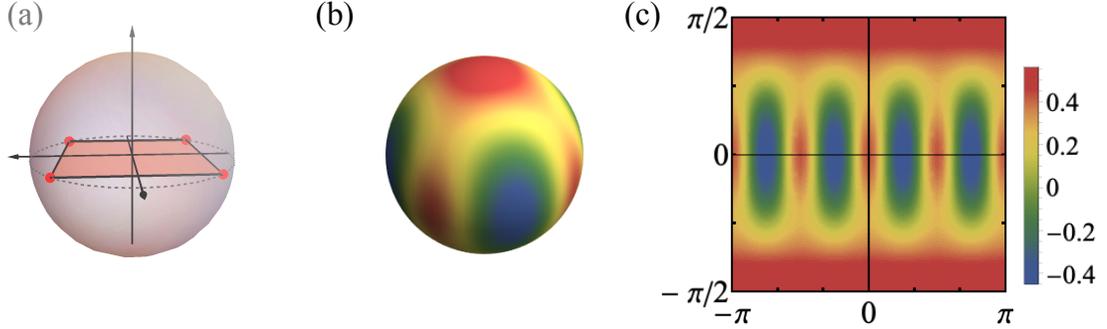


Figure 3.2: Three different visual representations of the N00N state  $1/\sqrt{2}(|4_H, 0_V\rangle + |0_H, 4_V\rangle)$ . (a) Majorana representation of the state. All points are placed equally spaced on the equator of Bloch sphere. (b) Wigner distribution on the sphere. The quasi-probability distribution is concentrated on the North and South Pole with four fringes along longitude lines of the sphere. (c) Equirectangular projection (longitude vs latitude) of the Wigner distribution.

To remedy this situation, we could imagine dividing our available photons into three different spin coherent states aligned with the  $x$ ,  $y$  or  $z$  axes. Because the quantum Fisher information is additive for uncorrelated measurements [35],  $\mathcal{C}$  can be found by inverting the sum of the different spin covariance matrices for each of these states. Assuming  $N$  is divisible by 3, the resulting matrix and the corresponding s-QCRB are

$$\sum_i C_{\psi_{coh}^{(i)}}(\hat{\mathbf{J}}) = \frac{N}{6} \times \mathbf{1}, \quad \mathcal{C} = \frac{9}{2N}. \quad (3.20)$$

As we can see, by splitting our photons into three batches, we retain sensitivity to every parameter and recover the expected shot-noise scaling from a classical state.

### N00N states

N00N states are entangled states of  $N$  photons. N00N states are notable for being the most sensitive states for measuring the phase difference between two arms of an interferometer, or equivalently for estimating the angle of a polarization rotation around a known axis[9–12]. N00N states were first experimentally studied in [58, 59]. The particular N00N state aligned with the  $z$ -axis, visualized in Fig. 3.2, can be written as

$$|\psi_{N00N}\rangle = \frac{1}{\sqrt{2}}(|N_H, 0_V\rangle + |0_H, N_V\rangle) \quad (3.21)$$

When it comes to estimating all parameters of a rotation however, we see from the following spin covariance matrix and its corresponding bound

$$C_{\psi_{N00N}}(\hat{\mathbf{J}}) = \begin{pmatrix} N/4 & 0 & 0 \\ 0 & N/4 & 0 \\ 0 & 0 & N^2/4 \end{pmatrix}, \quad (3.22)$$

$$\mathcal{C} = \frac{2}{N} + \frac{1}{N^2},$$

that the performance of N00N states still exhibits a classical scaling. This is because N00N states fail at performing the estimation of the other two parameters with a quantum advantage and their contributions to the s-qCRB dominate in the high  $N$  limit.

We can attain a quantum scaling, however, by splitting the N00N state into three batches, as with the spin coherent state. If the photons are split equally between N00N states aligned with the  $x$ ,  $y$  and  $z$ -axes, we get

$$\sum_i C_{\psi_{N00N}^{(i)}}(\hat{\mathbf{J}}) = \frac{1}{36} (N(N+6)) \times \mathbf{1}, \quad (3.23)$$

$$\mathcal{C} = \frac{27}{N(N+6)}.$$

While splitting photons into three batches recovers the advantageous Heisenberg scaling, it causes us to pick up a factor of 3 in the leading order of  $N$ . Informally, we can think of the scaling as  $\mathcal{O}(3 \times 1/(N/3)^2) = \mathcal{O}(3^3/N^2)$ .

### Second-order unpolarized states

It was proven in [50, 60, 61] that the optimal states for the task at hand are pure states that exhibit the following properties:

$$\langle \hat{\mathbf{J}} \rangle = 0 \quad \langle \hat{J}_l \hat{J}_m \rangle = \frac{N}{6} \left( \frac{N}{2} + 1 \right) \delta_{l,m}. \quad (3.24)$$

These states have the property that they are isotropic up to the second moment of the spin angular-momentum operators, hence their name. Indeed, no initial rotation would change the two properties above for any state that satisfies them. These states do not exist in all dimensions. It is, for example, impossible to satisfy these conditions with one, two, or three photons. There is one solution for four photons but then again none for five photons. The dimensionality of the spaces for which these states exist still constitutes a topic of research [62, 63] that may be intimately connected to spherical  $t$ -designs [64–67] and other problems for distributing points on the surface

of a sphere [68–72].

Multiple names have been used to describe these states and their usefulness has been theoretically demonstrated for a range of applications. They were initially studied for being the opposite of spin coherent states, earning the epithet “anticoherent” [62, 73, 74], making them in some sense the most quantum states. Since they are the furthest states from the set of spin coherent states, they have been termed “Queens of quantumness” [75]. In the context of polarization, they manifest “hidden polarization” [76], as their classical (first-order) polarization properties are ignorant of higher-order polarization features [77]. They minimize the cumulative multipoles of their polarization distribution, for which they earned the moniker “Kings of quantumness” [67, 78], and are optimal states for rotation sensing, explaining the sobriquet “Quantum rotosensors” [79, 80]. Because they sometimes exhibits symmetries of platonic solids, they have also been dubbed “Platonic Solid States” [50] or “Platonic States”, a name that we will use parsimoniously in this document.

We can quickly see that the states satisfying Eq. 3.24 outperform all the other schemes imagined here by looking at their spin covariance matrix and the s-QCRB:

$$C_{\psi_{plat}}(\hat{\mathbf{J}}) = \frac{1}{12} (N(N+2)) \times \mathbf{1}, \quad (3.25)$$

$$\mathcal{C} = \frac{9}{N(N+2)}.$$

The first thing to notice is that, in the leading order of  $N$ , these states recover the factor of 3 that was lost from splitting the  $N00N$  state photons into three batches. Indeed, our cost function is now bounded by something on the order of  $\mathcal{O}(9/N^2)$  instead of  $\mathcal{O}(27/N^2)$ . This is in line with what we found in Sec. 2.4, where the increase in performance between sequential and simultaneous scheme is of order  $d^2 - 1 = 3$  for  $SU(2)$ .

In Figure 3.3, we show the performance scalings for the different families of state presented here.

#### 3.1.4 The Tetrahedron State

The second-order unpolarized state with the smallest photon number ( $N = 4$ ) is the one we call the “tetrahedron state”. It is named as such because it has the same symmetries as the tetrahedron under rotation, which is apparent in its Majorana representation (depicted in Fig. 3.4(a)), where the points fall on the vertices of a

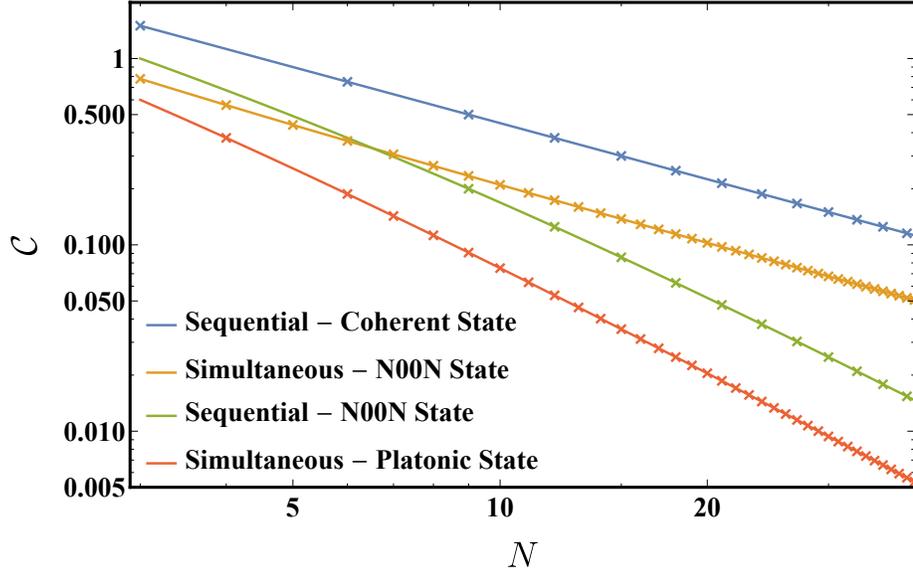


Figure 3.3: Scalar Quantum Cramér-Rao Bound  $\mathcal{C}$  of four different schemes vs Photon Number  $N$  used. The four different colours represent the four different schemes presented in Sec. 3.1.3. While simultaneous parameter estimation with N00N states and sequential parameter estimation with coherent states have shot-noise scalings ( $\mathcal{C} \propto 1/N$ ), sequential parameter estimation with N00N states and simultaneous parameter estimation with Platonic states reach the Heisenberg scaling ( $\mathcal{C} \propto 1/N^2$ ), with the latter beating the former by a factor of 3 for large enough  $N$ .

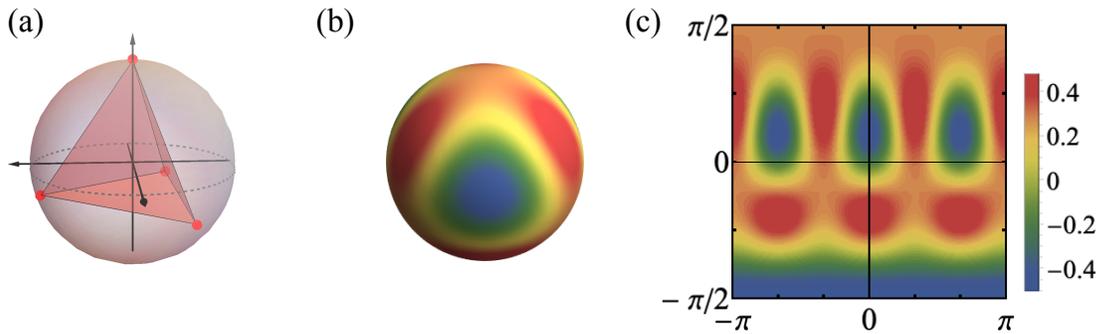


Figure 3.4: Three different visual representations of the tetrahedron state  $1/\sqrt{3}(|4_H, 0_V\rangle + \sqrt{2}|1_H, 3_V\rangle)$ . (a) Majorana representation of the state. The points are placed on the vertices of a tetrahedron inscribed in the Bloch sphere (b) Wigner distribution on the sphere. The quasi-probability distribution has fringes in the shape of a tetrahedron. (c) Equirectangular projection (longitude vs latitude) of the Wigner distribution.

tetrahedron inscribed in the Bloch sphere. The state can be written as

$$|\psi_{tetra}\rangle = \sqrt{\frac{1}{3}} |4_H, 0_V\rangle + \sqrt{\frac{2}{3}} |1_H, 3_V\rangle. \quad (3.26)$$

In this work, we create the tetrahedron state in the polarization state of four photons ideally in the same spatial and temporal mode, which guarantees the polarization to be full-symmetric. Second-order unpolarized states of dimension higher than the tetrahedron state have been previously studied experimentally [25] in the orbital angular momentum space of a single photon, which can be analogous to  $S_j$ . While the work was impactful, the creation of these states in this fashion requires no particle entanglement and we believe that the creation of the entangled tetrahedron state in the polarization of 4 photon remains a very interesting and new experiment.

## 3.2 Apparatus

In this section, we will talk about the apparatus we used to create and characterize the tetrahedron state. We will start with a high level description of the scheme before being a bit more technical as we take a deeper look at some parts of the apparatus. Some of the details included here are mostly for future group members who will wish to reuse sections of the apparatus for other purposes.

### 3.2.1 Concept

We create the tetrahedron state using an approach intuitive in the Majorana picture. This approach, conceptually illustrated in Figure 3.5, is to combine a 3-photon N00N state with a heralded single-photon. We first create a N00N state in the polarization of 3 photons in a single spatial/temporal mode. We label the two creation operators for the polarization modes  $\hat{a}_H^\dagger$  and  $\hat{a}_V^\dagger$ . At the same time, a heralded horizontally-polarized single-photon is created in a different spatial mode  $b$  with the creation operator  $\hat{b}_H^\dagger$ . The photon is created in pair with a photon in a spatial mode  $c$ , whose detection heralds the  $b$  photon. The creation of those two states will be discussed in Sec. 3.2.4 and 3.2.5.

Then, in a polarization interferometer named the Tetrahedron Assembly stage, the

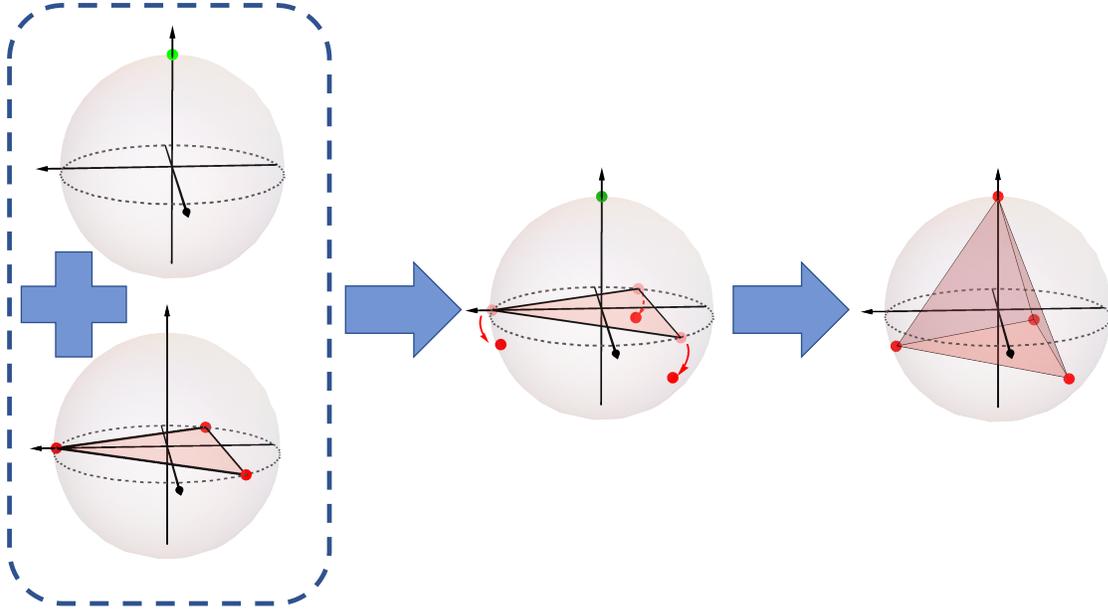


Figure 3.5: Conceptual illustration of the tetrahedron state's creation process using the Majorana representation. The H polarization of a 3-photon N00N state (three red points on the vertices of an equilateral triangle on the equator) is partially attenuated resulting in a triangle of points with lower latitude (red). Simultaneously, an H polarized heralded single-photon is probabilistically added (green). The result is a 4-photon state represented by four points on the vertices of a tetrahedron.

following mode transformation is performed:

$$\begin{aligned}
 \hat{a}_H^\dagger &\longrightarrow e^{i\phi} \hat{a}_V^\dagger \\
 \hat{a}_V^\dagger &\longrightarrow \frac{\hat{a}_H^\dagger}{\sqrt{2}} \\
 \hat{b}_H^\dagger &\longrightarrow \frac{\hat{a}_H^\dagger}{\sqrt{2}}.
 \end{aligned} \tag{3.27}$$

In this transformation, we attenuate the H polarization of the N00N state and probabilistically combine it with the H-polarized heralded single photon. This transformation results in the tetrahedron state.

$$\begin{aligned}
 |\psi\rangle &\propto \hat{b}_H^\dagger \left( \hat{a}_H^{\dagger 3} + \hat{a}_V^{\dagger 3} \right) |0\rangle \\
 &\longrightarrow \left( \frac{1}{4} \hat{a}_H^{\dagger 4} + \frac{1}{\sqrt{2}} e^{3i\phi} \hat{a}_H^\dagger \hat{a}_V^{\dagger 3} \right) |0\rangle \\
 &\propto \frac{1}{\sqrt{3}} |4_H, 0_V\rangle + \sqrt{\frac{2}{3}} e^{3i\phi} |1_H, 3_V\rangle
 \end{aligned} \tag{3.28}$$

As described, this non-unitary operation accepts three orthogonal modes as inputs and has two orthogonal modes as outputs. In reality, we rely on many photons leaving

the interferometer undetected through hidden output modes  $\hat{b}_{H/V}^\dagger$  to partially polarize the 3-photon N00N state. The detection of the tetrahedron state is conditioned on four photons being detected in modes  $a$  and one herald photon being detected in mode  $c$ . When photons are lost, less than 5 photons are simultaneously detected and these events are discarded, ideally not affecting the results.

### 3.2.2 Experimental Scheme

The full optical arrangement for the experiment is presented in Figure 3.6, with a description in its caption.

The experiment can be divided in four main parts: the N00N state source, the heralded single photon source, the tetrahedron assembly stage and the state tomography apparatus. The only part of the experiment not depicted in Figure 3.6 is some optics that we use to lock the N00N state source. It will be discussed more in Sec. 3.2.5.

### 3.2.3 Hardware

We use a commercial Ti-Saph mode-locked pulsed laser to power the experiment. The make of the laser is Coherent and the model is the Chameleon. The laser was originally purchased in 2013 but started failing in the summer of 2018. It was then exchanged for a nearly new model as part of a “repair” program offered by the company. This laser system is a unit made for a wide range of applications, which requires little technical skill to operate. It is a turnkey system, with much of the design being proprietary and hidden from the user. The limited information available to the public can be found in the user manual.

Maintenance of the system is easy and straightforward. The laser should preferably be turned on relatively frequently to ensure that the laser cavity is not so far out of alignment that the laser does not lase. Once the laser is in operation, the laser cavity is automatically aligned to maximize the power of the beam. Furthermore, the coolant needs to be regularly changed. It was recommended to me by a technician from Coherent to change it every month if using distilled water or every six months using the solution sold by Coherent for this purpose. I believe this solution contains a fungicide to help prevent the growth of algae in the cooling pipes.

The output of the laser system is a beam of light in free space. The beam is pulsed with a repetition rate of 80 MHz and a pulse duration of 140 fs. The centre wavelength of the pulses can be tuned to be anywhere between 680 nm and 1080 nm but for this experiment we chose to operate at around 807 nm to match the narrow-band filters we used. At 800 nm, the laser’s average power is maximal and specified to be at least

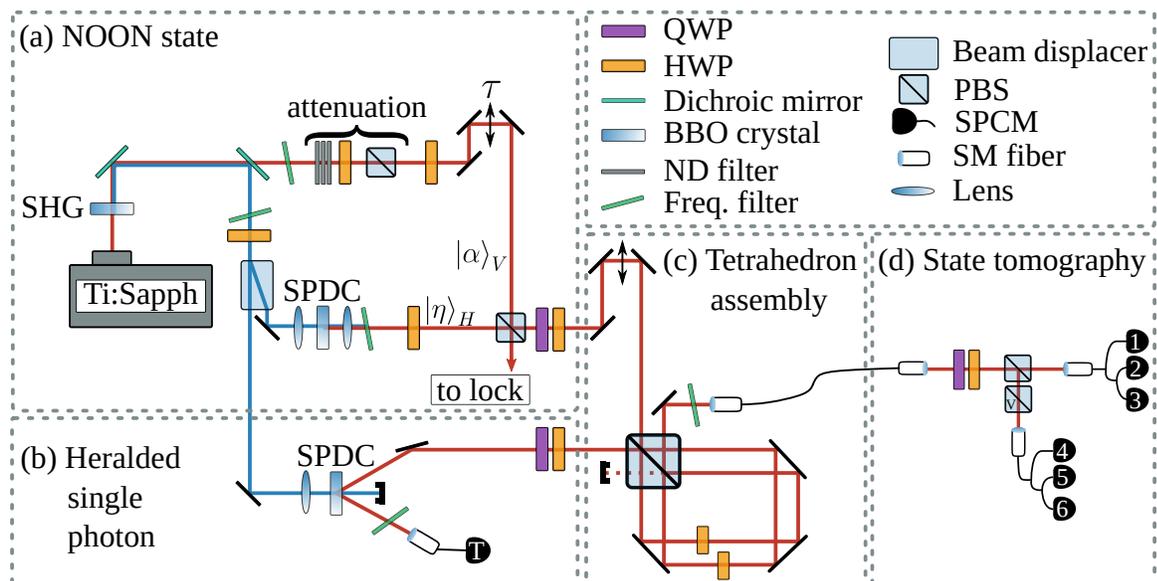


Figure 3.6: Experiment overview. (a) A 808 nm pulsed beam (red) is used to generate a 404 nm beam (blue) through second harmonic generation (SHG) in a 1 mm barium borate (BBO) crystal. Part of this blue beam is sent to generate co-propagating pairs of horizontally-polarized 808 nm photons (red) through type-I spontaneous parametric down-conversion (SPDC) in a 2 mm BBO crystal while the red beam is attenuated and rotated to a vertical polarization. The two beams are recombined on a polarizing beamsplitter (PBS) and polarization rotated to form a three-photon N00N state in the H/V basis. (b) Another blue beam then pumps a second SPDC (2 mm BBO) to generate another photon pair. Here, after a 3 nm narrow-band filter, a detection at a single photon counting module (SPCM) labelled “T” heralds the presence of a photon in the upper path. (c) Both the N00N state and heralded single photon enter a displaced polarization Sagnac interferometer. The heralded photon strictly takes the clockwise path while the N00N state is split equally amongst the paths. The half-waveplate (HWP) in the counter-clockwise path is set at  $45^\circ$  to transmit all the light in the upwards output while the HWP in the clockwise path is set at  $22.5^\circ$  to attenuate the V polarization of the N00N state by half and transmit the heralded single-photon half of the time (see Figure 3.5 for context). After being re-combined, all photons pass through a 3 nm narrow-band filter to make sure their spectra are the same. (d) The state is transmitted via single-mode (SM) fibre to a tomography apparatus for projection onto an arbitrary polarization axis. Each output path is coupled to a multi-mode fibre-splitter to allow for partial photon number counting capabilities (a maximum of three detections per polarization). Five-fold detection events between T and any four SPCMs in the tomography setup are recorded.

