

EFFICIENT CHARACTERIZATION AND MANIPULATION OF STATES OF LIGHT

ARMAN MANSOURI

THESIS SUBMITTED TO THE UNIVERSITY OF OTTAWA IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTORATE IN PHILOSOPHY, PHYSICS

DEPARTMENT OF PHYSICS
FACULTY OF SCIENCE
UNIVERSITY OF OTTAWA

UNDER THE SUPERVISION OF PROF. JEFF S. LUNDEEN



©ARMAN MANSOURI, OTTAWA, CANADA, 2026

Declaration of Authorship

I hereby certify that this thesis is entirely my own original work except where otherwise indicated. I am aware of the University's regulations concerning plagiarism, including those concerning consequent disciplinary actions. Any use of the works of any other author, in any form, is properly acknowledged at their point of use.

Abstract

Quantum science and technology have seen rapid progress in recent decades, finding applications in computation, information processing, communication, and sensing. Accurate characterization and manipulation of quantum states is essential to the successful implementation of these applications. Moreover, optimal quantum state estimation is closely related to fundamental questions, such as limits on how fast information can be extracted about quantum systems. In this thesis, we present several contributions to efficiently estimating and transforming quantum states. We discuss optimal quantum state estimation, both in the regime of few available quantum systems as well as the asymptotic regime. We numerically investigate how assuming a pure-state can affect the performance of tomographic methods, and where adaptive quantum state tomography is in fact advantageous. While collective measurements have been formally considered as optimal solutions to various problems in quantum information, their implementation has been largely neglected due to experimental challenges. We present a proof-of-concept experiment realizing Massar and Popescu's optimal collective measurement for quantum state estimation, utilizing two-photon interference and polarization-dependent loss to achieve non-maximally entangled state projections. We demonstrate that collective strategies can outperform the best separable measurements. Furthermore, we introduce and experimentally implement a fast and automated optical polarization compensation scheme using liquid crystal variable retarders and rotating quarter-wave plate tomography. Desired accuracy thresholds can be achieved within a few compensation steps. Finally, we use a matrix-model approach to investigate the realization of spatial unitary transformations via sequences of thin phase gratings and free-space propagation. We numerically demonstrate that the grating parameters and propagation distances can be optimized to achieve various two-beam and three-beam unitary transformations, such as Pauli gates and the discrete Fourier transform. Through experimental and numerical approaches, this thesis offers new perspectives on methods for quantum state estimation and unitary transformation.

Résumé

La science et la technologie quantiques ont connu d'importantes avancées au cours des dernières décennies, avec des applications dans le calcul, le traitement de l'information, la communication et la détection. La mise en œuvre réussie de ces applications repose sur une caractérisation et une manipulation précises des états quantiques. De plus, il existe une étroite relation entre l'estimation optimale de l'état quantique et des questions fondamentales, telles que les limites sur la vitesse d'extraction de l'information à propos des systèmes

quantiques. Dans cette thèse, nous présentons plusieurs contributions à l'estimation et à la transformation efficaces des états quantiques. Nous considérons l'estimation optimale des états quantiques dans le régime des quelques systèmes quantiques disponibles ainsi que dans le régime asymptotique. Nous mettons en évidence numériquement comment l'hypothèse d'un état pur affecte la performance des méthodes tomographiques, et où la tomographie adaptative est en fait avantageuse. Bien que les mesures collectives aient été formellement considérées comme des solutions optimales à divers problèmes liés à l'information quantique, leur mise en œuvre a été largement négligée en raison de défis expérimentaux. Nous présentons une expérience de preuve de concept réalisant la mesure collective optimale de Massar et Popescu pour l'estimation de l'état quantique, en utilisant l'interférence à deux photons et la perte dépendante de la polarisation pour obtenir des projections d'états non maximalelement intriquées. Nous démontrons que les stratégies collectives peuvent surpasser les meilleures mesures séparables. De plus, nous introduisons et mettons en œuvre expérimentalement un schéma de compensation de polarisation rapide et automatisé utilisant des retardateurs variables à cristaux liquides et la tomographie par une lame quart d'onde en rotation. En quelques étapes de compensation, on peut atteindre les seuils de précision souhaités. Enfin, nous utilisons une approche matricielle pour étudier la réalisation de transformations unitaires spatiales via des séquences de réseaux de phases minces et la propagation en espace libre. Nous illustrons numériquement que les paramètres du réseau et les distances de propagation peuvent être optimisés pour réaliser diverses transformations unitaires à deux et trois faisceaux, telles que les portes de Pauli et la transformée de Fourier discrète. Par des approches expérimentales et numériques, cette thèse offre de nouvelles perspectives sur les méthodes d'estimation d'état quantique et de transformation unitaire.

Related Publications

Chapters 3 and 4 of this thesis are closely adapted from the two manuscripts listed below. They are respectively:

- Reproduced from A. Mansouri, K. M. Jordan, R. A. Abrahao, and J. S. Lundeen, “Experimental demonstration of a multi-particle collective measurement for optimal quantum state estimation,” [Optics Express](#), vol. 34, p. 5718, Feb. 2026. Reproduced with permission.
- Reproduced from N. Braband, A. Mansouri, R. Fazili, S. Czischek, and J. S. Lundeen, “Fast and automated optical polarization compensation of fiber unitaries,” [Review of Scientific Instruments](#), vol. 96, p. 063101, June 2025, with the permission of AIP Publishing.

[Optica Publishing Group policies state](#): “Authors who wish to use all or part of papers they previously authored and published in Optica Publishing Group Journals within a thesis or dissertation may do so according to the license and reuse terms outlined [here](#).”

[AIP copyright policies state](#): “Authors do not need permission from AIP Publishing to: [...] reuse your own AIP Publishing article in your thesis or dissertation (please format your credit line: ‘Reproduced from [FULL CITATION], with the permission of AIP Publishing’).”

Other co-authored publication not discussed in this thesis (see abstract in [Appendix A](#)):

- R. A. Abrahao, A. Mansouri, and A. P. Lund, “Boson sampling with Gaussian input states: Toward efficient scaling and certification,” [Phys. Rev. A](#), vol. 110, p. 052437, Nov 2024. © 2024 American Physical Society. Reproduced with permission.

[APS copyright policies state](#): “Yes, the author has the right to use the article or a portion of the article in a thesis or dissertation without requesting permission from APS, provided the bibliographic citation and the APS copyright credit line are given on the appropriate pages.”

Acknowledgements

I thank my advisor, Prof. Jeff S. Lundeen, for his guidance, insightful comments, and the opportunity to carry out this research.

I thank all past and present members of the Lundeen group — including Thomas Bailey, Kyle Jordan, Riza Fazili, Michael Weil, Prof. Raphael Abrahao, Dr. Andrew Proppe, and Dr. Manuel Ferrer — for valuable discussions, and Michael Weil for his helpful suggestions on the thesis introduction. I also extend my appreciation to the members of the CERC group for their assistance.

I thank all the professors who helped with their guidance and teaching in furthering my knowledge of optics and quantum information, including Prof. Jeff Lundeen, Prof. Ebrahim Karimi, Prof. Robert Boyd, Prof. Stefanie Czischek, Prof. Anne Broadbent, and Prof. Xiaoyi Bao.

I am grateful to my family for their unwavering support throughout this journey.

Statement of Work

I acknowledge the following contributions to the chapters of this thesis and express my appreciation to all collaborators listed below:

Chapter 2

- Prof. Jeff S. Lundeen: supervision and feedback.
- Arman Mansouri: conceptualization, methodological design, investigation, software development, and simulation analysis.

Chapter 3

- Prof. Jeff S. Lundeen: supervision, feedback, conceptualization, methodological design, and manuscript preparation.
- Kyle M. Jordan: experimental assistance, software development, feedback, and manuscript preparation.
- Prof. Raphael A. Abrahao: experimental assistance, feedback, and manuscript review.
- Arman Mansouri: conceptualization, methodological design, assembly of experimental apparatus, investigation, software development, data analysis, and manuscript preparation.

Chapter 4

- Niklas Braband: assembly of experimental apparatus, investigation, software development, data analysis, and manuscript preparation.
- Riza Fazili: conceptualization, methodological design, experimental assistance, and manuscript preparation.
- Prof. Jeff S. Lundeen: supervision, feedback, and manuscript preparation.
- Prof. Stefanie Czischek: supervision, feedback, and manuscript preparation.
- Arman Mansouri: conceptualization, methodological design, experimental assistance, software development, and manuscript preparation.

Chapter 5

- Prof. Jeff S. Lundeen: supervision, feedback, conceptualization, and methodological design.
- Michael T. Weil: feedback, conceptualization, and methodological design.
- Dr. Manuel F. Ferrer: feedback, conceptualization, and methodological design.
- Arman Mansouri: conceptualization, investigation, methodological design, software development, and optimization analysis.

Table of Contents

List of Tables	xv
List of Figures	xvi
1 Introduction	1
1.1 Quantum States	2
1.2 Unitary Transformations	5
1.2.1 Some Notable Unitary Transformations	6
1.2.2 Polarization Transformations	9
1.2.3 Arbitrary Unitaries on Spatial Modes	12
1.3 Quantum Measurements	17
1.3.1 Projective Measurements and Observables	18
1.3.2 Positive Operator-Valued Measures (POVMs)	19
1.3.3 Separable Measurements	20
1.3.4 Adaptive Measurements and Local Operations and Classical Communication (LOCC)	21
1.3.5 Collective Measurements	21
1.4 Two-Photon Interference	22
1.4.1 Hong-Ou-Mandel Interference	22
1.4.2 General Two-Photon Polarization State Projector	24
1.5 Historical Notes on Quantum State Estimation	27

1.5.1	Back-action and Heisenberg Uncertainty Principle	27
1.5.2	No-Cloning Theorem	28
1.5.3	Estimating the State from an Ensemble	29
1.6	Quantum State Estimation	30
1.6.1	State Guessed in Correspondence to Each Outcome	31
1.6.2	Quantum State Tomography	31
1.6.3	Prior Distribution of States	36
1.7	Fidelity	37
1.7.1	Optimality in the Sense of Average Fidelity	39
1.7.2	Average Infidelity Scaling with Sample Size: Gill-Massar Bound	40
1.7.3	State Fidelity Values in Applications	40
1.7.4	Average Gate Fidelity	41
1.8	Outline	42
2	Optimal Quantum State Estimation	44
2.1	Motivation	44
2.2	Theoretical Framework	46
2.2.1	State Estimation Game	46
2.2.2	Massar and Popescu Optimal Collective Measurement	47
2.2.3	Extensions of the State Estimation Game	48
2.2.4	Asymptotic Infidelity and State Estimation	49
2.2.5	Adaptive Tomography	49
2.2.6	Stochastic Optimization-Based Tomography	52
2.3	Original Simulation Results	52
2.3.1	Assumptions about State Purity	53
2.3.2	Infidelity Scaling Under <i>a Priori</i> Pure-State Assumption	53
2.3.3	Infidelity Scaling Without <i>a Priori</i> Pure-State Assumption	54
2.3.4	Summary	56
2.4	Summary	57

3	Quantum State Estimation with Collective Measurements	58
3.1	Motivation	58
3.2	Theoretical Framework	59
3.2.1	Average Fidelity	59
3.2.2	Optimal Collective Measuring Strategy	60
3.2.3	General Massar and Popescu Game (GenMP Game)	62
3.2.4	Optimal LOCC Strategy	62
3.2.5	Tetrahedron Massar and Popescu Game (TetraMP Game)	63
3.2.6	Quantum State Tomography: Two-Copy Collective Measurements	64
3.3	Experimental Methods	65
3.3.1	Non-Maximally Entangled Projections	66
3.3.2	Spontaneous Parametric Down-Conversion Source	67
3.3.3	Two-photon Interference and Visibility	68
3.3.4	Projector Selection	70
3.3.5	Experimental Test of the Massar and Popescu Game	70
3.3.6	Summary	71
3.4	Experimental Results	72
3.4.1	GenMP Game Results	72
3.4.2	TetraMP Game Results	73
3.4.3	Tomographic Infidelity Scaling for Two-Copy Collective Measurements	74
3.5	Summary	79
4	Fast and Automated Compensation of Fiber Unitaries	80
4.1	Motivation	80
4.2	Theoretical Framework	81
4.2.1	Previous Work on Polarization Compensation	81
4.2.2	Rotating QWP Tomography Method	82

4.2.3	Arbitrary Polarization Manipulation Using Three Fixed Liquid Crystals	83
4.2.4	Fidelity in Terms of Stokes Parameters	85
4.2.5	Polarization Compensation Scheme	85
4.2.6	Summary	87
4.3	Experimental Methods	88
4.3.1	Liquid Crystal Characterization Setup	88
4.3.2	Rotating QWP Tomography Setup	89
4.3.3	Polarization Compensation Setup	90
4.4	Experimental Results	91
4.4.1	Liquid Crystal Characterization Results	91
4.4.2	Rotating QWP Tomography Results	92
4.4.3	Polarization Compensation Results	94
4.5	Summary	96
5	Spatial Unitary Transformations with a Sequence of Spaced Thin Phase Gratings	98
5.1	Motivation	98
5.2	Theoretical Framework	99
5.2.1	Raman-Nath and Bragg Diffraction Regimes	100
5.2.2	Amplitude Transmittance Method	101
5.2.3	Matrix Formalism	101
5.2.4	Non-Zero Fourier Coefficients of a Phase Grating	105
5.2.5	Phase Gratings and Free-space Propagation for Achieving Unitaries	107
5.2.6	Genetic Algorithm	108
5.2.7	Summary	109
5.3	Optimization Results for Sinusoidal Gratings	109
5.3.1	Sinusoidal Gratings: Variable Propagation Lengths	111

5.3.2	Sinusoidal Gratings: Fixed Propagation Length	117
5.4	Square-Wave Gratings	121
5.5	Summary	123
6	Conclusion	124
6.1	Future directions	125
	References	128
	APPENDICES	146
A	Abstract of “Boson Sampling with Gaussian Input States: Toward Efficient Scaling and Certification”	147
B	Separability of a Two-Qubit State	149
C	Group Axioms	151
D	Multi-particle Collective Measurements for Quantum State Estimation	152
D.1	Rotated Tetrahedron States	152
D.2	Data Collection	153
D.3	Reference states: Mitigating bias	154
E	Fast and Automated Compensation of Fiber Unitaries	155
E.1	Polarization Tomography	155
E.2	Polarization Transformation	157
E.3	Apparatus Specification	158
F	Spatial Unitaries with a Sequence of Thin Phase Gratings	159
F.1	Fourier Coefficient Calculation	159
F.1.1	Sinusoidal Phase Grating	159

F.1.2	Square-Wave Phase Grating	160
F.2	Optimized Parameter Values	160
F.3	Two Square-Wave Gratings: Zero Trace	161
F.4	Equidistant Sinusoidal Gratings: Performance for Other Target Unitaries	163

List of Tables

3.1	Randomly chosen pure reference states and the corresponding power-law fits for tomography using two-copy collective measurements.	75
3.2	Experimental two-copy collective measurement infidelity scalings for nearly-pure states	77
4.1	Number of polarization compensation steps to reach fidelity thresholds . .	95
F.1	Optimized parameter values for sinusoidal gratings	160

List of Figures

1.1	Bloch sphere	3
1.2	Quantum circuit representation of a controlled-NOT gate.	8
1.3	Mach-Zehnder Interferometer	13
1.4	Reck interferometric scheme for realizing arbitrary discrete unitaries	15
1.5	Multi-plane light converter for realizing unitaries	16
1.6	POVM multiport device	20
1.7	Example of an adaptive measurement	21
1.8	Collective vs. separable measurements	22
1.9	HOM Interference	23
1.10	General two-photon polarization state projector	25
1.11	Balanced homodyne detection setup for quantum state tomography	30
1.12	Stages of quantum state estimation	31
1.13	Poincaré sphere	32
1.14	Six-outcome Pauli tomography setup	33
2.1	Ensemble of identically prepared qubits.	46
2.2	Infidelity scaling of Pauli tomography with <i>a priori</i> pure-state assumption	54
2.3	Tomographic infidelity scaling <i>without</i> the pure-state assumption	55
3.1	Tetrahedron states	61
3.2	Optimal LOCC strategy states	63

3.3	Ensemble divided in pairs for two-copy collective measurements	64
3.4	Experimental setup for the Massar and Popescu game	65
3.5	Spontaneous parametric down-conversion	68
3.6	HOM interference scans	69
3.7	Fidelities for the GenMP Game	72
3.8	Tetrahedron Massar and Popescu game fidelities	74
3.9	Observed two-copy collective measurement infidelity scaling (pure states) .	76
3.10	Observed two-copy collective measurement infidelity scaling (nearly-pure states)	78
4.1	Arbitrary polarization transformation with three LCVRs	84
4.2	Algorithm for compensation of fiber unitaries	86
4.3	Characterization setup for liquid crystal variable retarders	89
4.4	Rotating quarter waveplate polarization tomography setup	90
4.5	Polarization compensation setup	90
4.6	Characterization curve for a liquid crystal variable retarder	92
4.7	Rotating quarter waveplate polarization tomography results	93
4.8	Polarization compensation: Experimental fidelity	94
5.1	Raman-Nath diffraction and increase in total number of beams	99
5.2	Sequence of thin phase gratings	107
5.3	Detailed diagram of phase gratings and free-space propagation	108
5.4	Genetic algorithm	109
5.5	Output of a sequence of sine phase gratings optimized for the Hadamard gate	112
5.6	Output of a sequence of sine phase gratings optimized for the symmetric beamsplitter	113
5.7	Output of a sequence of sine phase gratings optimized for the NOT gate .	113
5.8	Three-dimensional discrete Fourier transform gate fidelities	114
5.9	Output of a sequence of sine phase gratings optimized for a DFT	115

5.10	Three-dimensional clock gate fidelities	116
5.11	Output of a sequence of sine phase gratings optimized for the T gate . . .	117
5.12	NOT gate fidelities	118
5.13	Hadamard gate fidelities	119
5.14	DFT_3 gate fidelities	120
5.15	DFT_4 gate fidelities	120
5.16	DFT_5 gate fidelities	121
D.1	Variation of slope in log-log infidelity plot	154
F.1	T_2 gate fidelities	163
F.2	T_3 gate fidelities	163
F.3	Z_3 gate fidelities	164

Chapter 1

Introduction

Quantum information processing is a rapidly advancing field that promises to transform various areas, including communication, sensing, and computing [1–3]. At its core are quantum states, which encode the information being processed. Accurate characterization of the state of quantum systems is crucial, since it allows one to predict the outcome statistics for any measurement of the system. Unreliable preparation or characterization of the states can significantly degrade the performance of information processing schemes. However, characterizing quantum states is a challenging task due to their high dimensionality and the invasive nature of most quantum measurements. In contrast to classical systems, quantum measurements generally alter the state, making the measured system unsuitable for further use in state estimation. To estimate a general d -dimensional quantum state, at least $d^2 - 1$ operators need to be measured [4, 5], and a significant number of measured systems are necessary for each operator to achieve reliable estimates. Therefore, efficient measurement strategies are essential to overcome the challenges of statistical uncertainty and high dimensionality.

Collective measurements, which entangle the states of multiple particles through the measurement process, are a powerful tool for quantum state estimation and metrology. While theoretically optimal measurements have been proposed, their experimental implementation has largely remained a challenge. In Chapter 3, I discuss the experimental realization of a multi-particle collective measurement, demonstrating its feasibility and effectiveness in estimating quantum states. The results support the long-standing prediction that collective measurements can provide more accurate estimates of quantum states than local measurement strategies.

Optical polarization errors can impact the performance of polarization-sensitive devices and quantum communication protocols. In Chapter 4, I present a fast and automated scheme for polarization compensation, which we experimentally demonstrated. We showed that this scheme is capable of achieving the desired accuracy in a few steps, and can be completed in a few minutes.

Realizing unitary transformations on multiple beams is a crucial component of various quantum information processing and imaging applications. While integrated optics offer a promising platform for these transformations, free-space optics may be more accessible for building and testing. Recently, the multi-plane light converter scheme has been proposed as a method for realizing complex unitary transformations in free-space. However, most implementations of this scheme rely on complicated optimization procedures with many parameters, which can be inscrutable and require additional heuristics. In Chapter 5, I consider a simpler approach in a matrix model for a sequence of spaced thin phase gratings with free-space between them. I present numerical optimization results for the thin grating parameters needed to achieve common unitary transformations, demonstrating the potential of this approach for realizing spatial unitaries in free-space.

This introduction provides an overview of the relevant background and context, including the fundamentals of quantum states, unitary transformations, quantum measurements, two-photon interference, quantum state estimation, and fidelity. These concepts form the basis for the work presented in this thesis.

1.1 Quantum States

The state of a quantum system can be described by a normalized vector in a complex vector space equipped with an inner product (a Hilbert space). This state vector, called a ket, is denoted by $|\psi\rangle$. Addition of kets or their multiplication by a complex number, results in another ket (after normalization). The state of a qubit can be represented in terms of the orthonormal *computational* basis states $|0\rangle$ and $|1\rangle$ as [6]:

$$|\psi\rangle = \cos \theta/2 |0\rangle + e^{i\phi} \sin \theta/2 |1\rangle, \quad (1.1)$$

where $0 \leq \theta \leq \pi$ and $0 \leq \phi < 2\pi$. Geometrically, these states correspond to points on the surface of a unit-sphere, called the Bloch sphere (Fig. 1.1), with the computational basis states $|0\rangle$ and $|1\rangle$ located at the poles. Alternatively, the ket $|\psi\rangle$ can be represented by a complex two-dimensional column-vector of norm one:

$$|\psi\rangle \longrightarrow \begin{bmatrix} \cos \theta/2 \\ e^{i\phi} \sin \theta/2 \end{bmatrix}. \quad (1.2)$$

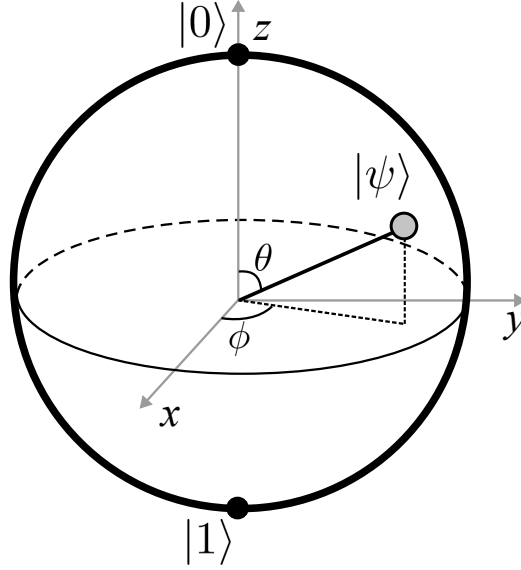


Figure 1.1: Bloch sphere. Any arbitrary pure qubit state $|\psi\rangle$ corresponds to a point on the surface of this unit sphere, specified by the angles θ and ϕ . Points inside the sphere correspond to mixed qubit states.

As an example, the polarization state of a photon can be represented in terms of the linear basis states $|H\rangle \equiv |0\rangle$ (horizontally polarized) and $|V\rangle \equiv |1\rangle$ (vertically polarized). It can also be expressed in terms of the diagonal and antidiagonal basis states $\{|D\rangle, |A\rangle\}$, or the left-circular and right-circular basis states $\{|L\rangle, |R\rangle\}$. These states are the following equal superpositions of $|H\rangle$ and $|V\rangle$ [7]:

$$|D\rangle \equiv \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle); \quad |A\rangle \equiv \frac{1}{\sqrt{2}}(|H\rangle - |V\rangle) \quad (1.3)$$

$$|L\rangle \equiv \frac{1}{\sqrt{2}}(|H\rangle + i|V\rangle); \quad |R\rangle \equiv \frac{1}{\sqrt{2}}(|H\rangle - i|V\rangle) \quad (1.4)$$

More generally, given an orthonormal basis $\{|\phi_i\rangle\}_{i=1}^d$, the arbitrary state $|\Psi\rangle$ can be written as [8]:

$$|\Psi\rangle = c_1 |\phi_1\rangle + c_2 |\phi_2\rangle + \cdots + c_d |\phi_d\rangle, \quad (1.5)$$

where the complex coefficients c_i are known as *probability amplitudes*. Eq. (1.5) represents a system that is in a coherent superposition of the states $|\phi_i\rangle$. Using Born's rule, the probability of $|\Psi\rangle$ to be found in the state $|\phi_i\rangle$ is given by $|c_i|^2$. Therefore, we must have $\sum_{i=1}^d |c_i|^2 = |\langle\Psi|\Psi\rangle|^2 = 1$. This explains the normalization requirement for state-vectors.

The vector space dual to the ket space, is a bra space. The bra corresponding to the ket $|\psi\rangle$ is denoted by $\langle\psi|$, and may be represented by a normalized row-vector. The

corresponding bra for a ket is obtained by taking the conjugate-transpose. The bra qubit is spanned by $\{|0\rangle, |1\rangle\}$. Using the bra-ket notation, the overlap between two states can be calculated as their inner-product. The modulus-squared of the overlap, $|\langle\psi|\phi\rangle|^2$, gives the probability of the state $|\phi\rangle$ to be found in $|\psi\rangle$. Their outer-product $|\psi\rangle\langle\phi|$, on the other hand, forms an operator, which can be seen as a matrix acting on the state-space.

The state-space of a composite physical system is the tensor product of the state-spaces of the component physical systems. For example, a general two-photon polarization state in spatial modes a and b can be expressed as:

$$\begin{aligned} |\psi_{2Q}\rangle &= c_{HH} |H_a\rangle \otimes |H_b\rangle + c_{HV} |H_a\rangle \otimes |V_b\rangle + c_{VH} |V_a\rangle \otimes |H_b\rangle + c_{VV} |V_a\rangle \otimes |V_b\rangle \\ &\equiv c_{HH} |H_a H_b\rangle + c_{HV} |H_a V_b\rangle + c_{VH} |V_a H_b\rangle + c_{VV} |V_a V_b\rangle, \end{aligned} \quad (1.6)$$

where c_{HH}, c_{HV}, c_{VH} , and c_{VV} are complex coefficients satisfying the normalization $|c_{HH}|^2 + |c_{HV}|^2 + |c_{VH}|^2 + |c_{VV}|^2 = 1$. The state $|\psi_{2Q}\rangle$ is called separable if it can be written as a *product-state* in the form $|\psi_{2Q}\rangle = |\psi_a\rangle \otimes |\psi_b\rangle$, where $|\psi_a\rangle$ and $|\psi_b\rangle$ are the states of the subsystems a and b , respectively. Bipartite states that do not fulfill this condition are entangled. Note that $|\psi_{2Q}\rangle$ is separable if and only if $c_{HH}c_{VV} = c_{HV}c_{VH}$ (see Appendix B).

Entanglement results in correlations between measurement outcomes of the system in any basis. It is one of the most striking features of quantum mechanics and is a unique resource playing a crucial role in many applications of quantum information and quantum computation, such as superdense coding and quantum teleportation.

The states discussed so far are known as *pure states*, and can describe coherent superpositions. In many quantum-computing algorithms, taking advantage of superposition allows probing systems in parallel, which can also help reduce the computational complexity of certain problems [6, 9].

A more general way to describe the state of a quantum system is using a density matrix ρ . The density matrix corresponding to the pure state $|\psi\rangle$ is simply $|\psi\rangle\langle\psi|$. Mathematically, this corresponds to an operator that projects onto the state $|\psi\rangle$, as $(|\psi\rangle\langle\psi|) |\phi\rangle = \langle\psi|\phi\rangle |\psi\rangle$.

More broadly, density matrices allow the representation of incoherent mixtures of pure states, such as partially polarized or unpolarized light. In general, the presence of noise, decoherence, or measurements in a quantum system can necessitate the use of density matrices to accurately represent its state. A density matrix can be defined as

$$\rho \equiv \sum_{i=1}^m p_i |\psi_i\rangle\langle\psi_i|, \quad (1.7)$$

where the probabilities p_i of the mixture sum to one. For example, the density matrix formalism allows the representation of an ensemble of photons, for which there is a $p_1 = 2/3$ probability of having $|\psi_1\rangle = |H\rangle$, and a $p_2 = 1/3$ probability of having $|\psi_2\rangle = |V\rangle$. The states $|\psi_i\rangle$ need not form an orthonormal basis, and Eq. (1.7) does not uniquely describe the ensemble corresponding to ρ [10–16].

From the definition Eq. (1.7), it follows that density matrices are Hermitian, positive semi-definite, and have trace unity. More explicitly, the conjugate-transpose $\rho^\dagger = \rho$, all eigenvalues p_i of ρ are non-negative, and $\text{tr}(\rho) = \sum_{i=1}^m p_i = 1$. Since ρ is diagonalizable, it can be written in terms of its unique eigenvalues and normalized eigenstates. From these properties, we can also conclude that a $d \times d$ density matrix has $d^2 - 1$ real free parameters.

The level of mixedness of the state is determined by the purity $\text{tr}(\rho^2)$, where $\text{tr}(\rho^2) = 1/d$ for a maximally mixed state, and $\text{tr}(\rho^2) = 1$ for a pure state. For a composite physical system having n subsystems with subsystem i prepared in ρ_i , the joint state is $\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n$ [6]. A mixed state ρ is separable if it can be expressed as a convex sum of product states; otherwise, it is entangled and of form [17, 18]:

$$\rho \neq \sum_i p_i \rho_1^i \otimes \cdots \otimes \rho_n^i. \quad (1.8)$$

1.2 Unitary Transformations

Unitary transformations are one of the most common types of operations in closed quantum systems. A unitary transformation U preserves the norm of states — meaning

$$\|U|\psi\rangle\| = \||\psi\rangle\|, \quad (1.9)$$

for any arbitrary $|\psi\rangle$ in the state-space — and more generally, the inner product, such that:

$$(\langle\phi|U^\dagger)(U|\psi\rangle) = \langle\phi|\psi\rangle, \quad (1.10)$$

for any arbitrary $|\psi\rangle$ and $|\phi\rangle$ in the state-space. From Eq. (1.10), we see that U is invertible, with its inverse given by $U^\dagger$:

$$UU^\dagger = U^\dagger U = \mathbb{I}. \quad (1.11)$$

The set $U(d)$ of unitary matrices of size d forms a group under matrix multiplication (group axioms shown in Appendix C). Moreover, for any unitary matrix U of finite size, the following important properties hold [19]:

- The rows of U form an orthonormal basis for the state-space.
- The columns of U form an orthonormal basis for the state-space.
- The matrix U is *normal*, meaning $U^\dagger U = U U^\dagger$.
- By the Spectral Theorem, U is diagonalizable. Therefore, U can be decomposed as $U = V D V^\dagger$, where D is diagonal and V is unitary.
- The eigenvalues of U have modulus one, and $|\det U| = 1$. The subset of unitaries with determinant 1 form the special unitary subgroup $SU(d)$.
- The matrix U can be written as the matrix-exponential $U = e^{iH}$, where H is a Hermitian matrix.

1.2.1 Some Notable Unitary Transformations

The time-evolution of a closed quantum system is described by a unitary transformation. That is, the state $|\psi\rangle$ of a system at reference time 0 is related to its state $|\psi'\rangle$ at some time t by a unitary operator as [20]:

$$|\psi'\rangle = e^{-i\hat{H}t/\hbar} |\psi\rangle, \quad (1.12)$$

where the Hermitian energy operator $\hat{H}$ is the *Hamiltonian* of the system and $\hbar = h/(2\pi)$ is the reduced Planck constant. Eq. (1.12) is valid provided $\hat{H}$ is time-independent.

Other prominent examples of unitaries are the translation operator T_x and the spin rotation operator $R_z(\phi)$. The operator $T_x(a) = e^{-ip_x a/\hbar}$, where p_x is the momentum operator along x , translates a position state $|x\rangle$ as [20]:

$$T(a) |x\rangle = |x + a\rangle. \quad (1.13)$$

The spin rotation operator $R_z(\phi) = e^{-iS_z\phi/\hbar}$, where S_z is the z spin angular momentum operator, causes a spin rotation of angle ϕ about the z -axis.

Unitary transformations can be used as quantum logic gates, with applications in quantum information processing, quantum communication, quantum computing, metrology,

and sensing. A ubiquitous example is the action of a beamsplitter. A 50:50 beamsplitter may be represented as:

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (1.14)$$

The matrix H above is also known as the Hadamard gate. Other notable 2×2 unitary transformations include the Hermitian, traceless, and involutory Pauli operators:

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}; \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (1.15)$$

Along with the identity matrix I_2 , Pauli operators form a basis for 2×2 Hermitian matrices.

There are two notable approaches to generalize Pauli matrices to higher dimensions. The generalized Gell-Mann matrices, acting on qudits, form a basis for Hermitian $d \times d$ matrices. On the other hand, Sylvester's generalized Pauli matrices X and Z are a non-Hermitian extension to d dimensions which maintain the relations

$$\sigma_x^2 = \sigma_z^2 = I, \quad \sigma_x \sigma_z = -\sigma_z \sigma_x. \quad (1.16)$$

The shift matrix

$$X = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \quad (1.17)$$

and the clock matrix

$$Z = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \omega & 0 & \cdots & 0 \\ 0 & 0 & \omega^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \omega^{d-1} \end{bmatrix} \quad (1.18)$$

where $\omega = e^{2\pi i/d}$ generalize σ_x and σ_z , respectively [21].

The discrete Fourier transform matrix,

$$\text{DFT}_d = \frac{1}{\sqrt{d}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{d-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(d-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{d-1} & \omega^{2(d-1)} & \cdots & \omega^{(d-1)(d-1)} \end{bmatrix} \quad (1.19)$$

where $\omega = e^{2\pi i/d}$, is a unitary operation with many interesting applications, for example as a subroutine in efficient quantum algorithms for the prime factorization and discrete logarithm problems [2].

A common type of gate used in quantum computation is a controlled gate, whose action depends on the state of another qubit (or other qubits). A notable example is the entangling two-qubit controlled-NOT gate, which acts as $\text{CNOT}(|x\rangle \otimes |y\rangle) = |x\rangle \otimes |y \oplus x\rangle$ for $|x\rangle, |y\rangle \in \{|0\rangle, |1\rangle\}$, where $\oplus$ represents addition modulo two. The target qubit $|y\rangle$ remains unchanged for the control qubit state $|x\rangle = |0\rangle$, whereas the σ_x (NOT) gate is performed on $|y\rangle$ for $|x\rangle = |1\rangle$. The control and target qubit states could of course be in a superposition of the computational basis states.

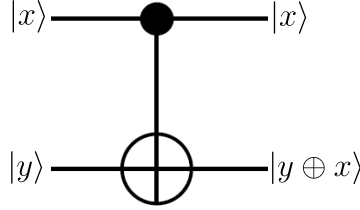


Figure 1.2: Quantum circuit representation of a controlled-NOT gate.

In matrix form, the CNOT gate in Fig. 1.2 can be expressed as

$$\text{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \quad (1.20)$$

Clifford gates map tensor products of Pauli operators to tensor products of Pauli operators under *conjugation*, meaning a unitary U is Clifford if

$$U(\sigma_{i_1} \otimes \cdots \otimes \sigma_{i_n})U^\dagger = \sigma_{j_1} \otimes \cdots \otimes \sigma_{j_n}. \quad (1.21)$$

This property plays a key role in the context of understanding achievable operations using a set of gates in quantum computing. Examples of Clifford gates are the Hadamard gate H , the CNOT gate, and the $\sqrt{\sigma_z}$ gate. A set of quantum gates is *universal* if finite products of its elements can approximate any unitary operation, in terms of yielding similar measurement outcome statistics [6, 22]. Clifford gates alone do not form a universal set of

quantum gates. However, the Clifford set $\{H, \text{CNOT}, \sqrt{\sigma_z}\}$ along with the non-Clifford T gate,

$$\mathsf{T}_2 = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}, \quad (1.22)$$

which simply acts as a phase-shift, do form a universal set. For a qutrit (three-level) system, the T gate is given by [23]

$$\mathsf{T}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{2\pi i/9} & 0 \\ 0 & 0 & e^{16\pi i/9} \end{bmatrix}. \quad (1.23)$$

1.2.2 Polarization Transformations

In this subsection, we introduce the mathematical framework used to describe and manipulate the polarization state of light, which is a key property of transverse electromagnetic waves that underlies their interaction with various materials and optical devices.

For a plane-wave propagating in the z -direction, the electric field may be decomposed into two orthogonal components $E_{x,0}$ and $E_{y,0}$, with phases δ_x and δ_y [24, 25]:

$$E_x(z, t) = E_{x,0} \cos(\omega t - kz + \delta_x), \quad (1.24)$$

$$E_y(z, t) = E_{y,0} \cos(\omega t - kz + \delta_y), \quad (1.25)$$

where ω is the angular frequency, $k = 2\pi/\lambda$ is the wave number, and λ is the wavelength.

Jones Formalism

The Jones vector provides a representation of the polarization state of a plane-wave in terms of the amplitudes and phases of its orthogonal components as [26]

$$\begin{bmatrix} E_{x,0} e^{i\delta_x} \\ E_{y,0} e^{i\delta_y} \end{bmatrix}. \quad (1.26)$$

It is convenient to use the normalized Jones vector — when we are only interested in the polarization state — so as to have a linear superposition of the polarization states $|H\rangle = [1 \ 0]^T$ and $|V\rangle = [0 \ 1]^T$ (discussed in Section 1.1). For a plane wave with an

electric field oscillating at an azimuth angle ψ relative to the x -axis and a phase difference $\delta = \delta_y - \delta_x$ between its components, the normalized Jones vector is given by [27]:

$$\mathbf{J}(\psi, \delta) = \begin{bmatrix} \cos \psi \\ e^{i\delta} \sin \psi \end{bmatrix}. \quad (1.27)$$

The polarization state can be altered with *birefringent* optical elements such as waveplates and general retarders, in which the refractive index depends on the polarization. In a uniaxial birefringent crystal, the two principal polarization states are the ordinary and extraordinary rays, whose electric field vectors are orthogonal and indicate the “slow” and “fast” axes of the medium. The refractive indices along these axes are n_s and n_f , respectively.

Unitary Jones matrices can describe the polarization transformation associated with waveplates and arbitrary phase retarders based on their action on Jones vectors. A linear phase retarder of thickness l with phase retardation $\Gamma = 2\pi l(n_s - n_f)/\lambda$ can be represented by the Jones matrix [27]:

$$\mathbf{J}_{\text{Lin. Ret.}} = \begin{bmatrix} e^{-i\Gamma/2} \cos^2 \theta + e^{i\Gamma/2} \sin^2 \theta & -i \sin(\Gamma/2) \sin(2\theta) \\ -i \sin(\Gamma/2) \sin(2\theta) & e^{-i\Gamma/2} \sin^2 \theta + e^{i\Gamma/2} \cos^2 \theta \end{bmatrix}, \quad (1.28)$$

where θ is the angle between the fast axis of the retarder and the x -axis. For a half-wave plate, $\Gamma = \pi$, while for a quarter-wave plate, $\Gamma = \pi/2$.

Stokes Parameters and Mueller Matrix Formalism

In many practical light sources, such as ensembles of randomly oriented atomic emitters, the polarization state fluctuates rapidly and irregularly and can therefore be partially polarized or fully unpolarized. The Stokes formalism provides a general description for any degree of polarization.

For a plane-wave propagating in the z -direction (see Eq. (1.24)-(1.25)), the Stokes parameters are defined as follows [28]:

$$S_0 = E_{x,0}^2 + E_{y,0}^2, \quad (1.29)$$

$$S_1 = E_{x,0}^2 - E_{y,0}^2, \quad (1.30)$$

$$S_2 = 2E_{x,0}E_{y,0} \cos(\delta), \quad (1.31)$$

$$S_3 = 2E_{x,0}E_{y,0} \sin(\delta), \quad (1.32)$$

where $\delta = \delta_y - \delta_x$. They are usually written as a vector, called the Stokes vector $\vec{S}$:

$$\vec{S} = \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = S_0 \begin{bmatrix} 1 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix}, \quad (1.33)$$

where S_0 describes the intensity of the wave, S_1 the intensity difference between horizontally (H) polarized light and vertically (V) polarized light, S_2 the intensity difference between diagonally (D) polarized light and anti-diagonally (A) polarized light, and S_3 the intensity difference between right-hand (R) circularly polarized light and left-hand (L) circularly polarized light. In Eq. (1.33), $S_i = S_i/S_0$ represent the normalized Stokes parameters.

The Stokes parameters satisfy the inequality:

$$S_0^2 \geq S_1^2 + S_2^2 + S_3^2. \quad (1.34)$$

When the light is fully polarized, Eq. (1.34) becomes an equality.

States of polarization can be represented using a sphere known as the *Poincaré* sphere, analogous to the Bloch sphere for qubit states. The three-component vector $\vec{S} = (S_1, S_2, S_3)$ describes the Cartesian coordinates of the polarization state within the Poincaré sphere. Fully polarized (pure) states lie on the surface, while partially polarized (mixed) states are represented by points inside the sphere.

The polarization of light can be changed using unitary optical elements such as waveplates and arbitrary retarders. It may also undergo a lossy transformation through polarizers or depolarizers. Any of these changes to the Stokes parameters can be described by a Mueller matrix $\mathbf{M}$ with elements m_{ij} as:

$$\vec{S}' = \begin{bmatrix} S'_0 \\ S'_1 \\ S'_2 \\ S'_3 \end{bmatrix} = \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{bmatrix} \cdot \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \mathbf{M} \cdot \vec{S}. \quad (1.35)$$

For a system of n optical elements, the Mueller matrices can be multiplied to get the Mueller matrix describing the entire system:

$$\vec{S}' = \mathbf{M}_n \cdot \dots \cdot \mathbf{M}_2 \cdot \mathbf{M}_1 \cdot \vec{S} = \mathbf{M}_{\text{sys}} \cdot \vec{S}. \quad (1.36)$$

The action of a liquid crystal variable retarder (LCVR) can be described by the following Mueller matrix:

$$\mathbf{M}_{\text{LCVR}}(\theta, \delta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2\theta + \sin^2 2\theta \cos \delta & \cos 2\theta \sin 2\theta (1 - \cos \delta) & -\sin 2\theta \sin \delta \\ 0 & \cos 2\theta \sin 2\theta (1 - \cos \delta) & \cos^2 2\theta \cos \delta + \sin^2 2\theta & \cos 2\theta \sin \delta \\ 0 & \sin 2\theta \sin \delta & -\cos 2\theta \sin \delta & \cos \delta \end{bmatrix}. \quad (1.37)$$

Here, θ is the angle of the liquid crystal birefringent fast axis relative to the horizontal axis, and δ is the relative phase retardance between the horizontal and vertical polarizations (which varies with voltage).

The Mueller matrix $\mathbf{M}$ corresponding to a Jones matrix $\mathbf{J}$ can be obtained using the relation [29]:

$$\mathbf{M} = \mathbf{A}(\mathbf{J} \otimes \mathbf{J}^*)\mathbf{A}^{-1}, \quad (1.38)$$

where $*$ indicates the complex-conjugate, and

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix}. \quad (1.39)$$

A thorough treatment of polarization theory can be found in Refs. [30, 31].

1.2.3 Arbitrary Unitaries on Spatial Modes

In this subsection, we discuss the realization of $d \times d$ unitary matrices transforming d input states into d output states. A lossless variable beamsplitter with a phase shifter at one output transforms an input state with modes (k_1, k_2) into the output state with modes (k'_1, k'_2) as follows [32]:

$$\begin{bmatrix} k'_1 \\ k'_2 \end{bmatrix} = \begin{bmatrix} e^{i\phi} \sin \omega & e^{i\phi} \cos \omega \\ \cos \omega & -\sin \omega \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}. \quad (1.40)$$

The parameter $0 \leq \omega < 2\pi$ describes the beamsplitter transmittance ($|t| = \cos \omega$) and the reflectivity ($|r| = \sin \omega$). The parameter ϕ is a phase shift that can be added at one output. Alternatively, the beamsplitter with variable reflectivity can be replaced by a Mach-Zehnder interferometer with symmetric 50 : 50 beamsplitters. Phases at the input ports (as well as a global phase at the output) can be chosen such that the beamsplitter matrix achieves any transformation in the $U(2)$ group [33, 34]. A variable beamsplitter

can alternatively be realized using a Mach-Zehnder interferometer (Fig. 1.3). As we will discuss next, the decomposition of $d \times d$ unitaries can be achieved through successive two-mode unitaries [32].

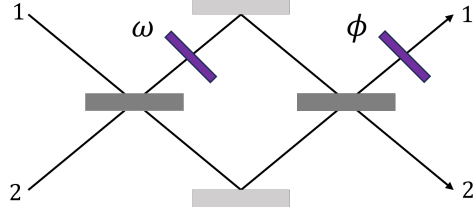


Figure 1.3: A variable beamsplitter can alternatively be realized using a Mach-Zehnder interferometer. A phase shift ϕ is added at one output. Input phases (and a global phase at the output) can be used to achieve any two-by-two unitary.

Reck Unitary Decomposition Scheme

Reck *et al.* [32] presented a deterministic algorithm to decompose a general unitary U_d of dimension d into the action of beamsplitters and phase shifters. In other words, any $d \times d$ unitary transformation can be realized using successive unitary transformations on two-dimensional subspaces.

Let $T_{pq}(\omega_{pq}, \phi_{pq})$ denote a $d \times d$ identity matrix with elements $I_{pp}, I_{pq}, I_{qp}, I_{qq}$ replaced by beamsplitter elements, as defined in Eq. (1.40). This matrix performs a unitary transformation on a two-dimensional subspace and leaves a $(d - 2)$ -dimensional subspace unchanged. To find a decomposition, we consider the problem in reverse. Suppose we successively multiply the desired unitary U_d on the right by $T_{dq}(\omega_{dq}, \phi_{dq})$ for $q = d - 1, \dots, 1$, realized with beamsplitter devices between ports d and q :

$$U_d(T_{d,d-1}T_{d,d-2} \cdots T_{d,1}) = \left[\begin{array}{c|c} U_{d-1} & 0 \\ \hline 0 & e^{i\alpha_d} \end{array} \right]. \quad (1.41)$$

The resulting matrix contains only zeros in the last row (and also in the last column due to unitarity) except for the diagonal element. Later transformations will not affect this row. The dimension of U_d has been reduced to $d - 1$. We may view $R(d) \equiv T_{d,d-1}T_{d,d-2} \cdots T_{d,1}$ as a generalized rotation of the last row of U_d in a d -dimensional vector space. From Eq. (1.40) for each pair of modes, $R(d)$ transforms a unit row vector into a general d -dimensional row

vector, such that:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}^T R^{-1}(d) = \begin{bmatrix} e^{-i\phi_1} \cos \omega_1 \\ -e^{-i\phi_2} \cos \omega_2 \sin \omega_1 \\ e^{-i\phi_3} \cos \omega_3 \sin \omega_2 \sin \omega_1 \\ \vdots \\ \mp e^{-i\phi_{d-1}} \cos \omega_{d-1} \cdots \sin \omega_1 \\ \pm \sin \omega_{d-1} \sin \omega_{d-2} \cdots \sin \omega_1 \end{bmatrix}^T. \quad (1.42)$$

In Eq. (1.42), the parameters ω_i and ϕ_i define general spherical coordinates in d -dimensions.

Recursively following this procedure of applying sequences of beamsplitter transformations (rotations) from the right, we obtain a diagonal matrix

$$D^{-1} = U_d R(d) R(d-1) \cdots R(2) \quad (1.43)$$

$$= U_d (T_{d,d-1} T_{d,d-2} \cdots T_{d,1}) (T_{d-1,d-1} T_{d-1,d-2} \cdots T_{d-1,1}) \cdots (T_{2,d-1} T_{2,d-2} \cdots T_{2,1}) \quad (1.44)$$

$$= U_d (T_{d,d-1} T_{d,d-2} \cdots T_{2,1}) \quad (1.45)$$

with elements that have modulus 1. We can then multiply the result by the diagonal matrix

$$D = \begin{bmatrix} e^{-i\alpha_1} & & \\ & \ddots & \\ & & e^{-i\alpha_d} \end{bmatrix}, \quad (1.46)$$

which is simply a phase shifter on each mode. This allows us to obtain the $d \times d$ identity matrix I_d :

$$U_d (T_{d,d-1} T_{d,d-2} \cdots T_{2,1}) D = I_d. \quad (1.47)$$

Having found a recipe for transforming U_d to I_d , we can simply invert Eq. (1.47) to get the decomposition of U_d into 2×2 beamsplitter and phase shift unitaries:

$$U_d = (T_{d,d-1} T_{d,d-2} \cdots T_{2,1} D)^{-1} \quad (1.48)$$

$$= (R(d) R(d-1) \cdots R(2) D)^{-1} \quad (1.49)$$

The reflectivity ω and phase ϕ of each of the $d(d-1)/2$ beamsplitter devices are chosen so that $R(p)$ rotates the p -th row of the unitary U_d into $[0, 0, \dots, e^{i\alpha_p}, \dots, 0]$, where the $e^{i\alpha_p}$ is in the p -th position.

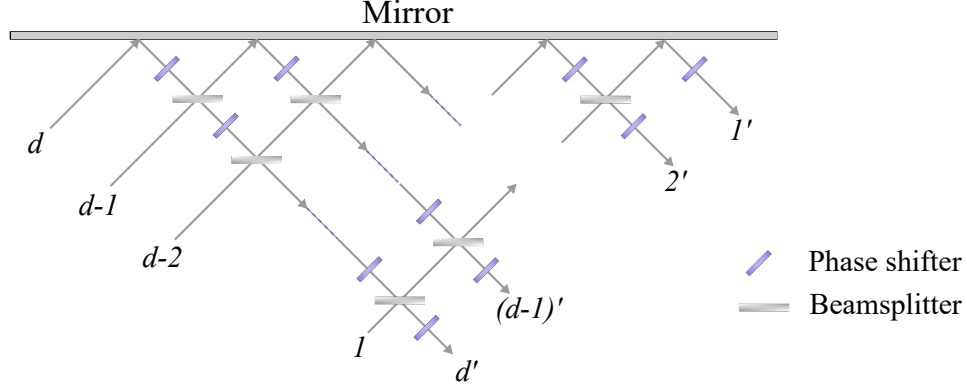


Figure 1.4: Reck *et al.*'s decomposition setup. A mesh of beamsplitters and phase shifters arranged in a triangular lattice form a large interferometer with d^2 arms that can implement any $d \times d$ unitary transformation.

The result of the decomposition is a mesh of beamsplitters in a large interferometer with $\sim d^2$ arms, as depicted in Fig. 1.4. Despite the powerful nature of the decomposition, its implementation requires phase stability in each of the arms to lengths much smaller than the optical wavelength. For this reason, most of the successful applications of the Reck (as well as the more compact square-lattice Clements [35]) decomposition schemes have been realized in integrated optics. Integrated waveguide devices with several evanescent couplers and thermo-optic phase shifters have been used to experimentally demonstrate important tasks in linear optical quantum computing: single-qubit and two-qubit quantum logic gates (e.g. Controlled-NOT), boson sampling, and random unitary transformations [36]. They have also been applied to matrix multiplication in optical neural networks on nanophotonic circuits [37]. The latter application makes use of singular-value decomposition to extend the range of accessible matrices from unitaries to arbitrary real-valued matrices.

Unitary Implementation through a Multi-Plane Light Converter

A multi-plane light converter (MPLC) is another possible way to implement spatial unitary transformations. It consists of alternating layers of phase-masks and free-space propagation (or optical Fourier transform), denoted by $\mathcal{P}$ and $\mathcal{D}$, respectively. Together, an MPLC consisting of P blocks of phase-masks and free-space propagation can perform a unitary transformation

$$U = \mathcal{D}_P \mathcal{P}_P \cdots \mathcal{D}_1 \mathcal{P}_1. \quad (1.50)$$

For illustration, an example of an MPLC is depicted in Fig. 1.5. This type of scheme facilitates the implementation of unitary transformations in free-space.

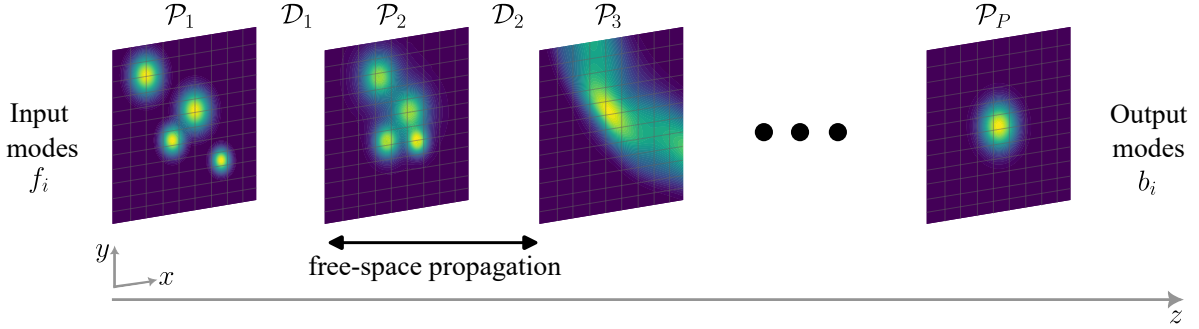


Figure 1.5: A multi-plane light converter (MPLC) uses alternating layers of phase-masks and free-space propagation to realize spatial unitaries in free-space between input modes f_i and output modes b_i .

In most applications of MPLC to date, numerical optimization has been used to determine the required phase planes for a particular transformation. A common optimization method used is a wavefront matching algorithm. Wavefront matching consists of propagating the input fields f_i forward and the output fields b_i backward through the system. The phase fronts of f_i and b_i need to match at every spatial coordinate for an input to excite an output. If the fields do not match, the phase error at each spatial coordinate is calculated and subtracted from that phase plane. In multi-plane systems, all f_i and b_i are moved together to the next plane or backwards together to the previous plane. This process iterates until the error is minimized [38].

Many implementations of MPLCs utilize liquid crystal spatial light modulators (SLMs) to apply the desired phase masks with high resolution. Wavefront matching and MPLCs have been used to demonstrate mode multiplexing and de-multiplexing between Gaussian beams and modes of a graded index fibre [38], mode-sorting for 210 Laguerre-Gaussian transverse spatial modes using only seven phase planes [39], and high-dimensional quantum gates using the orbital angular momentum degree of freedom of photons [40]. Different choices of basis for the modes f_i and b_i have been considered including the pixel basis of SLM devices, the Hermite-Gaussian basis [38], an array of Gaussian beams [41], or Laguerre-Gaussian beams [39].

Despite successful applications of wavefront matching, the complex optimization can be computationally slow, provide inscrutable results, and require additional heuristics for high performance. The wavefront-matching optimization algorithm can converge to local minima. Constraints on limiting the angle space or the size of the masks, avoiding updates in low power areas, forcing similarity between adjacent masks, zeroing masks, and averaging the optimization over multiple wavelengths have been considered to mitigate the issue of local minima [38]. Resulting phase masks can involve sudden phase jumps, but thresholding

and optimization over multiple wavelengths can help produce smoother masks. Moreover, several aspects in the design of general MPLC systems, such as the SLM pixel size, the use of free-space propagation or an optical Fourier transform, the choice of mode basis, and the mode size require fine tuning (often through trial and error). These challenges motivate the use of a more limited class of phase masks with fewer parameters such as gratings. We consider this approach using a matrix model in Chapter 5.

More recently, a deterministic algorithm for an exact implementation of any unitary using Fourier transforms and phase masks was proposed in Ref. [42]. This decomposition is achieved using $6d$ discrete Fourier transforms and $6d + 1$ diagonal phase masks.

1.3 Quantum Measurements

Quantum measurements can be described by a set of measurement operators $\{M_m\}$ acting on the state-space, which satisfy the completeness relation $\sum_m M_m^\dagger M_m = I$. The index m refers to the possible measurement outcomes. For a quantum system in state $|\psi\rangle$ before the measurement, the probability of obtaining the outcome m is [6]:

$$p(m) = \langle \psi | M_m^\dagger M_m | \psi \rangle, \quad (1.51)$$

and the system's state after the measurement is

$$|\psi_f\rangle = \frac{M_m |\psi\rangle}{\sqrt{\langle \psi | M_m^\dagger M_m | \psi \rangle}}. \quad (1.52)$$

The completeness relation equivalently means that the outcome probabilities sum to one:

$$\sum_m p(m) = \sum_m \langle \psi | M_m^\dagger M_m | \psi \rangle = 1. \quad (1.53)$$

Eq. (1.52) expresses a striking feature of measurements in quantum mechanics compared to classical systems: the quantum measurement *back-action*. The post-measurement state is generally altered through the measurement process. As I will explain, this implies that an unknown state $|\psi\rangle$ cannot be determined with a single measurement, or repeated measurements of the same system.

1.3.1 Projective Measurements and Observables

An important special case is that of projective measurements. A projective measurement can be described by an observable M , which is a Hermitian operator acting on the state-space of the system being measured, with the spectral decomposition [6]:

$$M = \sum_m m P_m, \quad (1.54)$$

where P_m is the projector onto the eigenspace of M corresponding to the real eigenvalue (outcome) m . Projectors P_m are Hermitian and by definition satisfy $P_m^2 = P_m$. Therefore, Eq. (1.51) for a projective measurement reduces to

$$p(m) = \langle \psi | P_m | \psi \rangle \quad (1.55)$$

and the post-measurement state is

$$|\psi_f\rangle = \frac{P_m |\psi\rangle}{\sqrt{p(m)}}. \quad (1.56)$$

Stated differently, projective measurements are the special case where the measurement operators M_m are orthonormal projectors, i.e., $M_m M_{m'} = \delta_{m,m'} M_m$. If the spectrum of M is non-degenerate, any projector would be a rank-1 projector $P_m = |m\rangle\langle m|$. This type of projective measurement is known as a von Neumann measurement [43]. An example of a rank-1 projector in optics is $P_H = |H\rangle\langle H|$, projecting onto horizontally polarized light using the transmission port of a polarizing beamsplitter. Whereas a rank-2 projector in the 2-dimensional polarization space is the identity $I = |H\rangle\langle H| + |V\rangle\langle V|$. Coupling into a single-mode fiber can be seen as a rank-1 projector $|u_0\rangle\langle u_0|$ in the spatial state-space, but the identity (rank-2) in the polarization state-space (as the fiber supports two polarization modes). Thus in the composite space, $|u_0\rangle\langle u_0| \otimes I$ is a rank-2 projector.

The action of an iris is to project onto a flat distribution of transverse-position eigenstates inside the diameter, and can be represented by a diagonal matrix in the transverse-position space. Unlike the single-mode fiber, it has a very large or infinite rank. It is still idempotent: placing a second identical iris yields the same output.

The expectation (or average) value of an observable M is given by the weighted average of the possible measurement outcomes m :

$$\langle M \rangle = \sum_m m p(m) \quad (1.57)$$

$$= \sum_m m \langle \psi | P_m | \psi \rangle \quad (1.58)$$

$$= \langle \psi | \left(\sum_m m P_m \right) | \psi \rangle \quad (1.59)$$

$$= \langle \psi | M | \psi \rangle. \quad (1.60)$$

Equivalently, if the system being measured is in a state described by the density matrix ρ , then [43]

$$\langle M \rangle = \text{Tr}(\rho M). \quad (1.61)$$

The standard deviation associated with observations of M is given by:

$$\Delta M = \sqrt{\langle (M - \langle M \rangle)^2 \rangle} \quad (1.62)$$

$$= \sqrt{\langle M^2 \rangle - \langle M \rangle^2}. \quad (1.63)$$

If a large number of experiments measuring M are performed on systems prepared in the state $|\psi\rangle$, a spread ΔM of measured values would be observed.

1.3.2 Positive Operator-Valued Measures (POVMs)

A positive operator valued measure (POVM) $\{\Pi_i\}$ is a set of positive semi-definite operators acting on the state-space that sum to the identity. POVMs allow the representation of generalized measurements in quantum mechanics, going beyond the paradigm of projective measurements, which leave the system's state in an operator eigenspace. Unlike the case of a projection-valued measure (PVM), the POVM elements $\{\Pi_i\}$ need not be orthogonal, i.e., $\Pi_i \Pi_j \neq \delta_{ij} \Pi_i$ in general. This implies that the number of POVM elements can be larger than the dimension of the state-space. Fig. 1.6 shows the representation of a POVM as a multiport device.

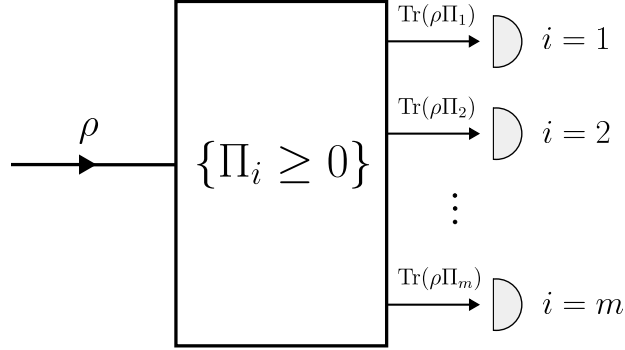


Figure 1.6: Multiport device representing a Positive Operator Valued Measure (POVM).

The probability of obtaining the outcome i is given by

$$\text{Prob}(i) = \text{tr}(\rho \Pi_i). \quad (1.64)$$

The example of coupling into a single-mode fiber corresponds to a fine-grained (low-rank) POVM element, whereas the example of the iris corresponds to a coarse (high-rank) POVM element.

An important result regarding POVMs is Naimark's dilation theorem [44], which gives a method to realize a POVM as a PVM acting on a larger Hilbert space.

1.3.3 Separable Measurements

Separable measurements are a class of measurements which can be factorized as a tensor product of individual measurements on the subsystems of a composite system:

$$M_{\text{sep}} = M^{(1)} \otimes M^{(2)} \dots \otimes M^{(N)}, \quad (1.65)$$

where $M^{(i)}$ represents a measurement of subsystem i .

We may otherwise define separable measurements from the perspective of POVMs. Recall that separable states are states which can be expressed as a convex sum of product states [17]. A POVM is separable if every POVM element is proportional to a separable state [45].

1.3.4 Adaptive Measurements and Local Operations and Classical Communication (LOCC)

Adaptive measurements use prior outcomes to guide the choice of future measurements through the course of the experiment. In other words, such measurements can access Local Operations and Classical Communication (LOCC). Adaptive measurements can be helpful in parameter estimation, as the optimal measurement often depends on the unknown state being probed. Although the true optimal measurement cannot be implemented exactly, one can estimate its optimal bases by performing preliminary measurements on a subset of the ensemble. After approximating the bases, we carry out a second measurement stage using that information on the remaining subset of the ensemble [46, 47]. An example showing this dependence of the measurement basis on a previous outcome is shown as a quantum circuit diagram in Fig. 1.7. Alternatively, adaptive strategies can reduce redundancy by selecting measurement configurations with the highest outcome uncertainty [48].

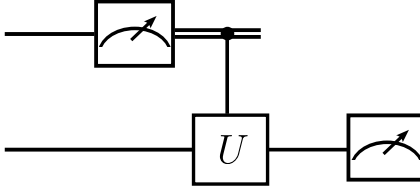


Figure 1.7: Example of an adaptive measurement, shown as a quantum circuit diagram. The choice of measurement basis for the second system, affected by the unitary U , depends on the measurement outcome for the first system.

1.3.5 Collective Measurements

Collective measurements probe many-body properties such as total spin (see Fig. 1.8 (b)) or permutation symmetry rather than properties such as single-particle spin. They arise naturally in many-body systems, for instance in the study of ferromagnetism, where the macroscopic field is a collective effect of the many spins. Collective measurements are non-local in the sense that they project onto entangled states and, thus, they enable applications that are impossible within the paradigm of Local Operations and Classical Communication (LOCC). The quantum circuit in Fig. 1.8 (a) contrasts collective measurements with separable measurements. Probing collective properties provides optimal solutions to several parameter and state estimation problems in quantum information [49–60].

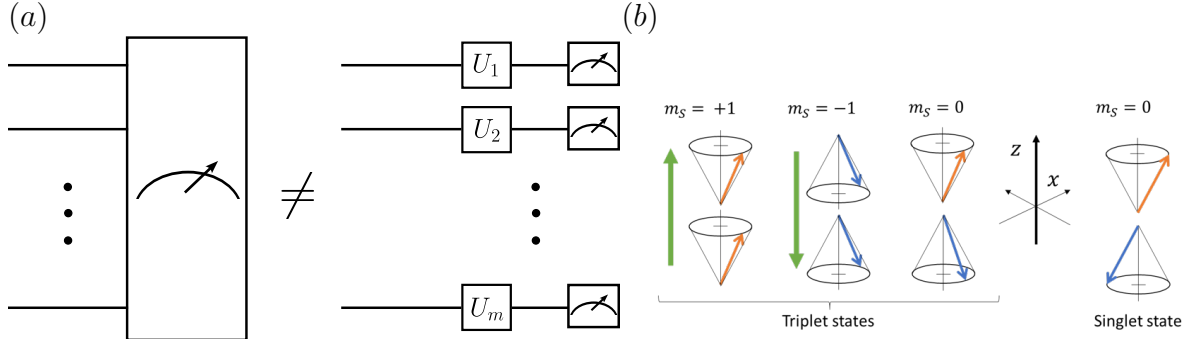


Figure 1.8: (a) Quantum circuit representation of a collective measurement. A collective measurement cannot be factorized into individual measurements of the subsystems, even with the addition of local unitaries. (b) Measuring the total spin of two spin-1/2 particles along the z -direction is an example of a collective measurement.

However, the majority of implementations in state estimation and metrology have so far been focused on the use of separable measurements, due to the practical challenges associated with the realization of collective measurements. Some photonic implementations have relied on encoding the qubits in different degrees of freedom of the same photon [61,62]. While this seemingly simplifies the task of performing collective measurements, the outcomes can be explained by classical optics. In particular, outcome statistics can be predicted based on first-order interference. Moreover, the states corresponding to the measurement projector do not allow useful Bell violations and quantum information processing tasks. Viable experimental methods for collective measurements will open up optimal sensing and state estimation applications [63–65]. In Chapter 3, we experimentally apply a collective measurement to quantum state estimation.

1.4 Two-Photon Interference

1.4.1 Hong-Ou-Mandel Interference

When two photons impinge onto different ports of a 50:50 beamsplitter (BS), one can expect four possible outcomes based on the transmission or reflection of the photons. However, when the photons are indistinguishable (identical in all degrees of freedom), the probability amplitudes for the two outcomes that result in a coincidence (one photon in each of the ports c and d in Fig. 1.9) sum to zero.

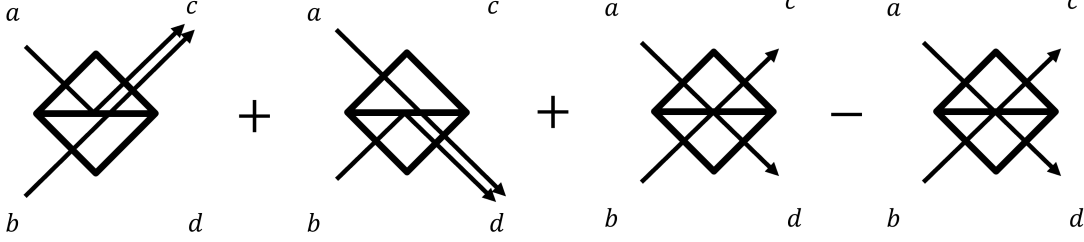


Figure 1.9: HOM interference. Indistinguishable photons enter opposite ports of a beamsplitter. The probability amplitudes for the photons to exit from opposite ports cancel, resulting in no coincidences for a 50:50 beamsplitter.

To describe two-photon interference, we use the quantum description of a beamsplitter according to the second quantization formalism. This is done by introducing a set of bosonic creation and annihilation operators $\hat{a}_i^\dagger$ and $\hat{a}_i$ respectively to represent electromagnetic fields in a mode i . These operators obey the commutation relations

$$[\hat{a}_i, \hat{a}_j^\dagger] = \hat{a}_i \hat{a}_j^\dagger - \hat{a}_j^\dagger \hat{a}_i = \delta_{ij}, \quad (1.66)$$

where δ_{ij} is the Kronecker delta ($\delta_{ii} = 1$ and $\delta_{ij} = 0$ otherwise). Applying the creation operator n times on the vacuum state $|\text{vac}\rangle$ results in an n -photon Fock state $|n\rangle_i$ in the mode i :

$$(\hat{a}_i^\dagger)^n |\text{vac}\rangle = \sqrt{n!} |n\rangle_i. \quad (1.67)$$

The notation $\hat{a}, \hat{b}, \hat{c}, \hat{d}$ can be used to denote the annihilation operators in modes a, b, c, d in Fig. 1.9, respectively. The action of a 50 : 50 beamsplitter can be expressed in terms of these field operators as:

$$\begin{cases} \hat{a} = \frac{1}{\sqrt{2}}(\hat{c} + \hat{d}) \\ \hat{b} = \frac{1}{\sqrt{2}}(\hat{c} - \hat{d}), \end{cases} \quad (1.68)$$

Two-photon interference happens for indistinguishable photons, which occurs when they are identical in all degrees of freedom, i.e., polarization, time-frequency, and position. Consider the action of the beamsplitter as described in Eq. (1.68) on two indistinguishable photons impinging onto opposite ports of a beamsplitter:

$$\hat{a}^\dagger \hat{b}^\dagger \xrightarrow{\text{BS}} \frac{1}{2} (\hat{c}^\dagger \hat{c}^\dagger - \hat{c}^\dagger \hat{d}^\dagger + \hat{c}^\dagger \hat{d}^\dagger - \hat{d}^\dagger \hat{d}^\dagger) |\text{vac}\rangle \quad (1.69)$$

$$= \frac{1}{2} [(\hat{c}^\dagger)^2 - (\hat{d}^\dagger)^2] |\text{vac}\rangle \quad (1.70)$$

$$= \frac{1}{\sqrt{2}} (|2\rangle_c - |2\rangle_d). \quad (1.71)$$

We see that the states corresponding to the photons exiting from opposite ports interfere destructively, meaning photons will only exit from the same output port.

This bunching effect is otherwise known as Hong-Ou-Mandel (HOM) interference, named after the group who experimentally demonstrated this effect with photons in 1987 [66]. Fearn-Loudon [67] as well as Rarity-Tapster [68, 69] independently investigated this phenomenon during the same period.

A standard form of the HOM interferometer uses pairs of photons that are identical in all degrees of freedom except for a relative time delay. The number of coincidences (simultaneous detections) at the output ports can then be measured to obtain a characteristic dip as a function of the delay between the two photons. The HOM visibility can be calculated as

$$\mathcal{V} = \frac{C(\infty) - C(0)}{C(\infty)}, \quad (1.72)$$

where $C(\infty)$ is the coincidence rate at a large relative delay (compared to the inverse of the photon bandwidth) and $C(0)$ is the coincidence rate at zero delay. When the maximum coincidence $C_{\max}$ occurs at large delay, the visibility is calculated as

$$\mathcal{V} = \frac{C_{\max} - C_{\min}}{C_{\max}}, \quad (1.73)$$

where $C_{\min} = C(0)$. A recent review of HOM interference and its numerous applications can be found in Ref. [70].

1.4.2 General Two-Photon Polarization State Projector

As discussed in Section 1.3.5, collective measurements involve projections onto entangled states. In this subsection, we discuss how we can realize such measurements. Suppose we have two photons in a separable two-photon polarization state $|\phi\rangle = |\vec{n}_a\rangle \otimes |\vec{p}_b\rangle$ — one photon in each of the spatial modes a and b — with $|\vec{n}\rangle$ and $|\vec{p}\rangle$ being arbitrary polarization qubit states. We may be interested in measuring the projector $P_m = |\psi_{2Q}\rangle\langle\psi_{2Q}|$, where

$$|\psi_{2Q}\rangle \equiv c_{HH} |H_a H_b\rangle + c_{HV} |H_a V_b\rangle + c_{VH} |V_a H_b\rangle + c_{VV} |V_a V_b\rangle \quad (1.74)$$

is a possibly entangled polarization state, with one photon in each of the spatial modes a and b . Specifically, we want to learn the probability $\langle P_m \rangle = |\langle\psi_{2Q}|\phi\rangle|^2$ for the projection of the input state $|\phi\rangle$ onto $|\psi_{2Q}\rangle$. We will examine a method for measuring P_m , which was proposed and demonstrated in Ref. [71]. Fig. 1.10 illustrates the components of this general

two-photon polarization state projector. This method will be applied experimentally in Chapter 3 to quantum state estimation.

In particular, projection onto an entangled state is achieved through two-photon interference. When two photons entering opposite input ports of a 50:50 beamsplitter are indistinguishable in all degrees of freedom except (possibly) polarization, coincidence detection at the beamsplitter output ports projects onto the singlet state $|S\rangle = \frac{1}{\sqrt{2}}(|HV\rangle - |VH\rangle)$ [72, 73]. This projection can be derived using a similar second-quantization calculation as in Section 1.4.1. The effect may also be inferred from the bosonic nature of photons, where symmetry constraints imply that if the spatial part of the state is antisymmetric, the polarization part must also be antisymmetric, ensuring that the overall two-photon state remains symmetric.

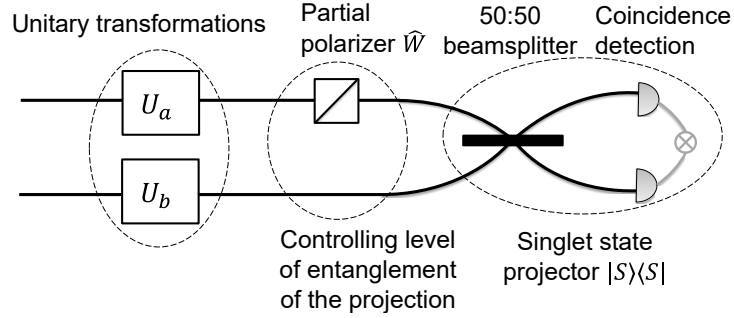


Figure 1.10: General two-photon polarization state projector.

The level of entanglement of a bipartite state can be quantified through various measures, such as the entanglement of formation and the concurrence. The concurrence of the arbitrary two-qubit state $|\psi_{2Q}\rangle$ is defined to be [74]:

$$\mathcal{C}(|\psi_{2Q}\rangle) = |\langle\psi_{2Q}|\tilde{\psi}_{2Q}\rangle|, \quad (1.75)$$

where the tilde represents the “spin-flip” operation $|\tilde{\psi}_{2Q}\rangle = (\sigma_y \otimes \sigma_y) |\psi_{2Q}^*\rangle$, and $|\psi_{2Q}^*\rangle$ is the complex-conjugate of $|\psi_{2Q}\rangle$. For a maximally-entangled state, $\mathcal{C} = 1$, whereas $\mathcal{C} = 0$ for a separable state. The intuition behind Eq. (1.75) is that in the separable case, the spin-flip operation maps each one-qubit state to its antipode (orthogonal state) on the Bloch sphere. Whereas for a maximally entangled state, the spin-flip only changes the global phase. To vary the concurrence of the projection state, polarization-dependent loss W can be introduced in one mode:

$$W = \begin{bmatrix} t_H & 0 \\ 0 & t_V \end{bmatrix} \otimes \mathbb{I}_{2 \times 2}, \quad (1.76)$$

where t_H, t_V represent complex loss amplitudes for the horizontal and vertical polarizations, respectively.

Unitary operations U_a and U_b , implemented through a half-wave plate (HWP) and quarter-wave plate (QWP) pair, are used to move between projections onto any state with the same concurrence. The required unitaries are found through the mapping between the representation of $|\psi\rangle$ in its Schmidt bases and the $\{|H\rangle, |V\rangle\}$ basis. The Schmidt decomposition allows writing the state of the bipartite system as the unique expression

$$|\psi_{2Q}\rangle = \lambda_1 |\zeta_a \theta_b\rangle - \lambda_2 |\zeta_a^\perp \theta_b^\perp\rangle, \quad (1.77)$$

where λ_i are non-negative real numbers with $\lambda_1^2 + \lambda_2^2 = 1$, and the sets $\{|\zeta_a\rangle, |\zeta_a^\perp\rangle\}$ and $\{|\theta_b\rangle, |\theta_b^\perp\rangle\}$ are the orthonormal Schmidt bases for a and b , respectively. The unitaries U_a and U_b are chosen such that they solve the equations $U_a |\zeta\rangle = |H\rangle$ and $U_b |\theta\rangle = |V\rangle$.