3.5 W. Our model over-performs and we get around 3.7 W at 807 nm.

For best and consistent results, it is best to let the laser run for at least an hour before taking meaningful measurements. It is also recommended to scan the frequency before starting the use. There is a built-in method in the software provided by Coherent to do this scan in an automated manner. This helps with the automated optimization of the alignment, all of which is obscured to the operator.

As for detectors, we use the very popular Single-Photon Counting Module (SPCM) from the manufacturer Excelitas. These used to be fabricated by Perkin-Elmer. For this experiment, we used nine different detectors from different years and with a variety of models.

All the detectors used are Single-Photon Avalanche Diode (SPAD) built with Silicon diodes, which work well with near infrared (NIR) light. The timing jitter is relatively low (350 ps) compared to the time between laser pulse (12 ns), which allows us to identify if two different detections or more come from the same laser pulse. The specifications for the current model quote the absolute detection efficiency to be  $\approx 62\%$ .

We use two home-built coincidence counting units built by Alan Stummer, who is now retired. Proper documentation regarding the coincidence box can be found on his website [81]. The device counts the coincidences between any of the eleven input channels. The inputs are TTLs, presumably coming from the SPCM outputs. There is an adjustable window for each channel, from 0 ns to 255 ns. Following a first detection, each of the photons falling in the coincidence window associated with this channel will be considered coincident. We can retrieve the number of coincidences in each of the  $2^{11} = 2048$  possible coincidence configurations.

### 3.2.4 SPDC

Spontaneous parametric down conversion (SPDC) [82] is a quantum non-linear optics phenomenon in which a high-energy photon is destroyed to create two low-energy photons. It is the time-reverse of a classical process called sum-frequency generation (SFG), where two low-frequency light beams interact in a non-linear material to generate a light at the sum of the two original frequency. Unlike SFG, which can be modelled by the classical Maxwell's equation in a nonlinear medium, SPDC is a strictly quantum process, requiring vacuum fluctuations, a feature of quantum theory. As a result, the quantum state of the light resulting from SPDC is quantum (meaning it cannot be described classically) and is a resource used in a plethora of quantum information schemes, including the violation of Bell's inequalities [83–85]. One application of SPDC is that it is used to create heralded single photons. Even though the

low-energy photons are created at random times during SPDC, they are only created in pairs, and so the detection of one of the photons guarantees the presence of another, thereby “heralding” it. SPDC is also used to create a “squeezed-state”, a state of light very useful in quantum metrology[15] as well as in quantum computing[86].

SPDC is at the heart of this experiment. Both the N00N state source(Fig. 3.6(a)) and the heralded single photon source(Fig. 3.6(b)) rely on this phenomenon. To mitigate the naturally low rate of pair creations, we typically use a high-power high-frequency beam to maximize the number of photons which could potentially decay into a photon pair. We call this beam our pump beam. Furthermore, we try to use as long of a crystal as possible, increasing the non-linear region in which SPDC can occur.

In this experiment, the high-frequency pump is generated with a 1 mm Barium-Borate crystal (BBO), in which we send the full output of the Chameleon laser at  $\sim 807$  nm. In that crystal, second harmonic generation (SHG) occurs[87]. This is a special case of SFG where both the inputs are of the same beam and the resulting high energy beam is at twice the original frequency. In SHG, the output intensity depends quadratically on the input intensity for low conversion efficiencies. The pulsed nature of our laser is very good for the conversion efficiency, effectively increasing the peak intensity by a factor of  $12 \text{ ns}/150 \text{ fs} \approx 8 \times 10^4$  from a CW beam of the same average power. We also focus this beam into the crystal with a 10 cm lens, further increasing the intensity in the crystal. Using the Chameleon output at full power ( $\approx 3.7 \text{ W}$ ), we can generate our pump beam at  $\approx 403.5 \text{ nm}$  (blue) with more than 600 mW of power.

Both SPDC processes in the experiment use 2 mm BBO crystal as the non-linear medium. We divide the blue beam into two to act as their pumps. In the N00N state source, we intend the two photons produced from the SPDC to be in the same mode, both in frequency and  $k$ -vector. The resulting state is what we call a vacuum squeezed state [82, 88]

$$|\eta\rangle_{a_H} = (1 - |2\eta|^2)^{\frac{1}{4}} \sum_i \frac{(-\eta)^i}{i!} \hat{a}_H^{\dagger 2i} |0\rangle = (1 - |2\eta|^2)^{\frac{1}{4}} \sum_i \frac{(-\eta)^i \sqrt{(2i)!}}{i!} |2i\rangle_{a_H}, \quad (3.29)$$

where  $\eta$  is the squeezing parameter. The probability of a pair being created during each pulse scales as  $\eta^2$  and is linear in the incoming pump power. In the heralded single photon source, we want the photons to be emitted in different directions because we would like to detect one of them immediately and send its partner to be used as

part of the tetrahedron state. The ideal result is a two-mode squeezed state[82]

$$|\mu\rangle_{b_H, c_H} = (1 - |\mu|^2)^{\frac{1}{2}} \sum_i \frac{(-\mu)^i}{i!} \hat{b}_H^{\dagger i} \hat{c}_H^{\dagger i} |0\rangle = (1 - |\mu|^2)^{\frac{1}{2}} \sum_i (-\mu)^i |i, i\rangle_{b_H, c_H}, \quad (3.30)$$

where  $\mu$  is the squeezing parameter. Going from a SPDC photon source where the pair of photons leave in the same direction to one where the pair leave in different directions can be done by tuning the angle between the extraordinary axis of the crystal and of the average pump beam's  $k$ -vector, thereby changing the phase-matching relations. In both sources, we want both photons of the same pair to be polarized equally in the H polarization. We therefore use Type I phase matching relation [87]. We align our pump's polarization to V and the optic axis of our BBO crystal to be in the plane formed by the average  $k$ -vector of the pump and its vertical polarization.

The equations above are idealized however. In the laboratory, the result of SPDC is a very multi-mode state with, in general, anti-correlations between frequency ( $\omega$ ) of the two photons in the photon pair and anti-correlations between their angle of emission relative to the average  $k$ -vector of the pump. These anti-correlations are consequence of the phase-matching relations, which ensure conservation of energy and momentum throughout the SPDC process. If we label the two photons of a pair the signal (s) and the idler (i), the energy conservation and phase-matching conditions are

$$\begin{aligned} \omega_s + \omega_i &= \omega_p \\ \mathbf{k}_s + \mathbf{k}_i &= \mathbf{k}_p. \end{aligned} \quad (3.31)$$

These relations and the resulting anti-correlations are not completely strict. The finite bandwidth of the pump due its pulsed nature, the finite phase-matching bandwidth of the crystal[87] caused by its finite length, and the non-zero range of the transverse  $k$ -vectors of the pump, which is focused into the crystal, all contribute to the partial reduction of the anti-correlations resulting from the phase-matching conditions.

These anti-correlations are a big problem for our experiment as they increase the distinguishability between photons from different pairs and different sources. To get rid of them, we filter our beams to suppress the anti-correlations. For the frequency anti-correlation, we introduce 3 nm filters after the SPDC. This is almost sufficient to get rid of the anti-correlation completely as 3 nm is very close to the wavelength width of our original transform limited 140 fs pulse. For the directional anti-correlation, we collect the photons in single-mode fibres, thereby ensuring that a single mode is collected from the N00N state source, and that only two modes are collected from the heralded source. Although this helps greatly with distinguishability, these filtering

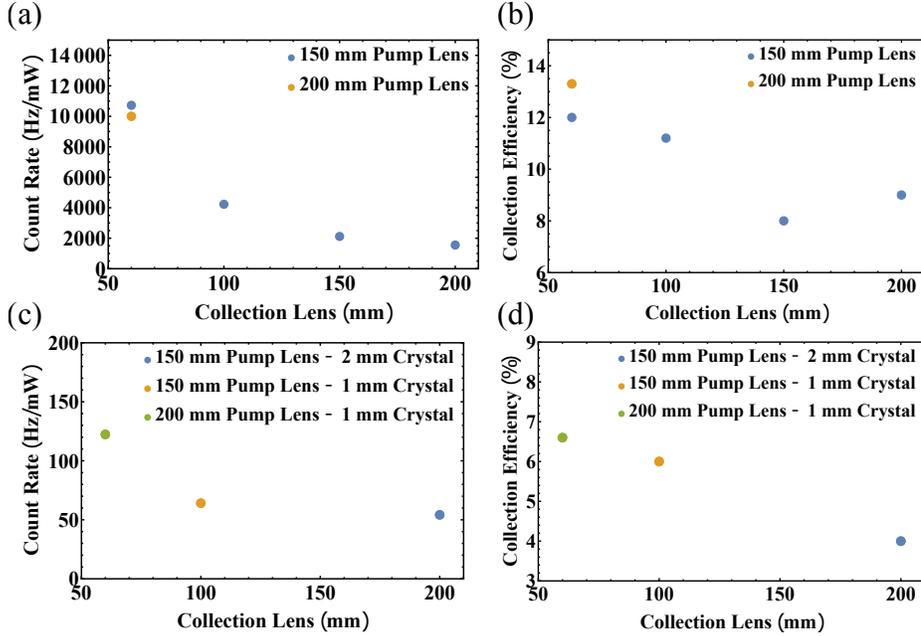


Figure 3.7: (a and c) Normalized coincidence count rate of the colinear SPDC source in the N00N state sources as a function of the collection lens focal length under various settings for the pump lens focal length and the crystal width. (b and d) Collection efficiency as a function of the collection lens focal length the same conditions. (c - d) are taken with the photons from the SPDC going through a 3 nm filter while in (a - b), the filter is removed.

processes increase photon losses after the SPDC. Loss is problematic for us in part because it decreases the state creation rate. Furthermore, loss leads to events when only one photon of a given pair is detected. This is particularly problematic as it decreases the quality of the squeezed states created through the SPDC. We will explain in more details what the repercussions are to our experiment in Section 3.4. Many theoretical discussions on how to filter while introducing the least amount of loss, both in frequency[89, 90], by changing the width of the filter, and in direction[90, 91], by changing the different lens shaping the size of the collection modes, can be found in the literature. While we kept these studies in mind when designing our apparatus, none seem to be exactly applicable to the source we needed and we took a more empirical approach. We tried various arrangements of crystal length, filter width, and lenses to focus the pump and the collection. The result can be found in Figure 3.7.

I wished we could have taken more points to fill this figure but the process ended up very time-consuming, with each point taking up to a long day of work, and we decided to cut the process short. I especially wish we had more points with the 2 mm crystal and the 3 nm filters, which we ended up using. In the end, for the N00N state source, we chose a filter width of 3 nm. We tried using two filters in succession with

different tilt angles to further reduce the window without much success. We chose to focus the pump beam with 200 mm lens before the 2 mm crystal. After the crystal, we first collimate the pairs with a 60 mm lens and then focus them into fibre with an asphere of 8 mm focal length. For the heralded sources, the lens focusing the pump had to be changed for a 150 mm lens and the lens used to collimate the collection mode was a 100 mm lens. The range of the focal length of the lens we tried was limited by physical constraints in our apparatus.

Right before we started data collection for our experiment, the N00N state source had a maximum coincidence rate of 53 kHz and a maximum heralding efficiency of 13.4% and the heralded source had a maximum coincidence rate of 32 kHz and a maximum heralding efficiency of 11.5% at full power. In a non-colinear geometry, the typical formula for the heralding efficiency is the coincidence detection rate divided by the product of the square root of each individual single detection rate. In our application, we found it more useful to report it as twice the coincidence detection rate over the sum of the single detection rate, replacing the geometric average of the single rate with the arithmetic one. Furthermore, for the N00N state source where both photons are emitted into the same mode, we first need to split these photons probabilistically in order to detect them with SPCMs. This decreases the efficiency which we need to readjust by multiplying by a factor depending on the arrangement used to split them in order for the measurement to be compared with that of the heralded source. For the numbers we just cited, three detectors were used and this factor is  $\approx 1.6$ , while in Fig. 3.7, only two detectors were used and the factor is  $\approx 2$ .

### 3.2.5 The N00N State Source

The N00N state source depicted in Figure 3.6(a) is the part of the apparatus we use to generate the 3-photon N00N state. The idea for this type of source dates back to 2007[92, 93], when it was realized that a remarkably simple apparatus was enough to create highly entangled N00N states. Since then, a few sources of this type have been built and used in a few experiments, some of which from our group[94, 95]. One has been demonstrated to create N00N states for up to 5 photons[96]. All that is required is to interfere a single-mode squeezed state with a coherent state[97, 98] at a beam-splitter. As a reminder a coherent state  $|\alpha\rangle$  can be expressed as

$$|\alpha\rangle_{aV} = e^{-\frac{1}{2}|\alpha|^2} \sum_i \frac{\alpha^i}{i!} \hat{a}_V^{\dagger i} |0\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_i \frac{\alpha^i}{\sqrt{i!}} |i\rangle_{aV}. \quad (3.32)$$

The coherent state is a classical state, commonly thought to be the quantum description of the state of a laser beam. A sketch of a conceptual N00N state source is

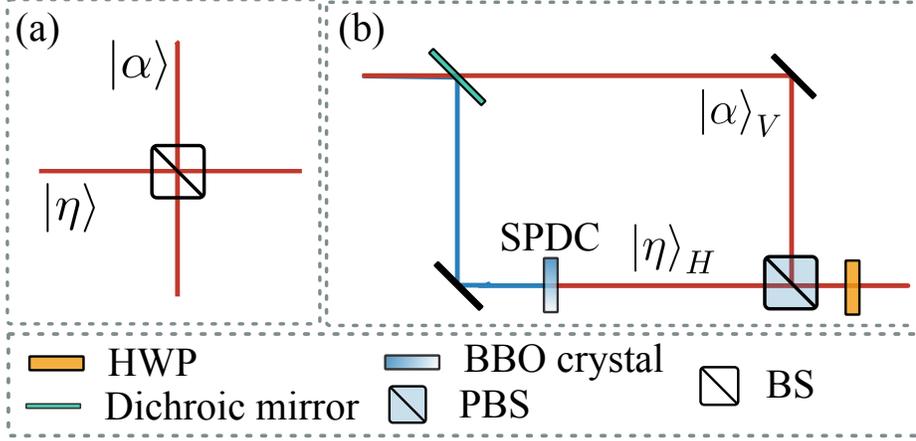


Figure 3.8: Sketch of an ideal N00N state source. In (a), the N00N state source is created in the two output mode of the BS. In (b), the N00N state source is created in the two polarization mode of the output beam.

depicted in Figure 3.8. The state generated is in a superposition of states with many total photon number  $N$ . What was realized is that for any  $N \geq 2$ , we can set the ratio of the coherent state amplitude  $\alpha$  and the squeezing parameter  $\eta$ , to make a state very close to a N00N state when post-selecting on the subspace with  $N$  total photons. For some values of  $N$ , like  $N = 2$  or  $N = 3$ , the N00N state can be reached exactly. In general, the state has at least 89% fidelity with the N00N state.

In our case, with  $N = 3$ , it is easy to see how we obtain a N00N state from this device. We can see that the only ways to obtain three photons in the final state is to get either three photons from the coherent state or two photons from the SPDC and one photon from the coherent state. Then, if we set

$$\alpha = \sqrt{2\eta}, \quad (3.33)$$

and perform the beam-splitter operation

$$\begin{aligned} \hat{a}_H^\dagger &\longrightarrow \frac{1}{\sqrt{2}}(\hat{a}_H^\dagger + \hat{a}_V^\dagger) \\ \hat{a}_V^\dagger &\longrightarrow \frac{1}{\sqrt{2}}(\hat{a}_H^\dagger - \hat{a}_V^\dagger), \end{aligned} \quad (3.34)$$

we get

$$(\hat{a}_H^{\dagger 2} \hat{a}_V^\dagger + \hat{a}_V^{\dagger 3}) |0\rangle \longrightarrow (\hat{a}_H^{\dagger 3} - \hat{a}_V^{\dagger 3}) |0\rangle \propto \frac{1}{\sqrt{2}} (|3_H, 0_V\rangle + |0_H, 3_V\rangle), \quad (3.35)$$

which is the N00N state. As it could be guessed from the mode labels, in practice we use polarization to encode the N00N state. The beam-splitter operation is performed

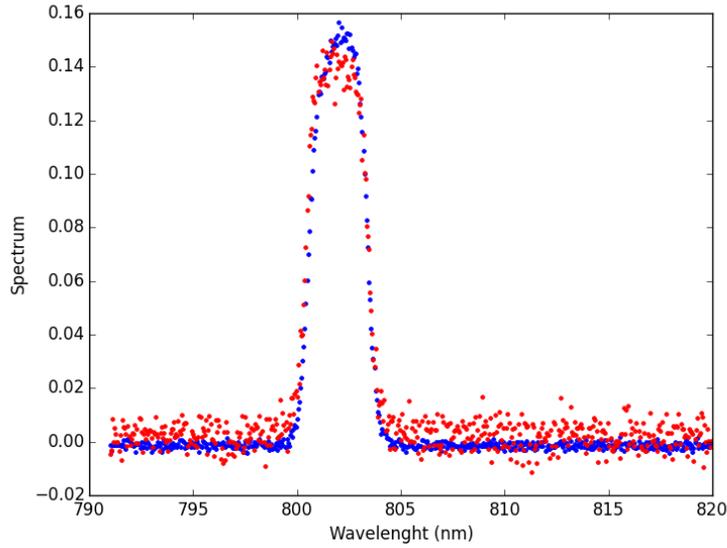


Figure 3.9: Normalized single photon spectra of the SPDC photons from the N00N state source (red) and of the coherent state from the N00N state source (blue). The spectra were taken with a grating spectrometer followed by a cooled CCD.

by a Half-Wave Plate (HWP) oriented at  $22.5^\circ$  from the H polarization axis. The coherent state  $|\alpha\rangle$  is the leftover red laser beam directly from the Chameleon after going through the SHG. The beam is separated from the blue beam with dichroic mirrors and frequency filters. It is then attenuated with many neutral density (ND) filters and its polarization is prepared to V with a HWP. The squeezed vacuum and the coherent state which are now of orthogonal polarization are then brought back into the same spatial mode with a polarizing beam-splitter (PBS) before the final aforementioned HWP.

### N00N state visibility

As we can see in Eq. 3.35 with the three photon example, it is important for the modes of the coherent state and the mode of the squeezed vacuum to only differ in polarization. If the modes with the creation operators  $\hat{a}_H^\dagger$  and  $\hat{a}_V^\dagger$  can be distinguished spatially or temporally, the interference that leads to the cancellation of many terms during the HWP rotation would not occur and the final state would be substantially different than a N00N state.

In order to make all these photons indistinguishable, we first need to make sure that their spectra are overlapped. In Fig. 3.9, we compare the single photon spectra of both the coherent state and of the SPDC after the 3 nm filters that we use for the experiment. We find an overlap exceeding 95%. The SPDC single photon spectrum

is much wider than the transmission spectrum of the filters and will therefore take the shape of the latter. The coherent state width however is comparable to that of the filter. It is very important to keep the filter's transmission spectrum centred on the spectrum of the coherent state. It can be changed by adjusting the tilt of the filter.

Then, we need to make sure that the photons from the coherent state arrive on the PBS at the same time as the photons from the SPDC. Because these pulses are very short ( $\approx 140$  fs), we cannot simply measure their arrival time to synchronize them. Instead we must use the interference effect that we aim to optimize. To do so, we look at the two-photon N00N state. This particular N00N state is interesting because it already looks like a N00N state before the final rotation by the HWP. The said rotation simply leaves the state unchanged. Once again, there are two ways to obtain two photons from both beams. We can get a single pair from the SPDC or a pair from the coherent state. If we tune the amplitude for these two pathways to be equal, by setting  $\alpha = \sqrt{2\eta}$ , we get the N00N state

$$|\psi_{(2-N00N)}\rangle = \frac{1}{\sqrt{2}} (|HH\rangle + |VV\rangle), \quad (3.36)$$

which we have expressed here in the first-quantized formalism. In the experiment, the coherent state timing on the PBS can be controlled with a motorized translation stage, and with a piezoelectric actuator for finer adjustments. Consider the temporal mode of the SPDC photon to be in  $|1\rangle$ , while the coherent state photons are in  $|2\rangle$ . The N00N state then becomes

$$|\psi_{(2-N00N)}\rangle = \frac{1}{\sqrt{2}} (|HH\rangle |11\rangle + |VV\rangle |22\rangle). \quad (3.37)$$

We use  $\tau$  to denote the difference in arrival time between the coherent state photons and the SPDC photons. If  $\tau \ll 150$  fs, then the temporal modes are almost overlapped and  $\langle 1|2\rangle \approx e^{i\omega\tau}$ . If  $\tau \gg 150$  fs, then the temporal modes are almost orthogonal  $\langle 1|2\rangle \approx 0$ . If we take the partial trace over the temporal modes, in the first case, we have a pure state

$$\rho = \frac{1}{2} (|HH\rangle + e^{i2\omega\tau} |VV\rangle) (\langle HH| + e^{-i2\omega\tau} \langle VV|), \quad (3.38)$$

while in the latter case, we have a mixed state that no longer depends on  $\tau$

$$\rho = \frac{1}{2} (|HH\rangle \langle HH| + |VV\rangle \langle VV|). \quad (3.39)$$

A simple projection on  $|D, D\rangle \langle D, D|$ , where  $|D\rangle = 1/\sqrt{2} (|H\rangle + |V\rangle)$  reveals the

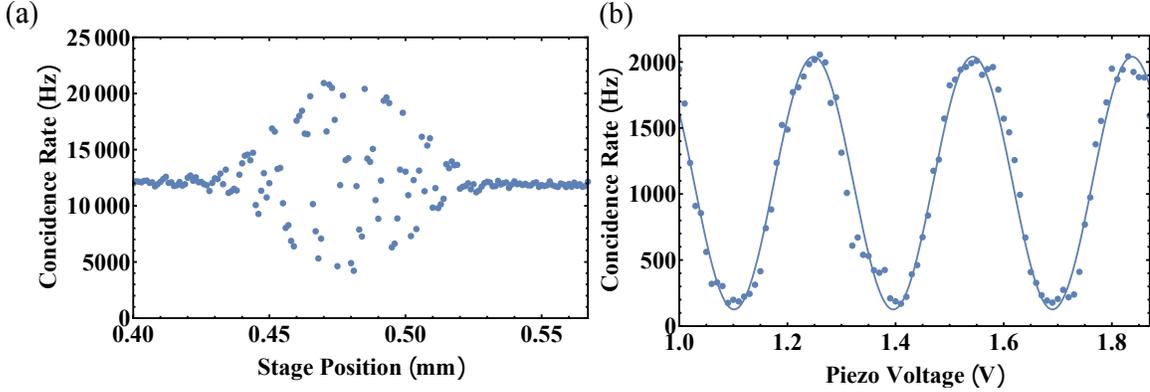


Figure 3.10: Coincidence rate after a  $\langle D, D \rangle$  polarization projection of the photons from the N00N state source as a function of (a) the translation stage position governing the coarse path length difference between the SPDC photon and the coherent state photons and (b) the voltage in the piezoelectric actuator governing the fine path length difference.

interference. In the first case

$$P_{DD} = \frac{1}{4} + \frac{1}{4} \cos(2\omega\tau), \quad (3.40)$$

while in the second case,  $P_{DD} = 1/4$ . In the laboratory, we exploit this interference to set the path length in the interferometer in the N00N state source. We use waveplates and a PBS to perform the projection and then scan the path length in steps of  $\approx 1 \mu\text{m}$  until we see a region where points deviate significantly from step to step. At this stage, we are probing coarsely enough that we cannot resolve the interference fringes and the points appear fall randomly within an envelope in which the fringes lie. An example of such one such scan can be seen in Fig. 3.10(a). We then set our motorized stage to the peak of the envelope and perform a scan with the piezo-electric actuator. This will map out the fringe pattern, which allows us to measure the visibility of the interference pattern. This can be seen in Fig. 3.10(b). Prior to the data collection for the experiment, we measured the visibility to be 88%. We think that the remaining imperfections might be due to a mismatch in the two-photon spectra of the SPDC and of the coherent state. Indeed, even if the single-photon spectra looked very closely matched, there could be anti-correlations in the spectrum of the photon pair which are not completely removed by the 3 nm filter.

### Lock

Once we set the phase between the coherent state and the SPDC, it is important that this phase remains the same throughout the experiment. Unfortunately, in practice, path lengths tend to perpetually shift due to air currents, vibrations and thermal expansions and contractions of optical components. Once we do everything

to passively reduce the impact of such effects by enclosing the N00N state source into a box, we still have significant drifts. The drifting speed varies but we would frequently measure drifts on the order of one fringe per 5-10 minutes. It is therefore important that we actively monitor these drifts and correct them in real time. To do this, we slightly change the polarization of the squeezed-vacuum state and of the coherent state before the PBS with HWPs. This has ideally little effect on the intended output state from the N00N state source but allows us to collect a bit of light from both states in the second output from the PBS, with their polarizations reversed. We then perform a polarization projection on both photons and measure the coincidence rate with a beam-splitter and two SPCM. The projection is chosen so that, while  $\tau = 0$ , the lock signal is in the middle of a fringe, the point most sensitive to small changes in  $\tau$ . Our lock is fairly slow, with a measurement every 2s. We use only an integral gain to calculate the correction voltage sent to the piezo-electric actuator. We set the gain to be as low as possible while keeping the lock-signal robustly around the set-point. Throughout the experiments, there were many times when the correction voltage exceeded the range of the piezo-electric actuator. When this happened, we automatically reset the actuator voltage to the middle of its range and let the interferometer re-lock on its own. We then excluded the data points which had been taken during the times when the interferometer was not in lock. Because the set point using this method is not on a zero-crossing of the lock signal, we are sensitive to drifts in the SPDC rate or in the coherent state rate. We would typically re-calibrate the set point every 8-12 hours of data collection to reduce the impact of these drifts.

### 3.2.6 The Tetrahedron Assembly

The tetrahedron assembly stage is where the 3-photon N00N state and the heralded singled photons are combined to create the tetrahedron. It is depicted in Figure 3.6(c). We already talked about the ideal transformation that is performed at this stage in Section 3.2.1, where we gave an overview of the tetrahedron state creation process. In short, it is a polarization displaced Sagnac interferometer with two inputs, each with its own polarization. We adjust a different HWP in each of the path allowing us to change which polarization from each path exits in which output. In Figure 3.6(c), the N00N state enters from the top. The H photons take the counter-clockwise path, go through a HWP at  $45^\circ$  and exit as V polarized through the top path. Meanwhile both the V photons from the N00N state source and the H heralded single photon coming from the left output take the clockwise path. In this path, there is a HWP at  $22.5^\circ$ , which means that when they hit the PBS on their exit, they interfere. The

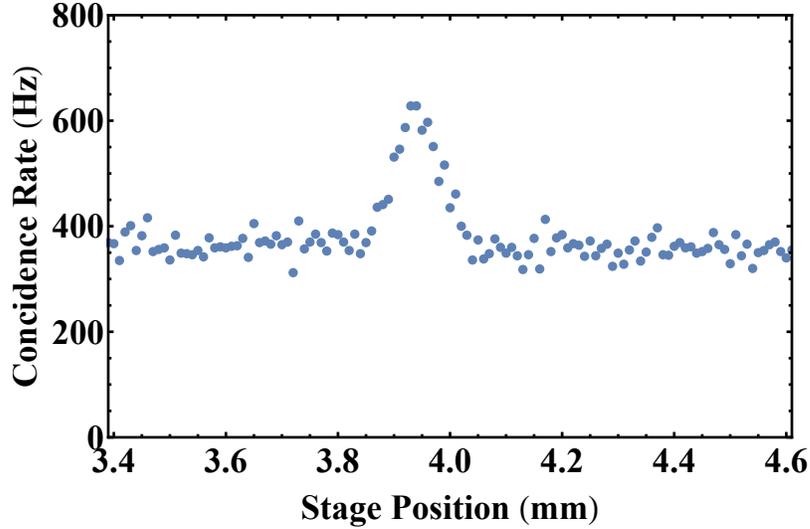


Figure 3.11: Triple coincidence rate between the herald detector and the pairs of detectors placed at the output of the tetrahedron assembly stage as a function of the translation stage position controlling the path length of the heralded single photon before the tetrahedron assembly. The peak we see is the result of a reverse HOM effect, where photon bunching is observed in a single output of the PBS in the tetrahedron assembly stage where the heralded single photon meet a coherent state photon from the N00N state source.

H polarization of this beams exits from the top and there we have our tetrahedron state. The V contribution leaves through the side port and is left undetected. The phase difference between the two paths, represented by  $\phi$  in Equations 3.27 and 3.29, is left uncontrolled and takes an arbitrary value. Thanks to the Sagnac geometry of this interferometer, it remains stable throughout the data collection without active locking.

Once again, it is important to make sure that the arrival time of the heralded photon and the N00N state at the PBS in the Sagnac interferometer is identical. We control this arrival time difference with a motorized translation stage in the path of the N00N state. This time however, the phase between the two beams is an irrelevant global phase, and so we need no fine control or active locking of this path difference. To make sure that we our path difference is aligned, we block the SPDC from the N00N state source and look for the Hong-Ou-Mandel[99] interference between the heralded single photon and the coherent state when they meet on the tetrahedron assembly's PBS. The effect comes from the single-photon contribution of the coherent state, which is dominant at low power. Because we only collect one of the paths, we look for a bump in coincidences between two detectors placed in that path, as opposed to a more traditional dip in coincidence between two detectors placed in different paths. This is sometimes called a reverse HOM effect. An example of such a scan can be seen in Fig. 3.11. All that remains is to set the translation stage to the position of

the peak.

### 3.2.7 The State Tomography Apparatus

To characterize the tetrahedron state we created, we perform state tomography in the section of the apparatus depicted in Fig.3.6(d). We will discuss the tomography protocol in much detail in Section 3.3.1 but we will briefly discuss the hardware used here. As we will discuss later, all we need is the ability to perform a polarization measurement in the same basis on all four photons of the tetrahedron state. We can perform these measurements by sending our tetrahedron state through a QWP and a HWP followed by a PBS. The QWP and HWP set the basis of the measurement and the PBS separates the two orthogonal polarizations in said basis. In the reflected path of the PBS, we set another PBS mounted vertically, which lets through most of the light reflected but removes some of the horizontally polarized light that leaks in small amount in the reflected port of the PBS. In each port of the PBS, we place what we call a probabilistic photon number resolving detector (PPNRD). They consist of a network of fibre beam-splitters, which probabilistically route photons onto different SPCMs. Here, we use 3 SPCMs in each of the port. Since there are 4 photons in the tetrahedron state, this allows us to register 3 different outcomes of the polarizations measurement in any of the basis. We label these by the projectors  $|3_T, 1_R\rangle\langle 3_T, 1_R|$ ,  $|2_T, 2_R\rangle\langle 2_T, 2_R|$  and  $|1_T, 3_R\rangle\langle 1_T, 3_R|$ .  $n_T$  and  $n_R$  indicate the number of photons that were transmitted and reflected at the PBS. This corresponds to a polarization measurement on the state prior to the tomography apparatus in a basis determined by the QWP and HWP's positions. Note that the projectors  $|4_T, 0_R\rangle\langle 4_T, 0_R|$  and  $|0_T, 4_R\rangle\langle 0_T, 4_R|$  cannot be measured as each of our PPNRD only have 3 SPCMs. It is important to note that the projector  $|n_T, n_R\rangle\langle n_T, n_R|$  is not a projector only on the symmetric basis. Any state with  $n_T$  transmitted photons and  $n_R$  reflected photons is part of this projector's subspace, including the non-full-symmetric ones. This may be a small abuse of notation as it looks like those kets are expressed in a second-quantized formalism.

## 3.3 Results and Discussion

### 3.3.1 Tomography with Hidden Differences

The Hilbert space for 4 photons' polarization is

$$\mathcal{H} = (S_{1/2})^{\otimes 4} = (S_2) \oplus (S_1)^{\oplus 3} \oplus (S_0)^{\oplus 2} \quad (3.41)$$

The change of basis between these two decompositions can be found in Appendix A. As explained, the tetrahedron state is fully contained in  $S_2$ , the full-symmetric sector. In practice, we can expect some mismatch in the modes of the photons from different sources, which allows other spin sectors to be populated. Importantly, the parasitic distinguishability between the photons cannot be accessed, i.e., we have no practical way of favorably selecting photons from one source over another, which means that the only measurements we can perform have projectors symmetric under particle exchange. As a result, only the parts of the density matrix depicted in Fig. 3.12 (a) can be reconstructed. They specifically consist of the spin sectors in the tensor sum decomposition of the Hilbert space. The coherences between these sectors are inaccessible. Furthermore, we are only sensitive to the sum of the different spin sectors with the same spin value. We make the choice to report their average in each of the different sectors. Importantly, although our knowledge of the full state of the system is fundamentally limited in this way, the accessible information is enough to predict the result of any symmetric measurement done on the same state.

The technique we use to reconstruct this part of the density matrix is called “Tomography with Hidden Differences” [52, 100–102]. It consists of a sequence of measurements on the polarization of our state akin to Stern-Gerlach measurements for a spin. The relevant section of the experimental apparatus is depicted in Fig. 3.6 (d) and has already been described in Sec. 3.2.6. We perform this polarization measurement in 13 different bases, chosen to be roughly uniformly spread on the Bloch sphere. In Figure 3.13, we illustrate these array of bases, each represented by a vector going through the centre of the faces of a small rhombicuboctahedron inscribed in the Bloch sphere. We chose this solid with 26 faces for nothing more than its apparent symmetry. For every vector in this set, the opposite vector, representing the orthogonal state, is also included. This pair of vector forms a redundant measurements, simply substituting  $T$  for  $R$  and vice-versa in the associated projectors. We keep this redundancy as a sanity check and hope to cancel out potential systematic bias between the two outcomes. With 13 measurement bases (26 waveplate settings), and with 3 projectors per basis, we can perform 39 different projections, which is enough to estimate the 35 linearly independent parameters of the accessible sections of the density matrix. For that purpose, we use a maximum likelihood method, following the method outlined in [103], except we use a Poisson distribution to model our counts. We further modify our model to account for the detector inefficiencies as well as to account for potential losses incurred in the PPNRDs due to multiple photons landing on the same SPCM. The inclusion of this loss in our model is critical since it affects different projectors differently. We discuss it in more details in Sec. 3.4.2.

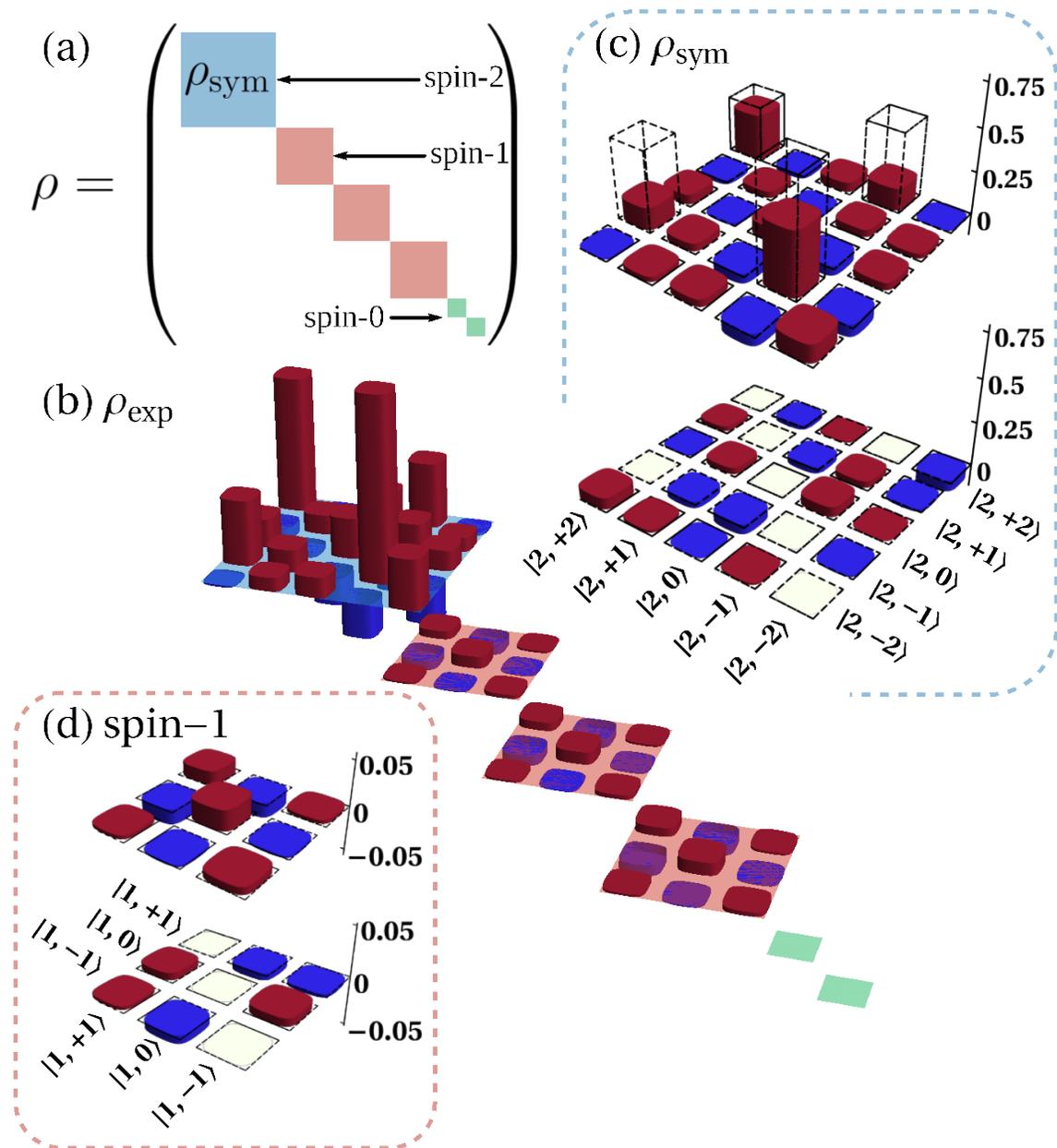


Figure 3.12: (a) Layout of the full density matrix.  $\rho_{\text{sym}}$  is the part of the state contained in full-symmetric  $S_2$  sector (blue). The  $S_1$  (red) and  $S_0$  (green) sectors are also shown. (b) The full tomographically reconstructed density matrix,  $\rho_{\text{exp}}$ . In hidden differences tomography, coherences between spin sectors are neglected and sectors of the same spins are chosen to be identical. (c) Real and imaginary parts of the reconstructed  $\rho_{\text{sym}}$ . The density matrix of the ideal tetrahedron state is shown for comparison (dashed). Elements of the density matrix are labeled with spin notation  $|j, m\rangle$ . (d) Real and imaginary parts of a  $S_1$  sector. The population in each  $S_0$  sector is less than  $10^{-10}$ .

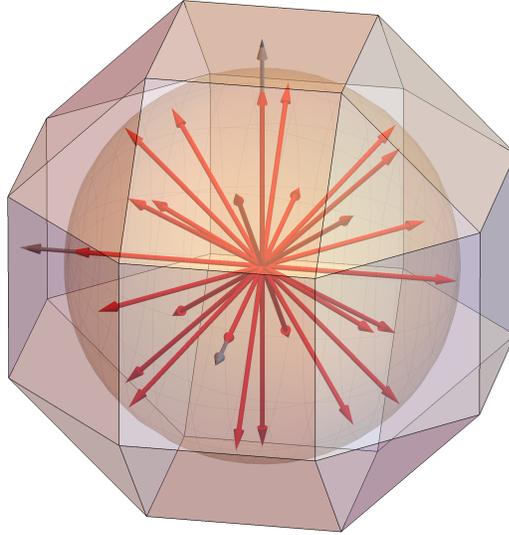


Figure 3.13: Illustration of the polarization basis measured during the tomography. All the axes go through the centre of the faces of a small rhombicuboctahedron. Vectors opposite of one another represent the same basis rotated so that the meaning of a transmission and reflection through the PBS are interchanged. For example, the vectors going through the North (South) poles indicate a measurement where the H (V) photons are transmitted through the PBS.

### 3.3.2 Results

Overall, during the experiment, 2434 successful events were recorded in 85 hours of counting time. These counting hours were distributed over the course of 8 days. The rest of the time was used to turn the tomography waveplates or was wasted while the interferometer in the N00N state source (Fig. 3.6 (a)) was out-of-lock. Every 8-12 hours, minor realignments were made to readjust the couplings in single-mode fibres and to reset the set-point of the locked interferometer.

Figure 3.12 shows the tomographic reconstruction of the density matrix  $\rho_{exp}$ . For ease of comparison with the theoretical tetrahedron state  $|\psi_{tetra}\rangle$ , a rotation  $e^{-i\hat{J}_z\phi}$  with arbitrary phase  $\phi$  was added to the reconstructed  $\rho_{exp}$ . We estimated  $\phi$  by maximizing the fidelity ( $F = \langle\psi_{tetra}|\rho_{exp}|\psi_{tetra}\rangle$ ) between the reconstructed and tetrahedron state. We find  $\phi = 0.135$  for a fidelity with the tetrahedron state of  $(0.46 \pm 0.02)$  and plot the appropriately rotated density matrix. Error bars on the entries of the density matrix were determined by a Monte-Carlo simulation of the tomographic reconstruction process, where the simulated number of detection events for each projection was drawn from a Poisson distribution centred on the actual measurement results. The full reconstructed density matrix and the corresponding error bars can be seen in Appendix B. Despite the low fidelity between the theoretical tetrahedron state and our reconstructed density matrix, it retains many of the notable features of

the former. As expected, most of the population ( $87 \pm 4\%$ ) is in the full-symmetric  $S_2$  subspace ( $\rho_{sym}$  in Fig. 3.12 (a)). The excess population in the other sectors is small but not quite consistent with zero, indicating that there were small mode mismatches between the different photons. As intended, the two basis elements with the biggest population in  $\rho_{sym}$  are  $|2, 2\rangle$  and  $|2, -1\rangle$ . The ratio between these populations however is  $(0.73 \pm 0.12)$ , which is significantly larger than the theoretical value of  $1/2$ . We unfortunately measure significant population ( $\approx 10\%$ ) in the states  $|2, 0\rangle$  and  $|2, -2\rangle$  that are meant to be empty. Finally, the biggest coherence measured is between  $|2, 2\rangle$  and  $|2, -1\rangle$ . Theoretically, this is the only non-zero coherence and it takes a value of  $\sqrt{2}/3$ . Here, we measure the significantly lower value of  $(0.13 \pm 0.02)$ , also much lower than  $(0.32 \pm 0.02)$ , the maximum possible value given the corresponding populations. The coherence quoted is perfectly real specifically because we allowed an optimal rotation around the H/V axis in the reconstructed density matrix. Most of these discrepancies can be explained by considering the contributions of events where six or more photons were created but were detected as a successful five-fold coincidence due to photon loss at the detectors or in the optical apparatus. In principle, the importance of these background terms could be made negligible by reducing the overall state creation rate, or by reducing the loss in the experiment. We dedicate Sec. 3.4 to discuss these sources of background.

### 3.3.3 Tetrahedral Symmetry

The most striking features of the tetrahedron state is its tetrahedral symmetry, which can be visualized in both its Majorana representation and in its Wigner function. The experimental state we created and reconstructed qualitatively shares these features. In Fig. 3.14, we can see the Wigner function of our state on the Bloch sphere. It can be compared to the ideal Wigner function depicted in Fig. 3.4. The symmetry may be more apparent on a 2D projection of the Wigner function. Because the projection deforms the shape of the features depending of their locations, we depict four different projections of the same state from different viewpoints in Fig. 3.15. For any state with tetrahedral symmetry, including the ideal tetrahedron state, these four different projections would look identical. In Figure 3.15, we see that the four projections are qualitatively very similar, with the main features, minima and maxima, being located in the same place.