The combined transformation performed by unitaries, followed by polarization-dependent loss, and coincidence detection at the output of a 50:50 beamsplitter is given by:

$$M = |S\rangle\langle S| W U_a U_b, \quad (1.78)$$

where U_a, U_b are unitaries in modes a and b , respectively. The POVM element associated with coincidence detection at the output of the 50:50 beamsplitter is then given by

$$\begin{aligned} \Pi &= M^\dagger M \\ &= U_b^\dagger U_a^\dagger W^\dagger |S\rangle\langle S| S \langle S| W U_a U_b \\ &= U_b^\dagger U_a^\dagger W |S\rangle\langle S| W U_a U_b \\ &= \eta |\psi_{2Q}\rangle\langle\psi_{2Q}|, \end{aligned} \quad (1.79)$$

where the projection state $|\psi_{2Q}\rangle = U_a^\dagger U_b^\dagger |\psi'_{2Q}\rangle$. Here, $|\psi'_{2Q}\rangle$ is given by

$$|\psi'_{2Q}\rangle = \eta^{-1/2} W |S\rangle = (\sqrt{1+\gamma} |HV\rangle - \sqrt{1-\gamma} |VH\rangle)/\sqrt{2}, \quad (1.80)$$

with $\gamma = (t_H^2 - t_V^2)/(t_H^2 + t_V^2)$ and projection efficiency $\eta = (t_H^2 + t_V^2)/2$. The quantity $|\gamma|$ is related to the concurrence of the projection state by $\mathcal{C}(|\psi'\rangle) = \sqrt{1-\gamma^2}$. We find that the probability associated with the POVM element Π for the input state $|\phi\rangle$ is given by

$$\begin{aligned} \langle\phi|\Pi|\phi\rangle &= \langle\phi| U_b^\dagger U_a^\dagger W^\dagger |S\rangle\langle S| W U_a U_b |\phi\rangle = |\langle S| W U_a U_b |\phi\rangle|^2 \\ &= \eta |\langle\psi_{2Q}|\phi\rangle|^2. \end{aligned} \quad (1.81)$$

In this way, one can measure the probability for the projection of any input state $|\phi\rangle$ onto any state $|\psi_{2Q}\rangle$. To maximize the efficiency of the projection, the total transmission

$t_H^2 + t_V^2$ should be maximized based on the corresponding value of $\gamma \in [-1, 1]$ (set by the concurrence). For $\gamma \geq 0$, this is achieved by having $t_H^2 = 1$ and $t_V^2 = (1 - \gamma)/(1 + \gamma)$. If $\gamma \leq 0$, choosing $t_V^2 = 1$ and $t_H^2 = (1 + \gamma)/(1 - \gamma)$ maximizes projection efficiency. The optimal efficiency is then given by:

$$\eta_{\text{opt}} = \frac{1}{1 + |\gamma|} = \frac{1}{1 + \sqrt{1 - \mathcal{C}^2}} \quad (1.82)$$

To summarize, the method described combines two-photon interference with polarization dependent loss and unitaries in each spatial mode to obtain the probability $|\langle \psi_{2Q} | \phi \rangle|^2$ for the projection of $|\phi\rangle$ onto the entangled state $|\psi_{2Q}\rangle$. Regardless of the level of entanglement, the projector $|\psi_{2Q}\rangle\langle\psi_{2Q}|$ can be measured with an appropriate choice of polarization dependent loss amplitudes t_H and t_V , as well as unitaries U_a and U_b . This enables the realization of various collective measurements.

1.5 Historical Notes on Quantum State Estimation

The goal of quantum state estimation is to completely characterize the state of a quantum system by estimating its density matrix. This characterization enables the prediction of outcome statistics for any measurements performed on the system. Although a physical mechanism may be expected to prepare a specific state, experimental imperfections can result in a state that deviates from the theoretical expectation. Consequently, state characterization is essential for ensuring functionality in quantum devices and information processing protocols, while also being fundamentally instructive to study.

1.5.1 Back-action and Heisenberg Uncertainty Principle

As we will discuss, the complete characterization of quantum systems in general requires the measurement of a set of non-commuting observables. This presents a fundamental challenge posed by back-action and the Heisenberg uncertainty principle, which constrain our ability to simultaneously and precisely probe different quantum system properties. The (generalized) Heisenberg uncertainty principle states that the observables A and B for a quantum system in state $|\psi\rangle$ must satisfy the inequality [75]:

$$\Delta A \Delta B \geq \frac{|\langle \psi | [A, B] | \psi \rangle|}{2}, \quad (1.83)$$

where $[A, B] = AB - BA$. The uncertainty principle implies that there is an inherent minimum uncertainty in simultaneous knowledge of two non-commuting observables for a single quantum system, e.g., position and momentum, energy and time, or Pauli operators σ_x and σ_y .

To illustrate the uncertainty principle, it helps to consider the back-action effect when measuring two observables on the same system. Let A have eigenstates $\{|a_i\rangle\}$ with corresponding eigenvalues a_i , and B have eigenstates $\{|b_i\rangle\}$ with corresponding eigenvalues b_i . Suppose that initially, the system is in the state $|\psi\rangle = |b_1\rangle$. A measurement of A is made, and the outcome a_1 is obtained. The post-measurement state is now $|a_1\rangle$. A subsequent measurement of B would yield an outcome b_1 with probability $\text{Prob}(b_1) = |\langle b_1 | a_1 \rangle|^2$. However, had B been measured first, we would obtain the outcome b_1 with certainty, as $|\langle b_1 | \psi \rangle|^2 = |\langle b_1 | b_1 \rangle|^2 = 1$. The probability of obtaining b_1 would only remain unaffected if the two observables shared a complete set of eigenstates (i.e., if $\{|a_i\rangle\} = \{|b_i\rangle\}$). Non-commuting observables do not share a complete set of eigenstates [20]. The act of measuring A thus results in a superposition of eigenstates of B and changes the outcome statistics.

Therefore, the effect of back-action in the measurement of non-commuting observables during state estimation prevents the use of a single quantum system for complete state characterization.

1.5.2 No-Cloning Theorem

The no-cloning theorem asserts that a general unknown quantum state cannot be copied perfectly and deterministically. Stated differently, no quantum device can be constructed which outputs $|\psi\rangle \otimes |\psi\rangle$, given an arbitrary $|\psi\rangle$ [6, 76].

This theorem was proved using the linearity property of quantum mechanics by Wootters and Zurek [77], and by Dieks [78]. Suppose there exists a unitary operation U that performs perfect cloning:

$$U(|\psi\rangle |\text{vac}\rangle |M\rangle) = |\psi\rangle |\psi\rangle |M(\psi)\rangle, \quad (1.84)$$

where $|M\rangle$ is the state of the auxiliary cloning apparatus. Such a device would achieve:

$$U(|0\rangle |\text{vac}\rangle |M\rangle) = |0\rangle |0\rangle |M(0)\rangle \quad U(|1\rangle |\text{vac}\rangle |M\rangle) = |1\rangle |1\rangle |M(1)\rangle \quad (1.85)$$

Now consider the superposition $|\phi\rangle = \alpha |0\rangle + \beta |1\rangle$. By linearity and using Eq. (1.85), we obtain:

$$U(\alpha |0\rangle + \beta |1\rangle) |\text{vac}\rangle |M\rangle = \alpha |00\rangle |M(0)\rangle + \beta |11\rangle |M(1)\rangle. \quad (1.86)$$

However, we could apply Eq. (1.84) to $|\phi\rangle$ directly, which gives:

$$\begin{aligned} U(|\phi\rangle |\text{vac}\rangle |M\rangle) &= |\phi\rangle |\phi\rangle |M(\phi)\rangle \\ &= (\alpha |0\rangle + \beta |1\rangle)(\alpha |0\rangle + \beta |1\rangle) |M(\phi)\rangle \\ &= (\alpha^2 |00\rangle + \alpha\beta |01\rangle + \beta\alpha |10\rangle + \beta^2 |11\rangle) |M(\phi)\rangle. \end{aligned} \quad (1.87)$$

We see that there is a contradiction, as the states on the right-hand side of Eq. (1.86) and Eq. (1.87) are different. This contradiction shows that there is no perfect cloning device for general states. This result was extended to the case of mixed states and non-unitary cloning devices in Ref. [79], and to composite systems in Ref. [80].

The no-cloning theorem has a profound consequence for the state estimation of an unknown state: *we cannot circumvent the measurement back-action effect by creating identical copies to acquire all the information necessary for describing the state.*

1.5.3 Estimating the State from an Ensemble

Despite the fundamental limitations due to measurement back-action, the Heisenberg uncertainty principle, and the no-cloning theorem, an ensemble of identically prepared quantum systems (when available) can be used for quantum state estimation. The problem of estimating the state of a quantum system from an ensemble was first addressed by Fano in 1957 [81], who used the term *quorum* to describe a set of observables sufficient to completely determine the density matrix. Park and Band [4] theoretically discussed a method based on multipolar expansion of tensors for generating a quorum of $(2J+1)^2 - 1$ observables to determine any spin J state. Royer formulated a method for measuring the Wigner function [82], a quasi-probability distribution over the phase-space of an optical mode which has a one-to-one correspondence to the density matrix [83].

In more practical terms, quantum state tomography was theoretically proposed by Vogel and Risken [84] in 1989 in the context of determining the Wigner function for an optical mode. Vogel and Risken's method involved using homodyne detection to measure the probability distribution for various rotated quadrature operators. Balanced homodyne detection (Fig. 1.11) consists of interfering the signal beam with a local oscillator (strong coherent state reference beam) at a 50:50 beamsplitter, and measuring the intensity difference between the output beams. Optical homodyne tomography¹ was first experimentally

¹Inspired by medical imaging, the term "tomography" in reference to this type of quantum state estimation was coined in Ref. [85]

demonstrated in 1993 in Ref. [85], where the density matrices of a light mode for squeezed states and the vacuum were measured.

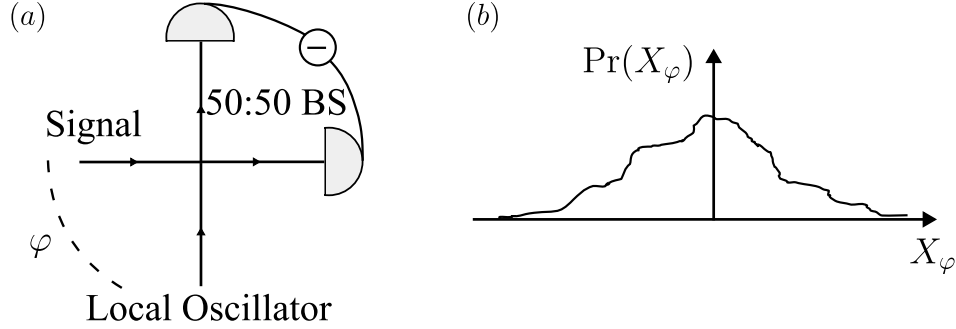


Figure 1.11: Balanced homodyne detection setup for quantum state tomography. (a) The signal field is interfered with a local oscillator (strong coherent state) at a 50:50 beamsplitter and the output photocurrent of the detectors is subtracted. (b) Cartoon of the marginal distribution $\text{Pr}(X_\varphi)$ of the rotated field quadrature X_φ (a slice of the Wigner function) for the signal obtained from balanced homodyne detection.

1.6 Quantum State Estimation

Quantum state estimation plays a crucial role in characterizing and understanding quantum systems. It involves using an ensemble of systems prepared in the same unknown state to perform known measurements. The measurement outcomes contain information about the unknown state. Depending on the size of the identical ensemble, a guessed state may correspond to each outcome (few available systems in the identically prepared ensemble), or a reconstruction algorithm may be used to obtain the state estimate from the measurement statistics (many systems available in the identically prepared ensemble). We briefly discuss each of these scenarios in this section.

Fig. 1.12 provides a diagram of the key elements of quantum state estimation. Some quantum state estimation strategies may involve adaptivity, where the choice of measurements is not entirely pre-determined. Instead, intermediate state estimates (or measurement outcomes) obtained throughout the experiment inform how the remaining systems in the identically prepared ensemble should be measured.

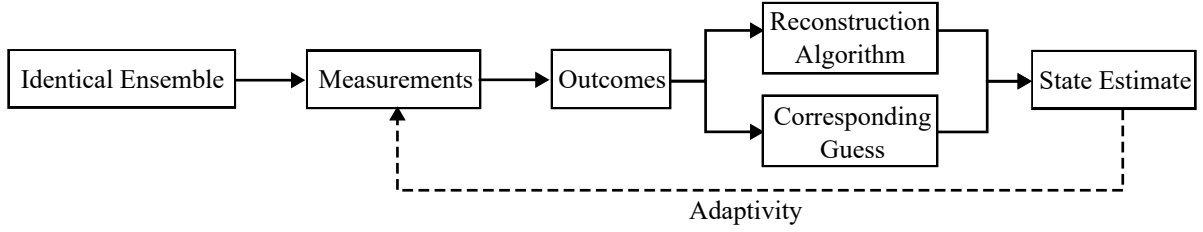


Figure 1.12: Key elements of quantum state estimation.

1.6.1 State Guessed in Correspondence to Each Outcome

When the number N of available systems in the identically prepared ensemble is too small for measuring the entire quorum, or acquiring sufficient statistics, the state may instead be guessed in correspondence to each measurement outcome. This was the approach adopted in Massar and Popescu’s seminal 1995 paper [50], where they showed that the optimal strategy for state estimation requires a collective measurement. For the case of two copies in the ensemble ($N = 2$), the optimal strategy proposed was a non-maximally entangling measurement with four possible outcomes. To each outcome corresponds a one-qubit guess, with the four possible guessed states being at the vertices of a tetrahedron on the Bloch sphere. In Chapters 2 and 3, this optimal measuring strategy, as well as its extensions to the case of mixed states and qudits, will be further discussed.

1.6.2 Quantum State Tomography

Given a sufficiently large ensemble of quantum systems prepared in the same state, quantum state tomography aims to determine the unknown state by a series of measurements. Each measurement operator provides information on a new dimension of the unknown state in the state-space, and additional statistics improve this knowledge. Quantum state tomography has been successfully used in a variety of quantum-mechanical systems [5, 85–90].

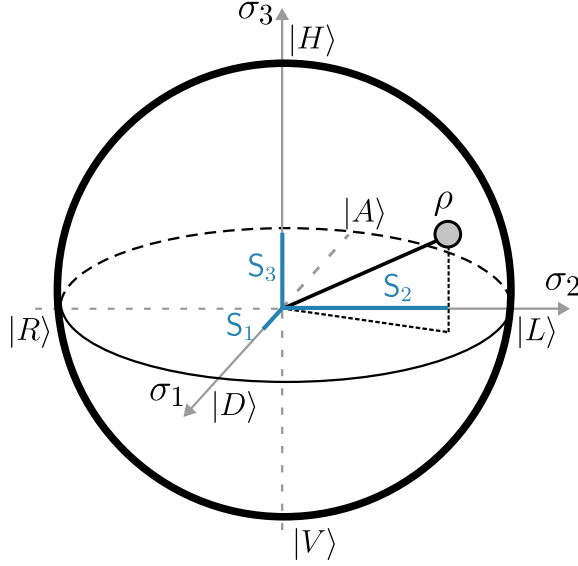


Figure 1.13: Poincaré sphere. The Stokes parameters correspond to the Cartesian coordinates for the state ρ in the $\langle\sigma_1\rangle, \langle\sigma_2\rangle, \langle\sigma_3\rangle$ space.

Any single qubit density matrix ρ can be uniquely represented by the Stokes parameters (S_1, S_2, S_3) , sometimes also called the Bloch vector, as [7, 91]:

$$\rho = \frac{1}{2} \sum_{i=0}^i S_i \sigma_i, \quad (1.88)$$

where $\sigma_0 = I$,

$$\sigma_1 = \sigma_x = |D\rangle\langle D| - |A\rangle\langle A| \quad (1.89)$$

$$\sigma_2 = \sigma_y = |L\rangle\langle L| - |R\rangle\langle R| \quad (1.90)$$

$$\sigma_3 = \sigma_z = |H\rangle\langle H| - |V\rangle\langle V|, \quad (1.91)$$

and the parameters S_i are given by $S_i \equiv \text{tr}(\sigma_i \rho)$ [92]. Fig. 1.13 shows the representation of an arbitrary state ρ and its corresponding Stokes parameters in the Poincaré sphere. Due to normalization, S_0 is always 1. The length $S_1^2 + S_2^2 + S_3^2 \leq 1$, being exactly one for pure states. We refer to the type of tomography that involves measuring the projections in the horizontal, diagonal, and circular bases as Pauli tomography. Analogous parameters to S_i can be defined for any three matrices $|\psi_i\rangle\langle\psi_i|$ provided $|\psi_i\rangle$ are linearly independent.

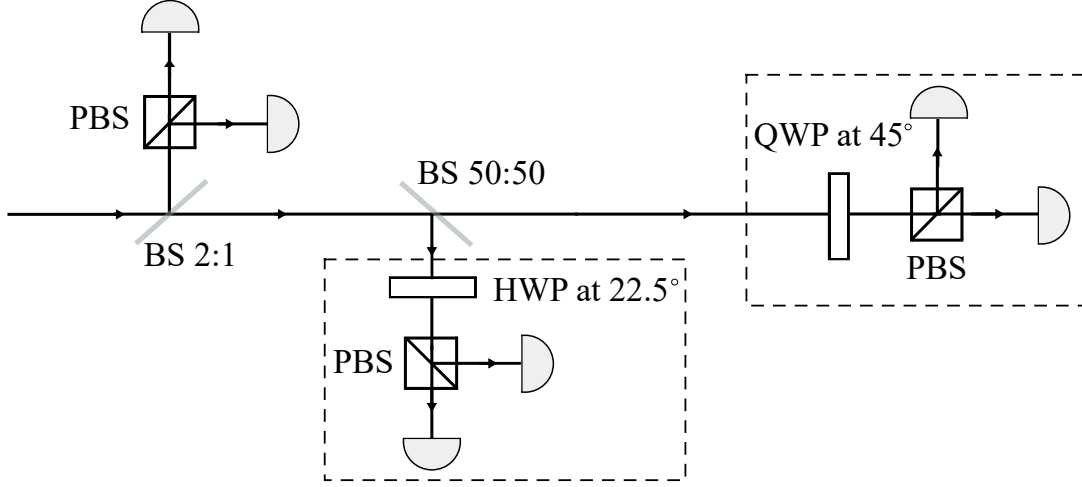


Figure 1.14: Six-outcome Pauli tomography setup. A 2:1 beamsplitter (BS) and a 50:50 BS evenly distribute the light among the setups for measuring each Pauli operator. A half-wave plate (HWP) and a quarter-wave plate (QWP) at the specified angles with respect to their fast axis, followed by a polarizing beamsplitter (PBS), allow the measurement of the six-outcome Pauli POVM.

The Stokes parameters $S_i \equiv \text{tr}(\sigma_i \rho)$ can be obtained from the following six-outcome POVM [93]:

$$\left\{ \frac{I \pm \sigma_1}{6}, \frac{I \pm \sigma_2}{6}, \frac{I \pm \sigma_3}{6} \right\} = \left\{ \frac{|D\rangle\langle D|}{3}, \frac{|A\rangle\langle A|}{3}, \frac{|L\rangle\langle L|}{3}, \frac{|R\rangle\langle R|}{3}, \frac{|H\rangle\langle H|}{3}, \frac{|V\rangle\langle V|}{3} \right\}, \quad (1.92)$$

which consists of projecting onto the eigenstates of the Pauli operators. The experimental setup for six-outcome Pauli tomography is shown in Fig. 1.14. The combination of a quarter-wave plate (QWP) and half-wave plate (HWP) placed at the indicated angles relative to their optic axis followed by a polarizing beamsplitter (PBS) and photo-detection allows the measurement of each Pauli observable.

Typical sources of imperfection in Pauli tomography of polarization qubits include waveplate retardance deviations, waveplate angle misalignment, PBS imperfect extinction, and detector dark noise. To account for some systematic errors, we could perform an initial characterization of the device imperfections using well-known reference states. The calibration results can then be used as an error-mitigation strategy to correct for systematic errors [94].

More generally, any n -qubit state ρ can be represented as [7]:

$$\rho = \frac{1}{2^n} \sum_{i_1, i_2, \dots, i_n=0}^3 S_{i_1, i_2, \dots, i_n} \sigma_{i_1} \otimes \sigma_{i_2} \cdots \otimes \sigma_{i_n}, \quad (1.93)$$

where $S_{i_1, i_2, \dots, i_n}$ are real parameters analogous to Stokes parameters, with $S_{0,0,\dots,0} = 1$, meaning there are $4^n - 1$ real parameters.

We note that for a general n -qubit state, we cannot reconstruct the entire state from local measurements without access to correlations (i.e., without classical communication). The generalized Stokes parameters that uniquely determine an n -qubit state require the measurement of correlations corresponding to expectations of the tensor products of Pauli observables $\sigma_{i_1} \otimes \dots \otimes \sigma_{i_n}$, which can be obtained using local operations and classical communication (e.g., via coincidence measurements in photonics).

What we can learn from local measurements without correlations are expectations of observables such as $\sigma_{i_1} \otimes \mathbb{I} \otimes \dots \otimes \mathbb{I}$, hence determining the reduced density matrices. This is sufficient in some cases, such as for separable states. However, if the n -qubit state is maximally entangled, all reduced density matrices are maximally mixed, so local measurements alone would yield data indistinguishable from those of a maximally mixed state.

Informationally Complete Set of Operators

In the case of exact quantum state tomography, the reconstructed state converges to the true state in the limit as the size of the identically prepared ensemble approaches infinity, as I will explain. Exact tomography in a d -dimensional state-space requires an informationally complete set of d^2 operators, otherwise known as a quorum [95, 96]. A quorum spans the space of self-adjoint operators acting on the state-space. In particular, any density matrix can be written as a linear combination of quorum operators. Therefore, measuring the expectation of these operators completely determines the state of the system. The requirement on the size of the quorum can be intuitively understood from the perspective of determining the $d^2 - 1$ real free parameters of the Hermitian and trace-one density matrix. In experiments, such as in optics, where detector efficiency and light intensity is not exactly known, an additional operator needs to be measured for trace normalization.

Quorums may be minimally complete (having d^2 outcomes) or overcomplete (with more than d^2 outcomes). Thus, the six-outcome Pauli POVM in Eq. (1.92), with projectors onto all eigenstates of all Pauli operators, is an example of an overcomplete quorum for one qubit states.

One example of a minimally complete quorum for one qubit tomography is the set of projections [5]:

$$\{\mu_0 = |H\rangle\langle H| + |V\rangle\langle V|, \mu_1 = |H\rangle\langle H|, \mu_2 = |A\rangle\langle A|, \mu_3 = |R\rangle\langle R|\}. \quad (1.94)$$

Another common minimally complete quorum for one qubit tomography is the set of projections:

$$\{\mu'_0 = |H\rangle\langle H|, \mu'_1 = |V\rangle\langle V|, \mu'_2 = |D\rangle\langle D|, \mu'_3 = |R\rangle\langle R|\}. \quad (1.95)$$

A quorum for an n -qubit system can be obtained from 4^n projection measurements $\mu_{i_1} \otimes \mu_{i_2} \otimes \cdots \otimes \mu_{i_n}$ ($i_k \in \{0, 1, 2, 3\}$ and $k \in \{1, 2, \dots, n\}$).

A symmetric informationally complete POVM Π (SIC-POVM) consists of d^2 linearly independent rank-one elements Π_j such that $\text{tr}(\Pi_j) = 1/d$, and [93]

$$\text{tr}(\Pi_j \Pi_k) = \frac{d\delta_{jk} + 1}{d^2(d+1)}. \quad (1.96)$$

The elements Π_j all have identical pair-wise overlap. Intuitively, they probe the state-space in a maximally symmetric and uniformly distributed way. An example of a minimally-complete POVM for one qubit states that is also a SIC-POVM is $\{|\uparrow_{n_i}\rangle\langle\uparrow_{n_i}|/2\}$, where the states $|\uparrow_{n_i}\rangle$ are at the vertices of a tetrahedron in the Bloch sphere:

$$\begin{aligned} |\uparrow_{n_1}\rangle &= |0\rangle \\ |\uparrow_{n_2}\rangle &= \frac{1}{\sqrt{3}}|0\rangle + \sqrt{\frac{2}{3}}|1\rangle \\ |\uparrow_{n_3}\rangle &= \frac{1}{\sqrt{3}}|0\rangle + e^{i2\pi/3}\sqrt{\frac{2}{3}}|1\rangle \\ |\uparrow_{n_4}\rangle &= \frac{1}{\sqrt{3}}|0\rangle + e^{i4\pi/3}\sqrt{\frac{2}{3}}|1\rangle. \end{aligned} \quad (1.97)$$

Reconstruction Algorithms

The reconstruction of the state from the measured data can proceed in various ways, including maximum likelihood estimation (MLE), linear inversion, or estimation of the expectations corresponding to $\mathbf{S}_i$. Suppose that an informationally-complete POVM $\{\Pi_j\}$ has been measured on an identical ensemble of size N prepared in the exact state ρ_{exact} , resulting in a sequence $\{n_j\}$ of detection occurrences ($\sum_j n_j = N$). Further, we define the frequencies $f_j = n_j/N$, such that $\sum_j f_j = 1$.

In linear inversion, we consider the Born probabilities p_j^{exact} for the exact state ρ_{exact} as being equal to the measured frequencies f_j . Born's rule applied to the quorum is then inverted to determine the density matrix. For over-complete measurements, this involves the use of a pseudo-inverse, which is equivalent to least-squares fitting of linear equations.

Linear inversion can allow unphysical states, in that the eigenvalues of the estimate ρ_{LI} , corresponding to probabilities of the mixture, can be negative and in some cases the purity $\text{tr}(\rho_{\text{LI}}^2) > 1$. This can be understood by considering, for instance, states lying on the boundary of the Bloch sphere. For such states, the uncertainty from estimation based on finite N — no matter how small — will always include states lying outside the Bloch sphere [93].

More commonly, MLE may be used as a reconstruction method. The likelihood for the sequence $\{n_j\}$ of measurement occurrences for an informationally-complete POVM $\{\Pi_j\}$ is given by [93, 97]:

$$\mathcal{L}(\rho; \{n_j\}) = \prod_j^n p_j^{n_j}, \quad (1.98)$$

where p_j represent the Born probabilities $p_j = \text{tr}(\rho \Pi_j)$, with ρ being a general density matrix. Although it is not concave, the likelihood has a unique maximum as a function of the Born probabilities. This maximum is in general found using an iterative numerical optimization. It may be preferable to use the log-likelihood function

$$\log \mathcal{L}(\rho; \{n_j\}) = \sum_j n_j \log p_j, \quad (1.99)$$

which is concave, as it is a (weighted) sum of the concave functions $\log p_j$. Both linear inversion and MLE provide *consistent* estimators: they approach the true state as the size N of the identically prepared ensemble increases, since the measurement frequencies f_j approach the exact probabilities p_j^{exact} . However, the true maximum of the likelihood function can lie outside the physically-allowed state-space of Hermitian positive semi-definite matrices with trace unity. Maximization while constraining the estimator to physical (positive semi-definite) density matrices can thus lead to rank-deficient reconstructions lying on the boundary of the physical state-space, which have at least one zero eigenvalue. Moreover, MLE may suffer from a lack of uncertainty quantification and computationally-intensive reconstruction for large quantum systems [93, 98, 99].

1.6.3 Prior Distribution of States

It is important to describe the prior distribution of the unknown input states over the state-space in identically prepared ensembles for state estimation. As we will see, this is an important consideration when examining the performance of state estimation schemes. In the absence of prior knowledge, we would like the state-space to be sampled uniformly at random with equal probability for the preparation of identical ensembles in each state.

We recall that an arbitrary pure qubit state can be represented in terms of the polar angle θ and azimuthal angle ϕ . However, drawing $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi)$ at random with uniform probability does not achieve our objective of uniformly sampling the surface of the Bloch sphere; it results in an oversampling near the poles.

Instead, our goal can be achieved by considering the unique uniform spherical probability measure $\mu(\theta, \phi) = \sin\theta d\theta d\phi/(4\pi)$, which is invariant under the action of any unitary operator. This corresponds to the probability distribution $P(\theta) = \sin\theta/(4\pi)$. Sampling from this distribution can be done by taking the amplitudes α and β of a state $|\vec{n}\rangle = 1/\sqrt{|\alpha|^2 + |\beta|^2}(\alpha|0\rangle + \beta|1\rangle)$ to be $\alpha = x_1 + iy_1$ and $\beta = x_2 + iy_2$, where x_i and y_i are independent and identically distributed normal random variables with mean zero and variance one [100].

Uniformly sampling the surface of the unit sphere can alternatively be approached from the perspective of a distribution over unitaries. This is rigorously established through the Haar measure defined over unitaries acting on the state-space, with the defining property that the measure is invariant under left and right multiplication by every unitary operator [100, 101]. Consequently, sampling uniformly at random from the surface of the unit sphere can be accomplished by applying a Haar-random unitary transformation to a reference state. This method yields the same distribution of states over the unit sphere as described previously. Given this close connection, we may use the terms ‘‘Haar-random state’’ to refer to states distributed according to $P(\theta) = \sin\theta/(4\pi)$.

The choice of prior distribution is less apparent for mixed qubit states, as there is no unique unitarily invariant distribution in this case. One common distribution considered for mixed states is the unitarily invariant Bures distribution, argued to represent maximal randomness of input states [102]. It has the radial probability distribution [103]

$$P_B(r) = \frac{4}{\pi} \frac{r^2}{\sqrt{1-r^2}}, \quad (1.100)$$

where $r = |\vec{S}| \leq 1$ is the length of the Bloch vector in the unit sphere. This distribution is a natural choice when considering the fidelity and the Bures distance as figures of merit, which will be introduced in the next section.

1.7 Fidelity

Quantum state estimation is frequently used to evaluate the quality of a prepared quantum state in experimental settings. Therefore, it is essential to have a means of quantifying the

proximity between the estimated state and the target state. Elsewhere, we may want to assess the performance of a state estimation scheme based on how the error with the true state scales as more systems are measured.

For the purpose of evaluating the closeness between two pure states $|\psi\rangle$ and $|\phi\rangle$, it is natural to consider the quantity

$$F(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi\rangle|^2, \quad (1.101)$$

which is the probability of one state to be found in the other. This is known as the fidelity between the two states.

A similar definition can be extended to mixed states, but it requires a more nuanced approach. For two mixed states ρ and σ , there are many possibilities for quantifying closeness. For example, one may first consider $\text{Tr}(\rho\sigma)$, but this quantity is less than one for a mixed state ρ when $\rho = \sigma$. Instead, the fidelity $F(\rho, \sigma)$ between two mixed states is given by [104, 105]

$$F(\rho, \sigma) = \text{tr} \left[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}} \right]^2, \quad (1.102)$$

ranging from zero to one, with $F(\rho, \sigma) = 1$ if and only if $\rho = \sigma$. The fidelity is symmetric, meaning $F(\rho, \sigma) = F(\sigma, \rho)$, and is also preserved by unitary evolution. When at least one of the states ρ or σ is pure, e.g., $\rho = |\psi\rangle\langle\psi|$, the fidelity does reduce to [105]:

$$F(\rho, \sigma) = \text{Tr}(\rho\sigma) = \langle\psi|\sigma|\psi\rangle. \quad (1.103)$$

In the context of state estimation, it is common to consider the infidelity $1 - F(\rho_{\text{exact}}, \rho_{\text{est}})$ to quantify the error in ρ_{est} compared to the exact state ρ_{exact} . While the fidelity is not a metric or a distance function, it is closely related to the Bures distance [106]

$$d_B^2(\rho_{\text{exact}}, \rho_{\text{est}}) = 2(1 - \sqrt{F(\rho_{\text{exact}}, \rho_{\text{est}})}). \quad (1.104)$$

For pure states, the Bures distance simply corresponds to the Euclidean distance between the states' Bloch vectors.

In the case of polarization qubit states, the fidelity in terms of Stokes parameters can be expressed as [104]:

$$F(\vec{S}_a, \vec{S}_b) = \frac{1}{2} \left[1 + \sum_{i=1}^3 S_{i,a} S_{i,b} + \sqrt{(1 - |\vec{S}_a|^2)(1 - |\vec{S}_b|^2)} \right]^2. \quad (1.105)$$

If one of the two states, for example $\vec{S}_b$ is fully polarized (pure), $|\vec{S}_b|^2 = 1$ and the fidelity reduces to:

$$F(\vec{S}_a, \vec{S}_b) = \frac{1}{2}(1 + \mathbf{S}_{1,a}\mathbf{S}_{1,b} + \mathbf{S}_{2,a}\mathbf{S}_{2,b} + \mathbf{S}_{3,a}\mathbf{S}_{3,b}). \quad (1.106)$$

Let α denote the angle between the Bloch vectors $\vec{S}_a$ and $\vec{S}_b$ for two pure states. Then from Eq. (1.106), their fidelity can be rewritten in terms of the angle as:

$$F(\vec{S}_a, \vec{S}_b) = \frac{1}{2}(1 + \cos \alpha) = \cos^2 \left(\frac{\alpha}{2} \right). \quad (1.107)$$

Thus the infidelity can also be expressed in terms of the angle α as

$$1 - F(\vec{S}_a, \vec{S}_b) = \sin^2 \left(\frac{\alpha}{2} \right). \quad (1.108)$$

1.7.1 Optimality in the Sense of Average Fidelity

It is natural to ask whether choosing a particular set of measurements is preferable over others for state estimation. For example, we may consider the situation where we have access to N identically prepared qubits. We know that they are all prepared in the same pure state, and want to best estimate the state $|\vec{n}\rangle$, where $\vec{n} = (\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3)$ is the Bloch vector. To quantify this, we may require that should we repeat this “game” for many rounds with the qubit states drawn from some distribution $P(\vec{n})$, the fidelity $F(|\vec{n}\rangle, |\vec{n}_g\rangle)$ between the true state $|\vec{n}\rangle$ and our guess $|\vec{n}_g\rangle$ be on average maximized. The average fidelity may be expressed as the following sum over the possible measurement outcomes g and integral over the state-space [50, 107]:

$$\begin{aligned} \bar{F} &= \int P(\vec{n}) d\vec{n} \sum_g P(g|\vec{n}) F(|\vec{n}\rangle, |\vec{n}_g\rangle) \\ &= \int P(\vec{n}) d\vec{n} \sum_g P(g|\vec{n}) \left(\frac{1 + \vec{n} \cdot \vec{n}_g}{2} \right), \end{aligned} \quad (1.109)$$

where $P(g|\vec{n})$ is the conditional probability of guessing $\vec{n}_g$ when the true direction is $\vec{n}$. This average fidelity accounts for each possible true state drawn from the distribution $P(n)$ and each possible measurement outcome g . In going to the second line, the fidelity $F(|\vec{n}\rangle, |\vec{n}_g\rangle)$ is expressed as a dot product between the Bloch vectors using Eq. (1.106).

Note that the projection states associated with probabilities $P(g|\vec{n})$ are, in general, different from the guesses $|\vec{n}_g\rangle$ themselves. In fact, as we shall see in Chapters 2 and 3, we may project onto an entangled state of multiple particles by performing a collective measurement. The guess $|\vec{n}_g\rangle$ of course remains a single-qubit state as we are estimating the unknown single-qubit state $|\vec{n}\rangle$.

1.7.2 Average Infidelity Scaling with Sample Size: Gill-Massar Bound

In the case where the number of available copies N is sufficiently large, state estimation may be considered from the point of view of multi-parameter estimation. For example, the parameters would be the polar and azimuthal angles θ and ϕ for a pure state, or the Stokes parameters S_1, S_2 , and S_3 for a mixed state. Through an analysis based on multi-parameter estimation theory, Gill and Massar [108] found that the optimal tomography of pure qubit states reaches an asymptotic average infidelity of $1 - \bar{F}_p = 1/N$. This can be attained using two-stage adaptive measurement schemes [46, 109] or random local von Neumann measurements [110]. As such, the performance gap between collective and LOCC strategies for pure qubit state estimation is bridged in the asymptotic limit [54, 109].

For mixed qubit states, a different bound applies. Specifically, when the set of available measurements is restricted to LOCCs, the average infidelity is bounded by $1 - \bar{F}_m = 9/(4N)$ [54, 108]. This bound, known as the Gill-Massar bound for mixed qubit states, is attainable using two-stage adaptive schemes [46, 54]. It is worth noting that for mixed qubit state estimation, the average fidelity attainable using collective measurements exceeds that of LOCC strategies, even in the asymptotic limit [54]. The constant in the Gill-Massar bound for mixed states applies for general prior distributions of states.

For specific priors, it has been shown in Ref. [110] that the asymptotic optimal infidelity (up to leading order in $1/N$) is given by

$$1 - F = 1 - \frac{3 + 2 \langle r \rangle}{4N}, \quad (1.110)$$

where $\langle r \rangle$ is the mean purity over the prior distribution, and is defined as

$$\langle r \rangle \equiv \int_0^1 dr w(r) r. \quad (1.111)$$

For example, for the Bures prior (Eq. 1.100), the average purity is $\langle r \rangle \approx 0.85$ and the coefficient is $(3 + 2 \langle r \rangle)/4 \approx 1.17$.

1.7.3 State Fidelity Values in Applications

The target minimum value of state fidelity depends on the application. For example, for two-qubit Bell pairs: (1) the provably entangled and distillable threshold for a Bell state

is $F > 1/2$ [17]; (2) in quantum teleportation, to outperform classical schemes, the shared Bell pair must satisfy $F > 2/3$ [111]; (3) the Bell inequality (CHSH) violation threshold is $F > 1/2(1 + 1/\sqrt{2}) \approx 0.854$ [112].

The following caveats should be considered when assessing reported state fidelity values: (1) the fidelity depends on the number of copies used to estimate the state, and is therefore more meaningful when reported alongside the measured number of copies; (2) the state of the art depends on the experimental platform (photonic, superconducting, trapped ion, or neutral atom); and (3) the achievable fidelity generally decreases with system size.

Refs. [46, 113, 114] have shown near-optimal infidelity scaling ($\sim O(1/N)$) for one-qubit and two-qubit polarization states by using adaptive tomography. As another example, focusing on one-qubit computational states $|0\rangle$ and $|1\rangle$, Ref. [115] reports an infidelity of $1 - F = (9.0 \pm 1.3) \times 10^{-5}$ from millions of system preparations of ion hyperfine qubits.

For superconducting multi-qubit systems in entangled states distributed across the modules of a quantum processor, Ref. [116] reports state fidelities of $(88.15 \pm 0.24)\%$ and $(75.18 \pm 0.11)\%$ with maximally entangled three-qubit and four-qubit GHZ states, respectively. In larger experiments, such as in Ref. [117] where a neutral-atom quantum computer was used to produce up to 60-qubit entangled states, the target has thus far been to demonstrate a fidelity $F > 0.5$. That is because even slight imperfections in each constituent gate accumulate rapidly across many qubits, reducing the final state fidelity.

1.7.4 Average Gate Fidelity

The performance of a trace-preserving, yet imperfect transformation $\tilde{U}$ can be compared to the theoretically expected unitary U using the average gate fidelity ranging from zero to one, with an average gate fidelity of one corresponding to perfect agreement. The average gate fidelity for a d -dimensional state-space can be calculated as [118]

$$\bar{F}_g(U, \tilde{U}) = \frac{\sum_{j=0}^{d^2-1} \text{tr}[(U U_j^\dagger U^\dagger)(\tilde{U} U_j \tilde{U}^\dagger)] + d^2}{d^2(d+1)}, \quad (1.112)$$

where $U_j/\sqrt{d}$ form an orthonormal operator basis. Orthonormality here means $\text{tr}(U_j^\dagger U_k) = \delta_{jk}d$. The set of operators $\{U_j\}$ can be found by computing $X^k Z^\ell$, $k, \ell \in \{0, \dots, d-1\}$, which are products of the shift and clock unitaries (Eq. (1.17)-(1.18)). The action of X

and Z on the computational basis states $|j\rangle$ is simply defined by

$$X|j\rangle \equiv |j \oplus 1\rangle \quad (1.113)$$

$$Z|j\rangle \equiv e^{2\pi i j/d} |j\rangle, \quad (1.114)$$

where $j \in \{0, \dots, d-1\}$ and $\oplus$ denotes addition modulo d . The average gate fidelity essentially gives the average probability that the approximation $\tilde{U}$ produces the correct output state compared to the ideal output, for a uniformly random input state.

For most purposes, we may only need to experimentally achieve a unitary transformation up to a global phase. The average gate fidelity is in this sense a suitable choice, as it is insensitive to global phases. That is, $\bar{F}_g(e^{i\gamma_1}U, e^{i\gamma_2}U) = 1$ for all unitaries U , and phases γ_1 and γ_2 .

In quantum computing, the physical error rate and gate fidelity of quantum gates set the number of physical qubits that would be needed to attain a given logical error rate when using surface error-correcting codes. For example, in Ref. [119], a gate fidelity above 99.9% (physical error rate below 10^{-3}) has been stated as a good target, as a logical qubit would need to contain $10^3 - 10^4$ physical qubits to achieve logical error rates of $10^{-14} - 10^{-15}$. This would be enough to perform Shor's factorization algorithm with reasonable success.

While there are other valid ways of quantifying similarity for unitaries, it is worth emphasizing that more intuitive formulas could fail in some cases. Consider for example the following formula, present in literature, based on the overlap between the transformed computational basis states:

$$\bar{f}_g = ? \frac{1}{d} \sum_{j=0}^{d-1} |\langle j | \tilde{U}^\dagger U | j \rangle|^2. \quad (1.115)$$

Consider the scenario where $U = \sigma_z$ and $\tilde{U} = \mathbb{I}$. Then it is easy to verify that the formula fails, as $\bar{f}_g = 1$ for two clearly distinct unitaries.

1.8 Outline

In what follows, Chapter 2 gives an overview of optimal quantum state estimation in the low- N and asymptotic regimes, addressing the role of collective measurements and adaptivity. My Monte Carlo simulation results highlight the effect of *a priori* knowledge

about state purity on the infidelity scaling observed for Pauli and adaptive single-qubit tomography.

In Chapter 3, I report on a proof-of-concept experimental realization of Massar and Popescu’s optimal collective measurement for quantum state estimation. Our results support the prediction that the average fidelity of a state estimation game using this multi-particle collective strategy can exceed the best separable measurement strategy. We also applied this collective measurement to quantum state tomography by dividing a large ensemble in pairs. The experimental tomographic infidelity scaling results observed using this collective strategy are discussed.

I present a fast and automated polarization compensation scheme of unwanted unitaries in Chapter 4, which we devised and experimentally demonstrated. I discuss the characterization of liquid crystal variable retarders (LCVRs), the rotating QWP tomography method, and our polarization compensation results. We successfully demonstrated polarization compensation with up to 99% fidelity within an average of fewer than six steps.

In Chapter 5, I discuss the use of a sequence of thin phase gratings and free-space propagation for achieving spatial unitary transformations. Our treatment, based on the use of a matrix model, aims to reduce the number of parameters compared to more intensive methods, such as a wavefront matching optimization algorithm. Our numerical optimization results show the potential of this approach for approximating unitaries such as Pauli gates, beamsplitters, and discrete Fourier transforms.

Finally, I summarize our main findings, and discuss future directions in Chapter 6.

Chapter 2

Optimal Quantum State Estimation

2.1 Motivation

The minimum number of required projection measurements for tomographic state estimation of a quantum system of size d scales as $d^2 - 1$ [4, 5], resulting in an exponential scaling with the number of systems. This can be intuitively understood by considering the number of real free parameters of a general density matrix ρ . Since ρ is an observable, it corresponds to a $d \times d$ Hermitian matrix acting on the d -dimensional state-space. Hermiticity implies the diagonal elements must be real-valued, and entries below the diagonal are entirely determined by entries above the diagonal: This reduces the number of real free parameters to:

$$\text{dim} = \text{Num. diagonal entries} + \frac{\text{Re and Im parts}}{2} \times \text{Num. entries above the diagonal} \quad (2.1)$$

$$= \text{Num. diagonal entries} + 2 \times \left(\frac{\text{Total entries excluding the diagonal}}{2} \right) \quad (2.2)$$

$$= d + 2 \left(\frac{d^2 - d}{2} \right) \quad (2.3)$$

$$= d^2. \quad (2.4)$$

Imposing the additional constraint of trace unity — which ensures the probabilities for measurement outcomes sum to one — means there is one fewer free parameter. This gives the total of $d^2 - 1$ real free parameters for the density matrix.

Additionally, the accurate characterization of quantum systems often requires significant statistics, relying on measuring many quantum systems per projection measurement.

Therefore, the resource-efficiency in the size N of the identically prepared ensemble is of great relevance.

Practically, this can help provide better characterization in the development of quantum devices and implementation of quantum information processing tasks. Even fault-tolerant quantum computing requires satisfying minimum accuracy thresholds [120–122]. Complete state estimation is the most informative way of diagnosing the types of imperfections in the prepared state. Additionally, it is of fundamental interest to seek optimal measurements, as the infidelity scaling with the size of the identically prepared ensemble is indicative of the time required for extracting information from quantum systems. Finding the limits of how efficiently information can be extracted with the use of different classes of measurements could further our understanding of the no-cloning theorem [107, 123, 124] and the no-signalling theorem [125].

We first examine optimal state estimation in the case where the size N of the identical ensemble may not be large (statistical). In this scenario, there may not be sufficient identically prepared systems for measuring or accurately estimating an informationally complete set of operators. Thus, we may proceed by providing a state guess corresponding to each measurement outcome rather than using a tomographic reconstruction. We discuss Massar and Popescu’s insight that the optimal measurement for pure qubit states — such that average fidelity is maximized — involves treating the ensemble as one entity. We briefly mention extensions of this work, treating optimal state estimation for more general quantum systems.

We then turn to the asymptotic case, where N is sufficiently large for tomographic estimation. The infidelity scaling with N is introduced as an indicator of the efficiency of tomographic methods. An overview of some methods in adaptive tomography is then given.

The original contributions in this chapter are Monte Carlo simulations that highlight the effect of the *a priori* pure state assumption during tomography of pure states. We discuss how this assumption affects the infidelity scaling of tomography methods, and in what circumstances adaptive tomography provides a scaling advantage.

2.2 Theoretical Framework

2.2.1 State Estimation Game

Consider the following state estimation game: Given N qubits, each in the same pure state $|\vec{n}\rangle$, we are tasked to measure them and provide a single guess $|\vec{n}_g\rangle$ of the unknown one qubit state. The number of available identically prepared systems N may not be large (statistical). The procedure is repeated many times, with $|\vec{n}\rangle$ given according to some prior distribution $P(\vec{n})$, e.g., uniformly distributed over the Bloch sphere. The question is to find the optimal measurement, which on average maximizes the fidelity $\mathcal{F}(|\vec{n}\rangle, |\vec{n}_g\rangle) = |\langle \vec{n} | \vec{n}_g \rangle|^2$, subject to unitarity constraints regarding the time-evolution of the interaction with the measured system. The average fidelity $\bar{\mathcal{F}}$ can be written as [50, 107]:

$$\begin{aligned}\bar{\mathcal{F}} &= \int P(\vec{n}) d\vec{n} \sum_g P(g|\vec{n}) \mathcal{F}(|\vec{n}\rangle, |\vec{n}_g\rangle) \\ &= \int P(\vec{n}) d\vec{n} \sum_g P(g|\vec{n}) \left(\frac{1 + \vec{n} \cdot \vec{n}_g}{2} \right),\end{aligned}\tag{2.5}$$

where $\vec{n}$ is the Bloch vector of the unknown state $|\vec{n}\rangle$, $\vec{n}_g$ is the Bloch vector associated with the guessed state, and $P(g|\vec{n})$ is the probability of guessing $\vec{n}_g$ when the true direction is $\vec{n}$.

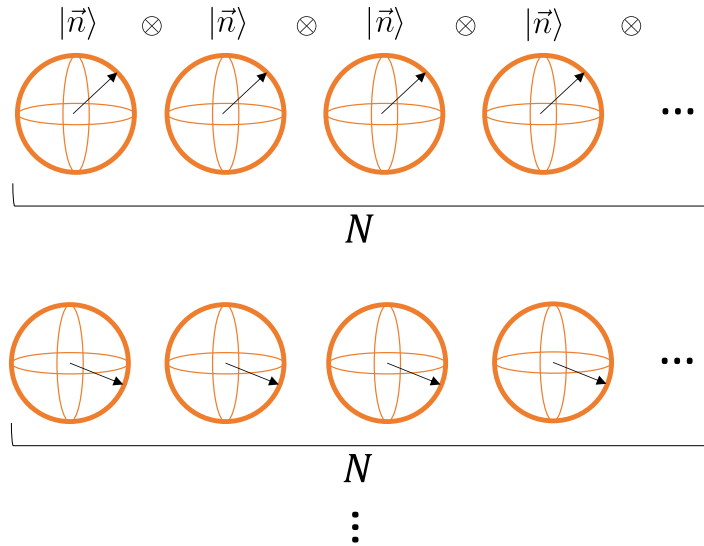


Figure 2.1: Ensemble of identically prepared qubits.

2.2.2 Massar and Popescu Optimal Collective Measurement

Massar and Popescu found that when the different directions $\vec{n}$ are uniformly distributed over the unit sphere, i.e., $P(\vec{n})d\vec{n} = \sin\theta d\theta d\phi/(4\pi)$, the optimal measuring strategy is to perform a collective measurement on the ensemble of N qubits. This strategy maximizes the average fidelity to $\bar{\mathcal{F}}_{\text{GenMP}}^{\text{Coll}} = (N+1)/(N+2)$ [50]. For the case of $N=2$, an optimal measuring strategy is to measure a nondegenerate operator whose eigenstates $|\text{MP}_i\rangle$ are the following superpositions of the singlet state $|S\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ and product tetrahedron states $|\uparrow_{n_i}\uparrow_{n_i}\rangle$:

$$\begin{aligned} |\text{MP}_1\rangle &= \frac{1}{2}|S\rangle + \frac{\sqrt{3}}{2}|\uparrow_{n_1}\uparrow_{n_1}\rangle \\ |\text{MP}_j\rangle &= -\frac{1}{2}|S\rangle + \frac{\sqrt{3}}{2}|\uparrow_{n_j}\uparrow_{n_j}\rangle, \quad j=2,3,4. \end{aligned} \tag{2.6}$$

The entangled states $|\text{MP}_i\rangle$ all have the same concurrence $\mathcal{C}(|\text{MP}_i\rangle) = 0.25$. In the Bloch sphere, the states $|\uparrow_{n_i}\rangle$ are located on the surface of the sphere at the vertices of a tetrahedron and can be chosen as:

$$\begin{aligned} |\uparrow_{n_1}\rangle &= |0\rangle; \\ |\uparrow_{n_j}\rangle &= \frac{1}{\sqrt{3}}|0\rangle + \sqrt{\frac{2}{3}}e^{i\alpha_j}|1\rangle, \quad j=2,3,4, \end{aligned} \tag{2.7}$$

with $\alpha_2, \alpha_3, \alpha_4 = 0, 2\pi/3, 4\pi/3$. Upon performing this collective measurement and obtaining the outcome i (corresponding to $|\text{MP}_i\rangle$), the tetrahedron state $|\uparrow_{n_i}\rangle$ is chosen as the guess $|\vec{n}_g\rangle$.

While Massar and Popescu did not provide an explicit optimal measurement with finite outcomes for arbitrary N , they showed that any optimal measurement for this problem (assuming more than one copy is available) requires treating the ensemble as one entity, and performing a collective measurement.

When allowing for an infinite set of possible outcomes, an optimal strategy for arbitrary N consists of measuring the POVM with elements proportional to the isotropically rotated spin-state projector $|j, m\rangle\langle j, m| = |N/2, N/2\rangle\langle N/2, N/2|$ [50, 126, 127]. The projection state $|N/2, N/2\rangle$ corresponds to a state of maximal spin $j = N/2$ with maximal spin component $m = N/2$ in some direction.

2.2.3 Extensions of the State Estimation Game

Pure Qubit State Estimation Game: Finite POVM

The state estimation game was further considered for arbitrary finite-dimensional quantum systems in a pure state in Ref. [127]. The authors provided an algorithm that specifies an optimal POVM with a *finite* number of outcomes (i.e., $(N + 1)^2$) for state estimation of an ensemble of N identically prepared quantum systems. This goes beyond the infinite continuous optimal measurement set previously proposed for the case of N qubits in Ref. [50]. A finite number of outcomes allows a physically realizable measurement. A further extension was made in Ref. [128], where optimal *minimal* measurements (with the minimum number of outcomes) were found.

Mixed Qubit State Estimation Game

The state estimation game based on an ensemble of N identically prepared mixed state qubits was treated in Ref. [129], where an optimal and minimal measuring strategy was found. The strategy involves projecting onto each subspace of total spin s , and within each subspace, onto spin eigenstates with maximal total spin component in some direction. In this way, a connection between the optimal measurements for mixed states is established with optimal measurements corresponding to a smaller number of copies of pure states. The best guesses are, in general, mixed states that depend on the prior distribution, as does the value of the maximum average fidelity.

Pure Qudit State Estimation Game

The state estimation problem was also considered for d -level pure state systems in Ref. [123] from the perspective of optimal cloning¹. The maximum average fidelity was found to be $\bar{F}_{max}(d) = (N + 1)/(N + d)$. In Ref. [132], optimal estimators with finite POVMs based on exact quadratures on a hypersphere were constructed.

Optimal measuring strategies for the state estimation game involve collective measurements. This is conceptually noteworthy, as the available ensemble is in a separable state, yet entangling the systems through the measurement helps extract more information than measuring each individually.

¹Although perfect cloning is impossible due to the no-cloning theorem, approximate copies of states can be created. Optimal cloning refers to approximate copying of states with the highest possible fidelity, as discussed for example in Refs. [130, 131]

2.2.4 Asymptotic Infidelity and State Estimation

In situations where a large identically prepared ensemble of size N is available, it is natural to use tomographic methods. To assess the performance, the asymptotic tomographic infidelity scaling with N can be studied. Using parameter estimation theory, Gill and Massar [108] derived the lower bound $1 - \bar{F}_p = 1/N$ (known as the Gill-Massar bound) on the infidelity for estimation of pure qubit states [108]. A similar result applies to mixed qubit states when measurements are restricted to being separable, with $1 - \bar{F}_m = (9/4)N^{-1}$ [54, 108, 133].

While collective measurements are optimal for state estimation, the gap in performance compared to separable measurements can be in part bridged in the asymptotic regime and with the use of adaptivity, particularly in the case of pure states. The asymptotic average fidelity of pure single-qubit state estimation using separable, adaptive, and collective measurements was theoretically studied in Ref. [109]. The authors found that fixed Pauli tomography produces an asymptotic infidelity of $(13/12)N^{-1}$. A two-stage adaptive scheme was outlined which leads to an asymptotic average infidelity of $1/N$, meaning it saturates the optimal collective measurement bound. The use of random local von Neumann measurements can also achieve an asymptotic average infidelity of $1/N$ [110].

The asymptotic case for mixed qubit states distributed according to the Bures prior was studied for separable, adaptive, and collective measurements in Ref. [54]. The authors provided a two-stage adaptive approach which was shown to improve the asymptotic infidelity scaling of $N^{-3/4}$ produced by fixed Pauli tomography to $(D^2/4)N^{-1}$ for $D = 2$ (in the plane) or $D = 3$ for general qubits. Therefore, unlike the case of pure states, there is a scaling advantage in using adaptive tomography. Additionally, the authors showed that there is a gap between the collective and adaptive average infidelity, even in the large N regime.

2.2.5 Adaptive Tomography

Two-Stage Adaptive Tomography

Two-stage adaptive tomography is a common approach that involves a single adaptation step. In stage one, standard Pauli tomography is performed on a subset of copies to obtain a preliminary estimate of the state parameters. In stage two, this estimate guides the choice of measurements on the remaining copies, yielding the final state estimate.

Different schemes design the stage-two measurements to optimize figures of merit as a function of the size N of the identically prepared ensemble. Their goal may be: 1. Saturating parameter-estimation bounds; 2. Optimizing asymptotic average infidelity scaling over all pure or mixed states; 3. Optimizing average infidelity scaling over all mixed states, while maintaining the same infidelity scaling for states of any purity. A common theme is to use the eigenstates of the preliminary estimate and their mutually unbiased bases for designing the second-stage measurements.

An important experimental demonstration of two-stage adaptive tomography for one-qubit states was performed in Ref. [46]. The proposed adaptive scheme achieves infidelity $1 - F = O(1/N)$ over the entire Bloch sphere, regardless of the purity of the unknown state. The infidelity scaling in this scheme applies to any arbitrary state, rather than being only achieved as an average over the state-space. In contrast, static state tomography using Pauli-measurements yields a scaling of $O(1/\sqrt{N})$ for nearly pure states [46, 103].

The first step of the scheme is to use measurements of σ_x, σ_y and σ_z on $N_0 = N/2$ copies to obtain a preliminary estimate ρ_0 using MLE. This preliminary estimate is then expressed as $\rho_0 = \lambda_1 |\psi_1\rangle\langle\psi_1| + \lambda_2 |\psi_2\rangle\langle\psi_2|$, where λ_i are the eigenvalues and $|\psi_i\rangle$ are the eigenstates of ρ_0 . These eigenstates are used to define the two additional bases $|\psi_{3/4}\rangle = (1/\sqrt{2})(|\psi_1\rangle \pm |\psi_2\rangle)$ and $|\psi_{5/6}\rangle = (1/\sqrt{2})(|\psi_1\rangle \pm i|\psi_2\rangle)$. The remaining $N - N_0 = N/2$ copies are measured in the three bases $\{|\psi_1\rangle, |\psi_2\rangle\}, \{|\psi_3\rangle, |\psi_4\rangle\}, \{|\psi_5\rangle, |\psi_6\rangle\}$, where $N/6$ measurements are performed per basis. MLE is then used to obtain a final state estimate.

The bottlenecks for scaling two-stage adaptive tomography to a larger number of qubits are both experimental and computational. Implementing additional mutually unbiased bases, when available, generally requires collective measurements in the case of n -qubit systems, which can be increasingly challenging to realize experimentally. Moreover, the number of measurement settings already scales exponentially with the number of qubits in standard tomography, so an increase in measurement settings compounds an experimental overhead. Additionally, MLE involves constrained optimization over a density matrix with $4^n - 1$ real parameters, and is already computationally demanding beyond a few qubits. Requiring an intermediate MLE estimate further exacerbates this computational challenge.

Bayesian Adaptive Tomography

A more elaborate adaptation method is to use Bayes' rule to update the evolving knowledge about the ensemble in the course of an experiment [48, 114, 134, 135]. The density matrix for the ensemble is written as a continuous mixture of pure states over the parameter space. A normalized probability distribution is used to represent the weights in the mixture. After

each measurement, this distribution is updated via Bayes' rule as

$$p(\rho|\mathcal{D}) \propto \mathcal{L}(\rho; \mathcal{D})p(\rho), \quad (2.8)$$

where $p(\rho|\mathcal{D})$ is the posterior distribution for a given observed data set $\mathcal{D}$, and $\mathcal{L}(\rho; \mathcal{D}) = \prod_{n=1}^N \mathbb{P}(\gamma_n|\rho, \alpha_n)$ is the likelihood for the data set $\mathcal{D}$. The likelihood is the product of the Born probabilities $\mathbb{P}(\gamma_n|\rho, \alpha_n)$ for the outcomes γ_n when measuring ρ in the measurement configurations α_n .

At each adaptive step, a figure of merit, such as information-gain or expected fidelity, is evaluated for candidate measurements. The measurement that maximizes this figure of merit is then selected, and the procedure repeats.

The benefit in using Bayesian inference for state reconstruction compared to MLE is that it does not suffer from rank-deficient estimates. Bayesian inference also provides a way to assign uncertainties. Not only does Bayesian inference provide error bars, but it captures the remaining uncertainty given the observed data $\mathcal{D}$. If a point estimate is needed, the Bayesian mean estimate can be used.

Bayesian adaptive schemes are naturally more computationally expensive, and unlike the two-stage schemes, could involve multiple adaptation steps. The challenge arises in the normalization of Eq. (2.8), which becomes analytically intractable when larger systems or many adaptive steps are involved. The normalization can be done through approximations based on Markov chain Monte Carlo (MCMC) sampling techniques. However, these approaches have computational cost $O(n)$ with n measurement configurations used, as they require evaluation of the full likelihood. To mitigate this computational problem, Ref. [48] introduced a fast sequential importance sampling algorithm with $O(1)$ likelihood evaluation cost. This involves approximating the posterior $p(\rho|\mathcal{D})$ in terms of discrete samples ρ_s from the distribution called *particles* with corresponding weights w_s such that $\sum_s w_s = 1$. Initial weights are chosen uniformly, and are then updated using Bayes' rule each time a measurement is made.

The drawback to this approximation of the posterior is that several weights w_s decay to zero over time, and the approximation becomes worse. The degree of weight decay can be quantified by the effective sample size $\mathcal{S}_{\text{eff}} = (\sum_{s=1}^{\mathcal{S}} w_s^2)^{-1}$, which ranges from 1 to $\mathcal{S}$. Monitoring the effective sample size and resampling when it falls below a threshold can help mitigate weight decay [114].

Nonetheless, periodic resampling to avoid weight decay increases the computation time, which motivates learned or approximate update rules (e.g., through artificial neural networks) instead of direct use of Bayes' rule.

Artificial Neural-Network Assisted Adaptive Tomography

Artificial neural networks offer a flexible framework for modelling complex correlations in measurement data [136–138]. By learning a mapping from previously acquired measurement outcomes to an internal representation of the state or its uncertainty, these methods could reduce the number of copies required to reach a desired reconstruction fidelity. Broadly speaking, two types of neural-network-assisted schemes have emerged: “query-by-committee” approaches [139], which maintain an ensemble of independently trained models whose mutual disagreement pinpoints the most informative measurement bases; and recurrent-network approximations of Bayesian updates [140], which learn to emulate costly particle-resampling steps and directly propose new measurements.

2.2.6 Stochastic Optimization-Based Tomography

Adaptive tomography may also be cast as the problem of minimizing an accessible objective — typically an infidelity or distance metric — over the space of quantum states. In these approaches, one maintains only a single candidate state (or a small set of parameters) and uses randomized perturbations to approximate gradients or to explore the state-space. By omitting full Bayesian updates or explicit information-gain calculations at every step, stochastic methods achieve low memory and runtime overhead, at the cost of introducing tunable step-size schedules and additional noise (e.g., random fluctuations in infidelity). Examples of this paradigm are self-guided tomography [141–143] and its extensions [133, 144].

2.3 Original Simulation Results

Being a source of both practical and fundamental interest, there is a significant amount of existing literature on qubit state tomography. While there is a wealth of information on this topic, there are competing claims of advantage for the use of particular tomographic measurement schemes.

Additionally, there is a variety of possible conditions and assumptions, such as prior knowledge about the purity and distribution of states, the state reconstruction method, and optimization with respect to different figures of merit. The performance of some schemes has been verified with limited evidence, for example, without considering the infidelity scaling, or only verifying the performance of one quantum state. Code accessibility in the case of simulations is limited in many cases.

The numerical simulations here aim to systematically examine pure qubit state estimation using two of the most experimentally practical measurement schemes, and clarify the performance effect of assuming a pure unknown state. In these original simulations, we numerically test the performance of single qubit tomography in terms of average infidelity scaling with N , the size of the identically prepared ensemble, for large N . Six-outcome Pauli tomography (Eq. (1.92)) and the two-stage adaptive tomography of Ref. [46], discussed in Subsection 2.2.5, are both examined. The reconstruction method used in both cases was MLE. To assess the performance, 50 pure input qubit states were generated randomly over the Bloch sphere by applying a Haar-random unitary [101] to the state $|H\rangle$. For each of the 50 states, tomography was simulated using $N = \{2^6, 2^7, \dots, 2^{19}, 2^{20}\} = \{64, 128, \dots, 524\,288, 1\,048\,576\}$. For each N (and for each state), the simulations were repeated 100 times.

2.3.1 Assumptions about State Purity

To elucidate the influence of the underlying assumptions in state estimation, original Monte Carlo simulations of one of the established adaptive schemes are compared to Pauli tomography for a wide range of pure states. The change in performance under the *a priori* pure state assumption in maximum-likelihood estimation is studied.

We use the Stokes parametrization to represent an arbitrary qubit (Eq. (1.88)):

$$\rho = \frac{1}{2} \sum_{i=0}^3 \mathbf{S}_i \sigma_i. \quad (2.9)$$

The *a priori* pure-state assumption then corresponds to imposing the constraint $\mathbf{S}_1^2 + \mathbf{S}_2^2 + \mathbf{S}_3^2 = 1$ when maximizing the log-likelihood $\log \mathcal{L}(\rho; \{n_j\})$ (Eq. (1.99)) in MLE.

2.3.2 Infidelity Scaling Under *a Priori* Pure-State Assumption

Six-outcome Pauli tomography simulations with the *a priori* pure-state assumption for 50 random pure states showed an average infidelity of $1 - \bar{F}_{\text{Pauli}}^{\text{sim, PA}} = (1.10 \pm 0.02)N^{-(1.005 \pm 0.002)}$. Fig. 2.2 shows the infidelity scaling observed for two typical states. This result is consistent with the theoretical calculations of Ref. [109], where the authors predicted an asymptotic average infidelity of $1 - \bar{F}_{\text{Pauli}}^{\text{theory}} = 1 - (13/12)N^{-1}$ for Pauli tomography of pure states.

These simulations confirm that in the presence of the *a priori* pure state assumption, adaptive tomography does not provide any scaling advantage, as Pauli tomography also

gives the optimal $O(1/N)$ asymptotic infidelity scaling. In fact, the constant is only $\sim 10\%$ larger than the optimal asymptotic fidelity $1 - \bar{F}_{\text{opt}} = 1/N$.

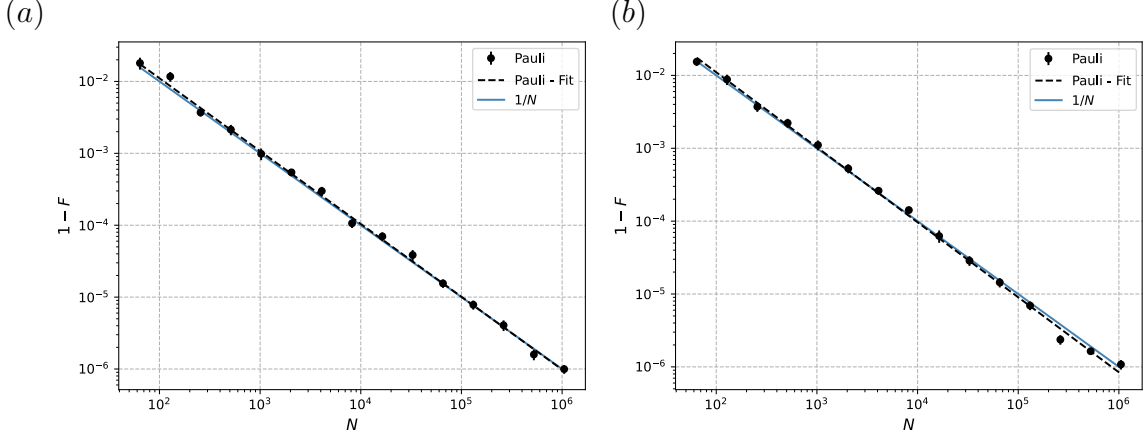


Figure 2.2: Typical Pauli tomographic infidelity scaling with size N of the identically prepared ensemble, *with* the pure-state assumption, is shown in black on log-log plots. Each data point is an average over 100 trials. The simulations were performed for 50 randomly generated pure states. The plots presented are for the random states: (a) $0.817|H\rangle + (0.16 - 0.555i)|V\rangle$ and (b) $0.429|H\rangle + (0.697 + 0.574i)|V\rangle$.

2.3.3 Infidelity Scaling Without *a Priori* Pure-State Assumption

Six-outcome Pauli tomography simulations without the *a priori* pure-state assumption for 50 random pure states showed an average infidelity of $1 - \bar{F}_{\text{Pauli}}^{\text{sim}} = (0.278 \pm 0.006)N^{-(0.528 \pm 0.002)}$. Under the same assumptions, the two-stage adaptive tomography method of Ref. [46] achieved an average infidelity of $1 - \bar{F}_{\text{2-stage}}^{\text{sim}} = (2.60 \pm 0.04)N^{-(1.000 \pm 0.001)}$. Fig. 2.3 shows the infidelity scaling results for two typical states.

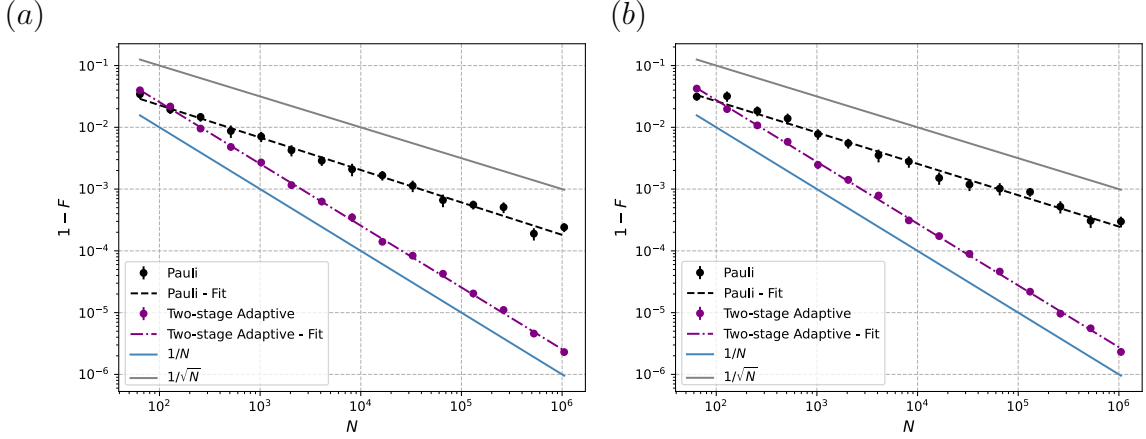


Figure 2.3: Tomographic infidelity scaling with the size N of the identically prepared ensemble, *without* the *a priori* pure-state assumption, shown on log-log plots. Each data point is an average over 100 trials. The simulations were performed for 50 randomly generated pure states. The plots presented are for the random states: (a) $0.817|H\rangle + (0.16 - 0.555i)|V\rangle$ and (b) $0.429|H\rangle + (0.697 + 0.574i)|V\rangle$.

Firstly, it is important to note that the infidelity scaling of Pauli tomography (Fig. 2.3) shows a nearly quadratic deterioration compared to Pauli tomography while assuming *a priori* pure states (Fig. 2.2). Intuitively, this deterioration of the average infidelity scaling for Pauli tomography in the absence of the pure-state assumption may be understood by examining what typically occurs in the state estimation of *nearly* pure states. Pauli tomography is a poor scheme for the estimation of such states. That is because minimizing infidelity requires accurately estimating the small eigenvalues of ρ_{exact} [46, 145]. The density matrix of a nearly pure qubit state has a small eigenvalue $\lambda_{\min}$ that can appear to be zero in the presence of statistical uncertainties, meaning a significant number of systems from the identical ensemble need to be used to reliably determine $\lambda_{\min}$. Previous work suggests that the point-wise infidelity of nearly-pure states follows a $O(1/\sqrt{N})$ scaling up to some $N = N_{\min}$ before a sudden transition to a $O(1/N)$ scaling [103]. The smaller the eigenvalue $\lambda_{\min}$, the higher is the minimum number of measured systems $N_{\min}$ before a transition to $O(1/N)$ occurs. Incidentally, this explains the average infidelity scaling of $O(N^{-3/4})$ for Pauli tomography over mixed states distributed according to the Bures prior [54]. This $N^{-3/4}$ scaling follows partly from the better performance for states of lower purity, as well as the aforementioned transition from $O(\sim 1/\sqrt{N})$ to $O(1/N)$ infidelity scaling for a nearly-pure state.

Meanwhile, for a pure exact state, the small eigenvalue $\lambda_{\min} = 0$. This means that using Pauli tomography, without assuming that the state is pure, there is no finite number $N_{\min}$

of systems from the identical ensemble which can be used to reliably determine whether $\lambda_{\min}$ is zero or not. Therefore, the $O(1/\sqrt{N})$ scaling in principle persists indefinitely. I believe this explains the dramatic difference of the infidelity scaling in the presence (or absence) of the pure-state assumption, despite the fact that neither the exact states nor the measurement scheme have changed.

The simulations in Fig. 2.3 also confirm that in the absence of the *a priori* pure state assumption, two-stage adaptive tomography improves the one qubit infidelity scaling from $O(\sim 1/(\sqrt{N}))$ to $O(1/N)$. For most states, the advantage in using the two-stage scheme starts to appear after N surpasses a few hundred. The advantage grows for higher N thanks to the quadratic improvement in infidelity scaling.

We also note that the two-stage adaptive infidelity shows a constant gap relative to the Gill-Massar bound for pure states, $1 - \bar{F}_p = 1/N$. Given the simulations in this section did not include any assumptions on the state purity, it is interesting to also compare the two-stage adaptive performance to the Gill-Massar bound for mixed states, $1 - \bar{F}_m = (9/4)N^{-1} = 2.25N^{-1}$ [54, 133]. In terms of the coefficient, the two-stage adaptive average infidelity of $1 - \bar{F}_{2\text{-stage}}^{\text{sim}} = (2.60 \pm 0.04)N^{-(1.000 \pm 0.001)}$ is considerably closer to the Gill-Massar bound for mixed states, relative to the one for pure states. Despite being closer, the scaling coefficient for the two-stage adaptive scheme still does not attain the Gill-Massar bound for mixed states $1 - \bar{F}_m = 2.25N^{-1}$.

2.3.4 Summary

Through Monte Carlo simulations, we examined the effect of the pure-state assumption based on prior knowledge on tomographic infidelity scaling. With this assumption, Pauli tomography achieves optimal scaling, with a coefficient about 10% higher than the Gill-Massar $1/N$ bound for pure states. This indicates that Pauli tomography performs nearly optimally in this scenario. However, when the assumption is not made, the average infidelity of Pauli tomography scales as $O(\sim 1/\sqrt{N})$.

The significant difference in average infidelity scaling between the two scenarios is noteworthy. While we usually consider the unknown state, measurement operators, and reconstruction algorithm in quantum state tomography as the key elements, this investigation shows that even with the same states and measurement operators — and using MLE as the reconstruction method — changing the assumptions about the state can greatly affect the infidelity scaling performance.

This behaviour may be understood by analogy with the state estimation of nearly-pure states using Pauli tomography. The small eigenvalue $\lambda_{\min}$ poses a challenge for such states,

since a minimum $N_{\min}$ is required for the point-wise fidelity to transition from $O(\sim 1/\sqrt{N})$ to $O(1/N)$. The same logic applied to exactly pure states — in the absence of the pure-state assumption — seemingly explains a near quadratic deterioration in average infidelity scaling.

Nonetheless, in practical experiments, it may not be possible to assume a pure exact state as the reference. In these situations, the two-stage adaptive scheme proposed by Ref. [46] provides a quadratic improvement in scaling compared to Pauli tomography.

2.4 Summary

In this chapter, we introduced optimal state estimation in more detail, emphasizing its goal of maximizing average fidelity (or minimizing average infidelity) as a function of the number of systems N in the identically prepared ensemble. We reviewed optimal state estimation for finite (and potentially non-statistical) N in various scenarios, noting the recurring role of collective measurements. We also explored optimal state tomography for asymptotic N and briefly discussed the expected performance of different schemes, including Pauli and adaptive tomography.

As an original contribution, we presented Monte Carlo simulations performed to highlight an important but often overlooked aspect of state estimation: the same exact states and measurements can result in significantly different estimation performance based on the *a priori* assumptions about purity. We suggested an explanation for this behaviour by drawing a connection to the state estimation of nearly-pure states and the role of a small eigenvalue of the density matrix.

In the absence of the pure-state assumption, we numerically verified the quadratic improvement in the average infidelity scaling of a two-stage adaptive tomography scheme introduced in Ref. [46]. Further comparisons were made with Gill-Massar bounds, showing a closer performance (in terms of the scaling coefficient) to the bound for mixed states.

Chapter 3

Quantum State Estimation with Collective Measurements

This chapter is closely adapted from the manuscript cited below,

Reproduced from A. Mansouri, K. M. Jordan, R. A. Abrahao, and J. S. Lundeen, “Experimental demonstration of a multi-particle collective measurement for optimal quantum state estimation,” [Optics Express](#), vol. 34, p. 5718, Feb. 2026. Reproduced with permission.

3.1 Motivation

An important and fundamental task in quantum information is quantum state tomography, which allows complete characterization of the state ρ of quantum systems through measurements on an identically prepared ensemble of size N_{ens} . This procedure involves the measurement of a set of observables sufficient for the complete determination of the density matrix, resulting in a state estimate ρ_{est} [95]. For a d -dimensional system, at least $d^2 - 1$ measurement operators need to be measured for a tomographic reconstruction [4, 5]. In practice, many quantum systems need to be measured for each projection to reduce statistical uncertainty, meaning a large total number of identically prepared systems are required for quantum state tomography. This poses a challenge for the efficient characterization of quantum systems. For convenience, we recall some important facts about quantum state estimation introduced in Sections 1.6-1.7 and Section 2.2.

A common way to assess the efficiency of a state characterization scheme is to consider the scaling of the infidelity $1 - F = 1 - [\text{Tr} \sqrt{\sqrt{\rho_{\text{est}}} \rho \sqrt{\rho_{\text{est}}}}]^2$ with the size N_{ens} of the identically prepared ensemble, which quantifies the error in the estimate ρ_{est} of the true state ρ . Tomographic scaling is important from a fundamental point of view, since it relates to how much information can be extracted. It is also significant when considering the practicality in experiments where the number of copies is limited.

Through a calculation based on the quantum Fisher information matrix and the quantum Cramér-Rao bound, Gill and Massar showed that the optimal infidelity for the tomography of a pure one-qubit state scales as $1/N_{\text{ens}}$ [108]. Massar and Popescu discussed a related state estimation task, where N qubits polarized in $|\vec{n}\rangle$ are provided in each round, and a single guess $|\vec{n}_g\rangle$ of the polarization vector is made [50].

We describe two experiments on state identification, starting with the game addressed by Massar and Popescu, and later applying the collective measurement to quantum state tomography. The Massar and Popescu game concerns state estimation in a scenario where the size of the ensemble may not be large (statistical). The state is thus identified through a guess corresponding to each measured outcome. We examine two variants of the game: one with uniformly distributed unknown states, and the other with the unknown state being one of four equiprobable states. When the number of available copies is sufficiently large, tomographic estimation becomes possible, with a continuous range of estimated states rather than discrete guesses. The optimal two-copy collective measurement of the Massar and Popescu game is then used to tomographically estimate the state from a large ensemble in pairs, allowing to examine its performance in the statistical regime.