To take a deeper look at the reconstructed state, we also analyze the performance of our state for single parameter estimation, and compare it to the ideal tetrahedron state. To estimate the rotation angle  $\theta$  of a rotation with a known rotation axis using an ideal tetrahedron state, a projection onto the original tetrahedron state is the ideal

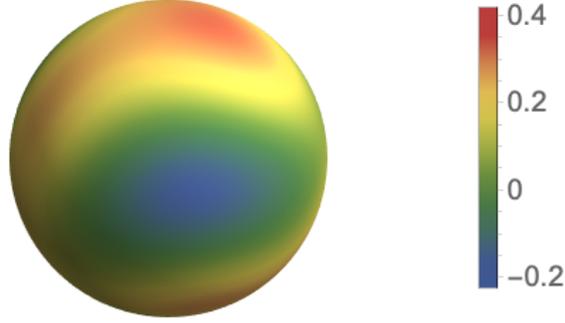


Figure 3.14: Wigner distribution of  $\rho_{sym}$ , our reconstructed experimentally created tetrahedron state.

measurement for small angles of rotation. The probability of a successful projection for an arbitrary state  $\rho$  is

$$P(\theta) = \langle \psi_{tetra} | e^{in \cdot \hat{J}\theta} \rho e^{-in \cdot \hat{J}\theta} | \psi_{tetra} \rangle. \quad (3.42)$$

In Figure 3.16, we plot  $P$  as a function of  $\theta$  for the  $x$ ,  $y$ , and  $z$ -axes, for our reconstructed state  $\rho_{exp}$ . We compare these curves with similar curves for the ideal tetrahedron state. As we can see, our state starts with a low probability of a successful projection at  $\theta = 0$  and this probability stays fairly low for all axes and for all  $\theta$ . This is a reflection of the low fidelity we have with the ideal tetrahedron state to begin with. We note however, that if we disregard this anomaly, the qualitative feature in these two sets of curves are similar. For all rotation axes, we clearly see an oscillation with a periodicity of  $2\pi/3$ , a distinct feature of the tetrahedron state, and once again, the maximia and minima for both states are aligned. The performance of such a measurement with our reconstructed state is worse. We plot the sensitivity of this measurement

$$\Delta\theta = \frac{\sqrt{P}\sqrt{1-P}}{\left| \frac{\partial P}{\partial \theta} \right|} \quad (3.43)$$

for multiple rotations as a function of the rotation angle  $\theta$  in Figure 3.17. The features here are a bit different, with divergences in the sensitivity appearing whenever a maximum or a minimum in  $P(\theta)$  takes a value lower than 1 or higher than 0 respectively. It is worthy to note that the tetrahedron state was not the optimal state for single

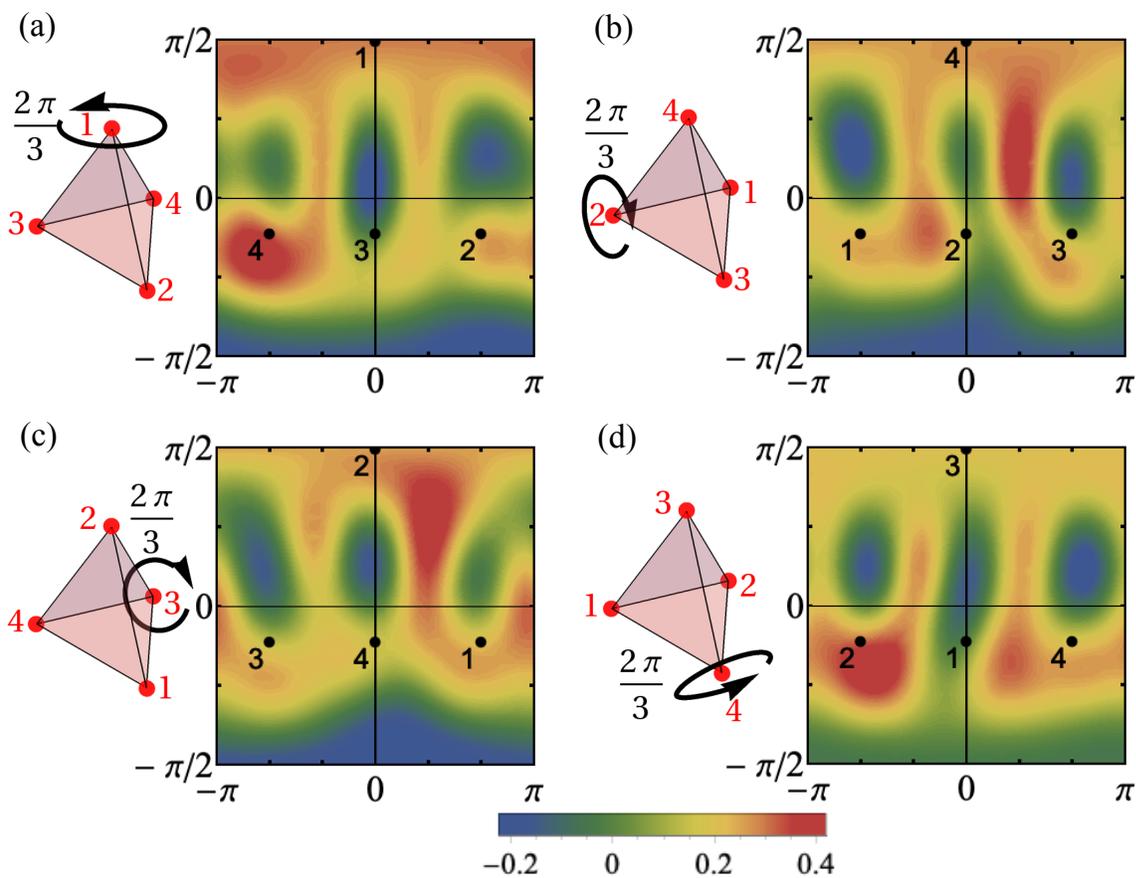


Figure 3.15: (a - d) Equirectangular projection (Longitude vs. Latitude map) of the Wigner function of  $\rho_{sym}$  rotated in 4 different ways to give different perspective on the Wigner function's features. The 4 different rotations are all anti-clockwise  $2\pi/3$  rotations around axes defined by each of the vertices of the ideal tetrahedron. All these rotations would leave  $|\psi_{tetra}\rangle$  unchanged. The similitude between these four figures is indicative that  $\rho_{sym}$  qualitatively possesses tetrahedral symmetry.

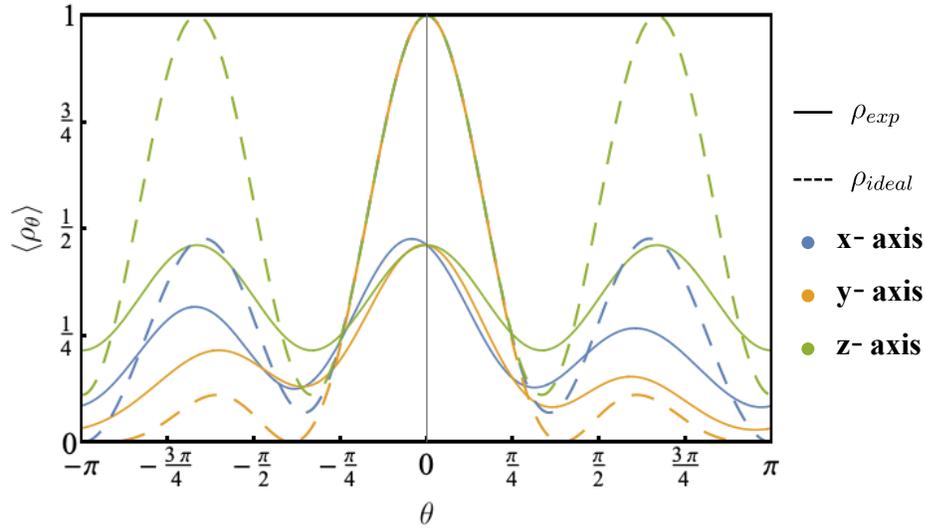


Figure 3.16: Projection of a  $\rho_{exp}$  and  $\rho_{ideal} = |\psi_{tetra}\rangle \langle \psi_{tetra}|$  onto the tetrahedron state  $\langle \psi_{tetra}|$  vs. the rotation angle  $\theta$ . The rotation is around three different known axes,  $x$ ,  $y$  and  $z$ . Both  $\rho_{exp}$  and  $\rho_{ideal}$  have three fringes per rotation, a feature of the tetrahedron state.

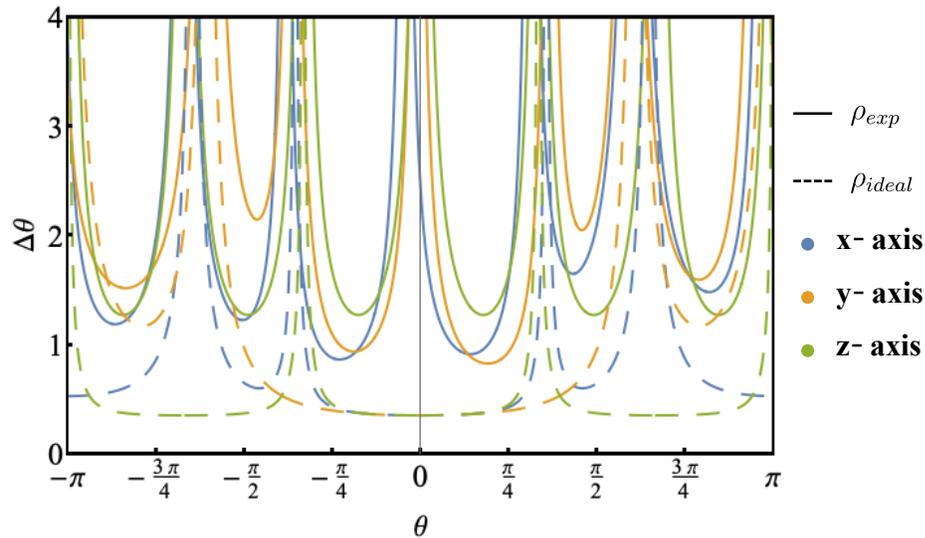


Figure 3.17: Sensitivity of a  $\rho_{exp}$  and  $\rho_{ideal} = |\psi_{tetra}\rangle \langle \psi_{tetra}|$  to the angle of a rotation around a known axis when projected back onto  $\langle \psi_{tetra}|$  vs. the rotation angle  $\theta$ . The rotation is around three different known axes,  $x$ ,  $y$  and  $z$ .

parameter estimation to begin with, nor the projection on a tetrahedron state the optimal measurement with our reconstructed state. We simply showed these plots to observe the qualitative resemblance between  $\rho_{exp}$  and  $|\psi_{tetra}\rangle$ .

### 3.3.4 Reconstructed Eigenstates

Because  $\rho_{exp}$  is mixed, it does not have a Majorana representation. A useful exercise however is to look at its eigenstates. Since the state lives in a 16-dimensional space, we can expect 16 different eigenstates. Since the tomography with hidden differences constrains all coherences between the different spin sector to 0, each eigenvector is fully contained in a single spin block. Furthermore, we have imposed that the three different  $S_1$  sectors and the two different  $S_0$  sectors of the density matrix be identical. This means that the eigenvectors contained in these blocks will be identical to the eigenvectors of the similar spin sectors. Overall, this means that we have 5 different eigenvectors for the  $S_2$  sector, 3 different eigenvectors repeated 3 times in each of the  $S_1$  sectors and an eigenvector for each of the two  $S_0$  sectors. The two  $S_0$  eigenvectors are the trivial ones since  $S_0$  has one dimension. In Table 3.3.4, we can see the Majorana representation of each of these eigenvectors with the associated eigenvalue. Note that for  $S_1$ , the Majorana representation does not strictly have the same interpretation as for  $S_2$ . These states are still formed of by four photons, despite their Majorana representation being comprised of two vectors. However, we can still use the Majorana representation to provide intuition on the behaviour of these states under rotations.

The biggest eigenstate, with eigenvalue  $\lambda = 0.490$ , looks remarkably like the tetrahedron state, with an overlap of 92%, confirming that almost half the time, we create a state very similar to the intended tetrahedron. Its Majorana representation is plotted in Fig 3.18 with a comparison to the tetrahedron state.

### 3.3.5 Experimental s-qCRB

Finally, we analyze how  $\rho_{exp}$  performs when characterizing SU(2) rotations. Unfortunately, we did not experimentally characterize any rotations. Rather, we look at the ideal performance for  $\rho_{exp}$ , the density matrix we tomographically reconstructed. We compute the QFI with the formula for mixed states[35]

$$[I]_{jk} = \sum_{lm} (\lambda_l + \lambda_m) \left( \frac{\lambda_l - \lambda_m}{\lambda_l + \lambda_m} \right)^2 \langle l | (\hat{H}_k - \langle \hat{H}_k \rangle) | m \rangle \langle m | (\hat{H}_j - \langle \hat{H}_j \rangle) | l \rangle \quad (3.44)$$

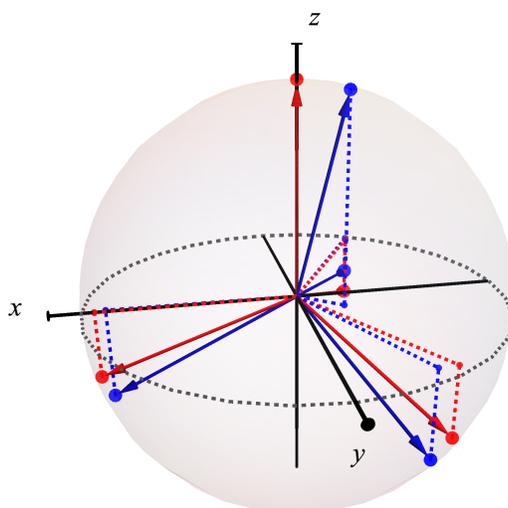


Figure 3.18: Majorana representation of the eigenstate of  $\rho_{exp}$  with the biggest eigenvalue (blue) and Majorana representation of  $|\psi_{tetra}\rangle$  (red).

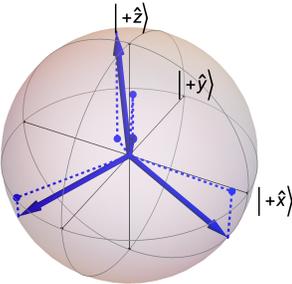
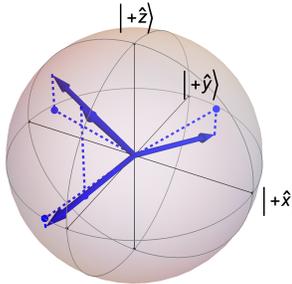
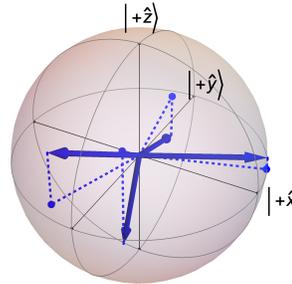
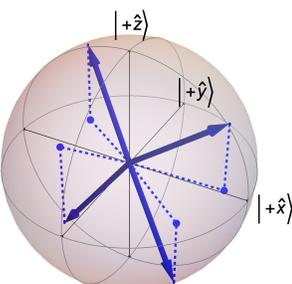
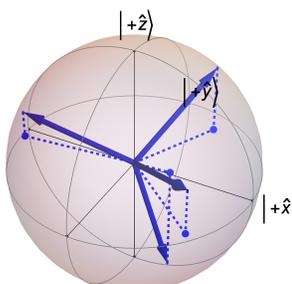
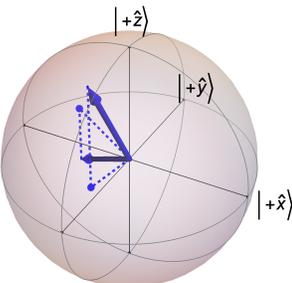
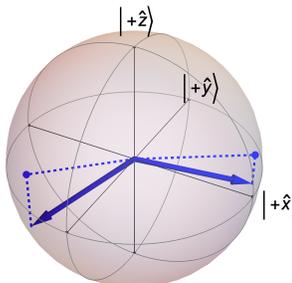
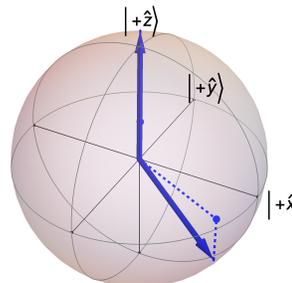
Spin-2 (Multiplicity 1)		
		
$\lambda = 0.490$	$\lambda = 0.255$	$\lambda = 0.116$
Spin-2 (Multiplicity 1)	Spin-0 (Multiplicity 2)	
		
$\lambda = 3.94 \times 10^{-3}$	$\lambda = 1.41 \times 10^{-6}$	N/A $\lambda = 6.76 \times 10^{-11}$
Spin-1 (Multiplicity 3)		
		
$\lambda = 3.94 \times 10^{-2}$	$\lambda = 5.46 \times 10^{-3}$	$\lambda = 7.72 \times 10^{-11}$

Table 3.1: Majorana representation of the different eigenstates of  $\rho_{exp}$  with their respective eigenvalues. There are 5 different eigenstates in  $S_2$  and three different eigenstates in  $S_1$ , repeated for each of the 3  $S_1$  sectors. The 2  $S_0$  eigenstates are one dimensional and can therefore not be depicted.

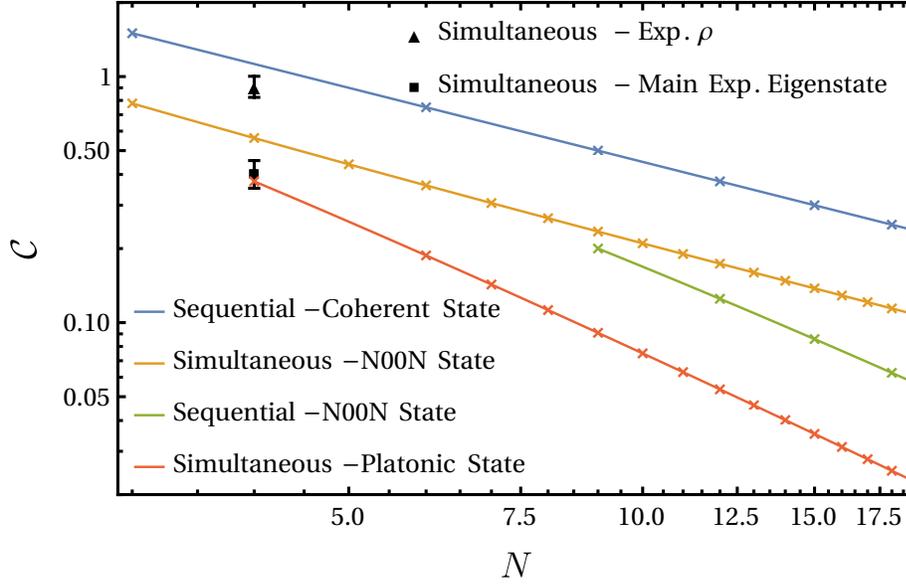


Figure 3.19: Scalar quantum Cramér-Rao for the estimation of  $SU(2)$  polarization rotation of  $\rho_{exp}$  and of its main eigenstate, both with a photon number of 4. These two s-qCRB are plotted with the performances of the multiple measurement strategies discussed in Sec. 3.1.3. We see that the main eigenstate of  $\rho_{exp}$  comes close to the optimal strategy while  $\rho_{exp}$  is too noisy.

where  $|m\rangle$  and  $\lambda_m$  are the eigenvectors of  $\rho_{exp}$ . We set both  $H$  and  $W$  to  $\mathbb{1}$ , which, as we discussed in Sec. 3.1.3, leads to the same s-qCRB as the s-qCRB calculated when  $W$  is the *natural* Cartan metric. We find the s-qCRB to be

$$\mathcal{C}_{\rho_{exp}} = 0.91 \pm 0.09. \quad (3.45)$$

As noted in Sec 3.1.3, the Tetrahedron state has a s-qCRB of 0.375. In Figure 3.19, we plot these values on a graph similar to Fig. 3.3 and compare our results with the different measurement strategies introduced in Sec. 3.1.3. We also calculate the s-qCRB of the main eigenstate of  $\rho_{exp}$ , which is  $(0.40 \pm 0.05)$ . This gives a hint of the kind of performance that could be achieved had we managed to reduce the noise levels in our experiment.

### 3.4 Background

The main source of background in this experiment happens when more photons than expected are created, either in the coherent state or in any of the SPDC processes. Even though the probability of creating high photon number states decreases exponentially with the number of photons, in this experiment, the likelihood of creating one or two photons above the five photons required for a successful detection event is significant. This type of background is common in all high photon number exper-

iments based on SPDC. Because these experiments rely on more than one pair of SPDC photons to be created, they have to operate in regimes which are inherently susceptible to an extra undesirable pair being created from time to time. In this experiment, we also use a coherent state in the state generation process, allowing for a single undesirable photon to be created. This is even more likely and makes the problem even worse.

Furthermore, the loss of photons throughout the experiment prevents us from discriminating the desirable events, where five photons are created (the tetrahedron photons and the herald), from the undesirable events, where more than five photons are created. Loss also significantly reduces the probability discrepancies between the desirable and undesirable events, increasing the levels of background.

It is important to properly model this background during the design phase of the experiment. As we will explain later, the background contributions can be made smaller at the price of an overall state creation rate reduction. The modelling can therefore help us tune the different rates to help us strike a balance between the quality of the state and its creation rate, which needs to be high enough to accumulate decent measurement statistics in a reasonable time.

### 3.4.1 Modelling the State Creation

In this section, we describe how we model the state creation process and calculate the quantum states of the two polarization modes right before entering the state tomography stage in Fig. 3.6(d). Our main aim with this model is to calculate the impact of the backgrounds coming from the accidental creation of too many photons. We neglect other experimental complications. For example, we won't worry about the potential temporal or spectral mode mismatch between photons of different sources. First, we assume that the coherent state  $|\alpha\rangle_{a_V}$ , the squeezed vacuum state  $|\eta\rangle_{a_H}$ , and the two-mode squeezed vacuum state  $|\mu\rangle_{b_H, c_H}$  are all ideal states of their kind defined by Equations 3.29, 3.30 and 3.32. We truncate the series expansions to exclude terms with more than 6 photons, which have very low probability. Second, we introduce loss. We model the loss by adding imaginary beam splitters to the optical diagram as depicted in Fig. 3.6. This adds three spatial modes  $d$ ,  $e$  and  $f$  to the experiment where the lost photons are imagined to go. We labelled the modes and input states in Fig. 3.20. This is a common method for modelling loss in quantum optics. The magnitude of the loss is encoded in the transmission and reflection properties of these imaginary beam splitters. Here, we denote the transmission probability through each beam splitters as  $t$ ,  $\tau_1$  and  $\tau_2$ . In Section 3.4.3, we will discuss how to experimentally measure  $\alpha$ ,  $\eta$ ,  $\mu$ ,  $t$ ,  $\tau_1$  and  $\tau_2$ , the free parameters of the models. We can model loss this

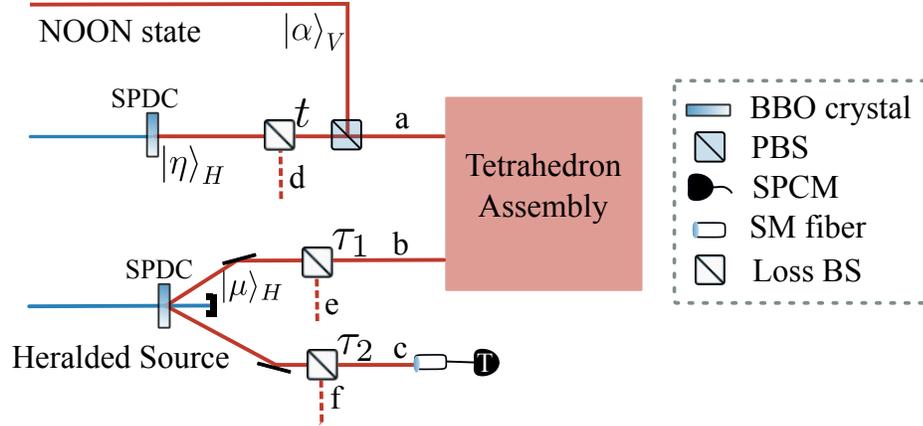


Figure 3.20: Schema of the experiment with the added virtual beam splitters modelling the loss channels of the experiments. These introduces the three new spatial modes  $d$ ,  $e$  and  $f$ , which are populated by lost photons.

way because all the linear losses commute with every stage of the experiment, which consist of linear optics transformations. For that reason, it does not matter where the loss occurs and we can concentrate it in these three points in our model. Importantly, this is only valid for linear loss, which applies identically to every photon. In Sec. 3.4.2, we will discuss a form of non-linear loss which happens in our probabilistic photon number resolving detectors in the state tomography apparatus. We will need to deal with that kind of loss differently. Finally, we apply the mode transformations for each of the optical elements used in the apparatus to the state  $|\eta\rangle_{a_H} \otimes |\alpha\rangle_{a_V} \otimes |\mu\rangle_{b_H, c_H}$ , including the virtual loss' BSs. At this stage, the model consists of a huge quantum state for the occupation of the modes  $a, b, c, d, e$  and  $f$ . The  $a$  mode is really two modes since it carries both polarizations  $a_H$  and  $a_V$ . The expression for this quantum state is too large to be shown here. We keep track of all the terms using the Wolfram Mathematica software. The ideal tetrahedron state has only 4 photons in the  $a$  mode, and 1 photon in the  $c$  mode. From this expansion, we select some of the prominent background terms. These consist of terms satisfying these same criteria but which also have an extra photon in either  $b, c, d, e$  or  $f$ . All these terms have 4 photons shared between  $a_H$  and  $a_V$  and we can express these 4 photon polarization states as density matrices in the symmetric sector  $S_2$ . The probability for each of these states to be created is a function of the parameters of the model. Once we trace over the population in the extra modes, these density matrices add incoherently. In the next section, we describe how we deal with two other terms which have 5 photons in  $a$  and 1 photon in  $c$ .