3.2 Theoretical Framework

3.2.1 Average Fidelity

From Section 2.2, we recall that the optimal measurement for the Massar and Popescu game is defined as one that, on average, maximizes the fidelity $\mathcal{F}(|\vec{n}\rangle, |\vec{n}_g\rangle) = |\langle \vec{n} | \vec{n}_g \rangle|^2$, subject to unitarity constraints regarding the time evolution of the interaction with the measured system. The procedure is repeated many times, with $|\vec{n}\rangle$ given according to some prior distribution $P(\vec{n})$, e.g., uniform over the unit sphere when assuming no prior

knowledge. The average fidelity $\bar{\mathcal{F}}$ can thus be written as [50, 107]:

$$\begin{aligned}\bar{\mathcal{F}} &= \int P(\vec{n}) d\vec{n} \sum_g P(g|\vec{n}) \mathcal{F}(|\vec{n}\rangle, |\vec{n}_g\rangle) \\ &= \int P(\vec{n}) d\vec{n} \sum_g P(g|\vec{n}) \left(\frac{1 + \vec{n} \cdot \vec{n}_g}{2} \right),\end{aligned}\tag{3.1}$$

where $\vec{n}$ is the Bloch vector of the unknown state $|\vec{n}\rangle$, $\vec{n}_g$ is the Bloch vector associated with the guessed state, and $P(g|\vec{n})$ is the probability of guessing $\vec{n}_g$ when the true direction is $\vec{n}$.

3.2.2 Optimal Collective Measuring Strategy

In Ref. [50], Massar and Popescu maximized the expression Eq. (3.1) with respect to the possible measurements and guessing strategies, while satisfying unitarity constraints regarding the time evolution of the interaction between the measured system and the measuring device. When the different directions $\vec{n}$ are uniformly distributed over the unit sphere, i.e., $P(\vec{n})$ is uniform, the maximum achievable average gate fidelity for N systems in the identical ensemble is $\bar{\mathcal{F}}_{\text{GenMP}}^{\text{Coll}} = (N + 1)/(N + 2)$. The authors derived this result by using a Lagrange multiplier for the unitarity conditions and taking advantage of the spherical symmetry of the problem. The optimal measuring strategy — that maximizes the average fidelity — is to perform a collective measurement on the ensemble of N qubits. For the case of $N = 2$, an optimal strategy is to measure a nondegenerate operator whose eigenstates are the following superpositions of the singlet state $|S\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ and product tetrahedron states $|\uparrow_{n_i}\uparrow_{n_i}\rangle$:

$$\begin{aligned}|\text{MP}_1\rangle &= \frac{1}{2}|S\rangle + \frac{\sqrt{3}}{2}|\uparrow_{n_1}\uparrow_{n_1}\rangle \\ |\text{MP}_j\rangle &= -\frac{1}{2}|S\rangle + \frac{\sqrt{3}}{2}|\uparrow_{n_j}\uparrow_{n_j}\rangle, \quad j = 2, 3, 4.\end{aligned}\tag{3.2}$$

In the Bloch sphere, the states $|\uparrow_{n_i}\rangle$ are located on the surface of the sphere at the vertices of a tetrahedron (Fig. 3.1) and can be chosen as:

$$\begin{aligned}|\uparrow_{n_1}\rangle &= |0\rangle; \\ |\uparrow_{n_j}\rangle &= \frac{1}{\sqrt{3}}|0\rangle + \sqrt{\frac{2}{3}}e^{i\alpha_j}|1\rangle, \quad j = 2, 3, 4,\end{aligned}\tag{3.3}$$

with $\alpha_2, \alpha_3, \alpha_4 = 0, 2\pi/3, 4\pi/3$. Upon performing this collective measurement and obtaining the outcome i (corresponding to $|\text{MP}_i\rangle$), the tetrahedron state $|\uparrow_{n_i}\rangle$ is chosen as the guess $|\vec{n}_g\rangle$. Projections onto $|\text{MP}_i\rangle$ require collective measurements, since each $|\text{MP}_i\rangle$ is entangled, albeit non-maximally. The entangled states $|\text{MP}_i\rangle$ all have the same concurrence $\mathcal{C}(|\text{MP}_i\rangle) = 0.25$.

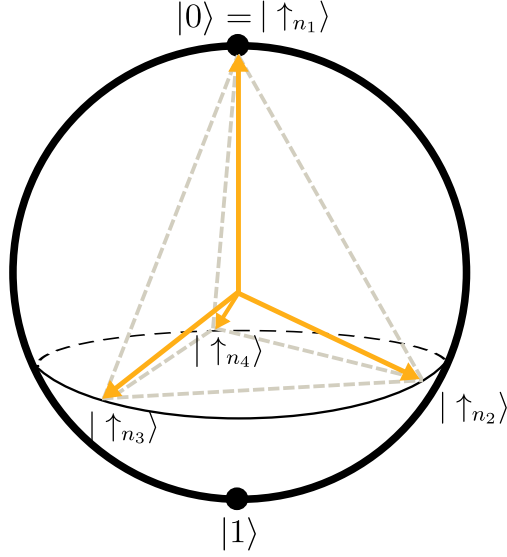


Figure 3.1: The guess states $|\uparrow_{n_i}\rangle$ for the optimal collective measuring strategy, associated with the entangled projection states $|\text{MP}_i\rangle$, are at the vertices of a tetrahedron in the Bloch sphere.

The reason for the non-maximally entangled nature of the projection states can be understood as follows. The input states $|\vec{n}\rangle \otimes |\vec{n}\rangle$ lie entirely in the symmetric (spin-1) subspace, and optimal estimation requires a POVM that probes this space uniformly without any preferred direction. The tetrahedron directions provide a uniform sampling of the Bloch sphere and are responsible for the optimal information extraction. However, the corresponding symmetric projectors do not resolve the identity (form a POVM) on the full four-dimensional state-space. To satisfy the POVM completeness and positivity conditions, each projection state includes a carefully chosen singlet component, which has zero overlap with $|\vec{n}\rangle \otimes |\vec{n}\rangle$ but is necessary to ensure that the POVM elements sum to the identity.

3.2.3 General Massar and Popescu Game (GenMP Game)

The theoretical average fidelity of the Massar and Popescu game in different scenarios can be directly calculated from Eq. (3.1). Consider again the scenario where the directions $\vec{n}$ for the identical ensemble of size $N = 2$ are uniformly distributed over the unit-sphere, i.e., $P(\vec{n})d\vec{n} = \sin\theta d\theta d\phi/(4\pi)$. We will refer to this case as the *General Massar and Popescu (GenMP)* game. For $N = 2$, the average fidelity using the optimal collective approach in this case is $\bar{\mathcal{F}}_{N=2, \text{GenMP}}^{\text{Coll}} = 3/4 = 0.75$. This can be compared with the average fidelity using the best LOCC approach, which is $\mathcal{F}_{N=2, \text{GenMP}}^{\text{LOCC}} = \frac{1}{6}(3 + \sqrt{2}) \approx 0.7357$ [146]. We will discuss this strategy next.

3.2.4 Optimal LOCC Strategy

The optimal solution for the GenMP game when restricted to LOCCs can be achieved by measuring one photon in the $\{|H\rangle, |V\rangle\}$ basis and the other in the $\{|D\rangle, |A\rangle\}$ basis [146]. The guessed state $|L_i\rangle$ on the unit sphere is taken to be at the bisector of the outcome states of the two photons:

$$\begin{aligned}
|D\rangle\langle D| \otimes |H\rangle\langle H| &\rightarrow |L_1\rangle = \cos\frac{\pi}{8}|H\rangle + \sin\frac{\pi}{8}|V\rangle, \\
|D\rangle\langle D| \otimes |V\rangle\langle V| &\rightarrow |L_2\rangle = \cos\frac{3\pi}{8}|H\rangle + \sin\frac{3\pi}{8}|V\rangle, \\
|A\rangle\langle A| \otimes |H\rangle\langle H| &\rightarrow |L_3\rangle = \cos\frac{\pi}{8}|H\rangle - \sin\frac{\pi}{8}|V\rangle, \\
|A\rangle\langle A| \otimes |V\rangle\langle V| &\rightarrow |L_4\rangle = \cos\frac{3\pi}{8}|H\rangle - \sin\frac{3\pi}{8}|V\rangle.
\end{aligned} \tag{3.4}$$

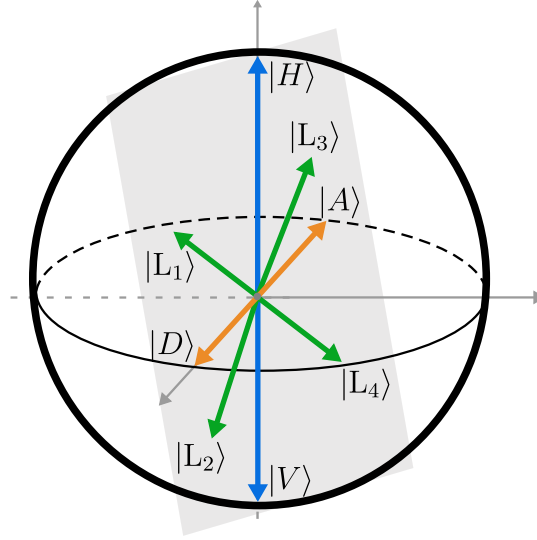


Figure 3.2: Optimal LOCC strategy states. The outcome states are shown in blue and orange, corresponding to each of the two photons. The guessed states $|L_i\rangle$ are at the bisectrix of the outcome states.

Experimentally, the LOCC optimal strategy was implemented by introducing a long delay between the photons in the two arms in Fig. 3.4 (to avoid the HOM interference underlying the singlet projection) and measuring coincidences with polarizers placed in front of single-photon detecting avalanche photodiodes.

3.2.5 Tetrahedron Massar and Popescu Game (TetraMP Game)

When the *a priori* distribution $P(n)d\vec{n}$ is made discrete in such a way that $\vec{n}$ is restricted to being one of the tetrahedron states $|\uparrow_{n_i}\rangle$, we have what we refer to as the Tetrahedron Massar and Popescu (TetraMP) game. The collective approach to the TetraMP game gives an average fidelity of $\bar{\mathcal{F}}_{N=2, \text{TetraMP}}^{\text{Coll}} = 5/6 \approx 0.8333$, conjectured to be optimal [107]. To the best of the authors' knowledge, the optimal LOCC approach is not known for the TetraMP game. Alternatively, the TetraMP game can be treated as a quantum state discrimination problem for the equiprobable set of states $\{|\uparrow_{n_i}\rangle\}$. While lower and upper bounds for multiple-state, minimum-error discrimination using single-copy measurements have been derived [147], they may not be achievable or tight (yielding a wide average fidelity range $[0.636, 0.908]$ in this case). More concretely, the collective strategy can be compared to the "pretty good" measurement [148, 149], expected to perform fairly well for the discrimination of any set of states. The pretty good measurement corresponds to the POVM $\{\frac{1}{2} |\uparrow_{n_i}\rangle\langle\uparrow_{n_i}|\}$

for the TetraMP game and yields an average fidelity of $\bar{\mathcal{F}}^{\text{PGM}} = 2/3 \approx 0.6667$, which is $\sim 16.7\%$ lower than $\bar{\mathcal{F}}_{N=2, \text{TetraMP}}^{\text{Coll}}$.

3.2.6 Quantum State Tomography: Two-Copy Collective Measurements

Projections onto the non-maximally entangled states $|\text{MP}_i\rangle$ can also be used for quantum state tomography. In this way, a large ensemble of size N_{ens} , potentially of thousands of photons, can be divided into pairs on which we perform collective measurements (Fig. 3.3). As the Massar and Popescu two-copy collective measurement allows extracting more information about a pair of photons than any separable measurement, we expect the tomographic infidelity scaling to also improve.

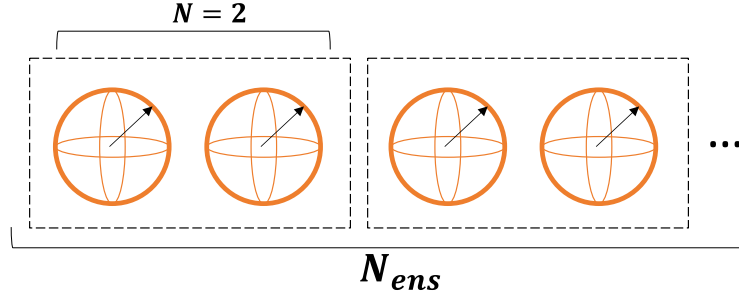


Figure 3.3: Larger ensembles are divided into pairs and measured using two-copy collective measurements for quantum state tomography.

The estimate of the state is found using a maximum-likelihood reconstruction algorithm. Additional constraints regarding purity and the tensor-product form $|\vec{n}\rangle \otimes |\vec{n}\rangle$ of the input state are imposed in an accelerated-projected-gradient maximum-likelihood algorithm [150]. The code for the implementation is available in [151]. On a standard consumer laptop (AMD Ryzen 3 5425U CPU at 2.7 GHz, 16 GB RAM), a single maximum-likelihood reconstruction required approximately 10 s of wall-clock time. The algorithm was implemented in Wolfram Mathematica, and runtimes depend on hyperparameters, such as the step size in the gradient descent, the momentum parameter, and the stopping criterion. More optimized implementations or different hardware may lead to different performance.

3.3 Experimental Methods

Attempts to implement the Massar and Popescu game have so far been limited to the use of photonic quantum walks, where various degrees of freedom of the same photon encode different qubits [61, 62]. Although this simulates the collective strategy using non-identical degrees of freedom of a single photon, such states cannot meaningfully violate Bell's inequalities or be applied to useful quantum computing. The same outcomes could be obtained with classical light of the same mean photon number, and the outcome statistics could be predicted without a quantum-mechanical treatment, based on first-order interference. Nonetheless, photonic quantum walks are valuable for simulating topological physics [152], quantum algorithms [153], as well as providing a testing ground for collective measurements.

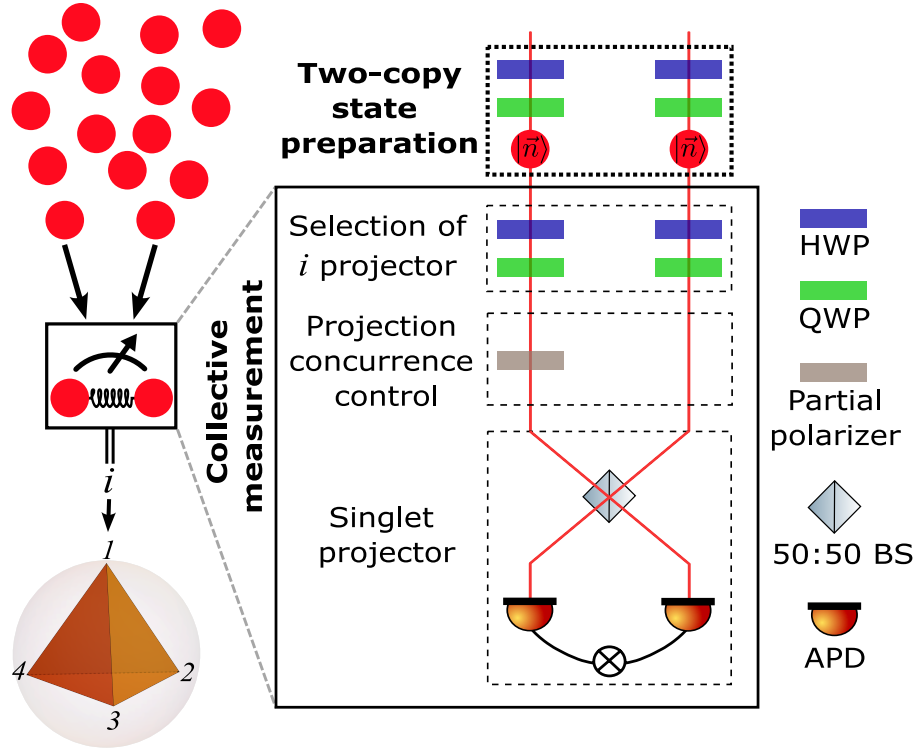


Figure 3.4: Scheme and experimental setup used for the Massar and Popescu game. The setup consists of half-wave plates (HWP), quarter-wave plates (QWP), a partial polarizer, a 50:50 beamsplitter (BS), and single-photon avalanche photodiode detectors (APD). See the main text for details.

Instead, we begin with two identical photons, which are our two copies in the ensemble

for the Massar and Popescu game. Our multi-particle implementation involves an entangling measurement of two-photon polarization states that combines Hong-Ou-Mandel (HOM) interference [66, 154] with polarization-dependent loss [155, 156], as shown in Fig. 3.4 [71].

3.3.1 Non-Maximally Entangled Projections

Whereas typical two-photon polarization measurements project onto either a separable or a maximally entangled state, we use a recently developed general two-photon polarization measurement [71] (detailed in Subsection 1.4.2) to project on the non-maximally entangled $|\text{MP}_i\rangle$ states. At the heart of our device is a projection onto the singlet state $|S\rangle\langle S|$: this is accomplished when two photons entering opposite input ports of a 50:50 beamsplitter exit opposite ports [72]. This part of the device, referred to as the singlet projector in Fig. 3.4, consists of a 50:50 BS, single-photon avalanche photodiode detectors, and a time tagger for coincidence counting.

The device varies the level of entanglement (i.e., concurrence) of the projection state by introducing a partial polarizer in one input mode, with transformation matrix

$$\hat{W} = \begin{bmatrix} t_H & 0 \\ 0 & t_V \end{bmatrix} \otimes \mathbb{I}_{2 \times 2}, \quad (3.5)$$

where t_H and t_V represent complex transmission amplitudes, here real-valued. With the partial polarizer shown in Fig. 3.4, we alter the projection onto $|S\rangle$ to have a lower entanglement, a concurrence $\mathcal{C}(|\text{MP}_i\rangle) = (2|t_H||t_V|)/(t_H^2 + t_V^2) = 0.25$ for all four $|\text{MP}_i\rangle$. The concurrence in turn determines the optimal efficiency of the projection $\eta_{\text{opt}} = 1/(1 + \sqrt{1 - \mathcal{C}^2})$, which is the probability of a successful projection given the incoming pair of photons were in the state $|\text{MP}_i\rangle$. To project onto a specific $|\text{MP}_i\rangle$, we add an appropriate unitary transformation in each input mode. As shown in Fig. 3.4, this selection of the i projector is achieved with HWP-QWP pairs. Altogether, this device projects onto any one of the states $|\text{MP}_i\rangle$.

Two-photon interference has been used to demonstrate collective measurements for parameter estimation and metrology [157, 158]. However, until the present work only basic two-photon collective measurements have been used in applications: projections on maximally entangled states. These projections are naturally implemented by a standard 50:50 beamsplitter. In this work, we project onto optimized collective states $|\text{MP}_i\rangle$, which are non-maximally entangled and non-trivial to implement.

The two-photon projections are implemented experimentally as shown in Fig. 3.4. After the input states are prepared, a HWP and QWP implement the appropriate unitary on each input photon. The concurrence $\mathcal{C} = 0.25$ corresponds to a polarization-dependent loss with extinction ratio $\sim 62 : 1$ between the vertical and horizontal polarization optical powers. Since this extinction ratio is significantly biased towards one polarization, the desired transformation was realized through a carefully chosen imperfect polarizing beamsplitter. We tilted the beamsplitter to fine-tune the extinction ratio.

The singlet state projection uses an in-fiber non-polarizing beamsplitter. The measured splitting ratio was 47:53. Unwanted polarization unitary transformations in the fiber and beamsplitter are compensated using a tilted quartz plate, a HWP, and a QWP placed in the arm without polarization-dependent loss, just before the entrance of the fiber to the non-polarizing beamsplitter. The output fibers of the beamsplitter traveled to single-photon detectors (avalanche photodiodes). A coincidence (i.e. simultaneous detections) signifies a successful projection on an $|\text{MP}_i\rangle$ state.

This implementation includes a 3/4 loss due to the sequential implementation of Massar and Popescu state projections, as well as a $1 - \eta(\mathcal{C} = 0.25) = 49\%$ loss per projection. This results in an overall $\sim 12.7\%$ efficiency.

Despite these efficiency limitations, the primary objective of this work was to demonstrate the feasibility, performance, and scaling properties of quantum state estimation with the Massar and Popescu entangling measurement on two particles. The insights gained will inform the viability and the level of performance that can be expected of this collective strategy when applied to useful systems in quantum science and technology involving multiple particles.

3.3.2 Spontaneous Parametric Down-Conversion Source

Spontaneous parametric down-conversion (SPDC) is a nonlinear optical process where a single pump photon of frequency ω_1 is converted into two photons of frequencies ω_2 and ω_3 . The process can occur when the phase-matching condition between the photon wavevectors k_i ,

$$\Delta k = k_1 - k_2 - k_3 = 0, \quad (3.6)$$

is satisfied. In Eq. (3.6), $k_i = n_i \omega_i / c$ — where n_i is the refractive index and c is the speed of light — and by energy conservation,

$$\omega_1 = \omega_2 + \omega_3. \quad (3.7)$$

Practically, this can be approximately achieved by pumping a crystal with second-order nonlinearity, and tuning the angle of the crystal axis relative to the pump polarization and wavevector. In Type II SPDC, used in this source, the daughter photons are orthogonally polarized. The case where $\omega_2 = \omega_3$ is known as degenerate SPDC [159].

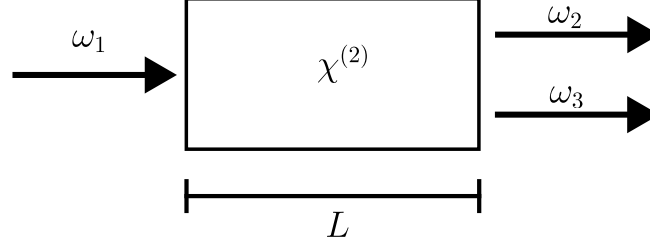


Figure 3.5: Spontaneous parametric down-conversion in a nonlinear material of length L with nonlinear susceptibility $\chi^{(2)}$. When phase-matching conditions are met, a pump photon of frequency ω_1 can be converted into photons of frequencies ω_2 and ω_3 through spontaneous parametric down-conversion.

The output produced from Type II SPDC is a two-mode squeezed state. In the limit of a weak pump field, most pump photons are simply transmitted through the crystal. Single photon pairs are created with some small probability, while the probability for higher order photon number states is negligible [160].

We used a narrow-linewidth (< 0.3 pm), 405 nm wavelength, 40 mW output power continuous wave laser to pump a 2 mm Type II BBO crystal. The orthogonally polarized 810 nm photon pairs were generated through collinear degenerate SPDC and then were split at a polarizing beamsplitter and then coupled into polarization-maintaining fibers. The photons were out-coupled into two arms, with one out-coupler on a translation stage to vary the relative delay between the photons. In each arm a half-wave plate (HWP) and a quarter-wave plate (QWP) were used to prepare the input polarization state $|\vec{n}\rangle \otimes |\vec{n}\rangle$, with the encoding $|0\rangle \equiv |H\rangle$ and $|1\rangle \equiv |V\rangle$.

An (810 ± 3) nm band-pass filter was used to ensure good spectral overlap between the photons. A 1 mm Type II BBO crystal placed after the down-conversion crystal was added to compensate for the transverse spatial walk-off.

3.3.3 Two-photon Interference and Visibility

As mentioned, two-photon interference is at the heart of projections onto entangled states. To ensure the presence of this effect and the indistinguishability of the photon pairs, HOM

interference scans were performed. Fig. 3.6(a) shows the coincidence dip observed when the photon pair was prepared in the state $|V\rangle \otimes |V\rangle$ just before the 50:50 beamsplitter. One out-coupler on a translation stage was used to vary the relative delay between the photons. The HOM visibility obtained from a Gaussian fit to this data was $\mathcal{V}_V = (98 \pm 1)\%$.

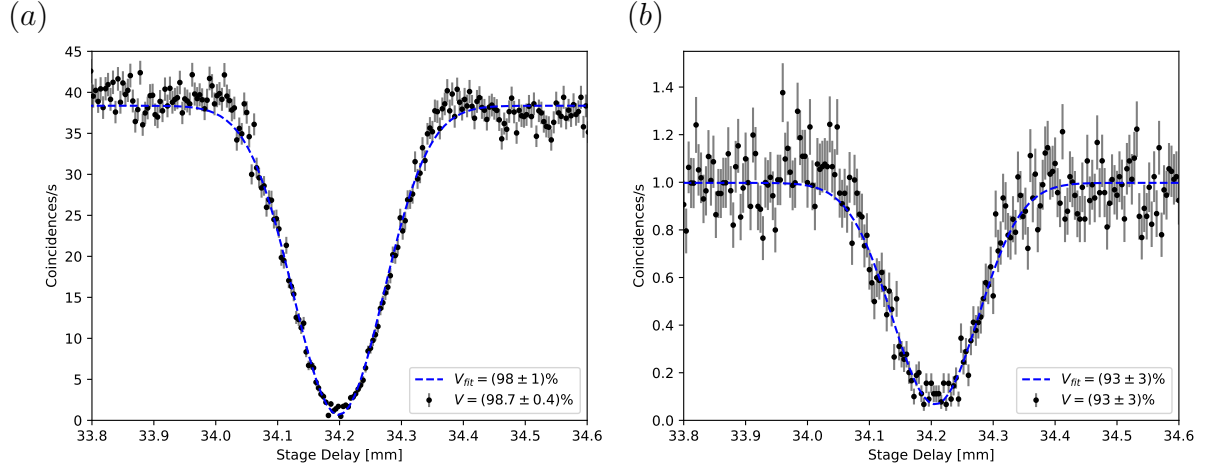


Figure 3.6: HOM interference scans for the input polarization state: (a) $|V\rangle \otimes |V\rangle$, with a HOM visibility $\mathcal{V} = (98.7 \pm 0.4)\%$ from the raw data and $\mathcal{V}_{\text{fit}} = (98 \pm 1)\%$ from the Gaussian fit; and (b) $|A\rangle \otimes |A\rangle$, with a HOM visibility $\mathcal{V} = \mathcal{V}_{\text{fit}} = (93 \pm 3)\%$.

The HOM interference scan for photons prepared in the state $|A\rangle \otimes |A\rangle$ — just before the 50:50 beamsplitter — is shown in Fig. 3.6(b). Due to the presence of the partial polarizer in Fig. 3.4, the baseline coincidence rate was significantly lower. A visibility of $\mathcal{V}_A = (93 \pm 3)\%$ was observed in this case. The discrepancy between the visibilities $\mathcal{V}_V$ and $\mathcal{V}_A$ can be in part be attributed to imperfect pre-compensation of the polarization changes in single-mode fibers, which were used to ensure spatial-mode overlap for HOM interference.

The width of the HOM dip is determined by the temporal indistinguishability of the interfering photons. More precisely, the dip envelope is given by the Fourier transform of the joint spectral intensity along the frequency-difference coordinate, which reflects the effective temporal extent of the two-photon wavepacket. In this case, the 0.3 pm pump bandwidth is narrow enough such that it does not significantly contribute to the dip width. In the absence of spectral filtering, this temporal width is dominated by group-velocity mismatch in the nonlinear crystal. For Type II down-conversion, the signal and idler propagate with different group velocities, resulting in a relative walk-off that scales as $|D|L$, the group-velocity mismatch $|D|$ of the two polarizations times the crystal length

L . With band-pass filters, the HOM width is broadened by the filter coherence time due to spectral narrowing. Using filter values $\lambda_0 = 810$ nm, $\Delta\lambda = 0.3$ pm, Gaussian profiles, and a BBO group velocity mismatch parameter $|D| = 0.2$ ps/mm, we find the HOM width to be:

$$\Delta x_{\text{HOM}} \sim c\sqrt{\tau_{\text{filter}}^2 + \tau_{\text{crystal}}^2} \approx 153 \text{ } \mu\text{m}; \quad \text{from} \quad (3.8)$$

$$\tau_{\text{filter}} = \frac{0.44}{\Delta\nu} = \frac{0.44}{(c/\lambda_0)^2(\Delta\lambda)} = 0.32 \text{ ps} \quad (3.9)$$

$$\tau_{\text{crystal}} \sim |D|L = 0.4 \text{ ps} \quad (3.10)$$

We see that Δx_{HOM} nominally agrees with the experimentally observed HOM full width at half maximum of $(164 \pm 5)\mu\text{m}$.

3.3.4 Projector Selection

To alternate between projection states with the same concurrence, unitary operations are applied prior to the partial polarizer and 50:50 beamsplitter using HWP-QWP pairs in each mode, as shown in Fig. 3.4. The appropriate unitaries for projecting onto each of the four Massar and Popescu states $|\text{MP}_i\rangle$ are found by performing a Schmidt decomposition of the $|\text{MP}_i\rangle$. As described in Subsection 1.4.2, the transformation relating the Schmidt bases to the $\{|H\rangle, |V\rangle\}$ basis determines this unitary operation.

3.3.5 Experimental Test of the Massar and Popescu Game

To test the implementation of the GenMP and TetraMP games, we considered a set S of 40 states. These states were generated randomly from a uniform distribution $P(\vec{n})$ over the Poincaré sphere by applying Haar random unitaries to the state $|H\rangle$. The states are of the form:

$$|\vec{n}\rangle = \cos \theta/2 |H\rangle + e^{i\phi} \sin \theta/2 |V\rangle, \quad (3.11)$$

where the polar angle $\theta \in [0, \pi]$ and the azimuthal angle $\phi \in [0, 2\pi)$ parameterize the state. The expected average fidelity for the set S when using collective measurements is $\bar{\mathcal{F}}_S^{\text{Th, Coll}} = 75.05\%$, while the expected fidelity for the LOCC measurement is $\bar{\mathcal{F}}_S^{\text{Th, LOCC}} = 73.68\%$. We measured the average fidelity for each state based on ~ 3500 detected photon pairs, that is, ~ 3500 trials of $N = 2$. The number of trials per state varied slightly in the experiment, although the same amount of time was spent on each state. These small variations are taken into account in the theoretical average fidelities $\bar{\mathcal{F}}_S^{\text{Th, Coll}}$ and

$\bar{\mathcal{F}}_S^{\text{Th, LOCC}}$. To obtain the experimental average fidelity over all states, corresponding to Eq. (3.1), we calculate the weighted average:

$$\bar{\mathcal{F}}_S^{\text{Exp}} = \sum_{|\vec{n}\rangle \in S} P(\vec{n}) \sum_{g=1}^4 P(g|\vec{n}) \mathcal{F}(|\vec{n}\rangle, |\vec{n}_g\rangle). \quad (3.12)$$

Although $P(\vec{n}) \approx 1/40$, we calculate $P(\vec{n})$ as the total coincidences measured for $|\vec{n}\rangle \otimes |\vec{n}\rangle$, divided by the total coincidences measured over all input states. This accounts for small variations in the number of trials per state. Now a frequency, $P(g|\vec{n})$ is calculated as the ratio of the coincidence counts for the projector corresponding to $|\vec{n}_g\rangle$ and the total coincidences for $|\vec{n}\rangle$. The guess $|\vec{n}_g\rangle$ is $|\uparrow_{n_i}\rangle$ (Eq. (3.3)) when using the collective strategy, and $|\text{L}_i\rangle$ (Eq. (3.4)) when using the LOCC approach.

3.3.6 Summary

Collinear Type II spontaneous parameter down-conversion allows us to generate photon pairs and separate them at a PBS. Product polarization states encoding the two copies are prepared using a PBS and waveplates. Massar and Popescu's optimal collective measurement is carried out using a recent two photon projection scheme [71] (detailed in Section 1.4.2) which utilizes two-photon interference and polarization dependent loss to control the concurrence of the projection state. We experimentally test the GenMP using 40 random states by comparing the state-specific and average fidelity to theory. In the following sections, we similarly perform a test of the TetraMP game, with the prior distribution of states limited to tetrahedron states.

3.4 Experimental Results

3.4.1 GenMP Game Results

In this section, we present experimental results for the GenMP game, where the unknown states are distributed uniformly on the unit sphere ($P(\vec{n})d\vec{n} = \sin\theta d\theta d\phi/(4\pi)$). In Fig. 3.7, we compare the experimentally measured fidelity with the theoretically expected one for each state using the collective measurement in (a) and the LOCC approach in (b). The left side of each circle represents the experimentally measured fidelity, while the right side represents the theoretically expected fidelity.

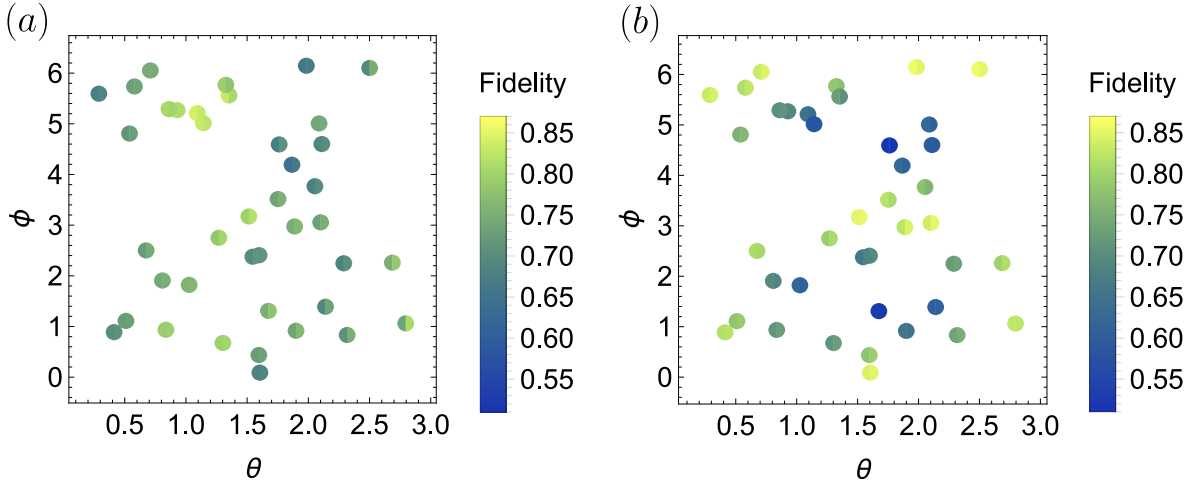


Figure 3.7: Scatter plots for the fidelity of the Massar and Popescu two-copy game: (a) using the collective approach; (b) using the LOCC approach; over the $\{\theta, \phi\}$ state-space. The left-side of each circle represents the experimentally measured fidelity, while the right-side represents the theoretically expected fidelity.

The experimental and theoretical values for the fidelity are generally in good agreement. The experimental average fidelities observed for the collective and LOCC schemes shown in Fig. 3.7 were $\bar{\mathcal{F}}_S^{\text{Exp, Coll}} = (73.5 \pm 0.2)\%$ and $\bar{\mathcal{F}}_S^{\text{Exp, LOCC}} = (73.4 \pm 0.2)\%$, respectively. The average fidelity using the collective strategy was found to be at least as high as that of the LOCC approach. The error bars are obtained by assuming Poisson statistics on the number of detected photons for each of the Massar and Popescu projectors and using Gaussian error propagation.

The deviations from the expected fidelities can be in part attributed to imperfect HOM interference. When the state prior to the 47:53 beamsplitter was prepared as $|V\rangle \otimes |V\rangle$,

the HOM visibility was $(98.7 \pm 0.4)\%$, while it was $(93 \pm 3)\%$ when the state entering the beamsplitter was $|A\rangle \otimes |A\rangle$. This suggests a slightly imperfect compensation of the polarization changes in single-mode fibers, which were used to ensure spatial-mode overlap for HOM interference. We can consider the effect of this additional polarization unitary $\tilde{U}$ that satisfies $|\langle V|\tilde{U}|V\rangle|^2 = 98.7\%$ and $|\langle A|\tilde{U}|A\rangle|^2 = 93\%$. Numerical optimization can be used to find $\tilde{U}$ that minimizes the mean-squared error between the experimentally measured and theoretically expected fidelities for the states in S using the collective strategy, subject to the constraints $|\langle V|\tilde{U}|V\rangle|^2 = 98.7\%$ and $|\langle A|\tilde{U}|A\rangle|^2 = 93\%$. The theoretically expected average fidelity in the presence of $\tilde{U}$ which best matches the experimentally measured values is about 74.4%, compared to 75.0% expected in its absence. The gap suggests that eliminating these systematic errors would at minimum increase the average fidelity by 0.5 %. Using this simple model, the average fidelity $\bar{\mathcal{F}}_S^{\text{Exp, Coll}}$ would at least increase to $\bar{\mathcal{F}}_{S, \text{Corrected}}^{\text{Exp, Coll}} = (74.0 \pm 0.2)\%$ if $\tilde{U}$ were eliminated. As the LOCC approach does not rely on HOM interference, $\bar{\mathcal{F}}_S^{\text{Exp, LOCC}} = (73.4 \pm 0.2)\%$ would remain unchanged. We see that when accounting for systematic errors, the collective strategy outperforms the LOCC approach.

3.4.2 TetraMP Game Results

In this section, we present experimental results for the TetraMP game, where the *a priori* distribution $P(n)d\vec{n}$ is such that the only possible $|\vec{n}\rangle$ are the four tetrahedron states $|\uparrow_{n_i}\rangle$ (with equal probability). The measured fidelities for the rotated tetrahedron states are presented in Fig. 3.8. The average fidelity for the TetraMP game was found to be $\bar{\mathcal{F}}_{\text{Tetra}}^{\text{Exp, Coll}} = (81.2 \pm 0.3)\%$. This was compared with a version of the measurement, performed with a long delay between the photons, with no other changes to the setup. This corresponds to a complete suppression of the entangling component of the collective measurement. The average fidelity in this case was measured to be $\bar{\mathcal{F}}_{\text{Tetra}}^{\text{Exp, Supp-Ent}} = (63.8 \pm 0.3)\%$. This gap indicates the importance of the entangling measurement in approaching the optimal bound, improving the fidelity by $\sim 17\%$.

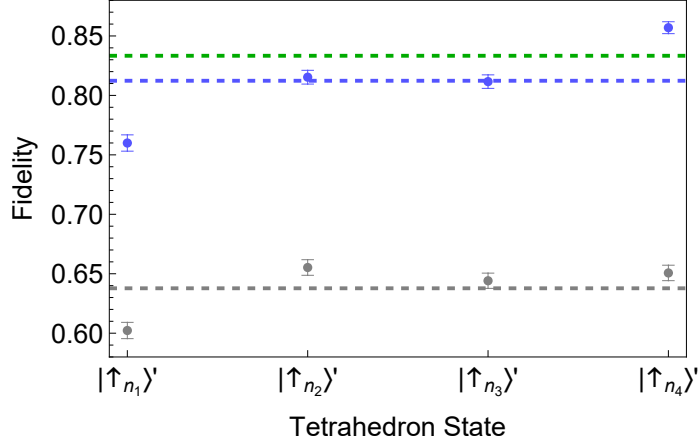


Figure 3.8: Fidelity of the Massar and Popescu two-copy game for the rotated tetrahedron states $|\uparrow_{n_i}'\rangle = \hat{R}|\uparrow_{n_i}\rangle$, where $\hat{R}$ is a fixed rotation in the Bloch sphere. These states are equivalent to Eq. (3.3) for the optimal collective measurement (see Appendix D.1 for details). The green dashed-line indicates the expected fidelity of $\mathcal{F}_{\text{Tetra}}^{\text{Coll}} = 5/6 \approx 0.8333$. The experimental collective measurement results are presented in blue, while the experimental measurement results with suppressed entanglement are in gray. The dashed lines represent the average for each data set. The error bars are obtained by assuming Poisson statistics and using Gaussian error propagation.

The deviations in tetrahedron state fidelity from the expected value of 83.33% can be in part attributed to the unitary $\tilde{U}$, which slightly distorts the measured frequencies by affecting the HOM interference statistics. These frequency deviations depend on the state prior to the beamsplitter rather than the prepared state $|\vec{n}\rangle \otimes |\vec{n}\rangle$. As a result, we are less likely to correctly guess the input $|\uparrow_{n_1}'\rangle$ and more likely to do so for the input $|\uparrow_{n_4}'\rangle$. This trend is also visible in Fig. 3.7(a), where states with a significant overlap with $|\uparrow_{n_1}'\rangle$ (in the plot's bottom right) exhibit lower fidelities than expected, while states with substantial overlap with $|\uparrow_{n_4}'\rangle$ (with $\theta \approx 1.2$ and $\phi \approx 5.2$) show higher than expected fidelity.

3.4.3 Tomographic Infidelity Scaling for Two-Copy Collective Measurements

To analyze the performance of the state tomography, the tomographic infidelity $1 - \bar{F}$ is plotted as a function of the size N_{ens} of the identically prepared ensemble. The exact state that we produce will differ from the nominal state due to experimental imperfections. Thus, we compare our state estimate to the pure state estimate with the highest N_{ens} gathered ($N_{\text{ens}} \sim 15000$) instead of the nominal state. Fig. 3.9 shows the scaling of the

infidelity observed for six randomly chosen states. Each data point is an average value over multiple reconstructions that used identically prepared ensembles of the same size N_{ens} . To mitigate bias toward the reference state at higher N_{ens} values, we applied a power-law fit to the initial eight data points (further details are provided in the Supplemental document). We observed that the infidelity for each state approximately scales as $O(1/N_{\text{ens}})$, the optimal infidelity scaling according to the Gill-Massar bound. The six random states examined had a combined average infidelity scaling of $1 - \bar{F} = (1.7 \pm 0.2)N_{\text{ens}}^{-(1.01 \pm 0.02)}$, where the fit uncertainties for the states have been propagated. Infidelity scaling results for all tested states can be found in Tab. 3.1. The coefficient could likely be improved with further attempts to minimize experimental imperfections and systematic errors. The optimal measurement (with $1 - \bar{F} = 1/N_{\text{ens}}$) requires an N_{ens} -copy collective measurement of *all* the systems in the identical ensemble, but it is encouraging to see that the two-copy collective measurement tomography does achieve a near-optimal scaling.

It is worth noting that for exactly pure one qubit states distributed according to the Haar measure, separable six-outcome Pauli tomography would theoretically allow an asymptotic average infidelity scaling of $(13/12)N^{-1}$ [103, 109] and the use of either two-stage adaptive tomography [109] or random local von Neumann measurements [110] would improve the asymptotic average infidelity scaling to $1/N$.

Reference $ \vec{n}\rangle$	$1 - \bar{F}$
(a) $0.549 H\rangle - (0.777 + 0.309i) V\rangle$	$(2.0 \pm 0.4)N_{\text{ens}}^{-1.02 \pm 0.04}$
(b) $0.883 H\rangle + (-0.216 + 0.416i) V\rangle$	$(1.5 \pm 0.4)N_{\text{ens}}^{-0.95 \pm 0.04}$
(c) $0.819 H\rangle + (0.565 - 0.098i) V\rangle$	$(1.5 \pm 0.2)N_{\text{ens}}^{-1.02 \pm 0.02}$
(d) $0.308 H\rangle - (0.916 + 0.256i) V\rangle$	$(1.6 \pm 0.4)N_{\text{ens}}^{-1.04 \pm 0.05}$
(e) $0.104 H\rangle - (0.459 + 0.882i) V\rangle$	$(1.8 \pm 0.6)N_{\text{ens}}^{-1.01 \pm 0.06}$
(f) $0.205 H\rangle + (0.513 + 0.834i) V\rangle$	$(1.8 \pm 0.3)N_{\text{ens}}^{-1.01 \pm 0.03}$

Table 3.1: Randomly chosen pure reference states and the corresponding power-law fits for tomography using two-copy collective measurements.

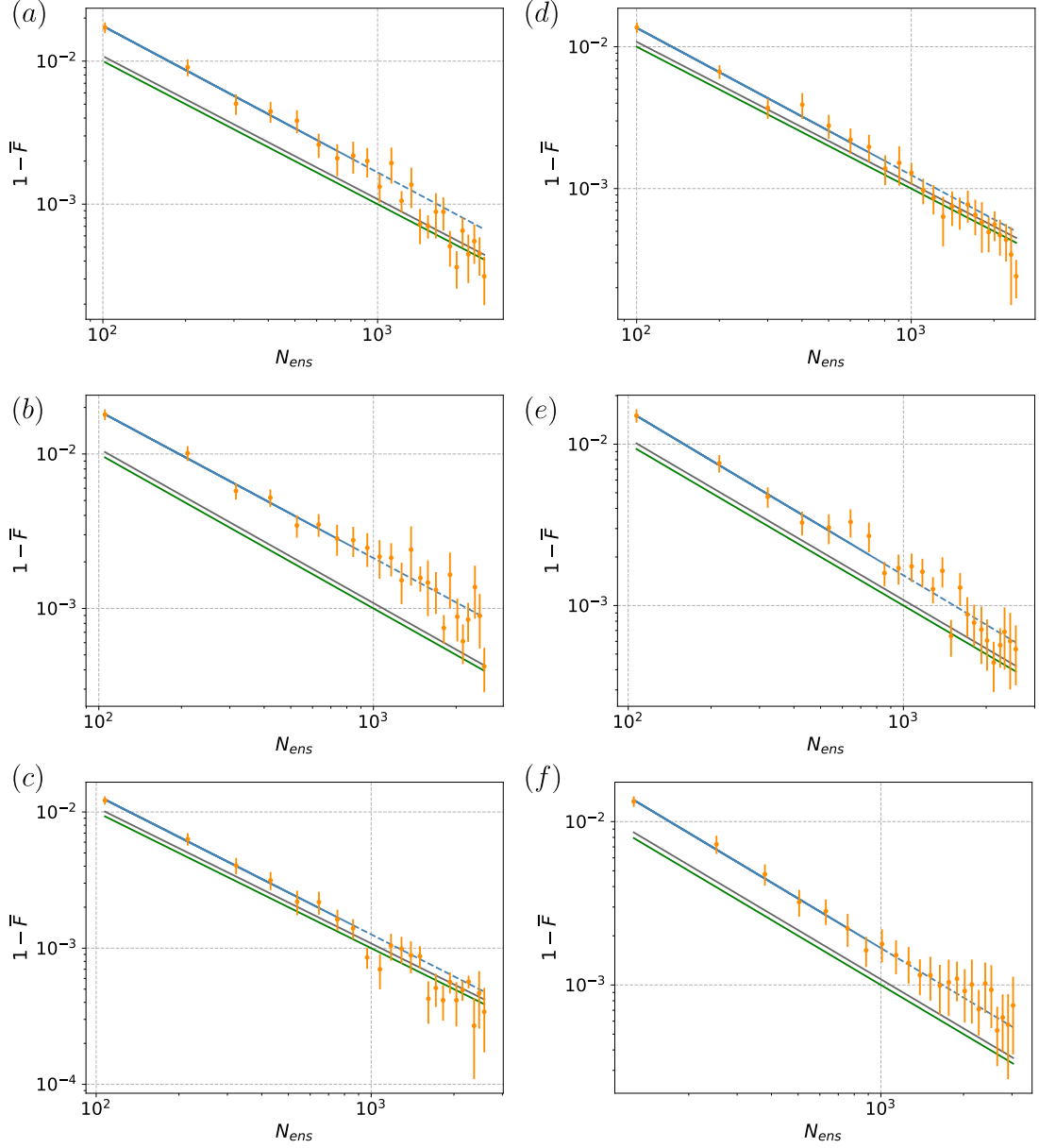


Figure 3.9: Scaling of the tomographic infidelity with the number of samples N_{ens} for the randomly-chosen states in Tab. 3.1 shown on a log-log plot. The average values over multiple experimental trials for the collective measurement tomography are shown in orange, with the blue line showing the power-law fits to the data, which are summarized in Tab. 3.1. The green line indicates the Gill-Massar bound $1 - \bar{F} = 1/N_{\text{ens}}$, while the grey line indicates the theoretical average infidelity of Pauli tomography for pure states, given by $1 - \bar{F} = (13/12)N_{\text{ens}}^{-1}$.

The scaling analysis could alternatively be performed with nearly-pure reference state estimates (purity $(99.91 \pm 0.02)\%$) obtained by averaging reconstructions with $N_{\text{ens}} \sim 2500$. The MLE output is still constrained to being a pure state. This constraint is used since $|\text{MP}_i\rangle$ states are mainly useful for pure-state tomography. In this case, the six randomly tested states show a combined average infidelity scaling of $1 - \bar{F} = (1.18 \pm 0.11)N_{\text{ens}}^{-(0.93 \pm 0.02)}$, where the fit uncertainties for the states have been propagated. These infidelity results are presented in Fig. 3.10 and Tab. 3.2. This suggests that the collective strategy is fairly robust in maintaining a near-optimal infidelity scaling. This is significant, as infidelity scaling for Pauli tomography generally deteriorates for nearly-pure states due to the challenges in accurately estimating a small eigenvalue of the density matrix [46, 103].

The goal of Section 3.4.3 was to demonstrate the experimental feasibility of quantum state tomography using multi-particle collective measurements, and to show that it is comparable to the Gill-Massar bound in terms of average infidelity scaling. Additionally, we observe this $\sim O(1/N)$ scaling with four outcomes as opposed to the six or more needed for the relevant Pauli, adaptive, and random tomography schemes [46, 48, 103, 109, 110, 113].

Reference ρ	$1 - \bar{F}$
(a) $0.302 H\rangle\langle H + (-0.426 + 0.169i) H\rangle\langle V $ $+ (-0.426 - 0.169i) V\rangle\langle H + 0.698 V\rangle\langle V $	$(1.6 \pm 0.3) N_{\text{ens}}^{-0.97 \pm 0.04}$
(b) $0.780 H\rangle\langle H - (0.191 + 0.367i) H\rangle\langle V $ $+ (-0.191 + 0.367i) V\rangle\langle H + 0.220 V\rangle\langle V $	$(1.2 \pm 0.3) N_{\text{ens}}^{-0.90 \pm 0.04}$
(c) $0.671 H\rangle\langle H + (0.462 + 0.081i) H\rangle\langle V $ $+ (0.462 - 0.081i) V\rangle\langle H + 0.329 V\rangle\langle V $	$(1.0 \pm 0.1) N_{\text{ens}}^{-0.95 \pm 0.02}$
(d) $0.095 H\rangle\langle H - (0.282 + 0.079i) H\rangle\langle V $ $+ (-0.282 + 0.079i) V\rangle\langle H + 0.905 V\rangle\langle V $	$(1.3 \pm 0.3) N_{\text{ens}}^{-0.99 \pm 0.04}$
(e) $0.011 H\rangle\langle H + (-0.048 + 0.092i) H\rangle\langle V $ $- (0.048 + 0.092i) V\rangle\langle H + 0.989i V\rangle\langle V $	$(1.0 \pm 0.3) N_{\text{ens}}^{-0.92 \pm 0.05}$
(f) $0.043 H\rangle\langle H + (0.105 - 0.171i) H\rangle\langle V $ $+ (0.105 + 0.171i) V\rangle\langle H + 0.957 V\rangle\langle V $	$(0.97 \pm 0.13) N_{\text{ens}}^{-0.87 \pm 0.02}$

Table 3.2: Randomly chosen nearly-pure reference states and the corresponding power-law fits for tomography using two-copy collective measurements.

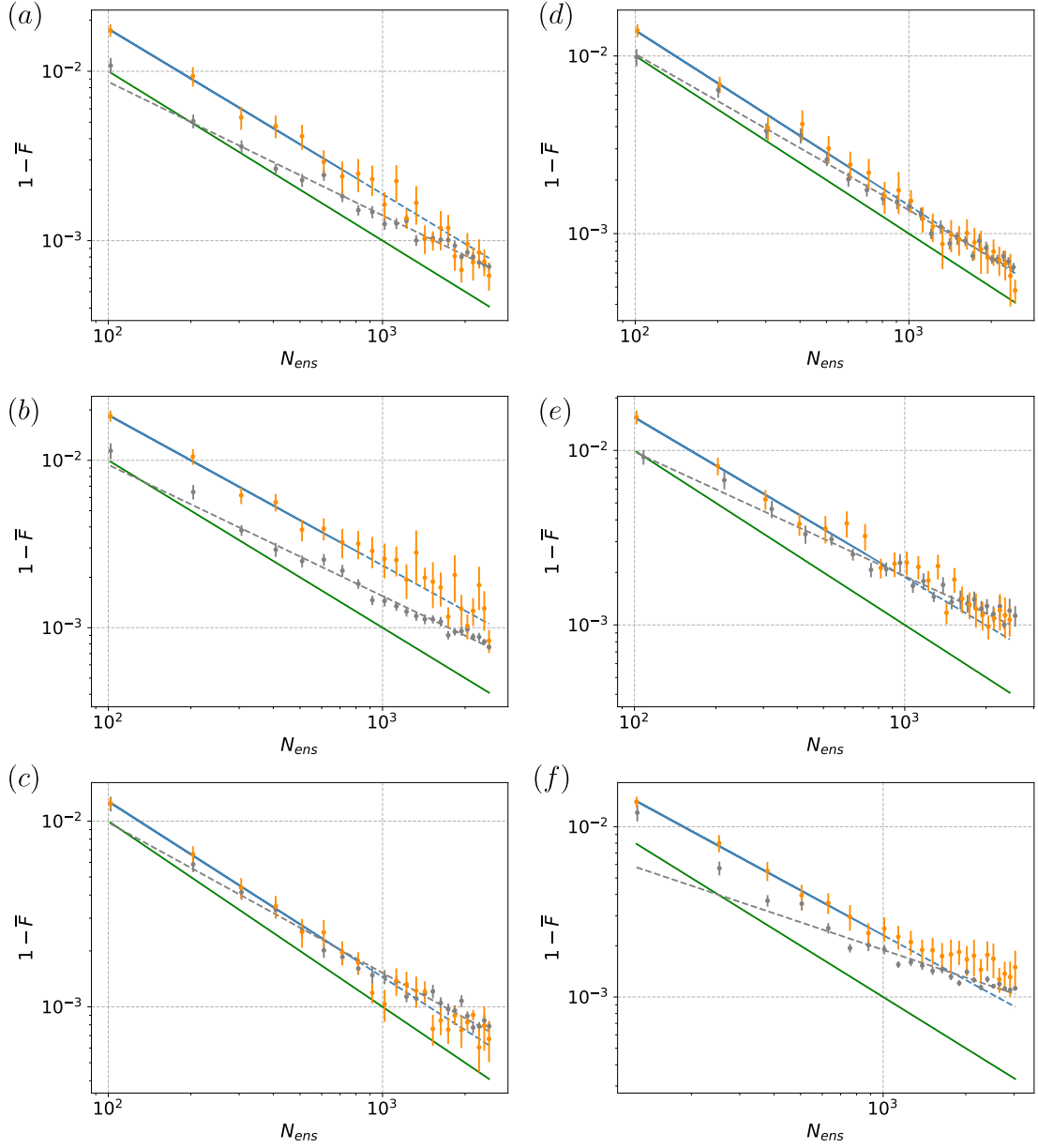


Figure 3.10: Scaling of the tomographic infidelity with the number of samples N_{ens} for the randomly-chosen states in Tab. 3.2 shown on a log-log plot. The average values over multiple experimental trials for the collective measurement tomography are shown in orange, with the blue line showing the power-law fits to the data, which are summarized in Tab. 3.2. The plotted error bars are calculated as the standard error of the mean over multiple trials. The green line indicates the Gill-Massar bound $1 - \bar{F} = 1/N_{\text{ens}}$. Pauli tomography simulation results with a pure MLE reconstructed state are shown in grey.

3.5 Summary

We performed a proof-of-concept demonstration of Massar and Popescu’s state estimation game using collective measurements on two particles. Our results suggest that the collective strategy can be in practice at least as effective as the best LOCC approach, and outperforms the local approach if systematic errors can be further reduced.

We further applied collective measurements to quantum state tomography, observing a near-optimal scaling of the infidelity with the size of the identically prepared ensemble. The effect of considering a pure reference input state as opposed to a nearly-pure state for tomographic comparison was examined. Although a pure reference results in a slightly improved tomographic scaling, the collective strategy approaches the optimal scaling in both cases. The significance of this can be seen in view of the quadratic deterioration in infidelity scaling of Pauli tomography for nearly-pure states [46].

Our results also emphasize the crucial role of entanglement in performing multi-particle collective measurements for state estimation, as evidenced by the results of the TetraMP game: our entangling measurements greatly surpassing those where entanglement was suppressed.

Our work opens the avenue to further implementations of collective measurements in the study of optimal strategies for parameter and state estimation tasks. Collective measurements can be particularly beneficial in systems where state-preparation is an expensive resource.

Chapter 4

Fast and Automated Compensation of Fiber Unitaries

This chapter is closely adapted from the paper cited below,

Reproduced from N. Braband, A. Mansouri, R. Fazili, S. Czischek, and J. S. Lundeen, “Fast and automated optical polarization compensation of fiber unitaries,” [Review of Scientific Instruments](#), vol. 96, p. 063101, June 2025, with the permission of AIP Publishing.

4.1 Motivation

The polarization of light plays a crucial role in a diverse range of applications, such as quantum cryptography, medical and biological imaging [161–163]. Accurate characterization and control of polarization states is a key requirement for effective operation of polarization-based information processing, sensing, and imaging schemes. In this work, we apply a fast and compact method for tomographic state estimation to a commonly encountered task in polarization control.

The polarization state can significantly alter as the light passes through various media, particularly in optical telecommunication fibers [164]. In these fibers, the routing geometry (e.g., bending or twisting) and manufacturing imperfections lead to birefringence, wherein the orientation of the slow and fast axes and the degree of birefringence vary along the fiber. In turn, this birefringence causes an unpredictable change in the input state of polarization (SOP) and introduces errors in the information encoded in polarization.

For example, an input bit might be encoded as horizontal (0 bit value) or vertical polarization (1 bit value) but the output might be right- or left-handed circularly polarized. Since the latter is an equal combination of horizontal and vertical polarization, the information is effectively completely randomized. This deleterious effect is compounded by its variability over time, influenced by movement and environmental changes such as temperature and pressure. The resulting error rate is particularly a problem for quantum key distribution (QKD) [165, 166], since when combined with photon loss and detector dark counts, it limits the range over which the cryptographic key can be transmitted. Polarization manipulation is also essential in various areas, with applications ranging from biomedical imaging to astronomical observation [167, 168]. Therefore, continuous polarization monitoring and rapid polarization manipulation would help maintain consistent performance in technologies reliant on accurate polarization states.

A straightforward method for characterizing an unknown polarization state is the rotating quarter-waveplate (QWP) technique [25]. This approach uses a quarter-wave plate and a polarizer to measure the four-element Stokes vector. By rotating the waveplate and analyzing the resultant light intensity changes, all components of the Stokes vector are determined. The speed of this method can vary depending on the rotation frequency of the waveplate, with faster rotation increasing the characterization speed. Once the SOP is known, a rapid and arbitrary polarization transformation can be achieved using three liquid crystal cells. This allows for the transformation of the SOP to a desired state, controlled solely by adjusting the operating voltages of the liquid crystals [169].

4.2 Theoretical Framework

4.2.1 Previous Work on Polarization Compensation

Existing polarization compensation methods often rely on stochastic techniques, interferometry, or quantum-bit-error-rate (QBER) minimization, which can be inefficient, particularly in long-range QKD systems where they result in low key rates [166, 170–174]. In Ref. [175], active compensation of polarization errors is achieved with adaptive optics via feedback to the polarization state generator, which is impractical in applications with limited control of the initial polarization. The work presented in [176] employs a polarimetry technique using a single liquid crystal cell, where both the voltage and angle are varied. However, the intermittent rotation of the liquid crystal, with frequent stops, poses a challenge to achieving fast and reliable compensation. Additionally, the work lacks a rigorous proof of the ability to change an arbitrary polarization to another, as demonstrated in [169].

Moreover, electronic digital signal processing [177], which involves compensation through measurement of the full field (both polarizations, amplitude, and phase) is not possible for the polarization qubits in QKD without state disturbance. For slow polarization drift (over several hours), calibration using reference signals can identify the relative basis rotation, while reference-frame-independent protocols would relax this requirement [178]. When the drift is faster or random in conventional QKD, the quantum bit error rates increase, reducing the secret key rate. In this case, active compensation is needed to mitigate the errors.

In this work we introduce a novel approach that relies on polarization tomography and directly solving for the compensation parameters, enabling fast polarization compensation with high fidelity. Specifically, we integrate the rotating quarter-waveplate polarimetry technique with polarization transformation by adjusting three liquid crystal cell voltages. Since our method does not rely on cryptographic figures of merit, it can be implemented on polarization-based QKD systems regardless of the range and key rates. Contrary to [176], we only vary the voltage of three liquid crystals set at fixed angles, quickly compensating for the effect of any arbitrary unitary transformation.

Moreover, there is a lack of comprehensive experimental compensation procedures, including required codes from detection to correction of polarization, which has hindered their broader adaptation. To address this, our work includes detailed technical information and essential codes, offering a helpful resource for those interested in implementing and exploring these methods. We believe this work will be beneficial in many systems requiring precise and continuous polarization control.

4.2.2 Rotating QWP Tomography Method

All Stokes parameters can be measured simultaneously using a rotating QWP, a polarizing beam-splitter (PBS) and a photo detector (PD) or power meter [25]. The incoming light first passes through the rotating QWP at an angle ϕ between the fast and the horizontal axes. The intensity $\tilde{S}_0$ transmitted by the PBS is then measured by the PD. The Stokes vector is thus transformed as follows:

$$\begin{aligned} \vec{\tilde{S}}(\phi) &= \mathbf{M}_{\text{PBS}} \cdot \mathbf{M}_{\text{QWP}}(\phi) \cdot \vec{S} \\ &= \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2(2\phi) & \sin(2\phi)\cos(2\phi) & -\sin(2\phi) \\ 0 & \sin(2\phi)\cos(2\phi) & \sin^2(2\phi) & \cos(2\phi) \\ 0 & \sin(2\phi) & -\cos(2\phi) & 0 \end{bmatrix} \cdot \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix}. \end{aligned} \quad (4.1)$$

From Eq. (4.1), the transmitted intensity $\tilde{S}_0$ is given by:

$$\begin{aligned}
I(\phi) &= \tilde{S}_0(\phi) \\
&= \frac{1}{2}(S_0 + S_1 \cos^2 2\phi + S_2 \sin 2\phi \cos 2\phi - S_3 \sin 2\phi) \\
&= \frac{1}{2}(A_0 + B_0 \sin 2\phi + C_0 \cos 4\phi + D_0 \sin 4\phi).
\end{aligned} \tag{4.2}$$

The intensity can be rewritten using

$$A_0 = S_0 + \frac{S_1}{2}, \quad B_0 = -S_3, \quad C_0 = \frac{S_1}{2}, \quad D_0 = \frac{S_2}{2}, \tag{4.3}$$

and can now be seen as a truncated Fourier series. As only a finite number of samples are measured for the intensity, each Fourier coefficient integral is turned into a sum with $\Delta\phi$ being the angle step size between two measurements and N the number of samples, resulting in:

$$A_0 = \frac{1}{\pi} \int_0^{2\pi} I(\phi) d\phi = \frac{2}{N} \sum_{n=1}^N I_n, \tag{4.4}$$

$$B_0 = \frac{2}{\pi} \int_0^{2\pi} I(\phi) \sin(2\phi) d\phi = \frac{4}{N} \sum_{n=1}^N I_n \sin(2n\Delta\phi), \tag{4.5}$$

$$C_0 = \frac{2}{\pi} \int_0^{2\pi} I(\phi) \cos(4\phi) d\phi = \frac{4}{N} \sum_{n=1}^N I_n \cos(4n\Delta\phi), \tag{4.6}$$

$$D_0 = \frac{2}{\pi} \int_0^{2\pi} I(\phi) \sin(4\phi) d\phi = \frac{4}{N} \sum_{n=1}^N I_n \sin(4n\Delta\phi). \tag{4.7}$$

4.2.3 Arbitrary Polarization Manipulation Using Three Fixed Liquid Crystals

As shown in Ref. [169], any polarization state can be transformed into another state using three full-wave LCVRs with their slow axes at 0° , 45° , and 0° respectively. Geometrically, this configuration corresponds to rotations with respect to the HV , DA , and HV axes in the Poincaré sphere, respectively. The applied LCVR voltage determines the retardance δ_i , i.e., the amount of rotation in the Poincaré sphere.

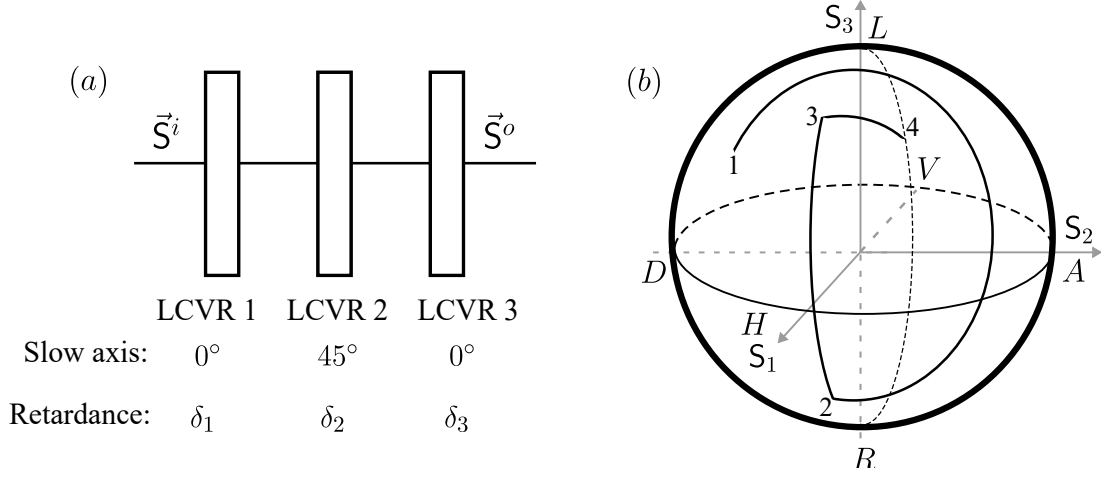


Figure 4.1: Arbitrary polarization transformation with three LCVRs. Part (a) shows the LCVR configuration, while part (b) shows the resulting rotations in the Poincaré sphere.

The normalized Stokes vector $\vec{S}^i$ of the input state can be written in terms of the polar angle θ and azimuthal angle ϕ as

$$\begin{aligned} S_1^i &= \cos(2\theta) \\ S_2^i &= \sin(2\theta) \cos(\phi) \\ S_3^i &= \sin(2\theta) \sin(\phi). \end{aligned} \quad (4.8)$$

If the first LCVR retardance is chosen such that $\delta_1 - \phi = \pi/2$, the output state $\vec{S}^o = \mathbf{M}_{\text{LCVR}}(0^\circ, \delta_3) \mathbf{M}_{\text{LCVR}}(45^\circ, \delta_2) \mathbf{M}_{\text{LCVR}}(0^\circ, \delta_1) \vec{S}^i$ can be written as

$$\begin{aligned} S_1^o &= \cos(\delta_2 - 2\theta) \\ S_2^o &= \sin \delta_3 \sin(\delta_2 - 2\theta) \\ S_3^o &= \cos \delta_3 \sin(\delta_2 - 2\theta), \end{aligned} \quad (4.9)$$

which we identify with the coordinates of a point on the Poincaré sphere with polar angle $\delta_2/2 - \theta$ and azimuthal angle δ_3 . Eq. (4.9) can be solved for the retardances δ_i [169]:

$$\begin{aligned} \delta_1 &= \pi/2 + \arctan(S_3^i/S_2^i) + m\pi \\ \delta_2 &= \arccos(S_1^i) + \arccos(S_1^o) \\ \delta_3 &= \arctan(S_2^o/S_3^o) + n\pi, \end{aligned} \quad (4.10)$$

where $m = 1$ when $S_3^i < 0$, $n = 1$ when $S_2^o < 0$, and $m = n = 0$ otherwise. Eq. (4.10) provides the required retardances to transform any input SOP $\vec{S}^i$ to any output SOP $\vec{S}^o$ using the LCVR configuration in Fig. 4.1.

4.2.4 Fidelity in Terms of Stokes Parameters

In terms of Stokes vectors of SOPs $\vec{S}_a$ and $\vec{S}_b$, the fidelity can be expressed as [104]:

$$F(\vec{S}_a, \vec{S}_b) = \frac{1}{2} \left[1 + \sum_{i=1}^3 S_{i,a} S_{i,b} + \sqrt{(1 - |\vec{S}_a|^2)(1 - |\vec{S}_b|^2)} \right]^2. \quad (4.11)$$

If one of the SOPs, for example $\vec{S}_b$ is fully polarized, $|\vec{S}_b|^2 = 1$ and the fidelity reduces to:

$$F(\vec{S}_a, \vec{S}_b) = \frac{1}{2} (1 + S_{1,a} S_{1,b} + S_{2,a} S_{2,b} + S_{3,a} S_{3,b}). \quad (4.12)$$

Given the target SOP that we try to match through polarization compensation is always fully polarized, we use Eq. (4.12) to calculate the closeness between the measured and the target SOPs.

4.2.5 Polarization Compensation Scheme

In optical fibers, random birefringence caused by thermal fluctuations as well as twisting and bending the fiber change the SOP of the light passing through the fiber. Three consecutive LCVRs at 0° , 45° , and 0° (relative to the slow axis for each LCVR) can be used to transform an arbitrary SOP to any other SOP [169]. This is important as it allows compensating for any unitary transformation of the SOP applied by the fiber. The Mueller matrix $\mathbf{M}_{\text{LCVRs}}$ for the three LCVRs is shown in Eq. (E.10) in Appendix E.2. For the polarization compensation, the algorithm sketched in Fig. 4.2 is used.

In the beginning, all retardances of the LCVRs are set to some arbitrary values δ'_i . Afterwards, the SOP $\vec{S}_{\text{meas}}$ of the light is measured. For any retardances, one can calculate the SOP of the light before the LCVRs from $\vec{S}_{\text{meas}}$. The disturbed SOP $\vec{S}_{\text{dis}}$ after propagating through a single-mode fiber is given by:

$$\vec{S}_{\text{dis}} = \mathbf{M}_{\text{LCVRs}}^{-1}(\delta'_1, \delta'_2, \delta'_3) \cdot \vec{S}_{\text{meas}}. \quad (4.13)$$

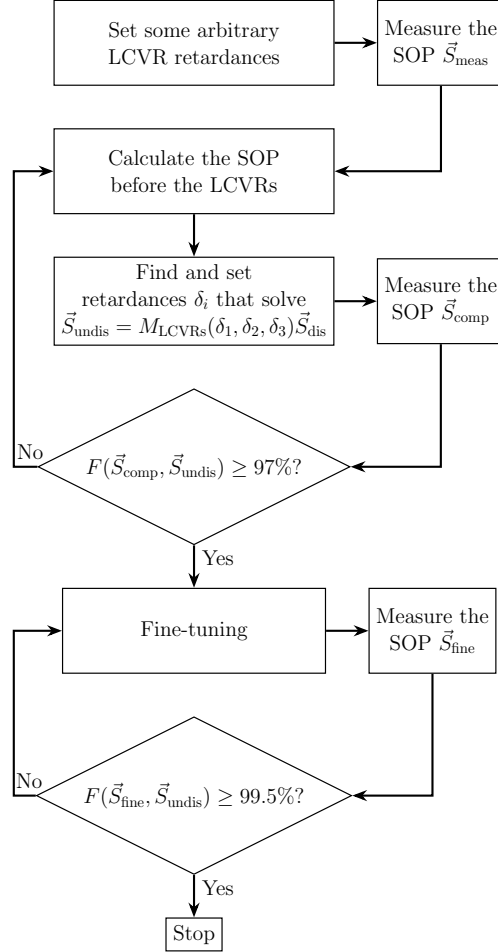


Figure 4.2: Algorithm for compensation of fiber unitaries. Here, SOP and LCVR stand for the state of polarization and liquid crystal variable retarders, respectively. We propose and implement this algorithm, which relies on polarization tomography and directly solving for the liquid crystal compensation retardances. Once a fidelity threshold is crossed, small variations of the retardances are used to reach the desired fidelity (here was 99.5 % fidelity).

Now that $\vec{S}_{\text{dis}}$ is known, one can calculate the required retardances for the LCVRs to compensate for the polarization change in the fiber. The system of equations for the compensation can be solved to get the required compensation retardances δ_i :

$$\vec{S}_{\text{undis}} = \mathbf{M}_{\text{LCVRs}}(\delta_1, \delta_2, \delta_3) \cdot \vec{S}_{\text{dis}}, \quad (4.14)$$

where $\vec{S}_{\text{undis}}$ is the desired undisturbed polarization state. The calculated compensation retardances are then set, and the polarization is measured again. The polarization at this

stage is $\vec{S}_{\text{comp}}$. If a 97 % fidelity threshold is satisfied, we move on to the fine-tuning stage. After each fine-tuning iteration, we measure the fine-tuned state $\vec{S}_{\text{fine}}$. If a 99.5 % fidelity threshold is met, the algorithm stops. These thresholds can be adjusted according to the desired accuracy.

Given our knowledge of $\vec{S}_{\text{undis}}$ and an estimate of $\vec{S}_{\text{dis}}$, Eq. (4.14) can be solved numerically using the Python `scipy.optimize.fsolve` function [179] (equivalent to the analytic approach of [169] discussed in Section 4.2.3). The values $\delta_1, \delta_2, \delta_3$ obtained from the numerical solver are not bounded and can be any real scalar. Because the LCVRs have a limited retardance range, the solution has to be shifted to be in the $[0.2\pi, 2.2\pi)$ range by adding or subtracting multiples of 2π . Each LCVR voltage is chosen such that its corresponding retardance is closest to the solution based on the characterization data.

Although Eq. (4.14) can be solved directly (both analytically and numerically), the compensation may be less accurate if the solution is in regions where the slope of the characterization curve is steeper (Fig. 4.6). In that case, the LCVR cannot be controlled as precisely. A small voltage error can cause a large error in the retardance. The accuracy of the tomographic estimate $\vec{S}_{\text{meas}}$ also affects the compensation. These issues are addressed by returning to the step labeled “Calculate the SOP before the LCVRs” in Fig. 4.2 (with different initial retardances) and repeating until the fidelity is $> 97\%$ and a better solution is found.

Once the fidelity is $> 97\%$, a fine tuning algorithm is used. One way to improve the result is to start with the first LCVR and make minor voltage adjustments. If fidelity improves, continue adjusting the voltage in the same direction. If not, revert to the previous voltage setting and move on to the next LCVR. Repeat this process until the 99.5 % threshold is met.

4.2.6 Summary

Unlike previous methods that rely on stochastic optimization — based on cryptographic figures of merit — we devised a compensation method which uses rotating QWP tomography and LCVRs. Rotating QWP tomography is compact and can allow rapid and continuous state characterization. Three fixed LCVRs with appropriate axis orientations can compensate for any state by varying the retardance of the LCVRs.

4.3 Experimental Methods

Liquid crystal variable retarders (LCVRs) are molecules with directional properties that can be aligned to exhibit uniaxial birefringence [180]. By applying a square wave AC voltage, the retardance of the optical element changes with the root-mean square value of the applied voltage. The action of an LCVR can be described by the following Mueller matrix:

$$\mathbf{M}_{\text{LCVR}}(\theta, \delta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2\theta + \sin^2 2\theta \cos \delta & \cos 2\theta \sin 2\theta (1 - \cos \delta) & -\sin 2\theta \sin \delta \\ 0 & \cos 2\theta \sin 2\theta (1 - \cos \delta) & \cos^2 2\theta \cos \delta + \sin^2 2\theta & \cos 2\theta \sin \delta \\ 0 & \sin 2\theta \sin \delta & -\cos 2\theta \sin \delta & \cos \delta \end{bmatrix}. \quad (4.15)$$

Here, θ is the angle of the liquid crystal birefringent fast axis relative to the horizontal axis, and δ is the relative phase retardance between the horizontal and vertical polarizations (which varies with voltage).

4.3.1 Liquid Crystal Characterization Setup

A characterization of each LCVR is needed to obtain the relation between the applied voltage and the retardance at a given wavelength [181]. This can be done using horizontally polarized input light $\vec{S}_H$. The light passes through a PBS, the LCVR at 45° , a HWP at 45° , and a second PBS. Both angles are measured between the fast axis of each component and the horizontal axis.

The Mueller matrix for a HWP with its fast axis forming an angle ϕ with the horizontal is given by:

$$\mathbf{M}_{\text{HWP}}(\phi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2(2\phi) - \sin^2(2\phi) & 2 \cos(2\phi) \sin(2\phi) & 0 \\ 0 & 2 \cos(2\phi) \sin(2\phi) & \sin^2(2\phi) - \cos^2(2\phi) & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \quad (4.16)$$

The equations for $\vec{S}(\phi, \delta)$ and $I(\phi, \delta)$ are given by:

$$\begin{aligned} \vec{S}(\phi, \delta) &= \mathbf{M}_{\text{PBS}} \cdot \mathbf{M}_{\text{HWP}}\left(\frac{\pi}{4}\right) \cdot \mathbf{M}_{\text{LCVR}}\left(\frac{\pi}{4}, \delta\right) \\ &\quad \cdot \mathbf{M}_{\text{PBS}} \cdot \mathbf{M}_{\text{HWP}}(\phi) \cdot \vec{S}_H, \end{aligned} \quad (4.17)$$

$$I(\phi, \delta) = S_0(\phi, \delta) = \frac{1}{2} \cos^2(2\phi) (1 - \cos(\delta)). \quad (4.18)$$

The maximum intensity is reached when $\cos(\delta) = -1$:

$$I_{\max} = \cos^2(2\phi), \quad (4.19)$$

$$I(\delta) = \frac{I_{\max}}{2}(1 - \cos(\delta)) \Leftrightarrow \delta = \arccos\left(1 - \frac{2I}{I_{\max}}\right). \quad (4.20)$$

This establishes the retardance-intensity relationship for a given LCVR voltage. By sweeping over the full voltage range, the complete retardance characterization curve as a function of voltage can be obtained.