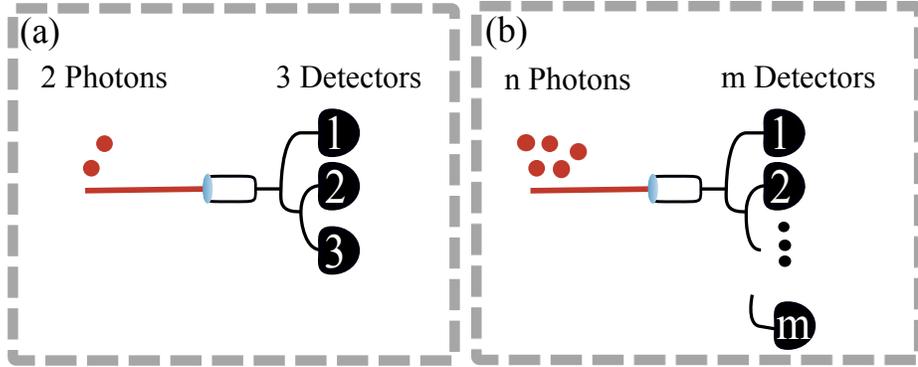


Figure 3.21: (a) Schematic description of both the transmitted and the reflected port of the PBS in the state tomography apparatus. The beam is split by optical fibre beam splitter onto three detectors. This allows the detection of a maximum of three photons per port. In this example, the two incoming photons can be both detected if they fall on different detectors or one photon can be lost if they both fall on the same detector. (b) A sketch of a general probabilistic photon number resolving detector, with  $m$  detectors.

### 3.4.2 Modelling the Loss in the PPNRDs

The above model accounts for all the losses which are identically experienced by each of the photons. In the state tomography apparatus, we use PPNRDs to mock up photon number resolving detector. They consist of networks of beam splitters with a SPCM at every output. In Figure 3.21(a), we see the exact configuration of the PPNRD we use at each of the output of the PBS, making the polarization measurement. They each have 2 BSs and 3 SPCMs. If a single photon enters one of them, it is guaranteed to be detected, whereas if four photons simultaneously enters one, at least one photon is guaranteed to be lost. This is a type of loss with a very non-linear behaviour and we need to model it differently than the linear loss.

More importantly, because of the non-linear loss, the expected detection probabilities are different for the different projectors, which have different numbers of transmitted and reflected photons. It is very important that we weight the different projectors accordingly when we perform the tomography, as we alluded to in Sec. 3.3.1.

Say we have  $n$  photons entering a general PPNRD, which has  $m$  SPCMs. This scenario is depicted in Figure 3.21(b). We introduce the array of probabilities  $\{P_i\}_{i=1}^m$ , where  $P_i$  is the probability that a single photon entering the PPNRD is detected by detector  $i$ . These probabilities need not have  $\sum_i^m P_i = 1$  if the detectors are not perfectly efficient. We classify the possible outcomes of such a scenario by an ordered list of up to  $m$  integers, representing the number of photons which could simultaneously fall on the same detectors. For example, if we have  $n = 3$  and  $m = 2$ , the two possible outcomes are  $\{3\}$  and  $\{2, 1\}$ , indicating that either all photons fall on

the same detector, or two photon fall on one detector and one on the other. Similarly, if  $n = 2$  and  $m = 3$ , the possible outcomes are  $\{2\}$  and  $\{1, 1\}$ . We call this set of outcomes  $\Lambda_m^n$ , and define it as

$$\Lambda_m^n = \left\{ \{n_i\} \mid 1 \leq i \leq m; \sum_i n_i = n; n_i \geq n_{i+1} \geq 1 \right\}. \quad (3.46)$$

For any  $\{n_i\} \in \Lambda_m^n$ , we can find the probability of that event with the formula

$$\begin{aligned} \Gamma(n, m, \{n_i\}, \{P_i\}) &= \left( \binom{n}{n_1} \binom{n-n_1}{n_2} \binom{n-n_1-n_2}{n_3} \dots \right) \\ &\times \sum_{\alpha \in S_{\{n_i\}}} \sum_{\sigma \subset \{P_i\}; |\sigma| = |\{n_i\}|} \prod_{i=1}^{|\{n_i\}|} \sigma_i^{\alpha_i}, \end{aligned} \quad (3.47)$$

where we have used  $|\cdot|$  to denote the length of a set and  $S_{\{n_i\}}$  to denote the set of all permutations of  $\{n_i\}$ . For example, if we have  $n = 2$  photons and  $m = 3$  detectors, as is the case in Fig 3.21(a), if we want to know the probability of the two photons being detected, we consider the case  $\{1, 1\}$  and find

$$\Gamma(2, 3, \{1, 1\}, \{P_i\}) = \binom{2}{1} (P_1 P_2 + P_2 P_3 + P_1 P_3) = 2P_1 P_2 + 2P_2 P_3 + 2P_1 P_3. \quad (3.48)$$

Similarly, if we want to find the probability that two photons land on the same detector (i.e., one photon is lost), we find

$$\Gamma(2, 3, \{2\}, \{P_i\}) = \sum_{i=1}^3 P_i^2 \quad (3.49)$$

Using these coefficients, we can calibrate the different projectors used in the tomography. In our experiment, we name the detectors in the transmitted port of the PBS, detectors 1, 2 and 3 and the detectors in the reflected port of the PBS, detectors 4, 5 and 6. We measure  $\{P_i\}$  by sending a coherent state to the transmitted port of the PBS and record the single rate at each detector. Then we change its polarization to send it to the reflected port and do the same. We then divide all those rates by the sum of the rates in the transmitted port to normalize them. In this way, we only account for the relative detection inefficiency between the different detectors. The overall contribution to the detection inefficiency is irrelevant to the tomography and can be attributed to linear loss in our model. We find the array

$$\{P_i\} = \{0.508318, 0.251097, 0.240585, 0.21236, 0.103793, 0.492823\} \quad (3.50)$$

for each of our 6 detectors. The various projectors in the tomography need to be weighted by the coefficients :

$$\begin{aligned}
|3_T, 1_R\rangle \langle 3_T, 1_R| &: \Gamma(3, 3, \{1, 1, 1\}, \{P_1, P_2, P_3\})\Gamma(1, 3, \{1\}, \{P_4, P_5, P_6\}) \\
&= 6P_1P_2P_3(P_4 + P_5 + P_6) = 0.14905 \\
|2_T, 2_R\rangle \langle 2_T, 2_R| &: \Gamma(2, 3, \{1, 1\}, \{P_1, P_2, P_3\})\Gamma(2, 3, \{1, 1\}, \{P_4, P_5, P_6\}) \\
&= 4(P_1P_2 + P_2P_3 + P_1P_3)(P_4P_5 + P_5P_6 + P_4P_6) = 0.220775 \\
|1_T, 3_R\rangle \langle 1_T, 3_R| &: \Gamma(1, 3, \{1\}, \{P_1, P_2, P_3\})\Gamma(3, 3, \{1, 1, 1\}, \{P_4, P_5, P_6\}) \\
&= 6P_4P_5P_6(P_1 + P_2 + P_3) = 0.0651753.
\end{aligned} \tag{3.51}$$

When trying to characterize the background of our experiment, we make a few simplifications to estimate the non-linear loss in our PPNRDs. First, we will only concern ourselves with the average loss. Since each of the projectors have different losses, it can be tricky to model. In effect, one would have different losses for each of the projectors, making the state effectively different in each of the measurement in the different bases. We simplify by averaging over all measurement bases and all projectors. We make the further assumption that our state's polarization averaged over each of the measurement basis is isotropic, meaning that each of the properly weighted projectors should on average succeed with the same probability. This allows us to simplify our model and consider the PBS where the measurement is done as a polarization independent BS. With this simplification, the entire tomography apparatus becomes a single PPNRD with 6 detectors. In the last section, our model of the linear loss in the experiment predicted the creation of the tetrahedron state and of a few other states where 4 photons entered the tomography apparatus. These should be weighted by the probability that they go through the measurement apparatus and get properly detected. This probability is  $\Gamma(4, 6, \{1, 1, 1, 1\}, \{P_i\}_{i=1}^6 / \sum P_i)$ . To this group of states, we should add the states with 5 photons entering the tomography apparatus, for which two of the photons end up on the same detector. These states will be weighted by the coefficient  $\Gamma(5, 6, \{2, 1, 1, 1\}, \{P_i\}_{i=1}^6 / \sum P_i)$ . In order to represent these states as a 4-photon density matrix, we pretend that the non-linear loss in the PPNRD is linear and choose a loss coefficient which would give  $\Gamma(5, 6, \{2, 1, 1, 1\}, \{P_i\}_{i=1}^6 / \sum P_i)$  as the probability of 1 photon out of 5 being lost. We model this loss with a BS and consider the 4 photon contribution of the transmitted state. Since either an  $H$  or a  $V$  photon can be lost in this fashion, we add two more background terms to our list. The polarization state of each of them can be once again represented by a density matrix in  $S_2$ .

### 3.4.3 Calibrating the model

To calibrate the model and find the values of  $\alpha$ ,  $\eta$ ,  $\mu$ ,  $t$ ,  $\tau_1$  and  $\tau_2$ , we performed a series of small measurements prior to the real experiment. The configuration of the optical arrangement was the same as throughout the experiment, except that some beams were blocked and some wave-plates were turned to better estimate the relevant parameters. We calibrate  $\eta$  and  $t$  together by blocking the coherent state and the heralded single photon source and by directing all the light from the squeezed in the N00N state source to the tomography apparatus. Inside the tomography apparatus, we make sure that the light is horizontally polarized so that it is transmitted at the PBS and sent to the PPNRD in that output. We then record the coincidence rate among any two pairs of detectors as well as the total single detection rate.

Using our model we calculate the expected rate of coincidence detection

$$\Omega_2 \approx R \times \eta^2 \times t^2 \times (1 - P_1^2 - P_2^2 - P_3^2), \quad (3.52)$$

where  $R$  is 80 MHz, the repetition rate of the laser. We predict the expected single detection rate to be

$$\Omega_1 \approx 2R \times \eta^2 t^2 \times (1 - P_1^2 - P_2^2 - P_3^2) + \eta^2 t^2 \times (P_1^2 - P_2^2 - P_3^2) + 2\eta^2 t(1 - t), \quad (3.53)$$

making sure to account for the fact that coincident detection also leads to two single detections. Taking the ratio of these equations, we get that

$$t = \frac{2\Omega_2/\Omega_1}{1 - \sum_i P_i^2 (1 - \Omega_2/\Omega_1)} \quad (3.54)$$

which allows us to in turn determine  $\eta$ ,

$$\eta \approx \sqrt{\frac{\Omega_2}{Rt^2 (1 - \sum_i P_i^2)}}. \quad (3.55)$$

A similar measurement can be made to calibrate  $\mu$ ,  $\tau_1$  and  $\tau_2$ . Mode  $b$  in Figure 3.20 is simply sent directly to the forward port of the tomography apparatus, with all other beams blocked. Then by measuring the coincidence rate between any of the three detectors in the PPNRD and the herald detector in mode  $c$  as well as the herald single detection rate and the total single detection rate in the PPNRD, we can easily find  $\mu$ ,  $\tau_1$  and  $\tau_2$ .

Finally, for the creation of the tetrahedron state, we typically match the coincidence rate from the coherent state to the coincidence rate from the squeezed vacuum, when these are successively sent directly to the forward port in the measurement

apparatus. This ensures that

$$\alpha \approx \sqrt{2\eta t}, \quad (3.56)$$

which is the condition necessary to the creation of the tetrahedron state in the first place.

### 3.4.4 Model Predictions and Discussion

By using the measured values of  $\eta = 0.078$ ,  $t = 0.16$ ,  $\mu = 0.14$ ,  $\tau_1 = 0.12$  and  $\tau_2 = 0.12$ , we can calculate the probability to create all the background terms that we previously gathered. These background terms are listed in Table 3.4.4. They are all the terms which were created with an extra photon, which subsequently was subsequently lost. Both an analytic and numerical expression for their probabilities are displayed as well as their normalized density matrices. These terms can be compared to the tetrahedron term displayed on the first line of the table. The dominant background term is on the sixth row of Table 3.4.4. This term occurs when an extra coherent state photon compensates for the loss of a photon from the SPDC in the N00N state source into mode  $d$ . This term is the dominant source of the reduction in the coherence between the  $|4_H, 0_V\rangle$  and  $|1_H, 3_V\rangle$  state in  $\rho_{exp}$ . It has a large and negative coherence between these basis elements to cancel the positive coherence of the tetrahedron state. Lowering  $\eta$  or increasing  $t$  would reduce the prominence of this term relative to the ideal state.

By mixing all these states, we can make a prediction of the final state being reconstructed by our tomography  $\rho_{pred}$ . This state is in  $S_2$  as losses give no mechanism to distinguish the photons. Its density matrix is

$$\rho_{pred} = \begin{pmatrix} 0.296 & -0.049 & 0.109 & 0.135 & 0.031 \\ -0.049 & 0.034 & -0.041 & 0.07 & -0.017 \\ 0.109 & -0.041 & 0.09 & -0.083 & 0.029 \\ 0.135 & 0.07 & -0.083 & 0.472 & -0.032 \\ 0.031 & -0.017 & 0.029 & -0.032 & 0.107 \end{pmatrix}. \quad (3.57)$$

We compare this state to  $\rho_{sym}$ , the re-normalized  $S_2$  portion of  $\rho_{exp}$ ,

$$\rho_{sym} = \begin{pmatrix} 0.315 \pm 0.034 & (-0.024 - 0.035i) \pm (0.024 + 0.03i) \\ (-0.024 + 0.035i) \pm (0.024 + 0.03i) & 0.045 \pm 0.015 \\ (0.061 - 0.024i) \pm (0.02 + 0.021i) & (-0.005 + 0.043i) \pm (0.017 + 0.018i) \\ (0.139 - 0.06i) \pm (0.023 + 0.026i) & (0.038 - 0.054i) \pm (0.019 + 0.02i) \\ (0.034 + 0.066i) \pm (0.013 + 0.045i) & (0.036 + 0.002i) \pm (0.016 + 0.017i) \\ \\ (0.061 + 0.024i) \pm (0.02 + 0.021i) & (0.139 + 0.06i)(\pm(0.023 + 0.026i)) \\ (-0.005 - 0.043i) \pm (0.017 + 0.018i) & (0.038 + 0.054i) \pm (0.019 + 0.02i) \\ 0.093 \pm 0.022 & (-0.072 + 0.055i) \pm (0.032 + 0.042i) \\ (-0.072 - 0.055i) \pm (0.032 + 0.042i) & 0.43 \pm 0.03 \\ (0.045 - 0.015i) \pm (0.019 + 0.015i) & (-0.077 + 0.02i) \pm (0.022 + 0.025i) \\ \\ & (0.034 - 0.066i) \pm (0.013 + 0.045i) \\ & (0.036 - 0.002i) \pm (0.016 + 0.017i) \\ & (0.045 + 0.015i) \pm (0.019 + 0.015i) \\ & (-0.077 - 0.02i) \pm (0.022 + 0.025i) \\ & 0.117 \pm 0.028 \end{pmatrix}. \quad (3.58)$$

The similitude between these two states indicates that we have a very good grasp of the imperfections of our experiment.

Even with the current losses, lowering  $\eta$  and  $\mu$  would in principle lead to a much better state, since the probabilities of the background terms always depend on higher powers of  $\eta$  or  $\mu$  than the probability of the ideal state. This would come at the cost of lowering the total tetrahedron creation rate. In our experiment, we felt that 8 days of data-taking was already substantial and did not wish to reduce the rate further.

### 3.5 Concluding Remarks

The tetrahedron state is a hard state to make. With our method, it requires the simultaneous creation of 5 photons. These are very unlikely events and happen infrequently. Many optical components were required for this experiment and they needed to be very well aligned throughout a long period of data collection. The state we have created,  $\rho_{exp}$ , is not perfect, but as we have explained, shares a lot of the qualitative features of the tetrahedron state. We think that we have also demonstrated a good

Analytical Probability	Numerical Probability	Density Matrix
$0.0830 \times \eta^3 \mu^2 t^3 \tau_1 \tau_2$	$2.96 \times 10^{-11}$	$\begin{bmatrix} 0.333 & 0. & 0. & 0.471 & 0. \\ 0. & 0. & 0. & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. \\ 0.471 & 0. & 0. & 0.667 & 0. \\ 0. & 0. & 0. & 0. & 0. \end{bmatrix}$
$0.581 \times \eta^4 \mu^2 t^4 \tau_1 \tau_2$	$2.44 \times 10^{-12}$	$\begin{bmatrix} 0.418 & -0.189 & 0.245 & -0.189 & 0.334 \\ -0.189 & 0.086 & -0.111 & 0.086 & -0.151 \\ 0.245 & -0.111 & 0.144 & -0.111 & 0.196 \\ -0.189 & 0.086 & -0.111 & 0.086 & -0.151 \\ 0.334 & -0.151 & 0.196 & -0.151 & 0.267 \end{bmatrix}$
$0.764 \times \eta^4 \mu^2 t^4 \tau_1 \tau_2$	$3.21 \times 10^{-12}$	$\begin{bmatrix} 0.016 & -0.034 & 0.04 & -0.115 & 0. \\ -0.034 & 0.073 & -0.084 & 0.244 & 0. \\ 0.04 & -0.084 & 0.098 & -0.282 & 0. \\ -0.115 & 0.244 & -0.282 & 0.813 & 0. \\ 0. & 0. & 0. & 0. & 0. \end{bmatrix}$
$0.0519 \times \eta^2 \mu^4 t^2 \tau_1^2 \tau_2^2$	$3.48 \times 10^{-13}$	$\begin{bmatrix} 0.6 & 0. & 0.49 & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. \\ 0.49 & 0. & 0.4 & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. \end{bmatrix}$
$0.147 \times \eta^4 \mu^2 t^4 \tau_1 \tau_2$	$6.18 \times 10^{-13}$	$\begin{bmatrix} 0.33 & -0.125 & 0.108 & 0. & -0.441 \\ -0.125 & 0.047 & -0.041 & 0. & 0.166 \\ 0.108 & -0.041 & 0.035 & 0. & -0.144 \\ 0. & 0. & 0. & 0. & 0. \\ -0.441 & 0.166 & -0.144 & 0. & 0.587 \end{bmatrix}$
$0.519 \times \eta^4 \mu^2 (1-t)t^3 \tau_1 \tau_2$	$1.24 \times 10^{-11}$	$\begin{bmatrix} 0.24 & -0.17 & 0.196 & -0.339 & 0. \\ -0.17 & 0.12 & -0.139 & 0.24 & 0. \\ 0.196 & -0.139 & 0.16 & -0.277 & 0. \\ -0.339 & 0.24 & -0.277 & 0.48 & 0. \\ 0. & 0. & 0. & 0. & 0. \end{bmatrix}$
$0.211 \times \eta^4 \mu^2 t^4 (1-\tau_1)\tau_2$	$7.18 \times 10^{-12}$	$\begin{bmatrix} 0.051 & -0.029 & 0.05 & -0.058 & 0.204 \\ -0.029 & 0.016 & -0.028 & 0.033 & -0.116 \\ 0.05 & -0.028 & 0.049 & -0.057 & 0.2 \\ -0.058 & 0.033 & -0.057 & 0.065 & -0.231 \\ 0.204 & -0.116 & 0.2 & -0.231 & 0.818 \end{bmatrix}$
$0.104 \times \eta^2 \mu^4 t^2 \tau_1^2 (1-\tau_2)\tau_2$	$5.62 \times 10^{-12}$	$\begin{bmatrix} 0.6 & 0. & 0.49 & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. \\ 0.49 & 0. & 0.4 & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. \end{bmatrix}$

Table 3.2: List of the different relevant states predicted to be produced incoherently by our experiment. The first row is tetrahedron term followed by background terms due to the loss of a single photon. The first column is an analytical expression for the probability of each of these states. The second column is the numerical evaluation of these probability using the measured value of the parameters of our model. The last column is the density matrix for these states. In order, the origin of loss of a single photon for these 7 background states are: in the tomography apparatus (H photon); in the tomography apparatus (V photon); in the herald mode  $c$  (with another photon present); in the  $b$  mode after the transformation in the tetrahedron assembly stage; in mode  $d$ ; in mode  $e$ ; and in mode  $f$ .

understanding of the origins of the imperfections of our experiment, which is a first step in trying to improve upon our results. Ultimately, the levels of backgrounds present in the experiment are not fundamental obstacles. With a better optical design leading to significant reductions in the photon loss, with high efficiency photon number resolving detectors, and with a more stable experiment, allowing for a longer period of data collection, some of the challenges we have experienced would be greatly reduced. We hope that this work will be regarded as a step forward in the pursuit of better measurements at the ultimate quantum limit.



## Chapter 4

# Beating Rayleigh’s Curse Using SPLICE

In this chapter, we present results previously published in 2017[104]. Most of the chapter is heavily drawn from this article.

### 4.1 Motivation

Any imaging device such as a microscope or telescope has a resolution limit, a minimum separation it can resolve between two objects or sources; this limit is typically defined by “Rayleigh’s criterion” [105], although in recent years there have been a number of high-profile techniques demonstrating that Rayleigh’s limit can be surpassed under particular sets of conditions [27, 28, 106]. As an electromagnetic wave, light is characterized by both an amplitude and a phase. Traditional imaging systems use lenses or mirrors to refocus this wave and project an image of the source onto a screen or camera, where the intensity (or rate of photon arrivals) is recorded at each position. (We refer to all such techniques as “image-plane counting” or IPC). Although the phase of the wave at the position of the optics plays a central role during the focusing, any information about the phase *in the image plane* is discarded. When light passes through finite-sized optical elements, diffraction smears out the spatial distribution of photons so that point sources map (via the point spread function or PSF) onto finite-sized spots at the image-plane. Thus, our ability to resolve the point sources is inhibited when their separation in the image plane,  $\delta$ , is comparable to or less than the width  $\sigma$  of the PSF.

The typical response to diffraction limits has been to build larger (or higher numerical-aperture) optics, thereby making the PSF sharper/narrower. In recent years, techniques have been developed in specific cases that address these limits in

more novel ways [27, 28, 94, 106–112]. Despite their success, these techniques require careful control of the source of illumination, which is not always possible in every imaging application (e.g. astronomy). In order to beat the diffraction limit for *fixed, mutually incoherent* sources, a paradigm shift arising from the realisation that there is a huge amount of information available in the phase discarded by IPC may prove revolutionary.

With Yang et al.[113], Sheng et al.[114], Paúr et al.[115], our paper [104], which we present here, shares the honour of being one of the first experimental study of this new paradigm. All these works were more or less published during the same period of time and were all done independently. Since then, many new meaningful results have sprouted from this line of research. At the time of the writing of this thesis, a very comprehensive list of meaningful references was maintained on Mankei Tsang's website[116]. He also wrote a review paper on the topic[117] in 2019. These references should help the interested readers navigate the progress of the field.

## 4.2 Theory

Inspired by ideas in quantum information and quantum metrology [3, 5, 59, 118, 119], Tsang et al [120] showed that whereas in IPC the Fisher Information,  $\mathcal{I}$ [29], vanishes quadratically with the separation  $\delta$  between two equal-intensity incoherent point sources of weak thermal light with gaussian PSF, it remains undiminished when the full electromagnetic field is considered. Later, these results were extended to more general types of sources [121, 122].  $\mathcal{I}$  is related to the performance of a statistical estimator by :

$$\text{Var}(\delta_{est}) \geq \frac{1}{\mathcal{I}} \left( 1 + \frac{\partial(\text{bias})}{\partial\delta_{actual}} \right), \quad (4.1)$$

where  $\delta_{est}$  is some estimator of  $\delta_{actual}$  and  $\text{bias} \equiv \langle \delta_{est} \rangle - \delta_{actual}$ [29].

In the case of an unbiased estimator, this limit is the CRB ( $\frac{1}{\mathcal{I}}$ ). The vanishing of  $\mathcal{I}$  as  $\delta \rightarrow 0$  suggests that for closely separated sources, the variance in an IPC-based estimate of  $\delta$  is cursed to diverge. That it is independent of  $\delta$  for the full field, on the other hand, appears to suggest that this divergence can be averted by using phase as well as intensity information.

One natural way to do this would be to use SPAtial mode DEmultiplexing (SPADE)[120, 123], in which incoming light is decomposed into its Hermite-Gauss (HG)[124] components and the amplitude of each is measured. This HG basis is centred exactly between the two PSF, and the width of the  $0^{th}$  mode( $TEM_{00}$ ) matches the width of the gaussian PSF. It can be shown that the full set of HG amplitudes contains

the same  $\mathcal{I}$  as the full EM field. A reduced version called binary SPADE prescribes discriminating only between the  $TEM_{00}$  mode and the sum of all other modes. For small  $\delta$ , only one other mode acquires significant amplitude in any case, so the  $\mathcal{I}$  available to binary SPADE becomes essentially equal to the full Fisher information. The method can be understood as follows: the projection always succeeds ( $P_{00} = 1$ ) when the two point-sources are overlapped ( $\delta = 0$ ), but has a failure probability  $1 - P_{00}$  which grows quadratically with  $\delta$ . Knowing the  $TEM_{00}$  component as a proportion of all HG amplitudes (i.e.  $P_{00}$  and  $1 - P_{00}$ ) allows one to deduce  $\delta$ . All the results above and the experimental work below deal with gaussian PSFs, which are of interest because they are frequently used as approximations for the Airy rings produced by circular apertures. However, recent results[125] have shown that in the small separation regime, a small number of suitable spatial projections of the electromagnetic field capturing the full Fisher information can also be generated for any PSF.

In this proof-of-principle experiment, we continue with gaussian PSFs. The spatial wavefunction of the EM field of the two sources is given by

$$\psi_{1/2}(x, y) = \exp\left(-\frac{(x \pm \delta/2)^2 + y^2}{4\sigma^2}\right) \quad (4.2)$$

Experimentally, merely capturing the  $TEM_{00}$  component (say, by coupling into a single mode fibre) without a normalization factor (which allows us to deduce  $1 - P_{00}$ ) provides no advantage over IPC. Practically speaking, the crucial information comes from a projection onto some mode *orthogonal* to  $TEM_{00}$  in order to estimate  $1 - P_{00}$ . While a mode such as  $TEM_{10}$  would contain all the information (for a separation in the x-direction in that example), the same scaling can be obtained by projecting onto any spatially antisymmetric field mode. As a proof of principle, we have designed and implemented an experimentally convenient method, SPLICE (Super-resolved Position Localisation by Inversion of Coherence along an Edge), which instead carries out one single technically straightforward projection onto the mode function

$$\psi_{\perp}(x, y) = \exp\left(-\frac{x^2 + y^2}{4\sigma^2}\right) \text{sign}(x). \quad (4.3)$$

This function is constructed such that its inner product with  $TEM_{00}$  vanishes. The probability that such a projection succeeds is:

$$P_{\perp} = \frac{1}{2} (|\langle \psi_1 | \psi_{\perp} \rangle|^2 + |\langle \psi_2 | \psi_{\perp} \rangle|^2) = e^{-2\Delta} \text{erf}^2 \sqrt{\Delta} \quad (4.4)$$

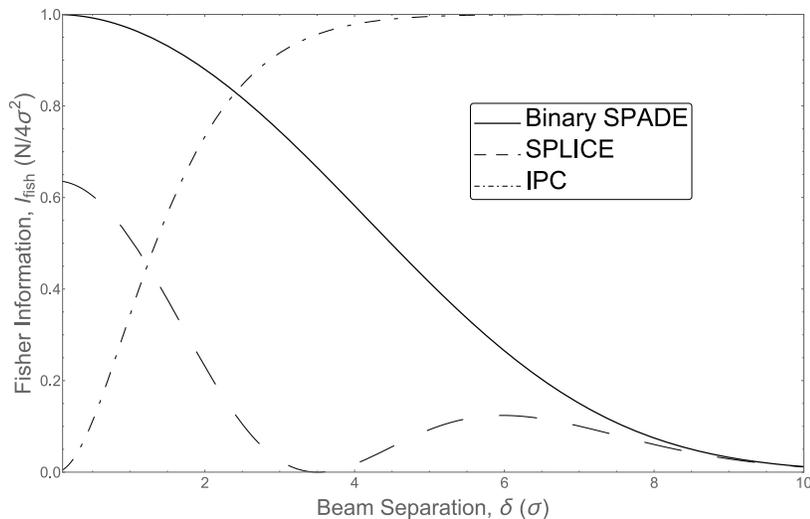


Figure 4.1: Theory plot of Fisher information for IPC, SPLICE and Binary SPADE vs beam separation  $\delta$ , normalized to units of  $N/4\sigma^2$  and  $\sigma$  respectively.

where  $\Delta = \delta^2/32\sigma^2$ , and  $\delta$  is the separation between point sources on the image plane, and  $\psi_{1/2}$  is the field from each source.

The per-photon Fisher information can be written as

$$\mathcal{I} = \frac{\left(e^{-\Delta}\sqrt{\pi\Delta}\text{erf}\sqrt{\Delta} - e^{-2\Delta}\right)^2}{2\pi\sigma^2} + \frac{\left(e^{-\Delta}\sqrt{\pi\Delta}\text{erf}^2\left(\sqrt{\Delta}\right) - e^{-2\Delta}\text{erf}\sqrt{\Delta}\right)^2}{2\pi\sigma^2\left(e^{2\Delta} - \text{erf}^2\sqrt{\Delta}\right)}, \quad (4.5)$$

where the first term comes from  $P_{\perp}$  and the second from  $1 - P_{\perp}$ . Crucially, as  $\Delta \rightarrow 0$ ,  $1 - P_{\perp}$  vanishes, meaning that an experimentally simple scheme for projecting only onto  $\psi_{\perp}$  does as well as a more complicated scheme which could measure multiple projections simultaneously. In Fig. 4.1., we plot the Fisher information for SPLICE in comparison with other methods. It is easy to see that it remains non-zero as  $\delta \rightarrow 0$ , evading Rayleigh's curse, and extracting nearly 2/3 of the total information available to full SPADE using an experimentally simple technique. More sophisticated methods relying on waveguides or cavities could be designed to approach 100% of the optimal  $\mathcal{I}$ .

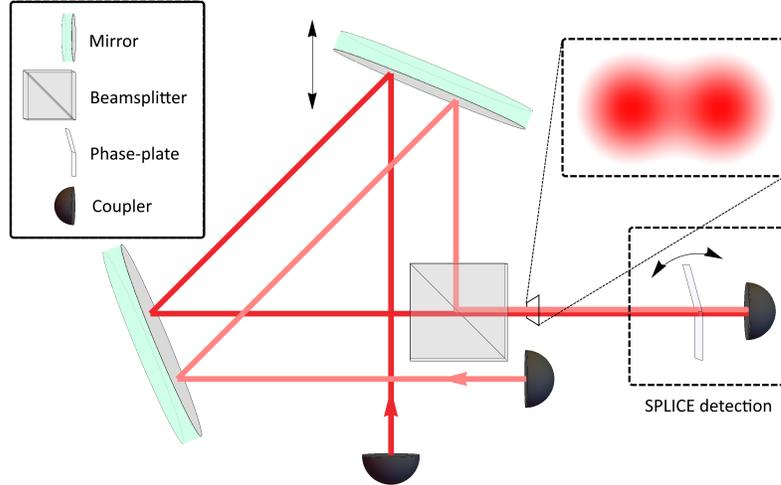


Figure 4.2: Shown is the experimental apparatus. In the lower right-hand box is a representation of SPLICE, the measurement scheme tested in this experiment. In the upper right-hand box is a sketch of the spatial profile of the electromagnetic field before the measurement. The rest of the figure depicts the device used to simulate the two light sources, which can be displaced around their centroid by the displacement of the top mirror.

### 4.3 Experiment and Results

In order to experimentally demonstrate improved performance over IPC, we used two mutually incoherent collimated  $TEM_{00}$  Gaussian beams in place of distant point sources and an imaging optical setup. The beams were directed through a Sagnac-like beam displacer shown in Fig. 4.2. By moving a mirror on a motorized translation stage as shown, we precisely control the separation  $\delta$  between the otherwise parallel beams. The separation is induced symmetrically, such that the geometrical centroid  $(x_0, y_0)$  remains static.

At zero separation, the beams are overlapped and are both coupled into single-mode  $TEM_{00}$  fibre (coupling efficiencies are 90% and 85% respectively). Collimation of the beams is such that their waists are closely matched immediately before the fibre coupler ( $\sigma = 434 \pm 3\mu m$  and  $420 \pm 7\mu m$ ) in order to emulate Gaussian point spread functions of distant sources. The projection onto  $\psi_{\perp}$  (the SPLICE measurement) is achieved by inserting a phase plate immediately in front of the coupler such that when  $\delta = 0$ , a semi-circular cross-section of the beams undergoes a  $\pi$ -phase shift whereas the other half experiences none. The phase plate consists of two transparent glass flat microscope cover slips, connected along one sharp rectangular edge. They are mounted such that one glass slip tilts relative to the other by pivoting along that edge. We then position this contraption such that the connected edge of the glass slips bisects the circular beam cross-sections when both beams are overlapped (i.e.

$\delta = 0$ ). Thus, we can impart different phase-shifts onto opposite halves of the beams by tilting one glass slip relative to the other (which we do, in order to minimize coupling into an otherwise well-aligned coupler). We typically achieve an extinction of  $\geq 99\%$ .

The light source used to create the two mutually incoherent beams is an 805-nm heralded single-photon source which relies on type-I spontaneous parametric down-conversion (SPDC) in a 2mm-thick BBO crystal. The crystal is pumped by 402.5 nm light obtained from a frequency-doubled 100-fs Ti:Sapph laser. One photon from the SPDC pair is used to herald the presence of a signal photon as a means of rejecting spurious background light and detector dark counts (our accidental coincidences average  $2 \pm 1$  counts/sec). Our SPDC source has a very low probability of producing more than one photon per coherence time ( $\approx 150$  fs). This resembles the regime investigated in the original theory proposal[120]. Furthermore, the low intensity of the source allows us to directly compare our experimental performance with the quantum limit shown in Fig. 4.1. It must be noted that the use of heralded single photons is not necessary to this scheme, which is independent of the photon statistics of the point sources. In order to emulate two point sources, the other photon is split at a 50/50 fibre-splitter and out-coupled to free-space. The two resulting beams are incoherent; they have splitter-to-coupler distances that differ by 5 cm whereas the SPDC photons are filtered to  $\Delta\lambda = 3$  nm (i.e. coherence length  $\approx 10$   $\mu$ m). ND filters were used to reduce the intensity imbalance between the beams to  $\approx (3 \pm 3)\%$ .

To compare the performance of our method (SPLICE) with a more traditional imaging setup relying on IPC, we replaced the phase plate with a 200  $\mu$ m slit that served as the image plane, coupling all the light transmitted through the slit into a multimode fibre. Scanning the slit, we were able to perform one-dimensional IPC.

With SPLICE, the separation of the incoherent beams was scanned, with the detectors counting for 1 second at each step. Two sets of SPLICE scans were performed, one at coarse intervals of  $\delta$  (spanning  $-1.96$  mm to  $+1.94$  mm, in steps of 0.1 mm). Another scan at finer intervals ( $-0.56$  mm  $\leq \delta \leq +0.44$  mm in steps of 0.04 mm) was performed to provide more data points in the region of low separation, where SPLICE provides an advantage. Data from nine repetitions of the coarse scan and fifteen of the fine scan were recorded. The raw data can be observed in Fig. 4.3.

Whereas the ideal functional form for the resulting count rate vs separation  $\delta$  is proportional to equation 4.4, we add a constant  $\gamma$  to account for residual background counts:

$$\text{SPLICE counts} = \alpha \exp\left(-\frac{(\delta - \delta_0)^2}{16\sigma_f^2}\right) \text{erf}^2\left(\frac{\delta}{4\sqrt{2}\sigma_f}\right) + \gamma \quad (4.6)$$

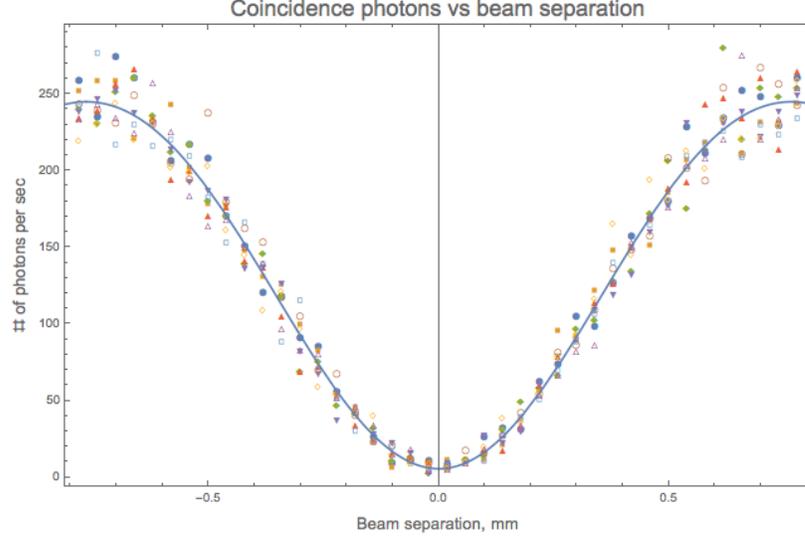


Figure 4.3: Raw data plot for SPLICE coarse scans. Dots are experimental photon coincidence counts plotted versus actual beam separation  $\delta$ . Solid overlay is a fit to equation 4.6.

A calibration curve was obtained from a least squares fit of this function to data from a longer run (2 sec count time instead of 1 sec for each  $\delta$ ). This calibration curve can be seen in Fig. 4.4. From a fit to the singles counts (counts which are not conditioned on the simultaneous detection of a heralding photon), the beam waist  $\sigma_f(0.46 \text{ mm})$  and  $\delta_0$  were extracted. Then, the normalisation  $\alpha(1206)$  was extracted from a fit to coincident counts. The background  $\gamma$  was fixed to an average of multiple values detected at a separation of  $\delta_0(2.73 \text{ mm})$ . This step might be thought of as being analogous to characterizing one's imaging optics before use. One might then use the curve as a "lookup table" from which  $\delta$  can be estimated from count rates. We performed such a lookup with the remainder of our data. The resulting estimates for  $\delta$  are plotted versus their actual values(as reported by motorized translation stage controllers) in Fig. 4.5(a).

The traditional image plane counting data were acquired using the configuration described above, scanning the  $200 \mu\text{m}$  slit between  $-1 \text{ mm}$  and  $+1 \text{ mm}$  of the centroid of the two beams, counting for 4 seconds at each step. Again, we repeated this for various separations  $\delta$  and in turn repeated the whole scan several times. As before, a set of coarse scans ( $-0.04 \text{ mm} \leq \delta \leq 1.56 \text{ mm}$  in steps of  $0.08 \text{ mm}$ , 16 repetitions) and a set of fine scans ( $-0.52 \text{ mm} \leq \delta \leq 0.44 \text{ mm}$  in steps of  $0.04 \text{ mm}$ , 17 repetitions) were performed. Estimation of  $\delta$  in this IPC comparison was done by least-squares fitting the resulting image plane intensity profile to

$$\text{IPC counts} = \alpha [\exp_- + \exp_+ + \gamma] \quad (4.7)$$

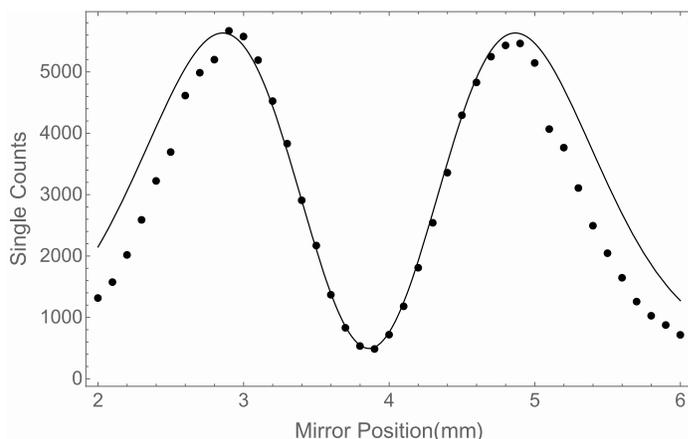


Figure 4.4: Calibration data plot for SPLICE scans. Dots are experimental photon single counts plotted versus actual beam separation  $\delta$ . Solid overlay is a fit to Eq. 4.6. From this fit, the value of the beamwaist is taken to be  $0.46\text{mm}$ . This value is larger than the measured value of the waist of the two beams ( $0.420\text{mm}$  and  $0.434\text{mm}$ ). Furthermore, the fit here overshoots the data for extreme values of  $\delta$ . We do not have an adequate explanation for this behavior. However, the important fact is that the fit is in good agreement with the data for central values of  $\delta$ , which is the important regime. This allowed us to use this parameter as a calibration to extract a separation from our SPLICE measurement. The other parameters for the calibration were taken from similar data but using coincidence counts.

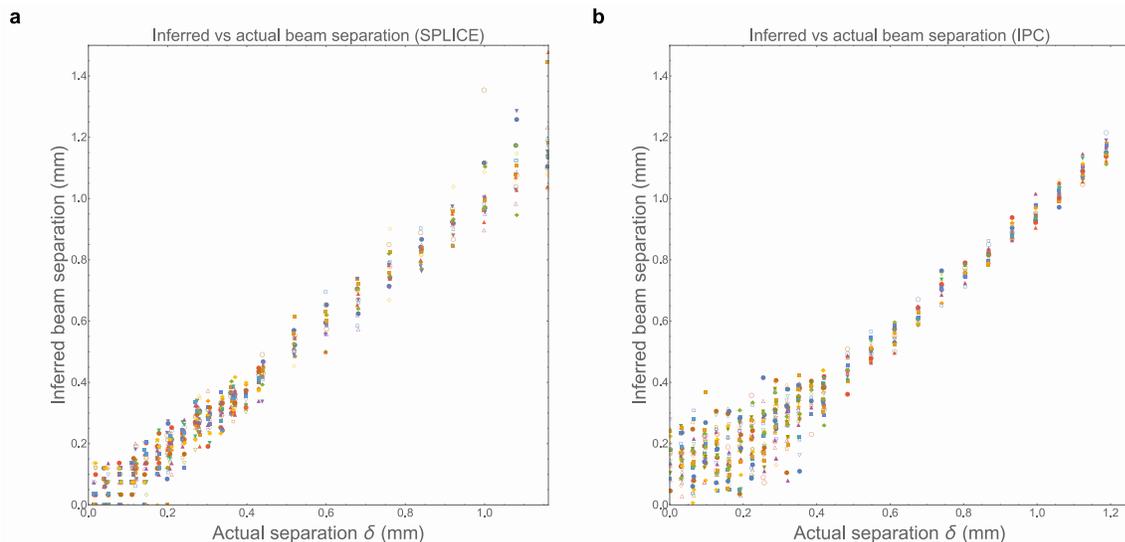


Figure 4.5: Inferred separation vs known actual separation for a) SPLICE (from “lookup” on calibration curve) and b) IPC. Note that the spread in the IPC estimates grows drastically as  $\delta \rightarrow 0$ , while the spread for SPLICE remains essentially constant.