Because of the arccosine relation between the intensity and the retardance, the result needs to be unwrapped. The final result should be a smooth curve with no jumps or sudden changes in the derivative. A detailed description of the phase unwrapping can be found in [182].

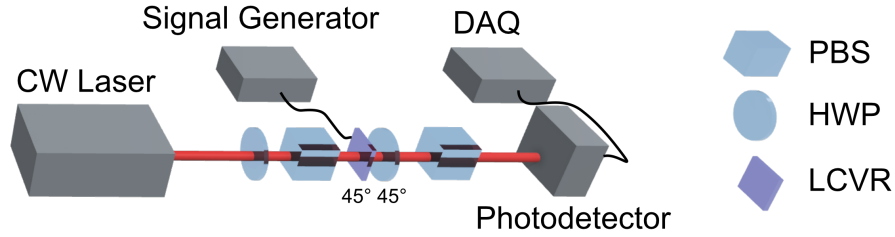


Figure 4.3: Setup to characterize liquid crystal variable retarders.

The setup for the LCVR characterization is shown in Fig. 4.3. One would need to characterize each LCVR for the wavelengths used. We used compensated full-wave compensated LCVRs to access the full range of required retardances. During the characterization, a square wave voltage with a frequency of 2 kHz was applied to the LCVR. The RMS voltage of the square-wave was swept from 0.1 V to 16 V with a step size of 0.01 V. Further details on the experimental setup and the error calculation can be found in Appendix E.

4.3.2 Rotating QWP Tomography Setup

The setup in Fig. 4.4 is used to verify the polarimetric method. For the measurement, one full rotation of the QWP is used. The rotation mount used was rotating at its fastest possible rate, which is 30° s^{-1} . The position of the QWP was measured via the rotation mount with the highest possible sampling rate, resulting in an angle step size of approximately

1.16°. Accordingly, 310 data points were used for one polarization measurement. To prevent drift and ensure accurate angle measurements, the rotation mount was re-homed and calibrated after every fifth rotation.

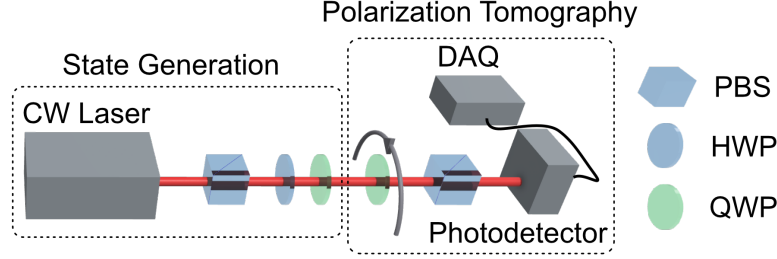


Figure 4.4: Polarization tomography setup consisting of a PBS to ensure an $\vec{S}_H$ state, a HWP and/or a QWP to prepare the different input states, and the SOP-measurement setup as described above. The photodetector voltages are recorded using a data acquisition device (DAQ).

4.3.3 Polarization Compensation Setup

The setup in Fig. 4.5 was used for polarization compensation. A manual polarization controller was used to vary the disturbance of the light. The three paddles of the fiber polarization controller approximately implement a QWP–HWP–QWP sequence, which can realize any rotation on the Poincaré sphere. The PD, the function generator, and the rotating QWP were automated and computer controlled.

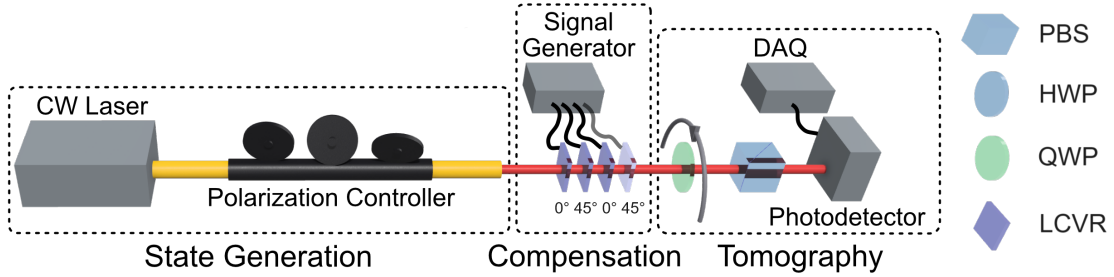


Figure 4.5: Setup for polarization compensation consisting of a polarization controller used to generate arbitrary disturbed input states, the LCVRs for the polarization compensation and the SOP-measurement setup; the last LCVR is optional and only used during the fine tuning.

Achieving the desired fidelity requires careful calibration of the liquid crystal cells, which can be influenced by several factors. Background noise and laser power fluctuations can

impact the accuracy of the calibration. Additionally, liquid crystals exhibit dependencies on temperature and wavelength, further affecting their response. These combined factors necessitate the fine-tuning of the liquid crystal cells to ensure optimal performance and achieve the target fidelity.

The system is tested by disturbing the polarization arbitrarily with a manual polarization controller in a single mode fiber (SMF) and compensating it to the arbitrarily chosen $\vec{S}_R$ polarization state. To prevent gimbal locking during the fine tuning [170], a fourth LCVR is added after the three LCVRs used for compensation. During the first part of the compensation, the retardance of the fourth LCVR does not change. It is only used in the fine tuning process.

Our scheme compensates unitary polarization noise arising from time-dependent birefringence in the fiber, such as temperature-induced index changes and mechanical bending. These correspond to slow polarization drift on sub-Hz to Hz timescales. The compensation speed is limited primarily by the polarization tomography stage. With a faster rotation mount/motor, the same approach could in principle track mechanical vibrations in the tens-of-Hz regime, provided the full compensation cycle is significantly faster than the drift rate.

4.4 Experimental Results

4.4.1 Liquid Crystal Characterization Results

As shown in Fig. 4.6, unwrapping the raw data leads to the expected curve for the nonlinear retardance-voltage relation of an LCVR [180–183]. Because the final curve exceeds a full-wave of retardance (2π), two unwrapping steps are needed. As the retardance range of the LCVR is at least a full-wave, any possible retardance can be accomplished.

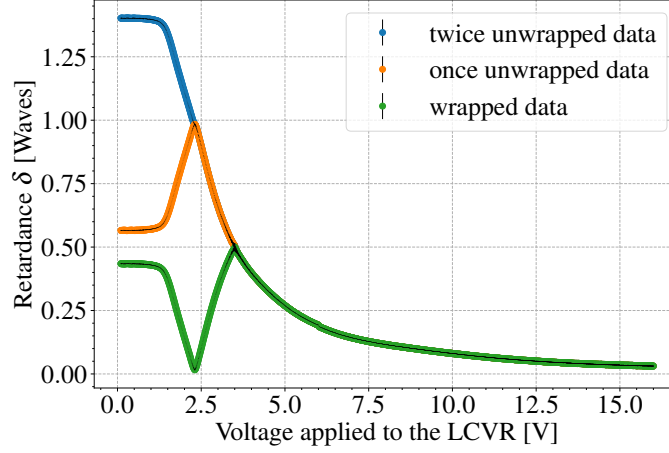


Figure 4.6: Example of the retardance-voltage relation of a LCVR at $\lambda=810$ nm; raw data (green) was unwrapped twice; errors calculated via the Gaussian error propagation (see Appendix E.2 for details).

4.4.2 Rotating QWP Tomography Results

The polarization states $\vec{S}_H$, $\vec{S}_V$, $\vec{S}_A$, $\vec{S}_D$, $\vec{S}_L$, and $\vec{S}_R$ were prepared using a QWP and a half-wave plate (HWP) as test states for the tomography system. We performed multiple rounds of state tomography for these SOPs, and calculated their corresponding fidelities, as shown in Fig. 4.7. The overall error for the polarimeter was estimated from the standard error in the mean of the measured fidelities of each SOP. Further details on the experimental setup can be found in the Appendix E.3.

The polarization tomography achieves average fidelities of 99.65 % and relatively small statistical fluctuations of 0.012 % (Fig. 4.7). The input laser power fluctuated by 0.1% over a trial duration. The error in the fidelity due to these power fluctuations is estimated to be 0.004% over 50 trials. The different mean fidelities for the input states are likely to be influenced by how well the specific input state was prepared. Therefore, the calculated standard error in the mean of the complete test measurement is most likely larger than the actual error of the polarimeter and thus a conservative value for the uncertainty.

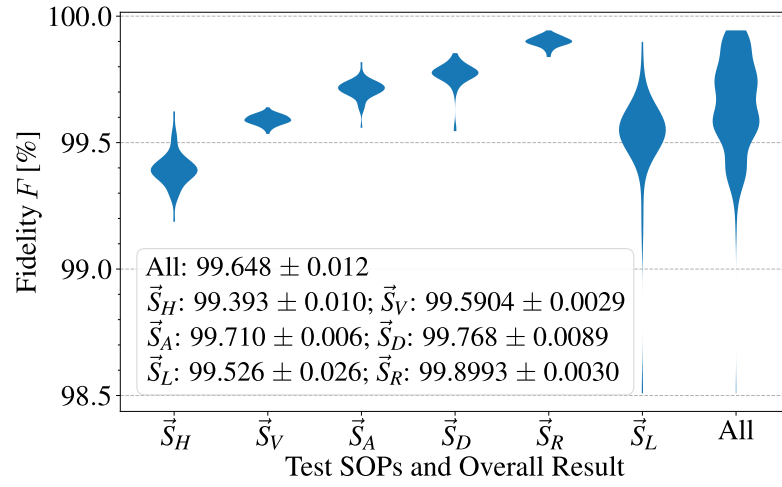


Figure 4.7: Results of the polarization tomography for six different input states are shown as a violin plot. This plot shows the fidelity distribution of the 50 data points for each input state, along with the overall distribution, average fidelity, and standard error of the mean.

4.4.3 Polarization Compensation Results

The compensation algorithm achieved fidelities over 97% within an average of 1.9 compensation steps (Fig. 4.8 and Tab. 4.1). Each compensation step involved measuring the SOP based on a complete QWP rotation. The subsequent fine-tuning improved accuracy until achieving fidelity above 99.5%. Given a maximum QWP rotation speed of 25°s^{-1} , a compensation step required approximately 30 seconds (when accounting for occasional homing of the rotation mount). Therefore, six compensation steps take 180 seconds and, on average, are sufficient to achieve over 99% fidelity. To account for variations due to different QWP rotation hardware, we chose to report our results in Fig. 8 in terms of compensation steps rather than absolute time.

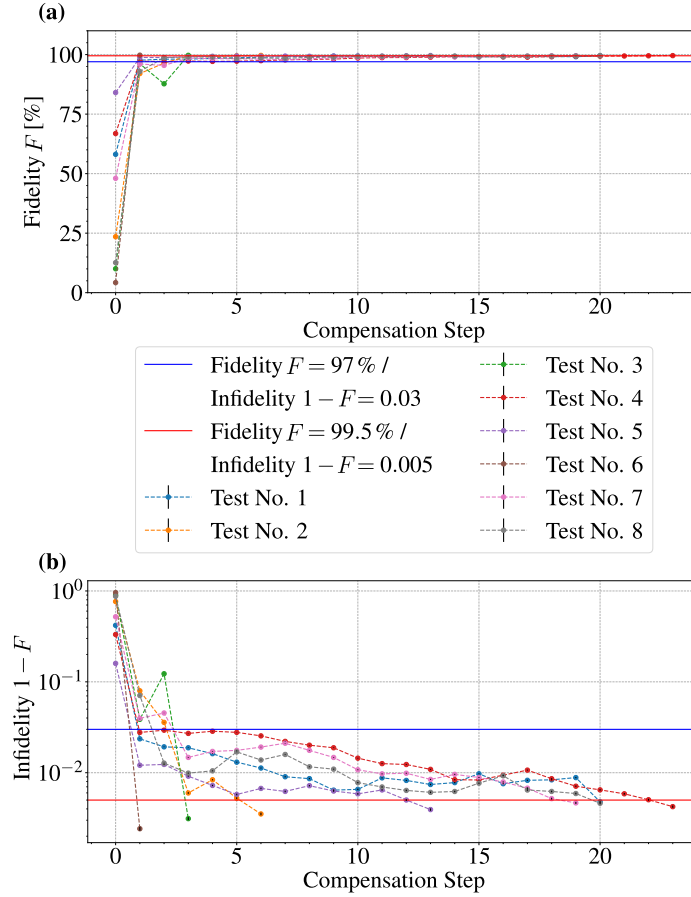


Figure 4.8: Polarization compensation with eight different arbitrary input states (configurations of polarization controller). (a) Fidelity plot for the full compensation process (b) Logarithmic plot of the infidelity.

Test number	Comp. steps until $F > 97\%$	Comp. steps until $F > 99\%$	Comp. steps until $F > 99.5\%$
1	1	7	20
2	3	3	6
3	3	3	3
4	1	14	23
5	1	3	13
6	1	1	1
7	3	11	19
8	2	3	20
Average	1.9	5.7	13.1

Table 4.1: Number of compensation steps until $F > 97\%$, $F > 99\%$ and $F > 99.5\%$. Eight tests with different polarization controller configurations were performed.

Our method is analytically direct, but in practice it runs in a feedback loop because a slight imprecision in the nonlinear retardance calibration can make the applied inverse imperfect. Disparities for the $F > 99.5\%$ threshold (e.g., 1 vs 23 steps) mainly reflect how far the initial fiber unitary is from the compensated state and how calibration errors affect that particular rotation. If the first inversion lands close, one step suffices, but if residual error is larger, several fine-tuning steps are needed.

The QBER is a crucial metric in QKD that quantifies the ratio of wrong bits to the total number of bits received. According to [184], the QBER comprises of contributions from the imperfect received state QBER_{opt} , dark counts QBER_{det} , and accidental coincidences QBER_{acc} :

$$\text{QBER} = \text{QBER}_{\text{opt}} + \text{QBER}_{\text{det}} + \text{QBER}_{\text{acc}}. \quad (4.21)$$

In polarization-based QKD, polarization compensation primarily reduces QBER_{opt} , which is the portion of the error due to unitary disturbances of the polarization state (e.g., from fiber propagation). Although Refs. [184, 185] derive this relationship in the context of entangled photon systems, we estimate the QBER using the visibility-fidelity-QBER linkage in the following way. Specifically, QBER_{opt} can be related to the polarization visibility V as:

$$\text{QBER}_{\text{opt}} = \frac{1 - V}{2}. \quad (4.22)$$

and the fidelity satisfies $F \geq V$ [21]. Combining these yields the QBER_{opt} upper bound:

$$\text{QBER}_{\text{opt}} = \frac{1 - F}{2}. \quad (4.23)$$

By assuming that the fidelity of the quantum state used in QKD is on par with our measured laser-based fidelity ($F = 99\%$), we estimate $\text{QBER}_{\text{opt}} \approx 0.5\%$. If the combined contributions of QBER_{det} and QBER_{acc} are around 1%, then the total QBER would be approximately 1.5% after an average of 5.7 compensation iterations. This relation provides a useful approximation for estimating QBER_{opt} based on measured polarization fidelity, offering practical insight into how improved polarization control can directly enhance QKD performance.

4.5 Summary

We have successfully proposed and implemented a scheme for active polarization control. Our method uses a single rotating QWP for polarization tomography, with rapid compensation achieved through LCVRs. We directly solve for the required retardances and apply further fine-tuning to reach higher fidelities. We demonstrated that our methodology can compensate for arbitrary polarization drifts within an average of fewer than six compensation iterations, achieving over 99 % fidelity.

The switching time of the LCVRs is below 150 ms, meaning the primary bottleneck was the rotation speed of the quarter-wave plate. Therefore, this method can achieve even quicker compensation by using a faster QWP rotation mount, such as the rotation mount used in [186]. Additionally, more sophisticated fine-tuning algorithms could further increase the compensation speed. This automated scheme could save researchers significant time and effort to compensate for polarization changes in fibers or other media due to environmental factors, and minimize downtime in QKD channels or polarization-based sensors.

To facilitate broader adoption, we have made the complete polarization compensation code publicly available, allowing researchers and students to implement this method with minimal effort. While our approach demonstrates robustness and efficiency, future work could focus on adapting the method to a wider range of wavelengths and urban fiber networks to improve its applicability.

Our polarization compensation method delivers four key benefits over existing approaches. First, it provides direct, deterministic compensation via polarization tomography

— bypassing time-consuming iterative or interferometric feedback and achieving >99% fidelity in under six iterations, reducing dead time in applications such as QKD. Second, a built-in fine-tuning step automatically corrects for practical disturbances (e.g., LCVR temperature shifts and laser power fluctuations), improving real-world robustness. Third, we present a complete implementation from tomography through active manipulation, supplying all code and technical details to enable immediate adoption. Finally, the use of a single rotating quarter-wave plate plus LCVRs yields a minimal, integrated hardware footprint that lowers system complexity while enhancing stability against environmental perturbations.

Beyond polarization-based QKD, our technique is equally relevant to other technologies. For instance, a recent review by He *et al.* [162] highlights the growing role of polarization optics in biomedical and clinical applications, where precise polarization control is pivotal for enhancing imaging resolution and contrast. Moreover, as demonstrated by Intaravanne *et al.* [163], advanced manipulation of three-dimensional polarization structures can be leveraged for quantum imaging, information encoding, and integrated photonic devices, where precisely tailored input polarization is crucial. Overall, this work provides a versatile, high-performance, and potentially scalable solution for polarization control, addressing critical challenges in modern optical systems.

Chapter 5

Spatial Unitary Transformations with a Sequence of Spaced Thin Phase Gratings

5.1 Motivation

Not only do we need careful characterization of quantum states, but we also need accurate and versatile methods to transform them. It is necessary to seek methods for implementing unitary operations, as they are the most crucial building blocks for transforming quantum states. Unitary transformations are essential in quantum information processing, while also having classical applications such as matrix multiplication in optical neural networks.

Among existing unitary implementation schemes are the interferometric Reck and Clements decompositions, which are mainly useful for integrated optics due to severe phase stability requirements.

The multi-plane light converter (MPLC) scheme is useful in free space for specific transformations but often relies on optimization over a very large parameter space (e.g., in wavefront matching algorithms). Despite the success of MPLCs when using wavefront matching algorithms, there is still a lack of knowledge about the possible underlying decomposition. In some cases, a surprisingly small number of phase planes can achieve unitaries of much larger dimension [39]. Perhaps this is made possible thanks to symmetries of specific transformations or the large number of degrees of freedom in each phase plane (i.e., the millions of pixels of an SLM). However, the large number of parameters has the

downsides of making optimization more computationally intensive and possibly obscuring any underlying unitary decomposition.

We consider simpler periodic phase gratings with a relatively small number of parameters in a matrix-based formalism. The action of each phase grating as well as free-space propagation can be easily described in the transverse k -space. In contrast, the phase planes obtained from wavefront matching can be far more complex and necessitate evolving the electric field at every point through the system. Being able to reduce the number of free parameters to at most a few dozen in our approach also has the potential to reveal the most crucial aspects of unitary decomposition. We explore the types of unitaries that can be realized numerically and analytically.

5.2 Theoretical Framework

Although a thin diffraction grating can perform useful transformations, such as spectral filtering, spatial mode conversion and beam shaping [187], passing each output beam through a phase grating again further increases the number of beams (Fig. 5.1). This growth in the size of the output Hilbert space at each pass through the device limits the transformations that can be realized. Specifically, unitary transformations between a few modes, such as the action of a beamsplitter, cannot be achieved with an individual thin grating. Instead, we investigate the types of unitary transformations attainable using a sequence of thin gratings with free-space propagation between them.

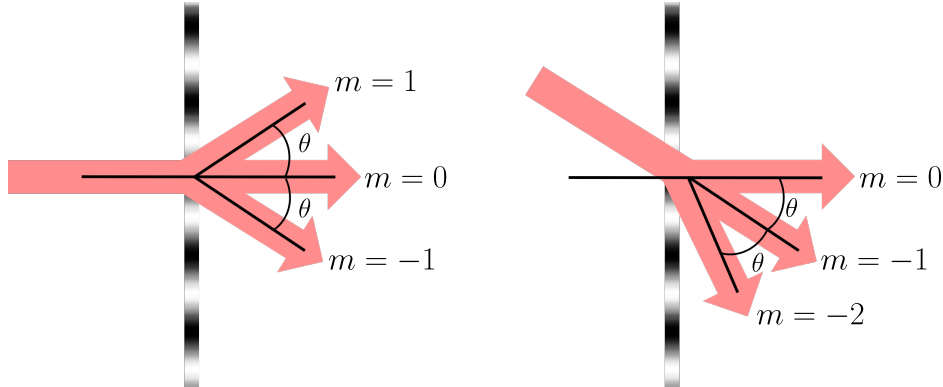


Figure 5.1: In Raman-Nath diffraction, the number of beams with a significant amplitude keeps increasing during consecutive passage through a thin sinusoidal phase grating. Each beam is diffracted into further orders. In terms of information processing, the Hilbert space dimension keeps increasing, preventing a unitary operation.

5.2.1 Raman-Nath and Bragg Diffraction Regimes

There are two primary regimes of diffraction: the Raman-Nath and the Bragg regimes. In Raman-Nath diffraction (Fig. 5.1), light is diffracted into multiple orders, with the angular distribution depending on the wavelength of the incident light, the grating wavenumber K , and the angle of incidence. For an incident plane wave e^{ikz} , the output orders have transverse wavenumbers $k_x \pm mK$, where m is an integer. The efficiencies of these various orders are influenced by the type of grating, including factors such as the grating profile. This regime arises from the interaction of light with a periodic structure, where the light waves are scattered by the grating, leading to the formation of multiple diffracted orders. Raman-Nath diffraction is commonly used in optical devices and communication systems.

In the Bragg regime, the incident light is coupled between a transmitted mode and a single diffracted mode. Diffraction occurs when light of wavelength λ incident at angle θ satisfies the Bragg condition $2d \sin(\theta) = m\lambda$, where d is the grating constant and m is an integer. This type of diffraction relies on the constructive interference of light waves scattered from parallel planes of atoms within a crystal. When the incident light meets the Bragg condition, the waves reflected from these planes interfere each other, resulting in a strong diffracted beam.

The thickness of the grating plays a significant role in determining the diffraction regime. Thin (surface) gratings typically give rise to Raman-Nath diffraction [188], while thick (volume) gratings are associated with Bragg diffraction [189].

Bragg diffraction in thick gratings limits coupling to two modes only, as required for a two-mode unitary. A thick grating can be modelled as a sequence of phase masks interleaved with free-space propagation. For instance, Ref. [190] applies a thin matrix decomposition method to thick gratings, taking the limit of an infinite number of gratings with infinitesimal thickness, such that the total grating thickness is preserved.

For the remainder of this chapter, we will be concerned with Raman-Nath diffraction and thin phase gratings. When combined with free-space propagation, this offers the potential for realizing general unitary transformations on a larger number of beams. Given the possibility of decomposing thick gratings into thin gratings interleaved with free space propagation, this approach also retains the ability to revert to Bragg diffraction and capture two-beam unitaries.

5.2.2 Amplitude Transmittance Method

The amplitude transmittance function $t_A(x)$ is defined as the ratio of the electric field amplitude over the output plane to the electric field amplitude incident on the input plane [187, 191]. For a thin phase grating with a periodic phase modulation $\phi(x)$, the amplitude transmittance $t_A(x) = e^{i\phi(x)}$ is itself periodic and has the Fourier series expansion:

$$e^{i\phi(x)} = \sum_{m=-\infty}^{\infty} c_m e^{imKx}, \quad (5.1)$$

where $K = 2\pi/\Lambda$ is the grating wavenumber and Λ is the grating period. The Fourier coefficients are then given by

$$c_m = \frac{1}{\Lambda} \int_{-\Lambda/2}^{\Lambda/2} e^{i\phi(x)} e^{-imKx} dx. \quad (5.2)$$

Because the exponentials in the series Eq. (5.1) are an expansion of diffracted plane waves, the Fourier coefficient c_m represents the diffracted wave amplitude for the order m . Moreover, the diffraction efficiency into the m^{th} order is $\eta_m = |c_m|^2$.

5.2.3 Matrix Formalism

We consider the action of thin sinusoidal phase gratings and free-space propagation in Fourier space using a matrix model similar to the one described in [189, 190]. We represent the amplitudes of incoming plane waves with transverse wavenumbers $k_x = mK$, for integer m , by a vector:

$$\tilde{\psi}(k_x) = \begin{bmatrix} \vdots \\ E(m = -1) \\ E(m = 0) \\ E(m = +1) \\ \vdots \end{bmatrix}, \quad (5.3)$$

where $E(m)$ is the amplitude of the plane wave with $k_x = mK$.

Grating Matrix

To find the matrix corresponding to the phase grating in transverse momentum space, we start with the amplitude transmittance $t_A(x) = e^{i\phi(x)}$. We know that the output of the

phase grating in transverse position (x) space can be obtained through multiplication of the input transverse field by the amplitude transmittance function as:

$$\psi'(x) = \psi(x)t_A(x). \quad (5.4)$$

We prefer to represent the electric field in transverse momentum space (k_x), where it is convenient to work with plane waves. To understand the action in k_x space, we take the Fourier transform of Eq. (5.4) and apply the convolution theorem:

$$\mathcal{F}[\psi(x)t_A(x)] = \tilde{\psi}'(k_x) \quad (5.5)$$

$$\mathcal{F}[\psi(x)] * \mathcal{F}[t_A(x)] = \tilde{\psi}'(k_x) \quad (5.6)$$

$$\tilde{\psi}(k_x) * \underbrace{\tilde{t}_{A,\theta_0}(k_x)}_{T_{A,\theta_0}} = \tilde{\psi}'(k_x), \quad (5.7)$$

where $\mathcal{F}[f(x)] = \tilde{f}(k_x)$ represents the Fourier transform, the symbol $*$ represents the convolution operation, and T_{A,θ_0} is the combined matrix for the convolution and Fourier transform of the grating function. Thus the output field $\tilde{\psi}'(k_x)$ is obtained by convolving the input field $\tilde{\psi}(k_x)$ with $\tilde{t}_{A,\theta_0}(k_x)$.

The convolution of two functions $f(x)$ and $h(x)$ is given by:

$$(f * h)(X) = \int_{-\infty}^{\infty} f(x)h(X - x)dx. \quad (5.8)$$

At each X , the output function $(f * h)(X)$ is the area under the curve of the product of the two functions with h shifted by X . The convolution of two discrete vectors f and h , with $f = [\cdots, f_{-2}, f_{-1}, f_0, f_1, f_2, \cdots]$ and $h = [\cdots, h_{-2}, h_{-1}, h_0, h_1, h_2, \cdots]$, is given by

$$(f * h)_n = \sum_{m=-\infty}^{\infty} f_m h_{n-m}. \quad (5.9)$$

We can rewrite Eq. (5.9) as a multiplication by a matrix that is constant along the diagonals — known as a *Toeplitz* matrix [192] — as:

$$f * h = \begin{bmatrix} \vdots & \vdots & \vdots & & & \\ \cdots & h_0 & h_{-1} & h_{-2} & \cdots & \\ \cdots & h_1 & h_0 & h_{-1} & h_{-2} & \cdots \\ \cdots & h_2 & h_1 & \ddots & \ddots & \ddots \\ & \vdots & h_2 & \ddots & h_0 & h_{-1} & \cdots \\ & & \vdots & \ddots & h_1 & h_0 & \cdots \end{bmatrix} \begin{bmatrix} \vdots \\ f_{-2} \\ f_{-1} \\ f_0 \\ f_1 \\ f_2 \\ \vdots \end{bmatrix}. \quad (5.10)$$

We now return to Eq. (5.7). The action of a thin grating with a periodic phase modulation $\phi(x)$ can be described by a Toeplitz matrix T_{A,θ_0} in k_x space — discretized in units of the grating wavenumber K . The matrix T_{A,θ_0} represents the convolution of the input with the Fourier transform of the grating's amplitude transmittance. It can be expressed as

$$T_{A,\theta_0} = \begin{bmatrix} \ddots & & \ddots & & \\ & c_0 & e^{-i\theta_0}c_{-1} & e^{-2i\theta_0}c_{-2} & \\ & e^{i\theta_0}c_1 & c_0 & e^{-i\theta_0}c_{-1} & \\ & e^{2i\theta_0}c_2 & e^{i\theta_0}c_1 & c_0 & \ddots \\ & & & & \ddots \end{bmatrix}, \quad (5.11)$$

where A is the amplitude (strength) of the phase modulation $\phi(x)$, θ_0 is the phase constant of $\phi(x)$, and the Fourier coefficients c_m are determined by Eq. (5.2). Note that due to the Toeplitz structure, the mn matrix element $(T_{A,\theta_0})_{mn} = e^{i(m-n)\theta_0}c_{m-n}(A)$.

Free-Space Propagation Matrix

A plane wave with wavenumber $k = 2\pi/\lambda$ — where λ is the wavelength — propagating in free-space for a length L accumulates a phase $e^{ik_z L} = e^{i\sqrt{k^2 - k_x^2}L}$. Since the transverse wavenumber k_x of diffracted plane waves will always be a multiple of the grating wavenumber K , we may write the propagation phase as $e^{i\sqrt{k^2 - (mK)^2}L}$, where m is the diffraction order. The diagonal matrix

$$D_L = \begin{bmatrix} \ddots & & \ddots & & \\ & e^{i\sqrt{k^2 - K^2}L} & 0 & 0 & \\ & 0 & e^{ikL} & 0 & \\ & 0 & 0 & e^{i\sqrt{k^2 - K^2}L} & \ddots \\ & & & & \ddots \end{bmatrix} \quad (5.12)$$

then provides the action of free-space propagation on the input state $\tilde{\psi}$ in k_x space, discretized in units of K .

Matrix for Sinusoidal and Square-Wave Phase Gratings

Sinusoidal and square wave phase gratings are natural choices in optics and their Fourier coefficients can be written in terms of familiar functions. The amplitude transmittance of a sinusoidal phase grating is given by

$$t_A^{\text{sin}}(x) = e^{iA \sin(Kx + \theta_0)}. \quad (5.13)$$

The Fourier coefficients of $t_A^{\text{sin}}(x)$ are given by (see Appendix F.1):

$$c_m^{\text{sin}} = e^{im\theta_0} J_m(A), \quad (5.14)$$

where $J_m(\cdot)$ is the m^{th} order Bessel function of the first kind.

The Toeplitz grating matrix for a sinusoidal phase grating is thus given by:

$$T_{A, \theta_0}^{\text{sin}} = \begin{bmatrix} \ddots & & & \ddots & & \\ & J_0(A) & e^{-i\theta_0} J_{-1}(A) & e^{-2i\theta_0} J_{-2}(A) & & \\ & e^{i\theta_0} J_1(A) & J_0(A) & e^{-i\theta_0} J_{-1}(A) & & \\ & e^{2i\theta_0} J_2(A) & e^{i\theta_0} J_1(A) & J_0(A) & \ddots & \\ & & & & \ddots & \ddots \end{bmatrix}. \quad (5.15)$$

The amplitude transmittance of a square-wave phase grating is given by

$$t_A^{\text{square}}(x) = e^{iA \text{sgn}(\sin(Kx + \theta_0))}, \quad (5.16)$$

where $\text{sgn}(\cdot)$ is the sign function. Using Eq. (5.2), the Fourier coefficients of the square-wave phase grating are found to be (see Appendix F.1):

$$c_m^{\text{square}} = \begin{cases} 0 & \text{if } m \text{ is even} \\ \frac{2}{m\pi} e^{im\theta_0} \sin(A) & \text{if } m \text{ is odd} \\ \cos(A) & \text{if } m = 0 \end{cases} \quad (5.17)$$

The Topelitz grating matrix for a square-wave phase grating is thus given by:

$$T_{A,\theta_0}^{\text{square}} = \frac{2}{\pi} \begin{bmatrix} \ddots & & & & & \\ & \ddots & & & & \\ & & \cos(A) & -e^{-i\theta_0} \sin(A) & 0 & -\frac{1}{3}e^{-3i\theta_0} \sin(A) & 0 \\ & e^{i\theta_0} \sin(A) & \cos(A) & -e^{-i\theta_0} \sin(A) & 0 & -\frac{1}{3}e^{-3i\theta_0} \sin(A) & \\ & 0 & e^{i\theta_0} \sin(A) & \cos(A) & -e^{-i\theta_0} \sin(A) & 0 & \\ & \frac{1}{3}e^{3i\theta_0} \sin(A) & 0 & e^{i\theta_0} \sin(A) & \cos(A) & -e^{-i\theta_0} \sin(A) & \\ & 0 & \frac{1}{3}e^{3i\theta_0} \sin(A) & 0 & e^{i\theta_0} \sin(A) & \cos(A) & \\ & & & & \ddots & & \ddots \end{bmatrix}. \quad (5.18)$$

5.2.4 Non-Zero Fourier Coefficients of a Phase Grating

Consider once again the Fourier series decomposition of the amplitude transmittance function $e^{i\phi(x)}$ for a thin phase grating:

$$e^{i\phi(x)} = \sum_{m=-N}^N c_m e^{iKmx}, \quad c_m = \frac{1}{\Lambda} \int_{-\Lambda/2}^{\Lambda/2} e^{i\phi(x)} e^{-imKx} dx, \quad (5.19)$$

where $\phi(x)$ is an arbitrary phase modulation, $K = 2\pi/\Lambda$ is the grating wavenumber, Λ is the grating period, and c_m are the Fourier coefficients. For the purposes of this calculation, we are interested in the case of finitely-many non-zero Fourier coefficients.

It is natural to ask if the phase-only nature of the grating poses any constraint on the Fourier coefficients c_m . A well-known constraint follows from Parseval's theorem regarding the unitarity of the Fourier transform. For phase gratings, this implies that $|c_m|^2$, also known as the efficiencies of the diffraction orders, must sum to one [192]:

$$\sum_{m=-N}^N |c_m|^2 = 1. \quad (5.20)$$

Additionally, we must have:

$$|e^{i\phi(x)}|^2 = \left| \sum_{m=-N}^N c_m e^{iKmx} \right|^2 = \left(\sum_{m=-N}^N c_m e^{iKmx} \right) \left(\sum_{m=-N}^N c_m e^{iKmx} \right)^* \quad (5.21)$$

$$= \underbrace{\sum_{m=-N}^N |c_m|^2}_{=1 \text{ by Parseval's Theorem}} + \sum_{m,n=-N}^N c_m^* c_n e^{iK(m-n)x} (1 - \delta_{mn}) \quad (5.22)$$

$$\stackrel{!}{=} 1 \quad (5.23)$$

$$\Rightarrow \sum_{m,n=-N}^N c_m^* c_n e^{iK(m-n)x} (1 - \delta_{mn}) = 0 \quad (5.24)$$

In the above, we have rewritten the product of the Fourier sum and its complex-conjugate as a constant sum over the efficiencies and a sum over the cross-terms. We will now state and prove a proposition.

Proposition: $2N$ of the $2N + 1$ Fourier coefficients must be zero for Eq. (5.24) to hold. That is, an arbitrary phase grating can have either one or an infinite number of non-zero Fourier components.

Proof. Consider the terms in Eq. (5.24) corresponding to e^{-2iNKx} . There is only one term, meaning we have:

$$c_N c_{-N}^* = 0 \quad (5.25)$$

Suppose $c_N \neq 0$. Then Eq. (5.25) implies that $c_{-N} = 0$.

Next, consider the terms in Eq. (5.24) corresponding to $e^{2i(N-1)Kx}$:

$$c_{-N} c_{N-1}^* + c_{-(N-1)} c_N^* = 0 \quad (5.26)$$

Since $c_{-N} = 0$, Eq. (5.26) reduces to:

$$c_{-(N-1)} c_N^* = 0 \quad (5.27)$$

However, since we assumed that $c_N \neq 0$, $c_N^* \neq 0$, meaning:

$$c_{-(N-1)} = 0 \quad (5.28)$$

Continuing on, we can consider the terms of Eq. (5.24) corresponding to $e^{-2i(N-2)Kx}$:

$$c_{N-2} c_{-N}^* + c_{N-1} c_{-(N-1)}^* + c_N c_{-(N-2)}^* = 0 \quad (5.29)$$

Since we showed that $c_{-N} = c_{-(N-1)} = 0$, Eq. (5.29) reduces to:

$$c_N c_{-(N-2)}^* = 0 \quad (5.30)$$

However, since we assumed that $c_N \neq 0$, it follows from Eq. (5.30) that $c_{-(N-2)} = 0$. Continuing in the same way, we can show that all Fourier coefficients but c_N must be zero.

Ref. [193] provides a proof of this proposition, which proceeds in a similar way as was presented here, but involves relabelling of the Fourier coefficient indices in Eq. 5.24 (which go from $-N$ to N) to range from 0 to $2N$.

5.2.5 Phase Gratings and Free-space Propagation for Achieving Unitaries

Using the matrix framework described in Section 5.2.3, we numerically investigate the transformations resulting from n layers of diffraction and free-space propagation $\tilde{U} = T_{A_n, \theta_n} D_{L_n} T_{A_{n-1}, \theta_{n-1}} D_{L_{n-1}} \cdots D_{L_1} T_{A_1, \theta_1}$.

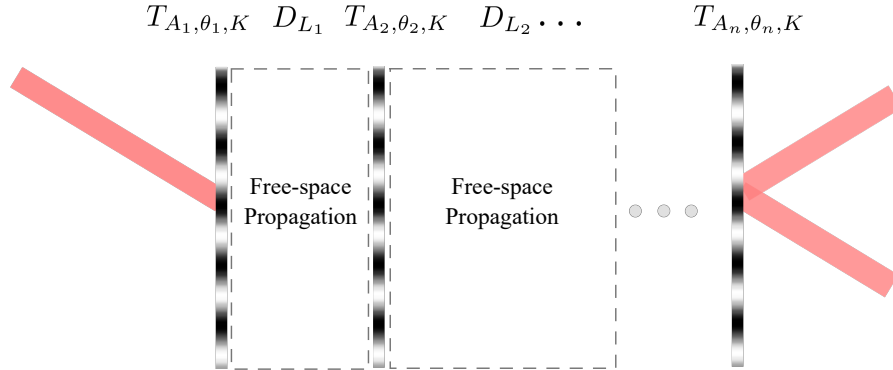


Figure 5.2: Sequence of thin sinusoidal phase gratings separated by free-space propagation. The grating parameters, including the phase modulation amplitudes A_i , phase constants θ_i , grating wavenumber K , as well as the free-space propagation lengths L_i can be optimized to approximate the desired unitary transformation.