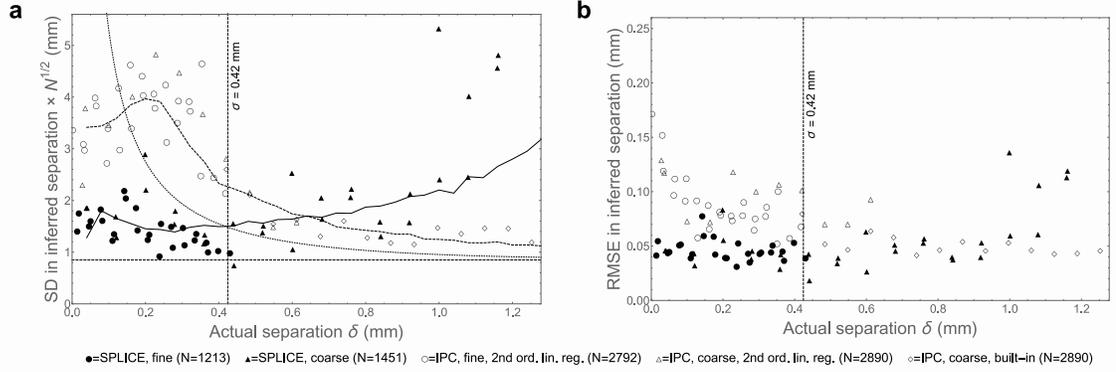


Figure 4.6: (a) Renormalized Standard Deviation (SD) and (b) Un-normalized Root Mean-Square Error (RMSE) in the estimated separation plotted as functions of actual separation for both IPC and SPLICE. We plot the computed SD multiplied by  $\sqrt{N}$  to compensate for the scaling of the uncertainty with the size of the data set. The solid and dashed curves are the corresponding Monte Carlo simulations. The dotted curve is the CRB for IPC and the dashed horizontal line represents the absolute fundamental limit of  $2\sigma/\sqrt{N}$ . The RMSE (unlike SD) is not similarly rescaled. It allows us to gauge absolute error relative to the known value of the parameter being estimated so that biases are accounted for. Note that two methods were used in the fitting of IPC data to equation 4.7; for small  $\delta$  ( $< 0.65$ mm), equation 4.7 was expanded to 2nd order and linear regression was performed whereas for large  $\delta$  ( $> 0.4$ mm), a nonlinear fitting routine built into Mathematica was used.

where  $\exp_{\pm} = \exp[-(x \pm \delta/2)^2/2\sigma^2]$ . Again, a calibration waist  $\sigma$  and background  $\gamma$  were obtained beforehand, leaving the scale  $\alpha$  and separation  $\delta$  as the only fitting parameters. In practice, the fitting procedure used to fit IPC data for small  $\delta$  was different from the one used to treat data for large  $\delta$ . For the latter, we simply used built-in numerical algorithms in Mathematica and NumPy. For small  $\delta$ 's however, the routines exhibited convergence and stability issues, forcing us to Taylor expand equation 4.7 to 2nd order in  $\delta$  and manually invert the resulting  $2 \times 2$  design matrix. The resulting estimated separations are plotted against actual separations in Fig. 4.5(b). As is immediately apparent, for separations below about 0.25 mm (approximately  $0.6\sigma$ ), the spread of the IPC data begins to grow, while that of the SPLICE data remains essentially constant.

Two key metrics for the performance of either method are the standard deviation or SD (i.e. "spread") and root-mean-square error (RMSE) of the estimated beam separation. The SD measures the *precision* of a dataset but not necessarily its accuracy, while the RMSE is sensitive to the accuracy since it quantifies the error relative to a known actual value and not simply the reported result. In Fig. 4.6, SD and RMSE are plotted versus known actual separations.

In order to ensure a reasonably even-footed comparison between IPC and SPLICE, the spreads in inferred separation plotted in Fig. 4.6(a) are scaled by  $\sqrt{N}$ . For IPC,

$N$  is simply the total photons that comprise an “image” on the image plane, which in our case is actually a set of photon counts, one at each position of the 200  $\mu\text{m}$  slit. For SPLICE, during a calibration run, we estimate  $N$  by counting at our detector over a 1 second window while both beams are centred (i.e.  $\delta = 0$ ) on the coupler into  $TEM_{00}$  fibre with the phase plate removed. Since our source intensity is stable, this gives us an estimate of the number of incident photons for subsequent measurements when  $\delta \neq 0$ .

The RMSE plotted in Fig. 4.6(b) is *not* similarly normalized because in addition to possible systematics, the inferred separation is biased relative to the actual separation when  $\delta$  is small (see Fig. 4.7 and Sec. 4.4). A priori, there is no reason to suspect either bias or systematics to scale as  $\sqrt{N}$ . Despite *not* normalizing and despite using approximately twice as many photons, the IPC method performs noticeably *worse* than SPLICE when  $\delta < 0.6$  mm.

## 4.4 Discussion

The attentive reader will note that while the spread is greater for IPC, it does not diverge as  $\delta \rightarrow 0$ . In fact, it would be implausible for the uncertainty on  $\delta$  to ever exceed  $\sigma$  (as is clear from inspection of Fig. 4.6(b) at small  $\delta$ ). The apparent discrepancy with the vanishing of the Fisher information can be understood by recognizing that the practically implemented IPC estimator is not unbiased. To better understand the bounds on the advantage that one can expect of SPLICE over IPC, we return to equation 4.1. Clearly, one needs to know the bias to evaluate the RHS. For SPLICE, a potential source of bias is the lookup procedure. If, for example, a less-than-perfect visibility results in a calibration curve that does not vanish at  $\delta = 0$ , then one might obtain “unphysical” datapoints that fall *under* the minima of the calibration curve, thereby resulting in a bias when a lookup is attempted. In our case, this is negligible since our visibility exceeds 99%. The CRB is therefore just the reciprocal of  $\mathcal{I}$ , implying a  $1/\sqrt{N}$  scaling in the spread of  $\delta_{est}$ .

With IPC, the least-squares estimate of  $\delta$  is heavily biased at small  $\delta$ . An intuitive way to understand this is to note that since the problem being addressed is the resolving of two *equal intensity* sources, the  $+\delta$  and  $-\delta$  cases are physically indistinguishable; therefore, what is really being estimated is the absolute value  $|\delta|$ . But as long as the spread in the estimated  $\delta$  is non-zero, the *mean* estimated  $|\delta|$  is never zero. Fig. 4.7 shows a plot of mean inferred  $\delta$  (averaged across all our datasets) vs actual  $\delta$ . Overlaid is a theory curve for IPC, which takes into account an expected bias at small  $\delta$ . The vanishing of the slope of this curve at low separations means that

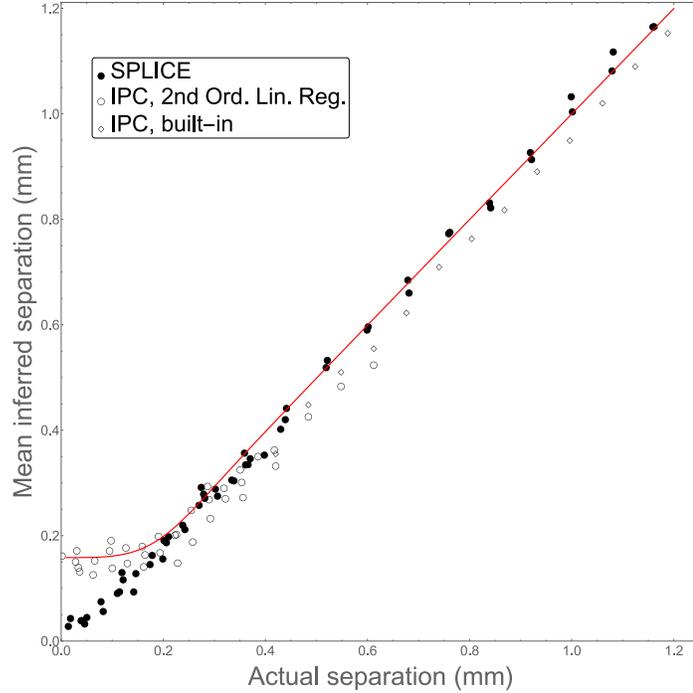


Figure 4.7: Mean estimated  $\delta$  for IPC and SPLICE plotted against known actual  $\delta$ . Two methods were used in the fitting of IPC data to equation 4.7; for small  $\delta$  ( $< 0.65mm$ ), equation 4.7 was expanded to 2nd order and linear regression was performed whereas for large  $\delta$  ( $> 0.4mm$ ), a nonlinear fitting routine built into Mathematica was used.

any attempt to invert it, generating an unbiased estimator, will introduce a diverging uncertainty.

We can notice in the form of the CRB for biased estimator presented in Eq. 4.1, that if the bias term for IPC falls to  $-1$  sufficiently quickly when  $\delta \rightarrow 0$ , the RHS of the inequality can tend to a finite value as both  $\mathcal{I} \rightarrow 0$  and  $\delta \rightarrow 0$ .

For IPC at small separations each image was fitted to a Taylor expansion of the detection probability  $p_i$  (the usual sum of two Gaussians) to 2nd order:

$$p_i \approx \frac{A}{\sigma\sqrt{2\pi}} \exp\left[-\frac{x_i^2}{2\sigma^2}\right] + \frac{A\delta^2}{8\sigma^5\sqrt{2\pi}} \exp\left[-\frac{x_i^2}{2\sigma^2}\right] (x_i^2 - \sigma^2)$$

Subscripts  $i$  were added in anticipation of an image consisting of many pixels at various values of some axis  $x$ . Performing a linear regression of a set of photon detection rates  $p_i$  yields parameters  $A$  and  $A\delta^2$ . Notice that the design matrix,  $M$ , in this case contains only  $x_i$ 's and  $\sigma$  and so is independent of photon number  $N$ . If we now assume that the noise at each pixel location is mutually independent, then:

$$\begin{bmatrix} \Delta A \\ \Delta (A\delta^2) \end{bmatrix} = \sqrt{\sum_j o_{ij}^2 (\Delta p_j)^2}$$

where  $o_{ij} = \left[ (M^T M)^{-1} M^T \right]_{ij}$ . Supposing that our *only* source of noise is Poissonian in nature, then  $\Delta p_j \sim \sqrt{p_j}$  so that  $\Delta A$  and  $\Delta (A\delta^2)$  both  $\sim \sqrt{N}$ .

Now elementary error propagation gives:

$$\Delta(\delta^2) = \sqrt{\left( \frac{\Delta(A\delta^2)}{A} \right)^2 + \left( \frac{A\delta^2}{\Delta A} \right)^2}$$

which in the small  $\delta$  limit reduces to

$$\Delta(\delta^2) \approx \frac{\Delta(A\delta^2)}{A} \sim 1/\sqrt{N}$$

Thus we expect the estimate of  $\delta^2$  from this method to have a spread that scales approximately as  $1/\sqrt{N}$ .

If there is a sufficiently large number of pixels in our image, the central limit theorem imposes a Gaussian distribution on  $\delta^2$ , with width  $s = \Delta(\delta^2) = \gamma/\sqrt{N}$  where  $\gamma$  is some constant of proportionality. Although at first glance negative values of  $\delta^2$  appear problematic, we can avoid having to censor parts of our data where this is the case by noting that they have a natural physical interpretation if we also allow  $\sigma \rightarrow -\sigma$  since the quadratic term is paired with an odd  $\sigma^5$  term. We can therefore compute the moments of the distribution relevant to the mean and spread of our estimate of  $\delta$ :

$$\begin{aligned} \langle \delta \rangle &= \frac{1}{s} \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} \sqrt{|x|} \exp \left[ -\frac{(x - \delta_{actual}^2)^2}{2s^2} \right] dx \\ \langle \delta^2 \rangle &= \frac{1}{s} \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} |x| \exp \left[ -\frac{(x - \delta_{actual}^2)^2}{2s^2} \right] dx \end{aligned}$$

In the limit where  $\delta_{actual} \rightarrow 0$ , we find the scalings:  $\langle \delta \rangle \sim \sqrt{s} \sim N^{-1/4}$  and  $\langle \delta^2 \rangle \sim s \sim N^{-1/2}$ . More crucially,  $\langle \delta \rangle$  can be shown to approach a constant value sufficiently quickly as  $\delta_{actual}$  goes to 0 for the CRB to converge to a finite value.  $\langle \delta \rangle$  is plotted in figure 4.7.

Note that the emergence of a bias in our estimate isn't specific to our treatment of the negative tail of  $\delta^2$ ; the same bias and scalings can be obtained even if we had opted for the lazier approach of censoring parts of our data that produce negative  $\delta^2$  values (tantamount to simply "chopping" rather than "folding" that tail of the distribution). Rather, the bias is more generally a consequence of performing the regression on  $\delta^2$  instead of  $\delta$ .

The bias vanishes if the two sources have unequal intensities. The breaking of this

symmetry introduces a term in  $p_i$  that is linear in  $\delta$ . If this term is much larger than the quadratic ( $\delta^2$ ) term, we can use  $\delta$  as a fit parameter instead, thereby obtaining an unbiased estimator. We explored scenarios with different intensities in [126], but focused on cases where the centroid of the beams and the intensity imbalances were unknown. This reintroduces the bias and brings a plethora of other challenges.

Thus while SPLICE does not offer an *infinite* advantage over IPC as a naive analysis might have us believe, it *does* nevertheless offer a substantial improvement in the absolute error and the scaling with photon number, while simultaneously eliminating the problem of bias.

## 4.5 Concluding Remarks

In summary, we have developed and demonstrated a simple technique that surpasses traditional imaging in its ability to resolve two closely spaced point-sources. For  $\delta < 0.2 \text{ mm} (0.47\sigma)$ , the average spread in the measured separation was approximately twice the quantum limit. Nearing zero separation, SPLICE has outperformed IPC by reducing the normalized standard deviation by a factor of 2 and the unnormalized total error by a factor of 3 despite the higher photon number used in IPC. Furthermore, unlike existing superresolution methods, ours requires no exotic illumination with particular coherence/quantum properties and is applicable to classical incoherent sources. Crucially, as a proof of principle, this technique highlights that the fundamental limits on the precision with which one can estimate the separation between two point sources ( $2\sigma/\sqrt{N}$ ) are independent of the separation itself. In traditional imaging techniques discarding the phase information (IPC), for separation below Rayleigh's criterion ( $\delta \lesssim \sigma$ ), the standard deviation in the measurement of separation goes as  $2\sqrt{2}\sigma^2/(\delta\sqrt{N})$  for the best unbiased estimator or as approximately  $\alpha\sigma/N^{1/4}$ , where  $\alpha$  is a numerical factor of the order of unity, when the bias becomes dominant. We expect that SPLICE and other related techniques that do not discard the phase information will be developed in the future for a broad range of imaging applications.



## Chapter 5

# Conclusion

In this thesis, I reported on three projects that I completed during my PhD.

In Chapter 2, we presented a new fundamental lower bound on the precision of quantum multi-parameter estimation measurements. It has the advantage of being a scalar bound independent of the quantities to be measured, making it an easy quantity to optimize for finding optimal measurement schemes. We are hopeful that this bound will pave the way for studying many multi-parameter estimation problems. Its application to the characterization of  $SU(d)$  unitaries, which we have considered, could be immediately applied to the estimation of unitaries acting identically on collections of qudits, or to the characterization of a general multiport interferometer, which could find application in boson sampling and other quantum information task with photons.

In Chapter 3, we reported on the creation in the laboratory of the tetrahedron state, the first of the platonic-solid states. It is a four-photon polarization-entangled state which is optimal for the characterization of  $SU(2)$  polarization rotation. This proof-of-principle experiment is another step in the direction of quantum schemes being used for real measurements. As we argued, it is our belief that the the quality of our state could be greatly increased with an effort to reduce the photon loss through the experiment. In order to scale up and create platonic-solid states with higher photon numbers, it may be necessary to use to a more naturally scalable infrastructure. Approaches like ours, using SPDC, lead to a state creation rate exponentially decreasing with the photon number. The exploration of alternative systems in which to create these platonic-solid states could be a very interesting line of future research.

In Chapter 4, we presented a new technique named SPLICE, which dramatically improves the precision of the measurement of the separation between two equal-intensity, incoherent point sources of light below Rayleigh's criterion. This is done by taking advantage of the information contained in the phase of the electromagnetic

field at the image plane of the imaging apparatus in a novel way. We demonstrated the efficacy of this technique in a proof-of-principle experiment. A next step is to relax the assumption that the point sources have equal intensities and to study the scenario where their intensities are both unequal and unknown. This is a much more interesting problem, closer to more realistic measurement settings. We have already published some theory work on the topic demonstrating that a technique like SPLICE with minor modifications retains part the advantage observed here. Our group is currently in the process of building a new experiment to test this proposal.

The field of quantum optics is changing rapidly. With the recent enthusiasm for quantum computing and the large financial investments that followed, a fraction of the activities that used to be done in an academic setting is now being tackled by private companies. As a PhD student, it can be challenging to navigate this climate. Many of the capabilities of the smaller academic groups are being outpaced by those of these large companies and collaborations. On the other hand, there are many interesting scientific endeavours left to explore which are outside their sphere of interest.

Quantum metrology is in essence a fundamentally scientific discipline. Measurements are at the heart of science and the improvement of their precision has always been a vector of scientific progress, with each improvement giving us a window to peer into a previously inaccessible part of the universe. I believe quantum metrology to be a noble and promising field and I am happy to have played a minor role in its development during my PhD.

# Bibliography

- <sup>1</sup>C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).
- <sup>2</sup>A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, 2nd (North Holland, Amsterdam, 2003).
- <sup>3</sup>V. Giovannetti, S. Lloyd, and L. Maccone, “Quantum-enhanced measurements: beating the standard quantum limit”, *Science* **306**, 1330–1336 (2004).
- <sup>4</sup>V. Giovannetti, S. Lloyd, and L. Maccone, “Quantum metrology”, *Physical review letters* **96**, 010401 (2006).
- <sup>5</sup>V. Giovannetti, S. Lloyd, and L. Maccone, “Advances in quantum metrology”, *Nature photonics* **5**, 222 (2011).
- <sup>6</sup>C. M. Caves, “Quantum-mechanical noise in an interferometer”, *Physical Review D* **23**, 1693 (1981).
- <sup>7</sup>M. Xiao, L.-A. Wu, and H. J. Kimble, “Precision measurement beyond the shot-noise limit”, [Phys. Rev. Lett.](#) **59**, 278–281 (1987).
- <sup>8</sup>M. G. A. Paris, “Quantum estimation for quantum technology”, [Int. J. Quantum Inform.](#) **7**, 125–137 (2009).
- <sup>9</sup>B. C. Sanders, “Quantum dynamics of the nonlinear rotator and the effects of continual spin measurement”, *Physical Review A* **40**, 2417 (1989).
- <sup>10</sup>A. N. Boto, P. Kok, D. S. Abrams, S. L. Braunstein, C. P. Williams, and J. P. Dowling, “Quantum interferometric optical lithography: exploiting entanglement to beat the diffraction limit”, *Physical Review Letters* **85**, 2733 (2000).
- <sup>11</sup>H. Lee, P. Kok, and J. P. Dowling, “A quantum rosetta stone for interferometry”, *Journal of Modern Optics* **49**, 2325–2338 (2002).
- <sup>12</sup>J. P. Dowling, “Quantum optical metrology—the lowdown on high-n00n states”, *Contemporary physics* **49**, 125–143 (2008).
- <sup>13</sup>J. P. Dowling, “Correlated input-port, matter-wave interferometer: quantum-noise limits to the atom-laser gyroscope”, [Phys. Rev. A](#) **57**, 4736–4746 (1998).
- <sup>14</sup>T. L. S. Collaboration., “A gravitational wave observatory operating beyond the quantum shot-noise limit”, *Nature Phys* **7**, 962–965 (2011).
- <sup>15</sup>J. Aasi, J. Abadie, B. Abbott, R. Abbott, T. Abbott, M. Abernathy, C. Adams, T. Adams, P. Addesso, R. Adhikari, et al., “Enhanced sensitivity of the ligo gravitational wave detector by using squeezed states of light”, *Nature Photonics* **7**, 613–619 (2013).

- <sup>16</sup>M. A. Taylor, J. Janousek, V. Daria, J. Knittel, B. Hage, H.-A. Bachor, and W. P. Bowen, “Biological measurement beyond the quantum limit”, [Nature Photonics](#) **7**, 229 (2013).
- <sup>17</sup>M. A. Taylor and W. P. Bowen, “Quantum metrology and its application in biology”, [Physics Reports](#) **615**, 1–59 (2016).
- <sup>18</sup>P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, “Quantum enhanced multiple phase estimation”, [Phys. Rev. Lett.](#) **111**, 070403 (2013).
- <sup>19</sup>M. Szczykulska, T. Baumgratz, and A. Datta, “Multi-parameter quantum metrology”, [Adv. Phys. X](#) **1**, 621–639 (2016).
- <sup>20</sup>J. S. Sidhu and P. Kok, “Geometric perspective on quantum parameter estimation”, [AVS Quantum Sci.](#) **2**, 014701 (2020).
- <sup>21</sup>F. Albarelli, M. Barbieri, M. G. Genoni, and I. Gianani, “A perspective on multiparameter quantum metrology: from theoretical tools to applications in quantum imaging”, [Phys. Lett. A](#) **384**, 126311 (2020).
- <sup>22</sup>E. Polino, M. Valeri, N. Spagnolo, and F. Sciarrino, “Photonic quantum metrology”, [AVS Quantum Sci.](#) **2**, 024703 (2020).
- <sup>23</sup>R. Demkowicz-Dobrzański, W. Górecki, and M. Guţă, “Multi-parameter estimation beyond quantum Fisher information”, [J. Phys. A: Math. Theor.](#) **53**, 363001 (2020).
- <sup>24</sup>A. Z. Goldberg, A. B. Klimov, G. Leuchs, and L. L. Sánchez-Soto, “Rotation sensing at the ultimate limit”, [J. Phys. Photonics](#) **3**, 022008 (2021).
- <sup>25</sup>F. Bouchard, P. de la Hoz, G. Björk, R. Boyd, M. Grassl, Z. Hradil, E. Karimi, A. Klimov, G. Leuchs, J. Řeháček, et al., “Quantum metrology at the limit with extremal majorana constellations”, [Optica](#) **4**, 1429–1432 (2017).
- <sup>26</sup>M. Tsang, R. Nair, and X. Lu, “Quantum theory of superresolution for two incoherent optical point sources”, arXiv preprint arXiv:1511.00552 (accepted for publication in PRX) (2015).
- <sup>27</sup>E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, “Imaging intracellular fluorescent proteins at nanometer resolution”, [Science](#) **313**, 1642–1645 (2006).
- <sup>28</sup>S. T. Hess, T. P. Girirajan, and M. D. Mason, “Ultra-high resolution imaging by fluorescence photoactivation localization microscopy”, [Biophysical journal](#) **91**, 4258–4272 (2006).
- <sup>29</sup>G. Casella and R. L. Berger, *Statistical inference*, Vol. 2 (Duxbury Pacific Grove, CA, 2002).
- <sup>30</sup>S. M. Kay, *Fundamentals of statistical signal processing: estimation theory* (Prentice-Hall, Inc., 1993).
- <sup>31</sup>A. C. Davison, *Statistical models*, Vol. 11 (Cambridge university press, 2003).
- <sup>32</sup>C. W. Helstrom, “Quantum detection and estimation theory”, [Journal of Statistical Physics](#) **1**, 231–252 (1969).
- <sup>33</sup>S. L. Braunstein and C. M. Caves, “Statistical distance and the geometry of quantum states”, [Phys. Rev. Lett.](#) **72**, 3439–3443 (1994).
- <sup>34</sup>D. Petz and C. Ghinea, “Introduction to quantum fisher information”, [Quantum Probability and Related Topics](#) **1**, 261–281 (2011).

- <sup>35</sup>J. S. Sidhu and P. Kok, “Geometric perspective on quantum parameter estimation”, *AVS Quantum Sci.* **2**, 014701 (2020).
- <sup>36</sup>L. Pezzè, M. A. Ciampini, N. Spagnolo, P. C. Humphreys, A. Datta, I. A. Walmsley, M. Barbieri, F. Sciarrino, and A. Smerzi, “Optimal measurements for simultaneous quantum estimation of multiple phases”, *Physical Review Letters* **119**, 130504 (2017).
- <sup>37</sup>R. Demkowicz-Dobrzański, W. Górecki, and M. Guţă, “Multi-parameter estimation beyond quantum fisher information”, *Journal of Physics A: Mathematical and Theoretical* **53**, 363001 (2020).
- <sup>38</sup>A. Carollo, B. Spagnolo, A. A. Dubkov, and D. Valenti, “On quantumness in multi-parameter quantum estimation”, *Journal of Statistical Mechanics: Theory and Experiment* **2019**, 094010 (2019).
- <sup>39</sup>J. S. Sidhu, Y. Ouyang, E. T. Campbell, and P. Kok, “Tight bounds on the simultaneous estimation of incompatible parameters”, *Phys. Rev. X* **11**, 011028 (2021).
- <sup>40</sup>M. Tsang, “The holevo cramér-rao bound is at most thrice the helstrom version”, arXiv preprint arXiv:1911.08359 (2019).
- <sup>41</sup>F. Albarelli, M. Tsang, and A. Datta, “Upper bounds on the holevo cramér-rao bound for multi-parameter quantum parametric and semiparametric estimation”, arXiv preprint arXiv:1911.11036 (2019).
- <sup>42</sup>A. Z. Goldberg, L. L. Sánchez-Soto, and H. Ferretti, “Intrinsic sensitivity limits for multiparameter quantum metrology”, *Phys. Rev. Lett.* **127**, 110501 (2021).
- <sup>43</sup>R. M. Wilcox, “Exponential operators and parameter differentiation in quantum physics”, *Journal of Mathematical Physics* **8**, 962–982 (1967).
- <sup>44</sup>L. Scharf, *Statistical signal processing. detection, estimation, and time analysis* (Addison Wesley, New York, 1991).
- <sup>45</sup>G. Chiribella, G. D’Ariano, P. Perinotti, and M. F. Sacchi, “Efficient use of quantum resources for the transmission of a reference frame”, *Physical review letters* **93**, 180503 (2004).
- <sup>46</sup>J. F. Cornwell, *Group theory in physics*, Vol. II (Academic, 1984).
- <sup>47</sup>F. T. Hioe and J. H. Eberly, “N-level coherence vector and higher conservation laws in quantum optics and quantum mechanics”, *Physical Review Letters* **47**, 838 (1981).
- <sup>48</sup>P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, “Quantum enhanced multiple phase estimation”, *Phys. Rev. Lett.* **111**, 070403 (2013).
- <sup>49</sup>L.-Z. Liu, Y.-Z. Zhang, Z.-D. Li, R. Zhang, X.-F. Yin, Y.-Y. Fei, L. Li, N.-L. Liu, F. Xu, Y.-A. Chen, et al., “Distributed quantum phase estimation with entangled photons”, *Nature Photonics* **15**, 137–142 (2021).
- <sup>50</sup>P. Kolenderski and R. Demkowicz-Dobrzanski, “Optimal state for keeping reference frames aligned and the platonic solids”, *Physical Review A* **78**, 052333 (2008).
- <sup>51</sup>A. Z. Goldberg and D. F. James, “Quantum-limited euler angle measurements using anticonherent states”, *Physical Review A* **98**, 032113 (2018).
- <sup>52</sup>R. Adamson, *Characterization of quantum states of light* (University of Toronto (Canada), 2009).

- <sup>53</sup>E. Majorana, “Atomi orientati in campo magnetico variabile”, *Il Nuovo Cimento* (1924-1942) **9**, 43–50 (1932).
- <sup>54</sup>G. S. Agarwal, “Relation between atomic coherent-state representation, state multipoles, and generalized phase-space distributions”, *Physical Review A* **24**, 2889 (1981).
- <sup>55</sup>J. P. Dowling, G. S. Agarwal, and W. P. Schleich, “Wigner distribution of a general angular-momentum state: applications to a collection of two-level atoms”, *Phys. Rev. A* **49**, 4101–4109 (1994).
- <sup>56</sup>A. B. Klimov, J. L. Romero, and H. De Guise, “Generalized su (2) covariant wigner functions and some of their applications”, *Journal of Physics A: Mathematical and Theoretical* **50**, 323001 (2017).
- <sup>57</sup>F. Arecchi, E. Courtens, R. Gilmore, and H. Thomas, “Atomic coherent states in quantum optics”, *Physical Review A* **6**, 2211 (1972).
- <sup>58</sup>M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg, “Super-resolving phase measurements with a multiphoton entangled state”, *Nature* **429**, 161–164 (2004).
- <sup>59</sup>P. Walther, J.-W. Pan, M. Aspelmeyer, R. Ursin, S. Gasparoni, and A. Zeilinger, “De broglie wavelength of a non-local four-photon state”, *Nature* **429**, 158–161 (2004).
- <sup>60</sup>A. Z. Goldberg, A. B. Klimov, M. Grassl, G. Leuchs, and L. L. Sánchez-Soto, “Extremal quantum states”, *AVS Quantum Science* **2**, 044701 (2020).
- <sup>61</sup>A. Z. Goldberg, A. B. Klimov, G. Leuchs, and L. L. Sánchez-Soto, “Rotation sensing at the ultimate limit”, *Journal of Physics: Photonics* **3**, 022008 (2021).
- <sup>62</sup>D. Baguette and J. Martin, “Anticoherence measures for pure spin states”, *Phys. Rev. A* **96**, 032304 (2017).
- <sup>63</sup>M. Grassl, *Extremal polarization states*, <http://polarization.markus-grassl.de/index.html> (visited on 10/01/2021).
- <sup>64</sup>R. H. Hardin and N. J. A. Sloane, “McLaren’s improved snub cube and other new spherical designs in three dimensions”, *Discrete & Computational Geometry* **15**, 429–441 (1996).
- <sup>65</sup>J. Crann, R. Pereira, and D. W. Kribs, “Spherical designs and anticoherent spin states”, *Journal of Physics A: Mathematical and Theoretical* **43**, 255307 (2010).
- <sup>66</sup>E. Bannai and M. Tagami, “A note on anticoherent spin states”, *Journal of Physics A: Mathematical and Theoretical* **44**, 342002 (2011).
- <sup>67</sup>G. Björk, M. Grassl, P. de la Hoz, G. Leuchs, and L. L. Sánchez-Soto, “Stars of the quantum universe: extremal constellations on the poincaré sphere”, *Physica Scripta* **90**, 108008 (2015).
- <sup>68</sup>J. Thomson, “Jj thomson, philos. mag. 7, 237 (1904).”, *Philos. Mag.* **7**, 237 (1904).
- <sup>69</sup>P. M. L. Tammes, “On the origin of number and arrangement of the places of exit on the surface of pollen-grains”, *Recueil des travaux botaniques néerlandais* **27**, 1–84 (1930).
- <sup>70</sup>L. L. Whyte, “Unique arrangements of points on a sphere”, *The American Mathematical Monthly* **59**, 606–611 (1952).
- <sup>71</sup>J. H. Conway, R. H. Hardin, and N. J. A. Sloane, “Packing lines, planes, etc.: packings in grassmannian spaces”, *Experimental Mathematics* **5**, 139–159 (1996).

- <sup>72</sup>E. B. Saff and A. B. J. Kuijlaars, “Distributing many points on a sphere”, *The Mathematical Intelligencer* **19**, 5–11 (1997).
- <sup>73</sup>J. Zimba, ““Anticoherent” spin states via the Majorana representation”, *EJTP* **3**, 143–156 (2006).
- <sup>74</sup>D. Baguette, F. Damanet, O. Giraud, and J. Martin, “Anticoherence of spin states with point-group symmetries”, *Phys. Rev. A* **92**, 052333 (2015).
- <sup>75</sup>O. Giraud, P. Braun, and D. Braun, “Quantifying quantumness and the quest for queens of quantum”, *New Journal of Physics* **12**, 063005 (2010).
- <sup>76</sup>D. Klyshko, “Multiphoton interference and polarization effects”, *Physics Letters A* **163**, 349–355 (1992).
- <sup>77</sup>P. de la Hoz, G. Björk, A. B. Klimov, G. Leuchs, and L. L. Sánchez-Soto, “Unpolarized states and hidden polarization”, *Phys. Rev. A* **90**, 043826 (2014).
- <sup>78</sup>G. Björk, A. B. Klimov, P. de la Hoz, M. Grassl, G. Leuchs, and L. L. Sánchez-Soto, “Extremal quantum states and their Majorana constellations”, *Physical Review A* **92**, 031801 (2015).
- <sup>79</sup>C. Chryssomalakos and H. Hernández-Coronado, “Optimal quantum roto-sensors”, *Phys. Rev. A* **95**, 052125 (2017).
- <sup>80</sup>J. Martin, S. Weigert, and O. Giraud, “Optimal Detection of Rotations about Unknown Axes by Coherent and Anticoherent States”, *Quantum* **4**, 285 (2020).
- <sup>81</sup>A. Stummer, *Alan stummer’s archives*, (2018) <https://www.physics.utoronto.ca/~astummer/> (visited on 10/01/2021).
- <sup>82</sup>C. Gerry and P. Knight, *Introductory quantum optics* (Cambridge University Press, 2004).
- <sup>83</sup>A. Aspect, J. Dalibard, and G. Roger, “Experimental test of bell’s inequalities using time-varying analyzers”, *Physical review letters* **49**, 1804 (1982).
- <sup>84</sup>M. Giustina, M. A. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-Å. Larsson, C. Abellán, et al., “Significant-loophole-free test of bell’s theorem with entangled photons”, *Physical review letters* **115**, 250401 (2015).
- <sup>85</sup>L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, et al., “Strong loophole-free test of local realism”, *Physical review letters* **115**, 250402 (2015).
- <sup>86</sup>C. Weedbrook, S. Pirandola, R. Garcia-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, “Gaussian quantum information”, *Reviews of Modern Physics* **84**, 621 (2012).
- <sup>87</sup>R. W. Boyd, edited by R. W. Boyd, Fourth Edition (Academic Press, 2020).
- <sup>88</sup>L.-A. Wu, H. J. Kimble, J. L. Hall, and H. Wu, “Generation of squeezed states by parametric down conversion”, *Phys. Rev. Lett.* **57**, 2520–2523 (1986).
- <sup>89</sup>E. Meyer-Scott, N. Montaut, J. Tiedau, L. Sansoni, H. Herrmann, T. J. Bartley, and C. Silberhorn, “Limits on the heralding efficiencies and spectral purities of spectrally filtered single photons from photon-pair sources”, *Physical Review A* **95**, 061803 (2017).
- <sup>90</sup>T. Guerreiro, A. Martin, B. Sanguinetti, N. Bruno, H. Zbinden, and R. Thew, “High efficiency coupling of photon pairs in practice”, *Optics express* **21**, 27641–27651 (2013).

- <sup>91</sup>R. S. Bennink, “Optimal collinear gaussian beams for spontaneous parametric down-conversion”, *Physical Review A* **81**, 053805 (2010).
- <sup>92</sup>H. F. Hofmann and T. Ono, “High-photon-number path entanglement in the interference of spontaneously down-converted photon pairs with coherent laser light”, *Physical Review A* **76**, 031806 (2007).
- <sup>93</sup>L. Pezzé and A. Smerzi, “Mach-zehnder interferometry at the heisenberg limit with coherent and squeezed-vacuum light”, *Physical review letters* **100**, 073601 (2008).
- <sup>94</sup>L. A. Rozema, J. D. Bateman, D. H. Mahler, R. Okamoto, A. Feizpour, A. Hayat, and A. M. Steinberg, “Scalable spatial superresolution using entangled photons”, *Phys. Rev. Lett.* **112**, 223602 (2014).
- <sup>95</sup>L. A. Rozema, D. H. Mahler, R. Blume-Kohout, and A. M. Steinberg, “Optimizing the choice of spin-squeezed states for detecting and characterizing quantum processes”, *Physical Review X* **4**, 041025 (2014).
- <sup>96</sup>I. Afek, O. Ambar, and Y. Silberberg, “High-noon states by mixing quantum and classical light”, *Science* **328**, 879–881 (2010).
- <sup>97</sup>R. J. Glauber, “The quantum theory of optical coherence”, *Physical Review* **130**, 2529–2539 (1963).
- <sup>98</sup>R. J. Glauber, “Coherent and incoherent states of the radiation field”, *Physical Review* **131**, 2766–2788 (1963).
- <sup>99</sup>C.-K. Hong, Z.-Y. Ou, and L. Mandel, “Measurement of subpicosecond time intervals between two photons by interference”, *Physical review letters* **59**, 2044 (1987).
- <sup>100</sup>R. Adamson, L. Shalm, M. Mitchell, and A. Steinberg, “Multiparticle state tomography: hidden differences”, *Physical review letters* **98**, 043601 (2007).
- <sup>101</sup>R. Adamson, P. Turner, M. Mitchell, and A. Steinberg, “Detecting hidden differences via permutation symmetries”, *Physical Review A* **78**, 033832 (2008).
- <sup>102</sup>L. Shalm, R. Adamson, and A. Steinberg, “Squeezing and over-squeezing of triphotons”, *Nature* **457**, 67–70 (2009).
- <sup>103</sup>D. F. James, P. G. Kwiat, W. J. Munro, and A. G. White, “On the measurement of qubits”, in *Asymptotic theory of quantum statistical inference: selected papers* (World Scientific, 2005), pp. 509–538.
- <sup>104</sup>W.-K. Tham, H. Ferretti, and A. M. Steinberg, “Beating rayleigh’s curse by imaging using phase information”, *Phys. Rev. Lett.* **118**, 070801 (2017).
- <sup>105</sup>Lord Rayleigh, “XXXI. Investigations in optics, with special reference to the spectroscope”, *Philosophical Magazine Series 5* **8**, 261–274 (1879).
- <sup>106</sup>S. W. Hell and J. Wichmann, “Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy”, *Optics letters* **19**, 780–782 (1994).
- <sup>107</sup>P. R. Hemmer and T. Zapata, “The universal scaling laws that determine the achievable resolution in different schemes for super-resolution imaging”, *Journal of Optics* **14**, 083002 (2012).

- <sup>108</sup>V. Giovannetti, S. Lloyd, L. Maccone, and J. H. Shapiro, “Sub-rayleigh-diffraction-bound quantum imaging”, *Physical Review A* **79**, 013827 (2009).
- <sup>109</sup>M. Tsang, “Quantum imaging beyond the diffraction limit by optical centroid measurements”, *Physical review letters* **102**, 253601 (2009).
- <sup>110</sup>H. Shin, K. W. C. Chan, H. J. Chang, and R. W. Boyd, “Quantum spatial superresolution by optical centroid measurements”, *Phys. Rev. Lett.* **107**, 083603 (2011).
- <sup>111</sup>F. Tamburini, G. Anzolin, G. Umbricco, A. Bianchini, and C. Barbieri, “Overcoming the rayleigh criterion limit with optical vortices”, *Phys. Rev. Lett.* **97**, 163903 (2006).
- <sup>112</sup>O. Schwartz, J. M. Levitt, R. Tenne, S. Itzhakov, Z. Deutsch, and D. Oron, “Superresolution microscopy with quantum emitters”, *Nano letters* **13**, 5832–5836 (2013).
- <sup>113</sup>F. Yang, A. Tashchilina, E. S. Moiseev, C. Simon, and A. I. Lvovsky, “Far-field linear optical superresolution via heterodyne detection in a higher-order local oscillator mode”, *Optica* **3**, 1148–1152 (2016).
- <sup>114</sup>Z. S. Tang, K. Durak, and A. Ling, “Fault-tolerant and finite-error localization for point emitters within the diffraction limit”, *Optics express* **24**, 22004–22012 (2016).
- <sup>115</sup>M. Paúr, B. Stoklasa, Z. Hradil, L. L. Sánchez-Soto, and J. Rehacek, “Achieving the ultimate optical resolution”, *Optica* **3**, 1144–1147 (2016).
- <sup>116</sup>M. Tsang, *Papers on quantum-inspired superresolution*, (2021) <https://blog.nus.edu.sg/mankei/superresolution/> (visited on 10/01/2021).
- <sup>117</sup>M. Tsang, “Resolving starlight: a quantum perspective”, *Contemporary Physics* **60**, 279–298 (2019).
- <sup>118</sup>M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg, “Super-resolving phase measurements with a multiphoton entangled state”, *Nature* **429**, 161–164 (2004).
- <sup>119</sup>T. Nagata, R. Okamoto, J. L. O’Brien, K. Sasaki, and S. Takeuchi, “Beating the standard quantum limit with four-entangled photons”, *Science* **316**, 726–729 (2007).
- <sup>120</sup>M. Tsang, R. Nair, and X.-M. Lu, “Quantum theory of superresolution for two incoherent optical point sources”, *Phys. Rev. X* **6**, 031033 (2016).
- <sup>121</sup>R. Nair and M. Tsang, “Far-field superresolution of thermal electromagnetic sources at the quantum limit”, *Phys. Rev. Lett.* **117**, 190801 (2016).
- <sup>122</sup>C. Lupo and S. Pirandola, “Ultimate precision bound of quantum and subwavelength imaging”, *Phys. Rev. Lett.* **117**, 190802 (2016).
- <sup>123</sup>D. Richardson, J. Fini, and L. Nelson, “Space-division multiplexing in optical fibres”, *Nature Photonics* **7**, 354–362 (2013).
- <sup>124</sup>A. E. Siegman, *Lasers*, edited by C. University Science Books Mill Valley (1986), p. 798.
- <sup>125</sup>J. Rehacek, M. Paúr, B. Stoklasa, Z. Hradil, and L. L. Sánchez-Soto, “Optimal measurements for resolution beyond the rayleigh limit”, *Opt. Lett.* **42**, 231–234 (2017).
- <sup>126</sup>K. A. Bonsma-Fisher, W.-K. Tham, H. Ferretti, and A. M. Steinberg, “Realistic sub-rayleigh imaging with phase-sensitive measurements”, *New Journal of Physics* **21**, 093010 (2019).



## Appendix A

# Clebsch-Gordan Matrix

We give the Clebsch-Gordan coefficient matrix, which is used to change the state of 4 photons' polarization from a basis natural for the Hilbert Space decomposition  $(S_{1/2})^{\otimes 4}$  to a basis natural for the Hilbert Space decomposition  $(S_2) \oplus (S_1)^{\oplus 3} \oplus (S_0)^{\oplus 2}$ . The basis elements for both bases are displayed in order below. We use the H/V notation for the first basis and the spin  $|j^{(\gamma)}, m\rangle$  notation for the latter. The index  $\gamma$  ( $1 \leq \gamma \leq d_j$ ) is used to denote the multiplicities of the spin sectors when necessary.

$ HHHH\rangle$	$ 2, 2\rangle$
$ HHHV\rangle$	$ 2, 1\rangle$
$ HHVH\rangle$	$ 2, 0\rangle$
$ HHVV\rangle$	$ 2, -1\rangle$
$ HVHH\rangle$	$ 2, -2\rangle$
$ HVHV\rangle$	$ 1^{(1)}, 1\rangle$
$ HVVH\rangle$	$ 1^{(1)}, 0\rangle$
$ HVVV\rangle$	$ 1^{(1)}, -1\rangle$
$ VHHH\rangle$	$ 1^{(2)}, 1\rangle$
$ VHHV\rangle$	$ 1^{(2)}, 0\rangle$
$ VHVH\rangle$	$ 1^{(2)}, -1\rangle$
$ VHVV\rangle$	$ 1^{(3)}, 1\rangle$
$ VVHH\rangle$	$ 1^{(3)}, 0\rangle$
$ VVHV\rangle$	$ 1^{(3)}, -1\rangle$
$ VVVH\rangle$	$ 0^{(1)}, 0\rangle$
$ VVVV\rangle$	$ 0^{(2)}, 0\rangle$



## Appendix B

### Density Matrix $\rho_{exp}$

We give the full reconstructed density matrix  $\rho_{exp}$ . To fit it on the page, we only give the non-empty spin sectors. Refer to Fig. 3.12(a) to see the position of those sectors in the full  $16 \times 16$  density matrix. We only display one copy of the repeated sectors  $S_1$  and  $S_0$ .

$$\rho_{exp}^{(S_2)} = \begin{pmatrix} 0.27 \pm 0.03 & (-0.02 - 0.03i) \pm (0.02 + 0.03i) \\ (-0.02 + 0.03i) \pm (0.02 + 0.03i) & 0.039 \pm 0.013 \\ (0.056 - 0.006i) \pm (0.017 + 0.019i) & (-0.010 + 0.036i) \pm (0.014 + 0.016i) \\ (0.131 \pm 0.019) & (0.044 - 0.036i) \pm (0.017 + 0.018i) \\ (-0.00 + 0.06i) \pm (0.02 + 0.03i) & (0.028 + 0.013i) \pm (0.014 + 0.016i) \end{pmatrix}$$

$$\begin{pmatrix} (0.056 + 0.006i) \pm (0.017 + 0.019i) & 0.131 \pm 0.019 \\ (-0.010 - 0.036i) \pm (0.014 + 0.016i) & (0.0440 + 0.036i) \pm (0.017 + 0.018i) \\ 0.08 \pm 0.02 & (-0.06 + 0.06i) \pm (0.03 + 0.04i) \\ (-0.06 - 0.06i) \pm (0.03 + 0.04i) & 0.37 \pm 0.03 \\ (0.041 - 0.002i) \pm (0.016 + 0.014i) & (-0.068 + 0.01i) \pm (0.019 + 0.02i) \end{pmatrix}$$

$$\begin{pmatrix} (-0.00 - 0.06i) \pm (0.02 + 0.03i) \\ (0.028 - 0.013i) \pm (0.014 + 0.016i) \\ (0.041 + 0.002i) \pm (0.016 + 0.014i) \\ (-0.068 - 0.01i) \pm (0.019 + 0.02i) \\ 0.10 \pm 0.02 \end{pmatrix}$$

$$\rho_{exp}^{(S_1)} = \begin{pmatrix} 0.0153 \pm 0.006 & (-0.016 - 0.007i) \pm (0.005 + 0.004i) \\ (-0.016 + 0.007i) \pm (0.005 + 0.004i) & 0.022 \pm 0.007 \\ (0.004 + 0.003i) \pm (0.004 + 0.004i) & (-0.002 - 0.008i) \pm (0.004 + 0.004i) \\ & (0.004 - 0.003i) \pm (0.004 + 0.004i) \\ & (-0.002 + 0.008i) \pm (0.004 + 0.004i) \\ & 0.007 \pm 0.005 \end{pmatrix}$$

$$\rho_{exp}^{(S_0)} = (6.76 \times 10^{-11}) \quad (\text{B.1})$$