The parameters A_i , θ_i , L_i , and K are optimized to approximate unitary transformations in the two-dimensional K subspace with $m = \pm 1$. In other examples, unitaries in the three-dimensional subspace with $m = -1, 0, +1$ are targeted. Fig. 5.3 illustrates the grating parameters and the propagation of the diffraction orders in more detail for two spaced sinusoidal phase gratings.

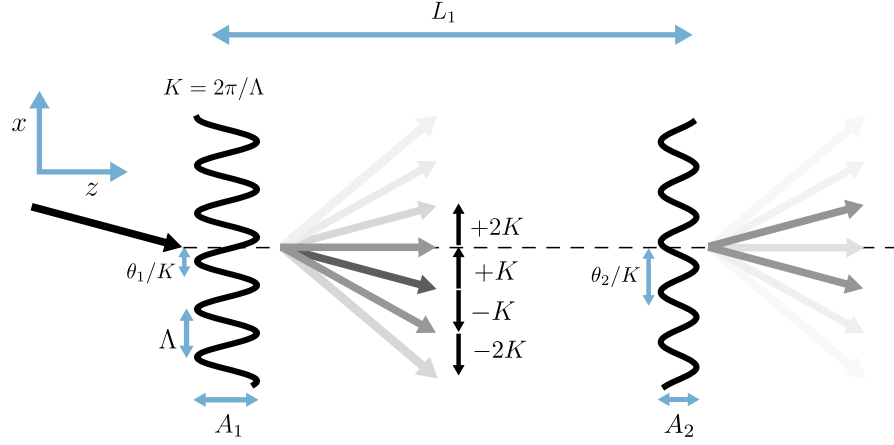


Figure 5.3: Diagram showing the various parameters of sinusoidal phase gratings and free-space propagation. The diffraction orders have transverse wavenumbers $k_x \pm mK$ where K is the grating wavenumber. An appropriate choice of A_i , θ_i , K , and L_i with sufficient layers can yield the desired couplings among the diffraction orders.

5.2.6 Genetic Algorithm

The genetic algorithm is a global optimization method inspired by natural selection. A population of candidate solutions are evolved toward better performance based on a fitness criterion through three steps: selection, crossover, and mutation. Selection involves probabilistically choosing parents, while favouring high-fitness solutions. During crossover, a subset of parameters (genes) for two parents are combined to form a “child”. Finally, mutations make random changes to the individual genes with low probability. Fitness evaluation is performed on the new population of solutions, and the procedure is repeated until convergence.

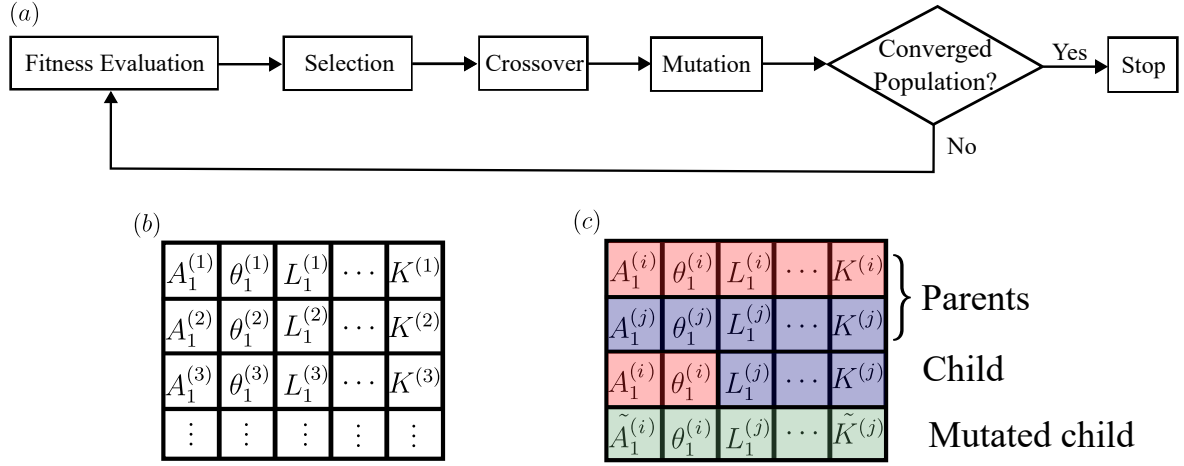


Figure 5.4: The genetic algorithm. Part (a) is a flowchart highlighting the key steps in each iteration of the genetic algorithm. Part (b) shows the initial population, with subscripts used to index the grating layer, and superscripts used to indicate different values. Part (c) depicts crossover between selected parents, followed by mutation.

5.2.7 Summary

A single thin phase grating cannot implement a unitary transformation, but we investigate whether multiple thin gratings separated by free-space can achieve this. The matrix formalism in the transverse k space provides a simple framework for exploring which unitaries can be achieved with a sequence of spaced thin gratings.

5.3 Optimization Results for Sinusoidal Gratings

We derived analytical expressions for the matrix representations of the free-space propagation operator, the grating and convolution operator, and will now consider their multiplication for a multilayer sequence. However, determining system parameters that approximate arbitrary unitaries will be performed numerically. The analytical challenge comes from the summation over propagation terms (even under the paraxial approximation) and the appearance of Bessel functions of multiple orders from the grating interaction. With the addition of more layers, the number of contributing terms grows rapidly, leading to expressions that become analytically challenging. This combinatorial growth suggests closed-form unitary expressions may be impractical.

The presence of Bessel functions and complex exponentials, related to diffraction and free-space propagation, imply that the parameter-space is highly oscillatory. Using a powerful global optimizer such as the genetic algorithm helps mitigate this challenge. It allows us to explore the types of unitaries and level of performance that can be achieved by varying the amplitudes, phases, propagation distances, and wavenumber of the phase gratings. The genetic algorithm was implemented in python using the PyGAD library [194].

In particular, the parameters optimized are the grating phase modulation amplitudes A_i , grating phase constants θ_i , free-space propagation lengths L_i , and grating wavenumber K . Fitness evaluation was based on the similarity of $\tilde{U} = T_{A_n, \theta_n} D_{L_n} \cdots D_{L_1} T_{A_1, \theta_1}$ to the target exact unitary U . For faster optimization, instead of the average gate fidelity $\bar{F}_g(U, \tilde{U})$, the quantity $\mathcal{S}(U, \tilde{U}) = |\text{tr}(U^\dagger \tilde{U})|/d$ based on the Hilbert-Schmidt inner product [19] was calculated and used as the primary term in the fitness function. Extensive testing suggests this quantity is a good measure of relative closeness despite being a slight overestimate of the gate fidelity in some cases.

Depending on the target transformation, small penalty terms were occasionally added for fine-tuning the performance or prioritizing a given aspect of the transformation. For example, for the Hadamard gate (Eq. (1.14)), adding a penalty term $-1/10(|\tilde{U}_{11}| - |\tilde{U}_{12}| + |\tilde{U}_{21}| - |\tilde{U}_{22}|)$ to the fitness function helped favour optimization with more even splitting. While the similarity $\mathcal{S}$ does encapsulate different types of deviations between $\tilde{U}$ and U , penalty terms allow us to control the balance between the requirements of being unitary (i.e., minimizing loss to other diffraction orders) and achieving the correct transformation on the light that remains in the desired Hilbert-space.

During optimization for unitaries of dimension $d = 2, 3$, only contributions from the seven diffraction orders $m = -3, -2, \dots, +2, +3$ were used, as they provide a fairly accurate estimate of the performance. For the same reason, when optimizing unitaries with dimension $d = 4, 5$, only contributions from the nine diffraction orders $m = -4, -3, \dots, +3, +4$ were calculated. After optimization, the level of performance with the obtained parameters was verified by including contributions from a larger number of diffraction orders.

The performance observed using the optimized values for A_i, θ_i, L_i , and K in terms of average gate fidelity (Eq. (1.112)) including contributions from 51 diffraction orders are discussed in Sections 5.3.1-5.3.2 and Section 5.4. All optimizations were performed for $\lambda = 1000$ nm wavelength light.

An initial subroutine focused on determining parameter guesses in a genetic algorithm using early stopping techniques. It started with 10 rounds of 15 trials, where increasingly larger numbers of generations were tested. The best results from earlier runs helped adjust

the initial guesses for the next iterations, aiming to find the most effective parameters and escape local minima. This was performed to obtain preliminary estimates of suitable system parameters.

These preliminary estimates $\vec{\mu}_0$ were then used as the mean values to initialize the population for 20 optimization trials, each lasting 750 generations. The initial population was Gaussian-distributed about the preliminary estimate $\vec{\mu}_0$ with standard-deviations $\sigma_i = 0.3\mu_0^{(i)}$. Key hyperparameters guided the optimization: the mutation probability was set to 1 divided by the number of genes, and the crossover probability was set at 0.7. The population size was determined to be 10 times the number of genes for sufficient variety. In the last generation, the fittest individual corresponding to the parameters achieving the highest similarity with the target was chosen as the solution. The entire procedure was repeated 20 times with different initializations of the population to ensure reproducibility.

In Fig. 5.8 and Fig. 5.10, the highest average gate fidelity achieved is shown in red, while the average of the ten trials is shown in blue.

5.3.1 Sinusoidal Gratings: Variable Propagation Lengths

Optimization was performed for several two-beam unitary transformations. Fig. 5.5 shows the output intensity distribution calculated for a sequence of five variably spaced thin sinusoidal gratings optimized for the Hadamard gate (Eq. (1.14)) between the diffraction orders $m = -1, +1$. Nearly all output light is evenly split between these two orders, and we have the expected relative phase between the output beams. An average gate fidelity of 97.4% was achieved in this case. The corresponding parameter values can be found in Appendix F.2.

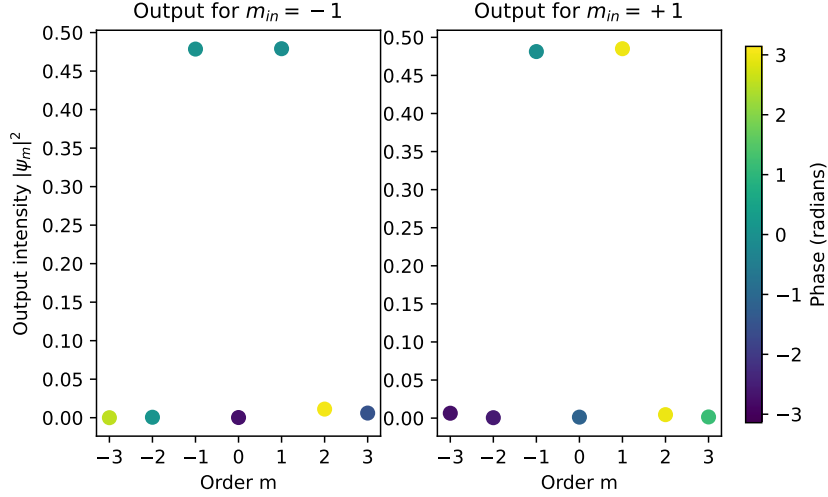


Figure 5.5: Output intensity (normalized by input intensity) and phase of a sequence of five sinusoidal gratings optimized for the **Hadamard** gate between diffraction orders $m_{in} = -1, +1$. The sub-figures show the output distribution for each input order.

To test the capacity for achieving other relative phases, optimization was also performed for the symmetric beamsplitter transformation:

$$\text{BS} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \quad (5.31)$$

as well as the σ_x (NOT) gate. The corresponding output distributions, when using five variably-spaced sinusoidal gratings, are shown in Fig. 5.6 and Fig. 5.7, respectively. Average gate fidelities of 97.8% and 99.7% were achieved for the symmetric beamsplitter transformation and σ_x (NOT) gate, respectively. The corresponding parameter values can be found in Appendix F.2.

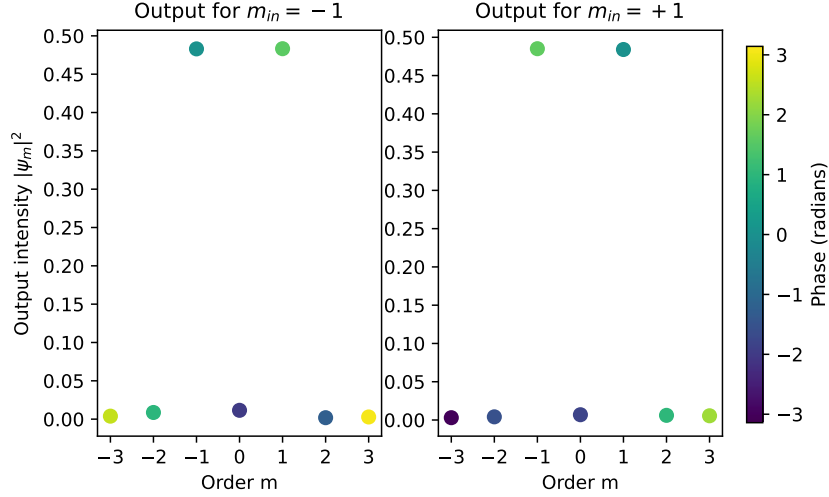


Figure 5.6: Output intensity (normalized by input intensity) and phase of a sequence of five sinusoidal gratings optimized for the **symmetric beamsplitter** transformation between diffraction orders $m_{in} = -1, +1$. The sub-figures show the output distribution for each input order.

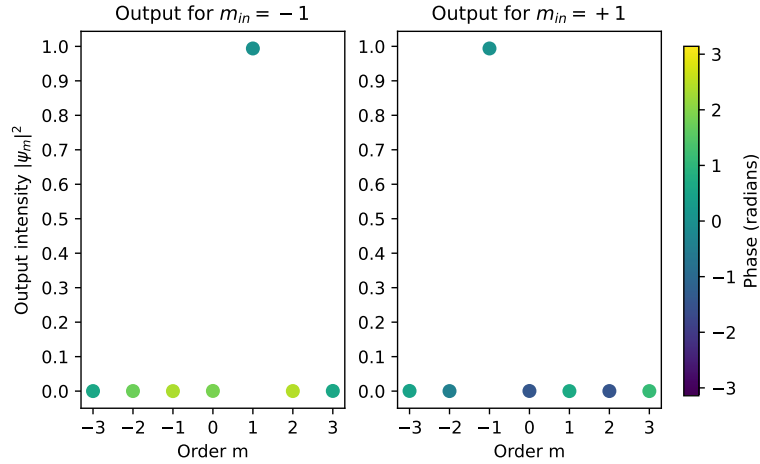


Figure 5.7: Output intensity (normalized by input intensity) and phase of a sequence of five sinusoidal gratings optimized for the σ_x (**NOT**) gate between diffraction orders $m_{in} = -1, +1$. The sub-figures show the output distribution for each input order.

Fig. 5.8 shows the average gate fidelity with the number of diffraction-propagation layers for a sequence of variously spaced thin sinusoidal gratings optimized for the three-dimensional discrete Fourier transform (Eq. (1.19)) between diffraction orders $m = -1, 0, +1$.

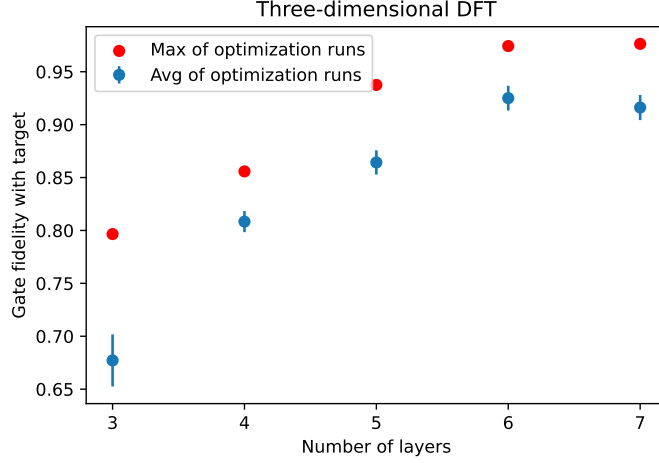


Figure 5.8: Gate fidelities with the three-dimensional **discrete Fourier transform** from parameter optimization of a sequence of thin sinusoidal gratings. The average gate fidelity based on 10 genetic algorithm optimization runs is shown in blue (error bars correspond to the standard error of the mean). The maximum gate fidelity among the optimization runs is shown in red.

To be more explicit, the target transformation is given by:

$$\text{DFT}_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega^4 \end{bmatrix}, \quad (5.32)$$

where $\omega = e^{2\pi i/3}$. Fig. 5.8 suggests that additional layers improve the performance, with a significant rise from three to five layers. Using six layers, the maximum gate fidelity achieved exceeds 95%. This includes losses to other diffraction orders.

Fig. 5.9 shows an example of the calculated output distribution for a sequence of six spaced thin gratings optimized for the three-dimensional discrete Fourier transform. We see that only a small fraction of light is directed to diffraction orders other than $m = -1, 0, +1$. The incident light is approximately evenly split between these three diffraction orders, and we have the desired phase relation between the three orders $m = -1, 0, +1$.

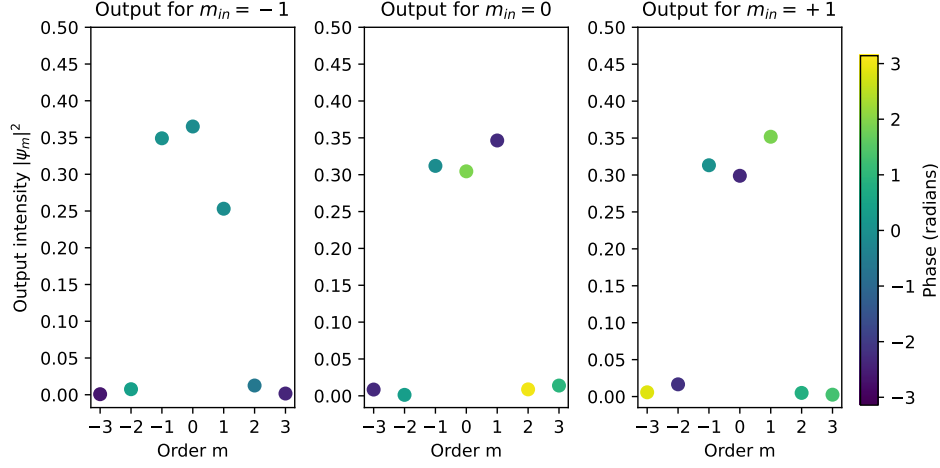


Figure 5.9: Output intensity (normalized by input intensity) and phase of a sequence of six sinusoidal parameters optimized for the three-dimensional **discrete Fourier transform**. The sub-figures show the output distribution for the $m_{in} = -1, 0, +1$ input orders respectively.

Fig. 5.10 shows the average gate fidelity with the number of diffraction-propagation layers for a sequence of variously spaced thin sinusoidal gratings optimized for the three-dimensional clock transformation (Eq. (1.18)) between $m = -1, 0, +1$:

$$Z_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{bmatrix}, \quad (5.33)$$

where $\omega = e^{2\pi i/3}$. Once again, the results suggest that the performance improves with additional of diffraction-propagation layers. The maximum gate fidelity achieved exceeds 92% with seven layers.

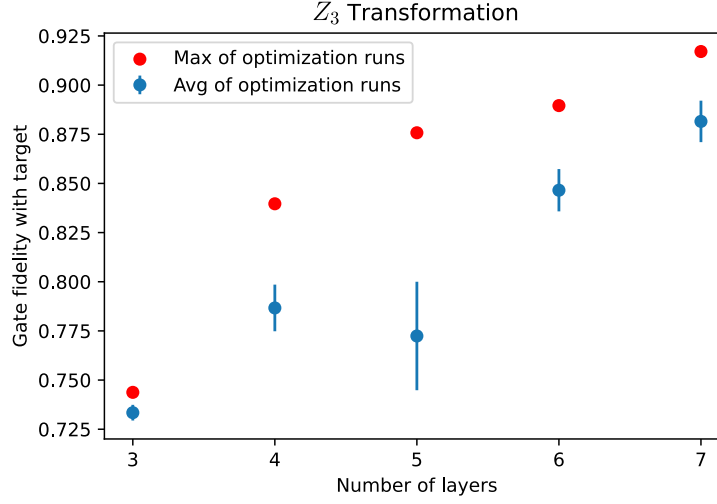


Figure 5.10: Gate fidelities with the three-dimensional Z_3 (**clock**) gate from parameter optimization of a sequence of thin sinusoidal gratings. The average gate fidelity based on 10 genetic algorithm optimization runs is shown in blue (error bars correspond to the standard error of the mean). The maximum gate fidelity among the optimization runs is shown in red.

To test whether non-Clifford gates could also be targeted using sinusoidal gratings and free-space propagation, we also considered optimizing for the T_3 gate (Eq. 1.23),

$$T_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{2\pi i/9} & 0 \\ 0 & 0 & e^{16\pi i/9} \end{bmatrix}. \quad (5.34)$$

Using seven layers, the optimization achieved a gate fidelity of 99%. The output distribution for this case is shown in Fig. 5.11. This demonstrates the versatility of our approach for achieving different classes of unitaries.

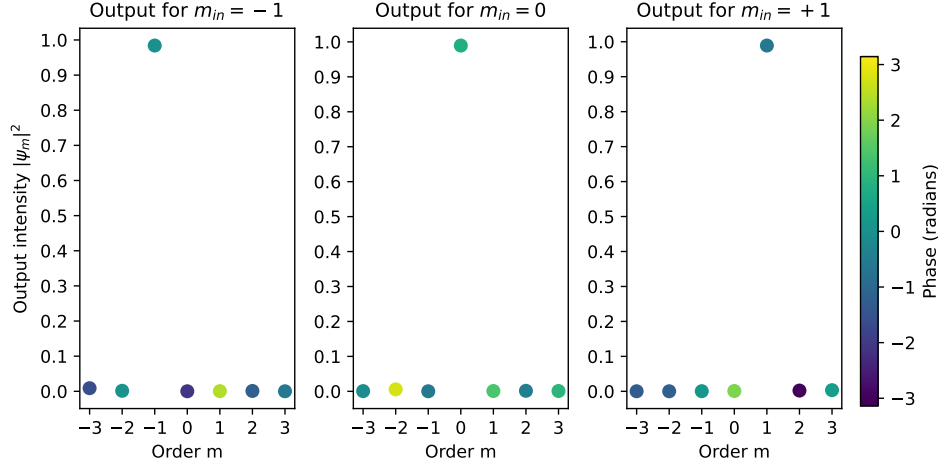


Figure 5.11: Output intensity (normalized by input intensity) and phase of a sequence of six sinusoidal parameters optimized for the **non-Clifford** T_3 gate. The sub-figures show the output distribution for the $m_{in} = -1, 0, +1$ input orders respectively.

5.3.2 Sinusoidal Gratings: Fixed Propagation Length

Many implementations of MPLCs rely on repeated reflections between a mirror and a reflective SLM. Therefore, it is practically preferable to have equally spaced phase gratings, i.e., $L_i = L$ for all layers. In this case, it is also helpful to consider a combined parametrization of the grating wavenumber K and the propagation length L to circumvent experimental limitations related to the resolution of the SLM and the size of the setup.

This can be done through the following paraxial approximation of the propagation phase:

$$\sqrt{k^2 - (mK)^2}L \approx kL - \frac{1}{2}m^2 \frac{(K^2L)}{k} \equiv kL - \frac{m^2}{2k}v, \quad (5.35)$$

where m is the diffraction order and we have used a binomial expansion in $(mK)^2L/k^2$. The first term, kL , adds an overall phase to each diffraction order. Thus, it can be neglected for purposes of achieving a unitary up to an overall phase. We can simply optimize the combined parameter $v \equiv K^2L$.

Two-Mode Unitaries

Optimization for equally-spaced sinusoidal thin phase gratings was performed using the genetic algorithm with the same methodology as discussed in Section 5.3. The average gate fidelity obtained for the σ_x (NOT) gate — between the $m = -1$ and $m = +1$ diffraction orders — with various number of layers is shown in Fig. 5.12. The plot shows improved peak performance for a higher number of layers, with a dramatic improvement (more than 20%) from two to three layers. The maximum gate fidelity obtained approaches 90% with five layers and exceeds 99.5% with seven layers. It further reaches $\bar{F}_g = 99.9\%$ using 11 layers.

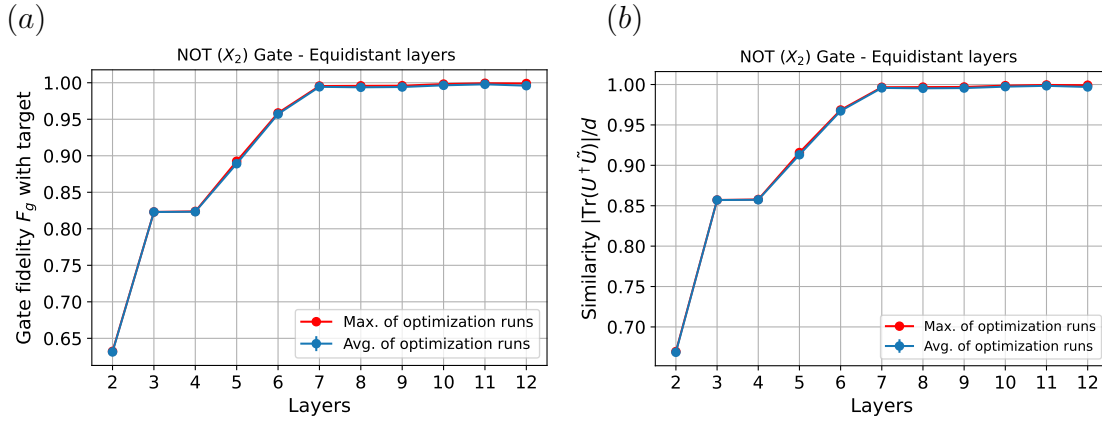


Figure 5.12: (a) Gate fidelity and (b) similarity with the X_2 (NOT) gate from parameter optimization of a sequence of thin equidistant sinusoidal gratings. The average gate fidelity based on 20 genetic algorithm optimization runs is shown in blue (error bars correspond to the standard error of the mean). The maximum gate fidelity among the optimization runs is shown in red.

Some intriguing patterns were observed in the optimized parameters for the NOT transformation using equidistant phase gratings, which persisted for each number of total layers considered. The amplitudes A_i of phase modulation showed a nearly symmetric variation with the number of layers, with the maximum occurring near the middle layer. Moreover, the phase constants θ_i of the gratings alternated by approximately π between consecutive layers.

Optimization of equidistant layers was also performed for the Hadamard gate between the $m = -1$ and $m = +1$ diffraction orders. The gate fidelities achieved with the number of layers are shown in Fig. 5.13. The maximum gate fidelity across 20 optimization runs increased significantly with each added layer up to five layers, where it reached 95%. Gradual improvements were observed thereafter, reaching $\bar{F}_g > 99.5\%$ with 10 layers.

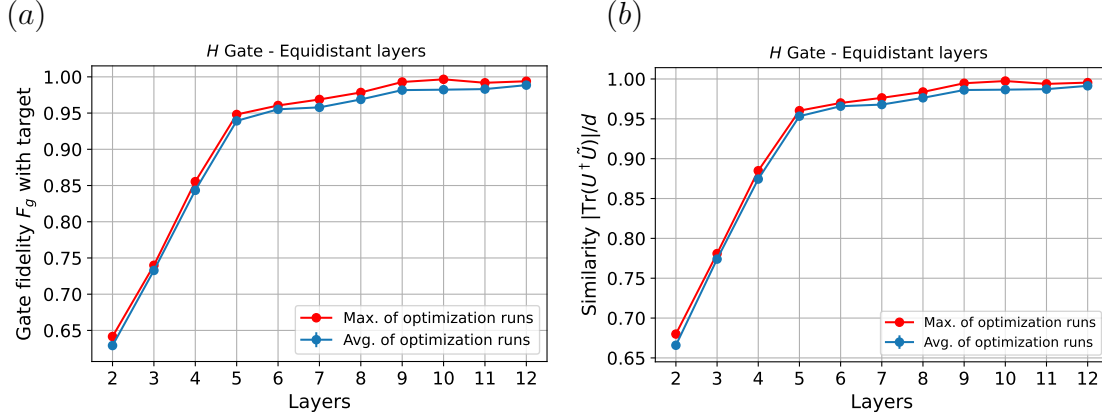


Figure 5.13: (a) Gate fidelity and (b) similarity with the H (**Hadamard**) gate from parameter optimization of a sequence of thin equidistant sinusoidal gratings. The average gate fidelity based on 20 genetic algorithm optimization runs is shown in blue (error bars correspond to the standard error of the mean). The maximum gate fidelity among the optimization runs is shown in red.

In a practical implementation, a particular parameter set could be chosen over another to prioritize accuracy in splitting, relative phase, or minimizing optical loss to other modes.

Unitaries on More than Two Modes

We optimized the layered optical system combining phase gratings with free-space propagation to implement unitaries with dimension greater than two. This optimization assesses the system's capacity to capture more conditions regarding coupling between diffraction orders.

Figures 5.14(a) and (b) show the performance of the three-dimensional DFT gate in terms of average gate fidelity and similarity. Performance improves significantly from 2–5 layers, then plateaus gradually. A similarity and average gate fidelity above 99% is achieved with 10 layers.

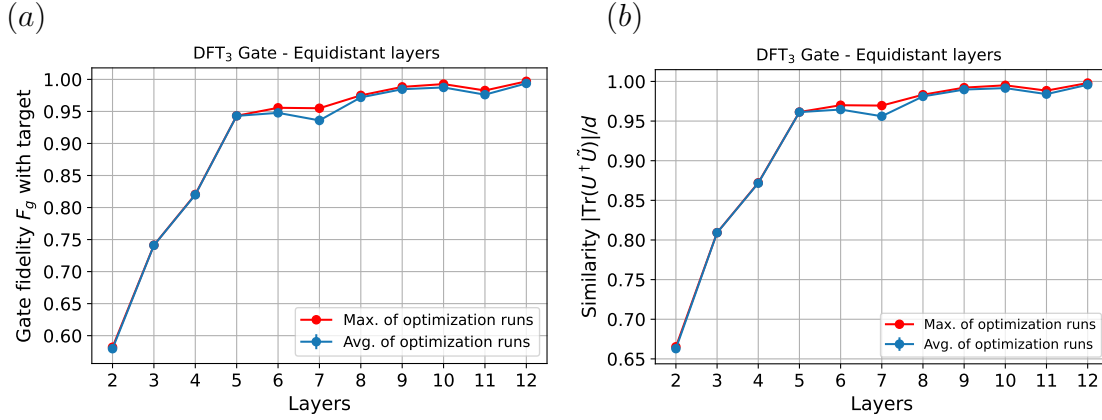


Figure 5.14: (a) Gate fidelity and (b) similarity with the $\mathbf{DFT}_3$ gate from parameter optimization of a sequence of thin equidistant sinusoidal gratings. The average gate fidelity based on 20 genetic algorithm optimization runs is shown in blue (error bars correspond to the standard error of the mean). The maximum gate fidelity among the optimization runs is shown in red.

The four-dimensional and five-dimensional DFT gates (Fig. 5.15 and Fig. 5.16, respectively) require 16–17 layers to reach 98% similarity and average fidelity. This higher layer count compared to $d = 2$ or $d = 3$ reflects the increased number of parameters in higher-dimensional unitaries. More layers are needed until the system parameters approach the unitary’s parameter count.

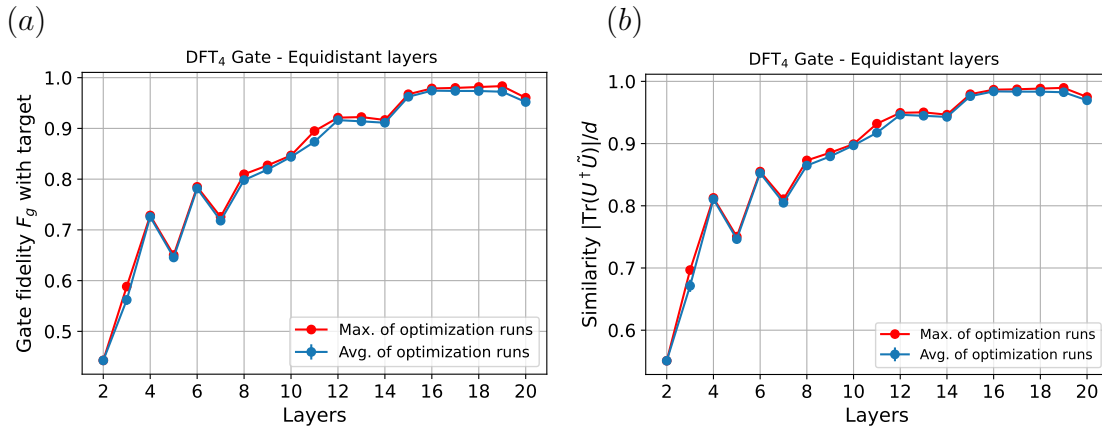


Figure 5.15: (a) Gate fidelity and (b) similarity with the $\mathbf{DFT}_4$ gate from parameter optimization of a sequence of thin equidistant sinusoidal gratings. The average gate fidelity based on 20 genetic algorithm optimization runs is shown in blue (error bars correspond to the standard error of the mean). The maximum gate fidelity among the optimization runs is shown in red.

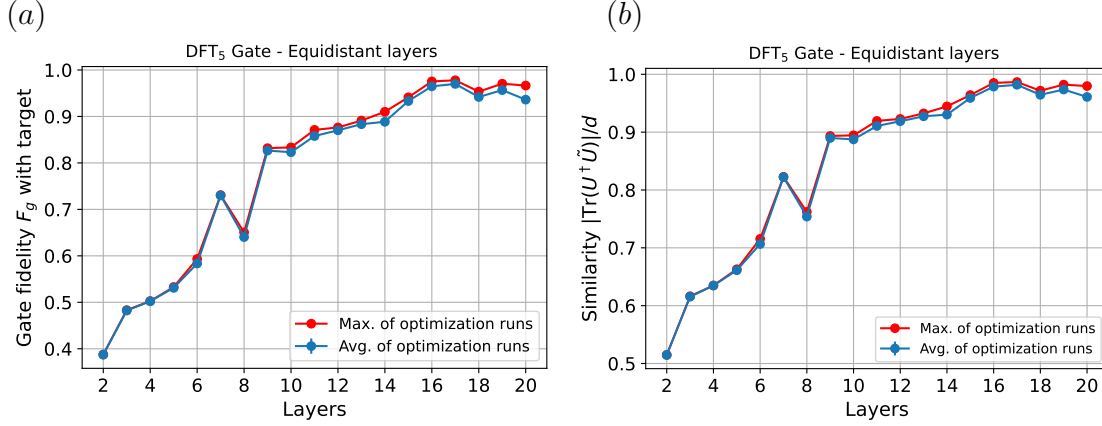


Figure 5.16: (a) Gate fidelity and (b) similarity with the **DFT₅** gate from parameter optimization of a sequence of thin equidistant sinusoidal gratings. The average gate fidelity based on 20 genetic algorithm optimization runs is shown in blue (error bars correspond to the standard error of the mean). The maximum gate fidelity among the optimization runs is shown in red.

The performance as a function of layers for other examples of target unitaries can be found in Appendix F.

5.4 Square-Wave Gratings

Two spaced thin square-wave phase gratings were considered for achieving various unitary transformations. The matrix corresponding to this sequence of layers is $\tilde{U} = T_{A_2, \theta_2}^{\text{square}} D_L T_{A_1, \theta_1}^{\text{square}}$. In this case, the following closed-form expressions for the matrix elements coupling $m = -1$ and $m = +1$, including contributions up to an arbitrary number of diffraction orders, were found:

$$\begin{aligned}
 \tilde{U}_{-1,-1} &= -\frac{4}{\pi^2} e^{ikL} \sin(A_1) \sin(A_2) e^{i(\theta_1 - \theta_2)} + \cos(A_1) \cos(A_2) e^{iL\sqrt{k^2 - K^2}} \\
 &\quad - \frac{4}{\pi^2} \sin(A_1) \sin(A_2) \sum_{n=1}^{n_{\max}} \left(\frac{e^{-i(2n+1)(\theta_2 - \theta_1)}}{(2n+1)^2} + \frac{e^{i(2n-1)(\theta_2 - \theta_1)}}{(2n-1)^2} \right) e^{iL\sqrt{k^2 - 4n^2 K^2}}
 \end{aligned} \tag{5.36}$$

$$\begin{aligned}
& \tilde{U}_{-1,+1} \\
&= \frac{4}{\pi^2} e^{ikL} \sin(A_1) \sin(A_2) e^{-i(\theta_1+\theta_2)} \\
&\quad - \frac{4}{\pi^2} \sin(A_1) \sin(A_2) \sum_{n=1}^{n_{\max}} \left(\frac{e^{i[(2n-1)\theta_2-(2n+1)\theta_1]} + e^{i[(2n-1)\theta_1-(2n+1)\theta_2]}}{(2n-1)(2n+1)} \right) e^{iL\sqrt{k^2-4n^2K^2}}
\end{aligned} \tag{5.37}$$

$$\begin{aligned}
& \tilde{U}_{+1,-1} \\
&= \tilde{U}_{-1,+1}(\theta_i \rightarrow -\theta_i) \\
&= \frac{4}{\pi^2} e^{ikL} \sin(A_1) \sin(A_2) e^{i(\theta_1+\theta_2)} \\
&\quad - \frac{4}{\pi^2} \sin(A_1) \sin(A_2) \sum_{n=1}^{n_{\max}} \left(\frac{e^{-i[(2n-1)\theta_2-(2n+1)\theta_1]} + e^{-i[(2n-1)\theta_1-(2n+1)\theta_2]}}{(2n-1)(2n+1)} \right) e^{iL\sqrt{k^2-4n^2K^2}}
\end{aligned} \tag{5.38}$$

$$\begin{aligned}
& \tilde{U}_{+1,+1} \\
&= -\frac{4}{\pi^2} e^{ikL} \sin(A_1) \sin(A_2) e^{i(\theta_2-\theta_1)} + \cos(A_1) \cos(A_2) e^{iL\sqrt{k^2-K^2}} \\
&\quad - \frac{4}{\pi^2} \sin(A_1) \sin(A_2) \sum_{n=1}^{n_{\max}} \left(\frac{e^{i(2n+1)(\theta_2-\theta_1)}}{(2n+1)^2} + \frac{e^{-i(2n-1)(\theta_2-\theta_1)}}{(2n-1)^2} \right) e^{iL\sqrt{k^2-4n^2K^2}}
\end{aligned} \tag{5.39}$$

One class of unitaries examined for this type of gratings were the traceless Pauli matrices. One can show (see Appendix F.3) that for $A_1 = \pi/2$, $\theta_1 = 0$, and $\theta_2 = \pi/2$, the trace $\tilde{U}_{-1,-1} + \tilde{U}_{+1,+1} = 0$. That leaves the amplitude of phase modulation A_2 , the free-space propagation length L , and wavenumber K as free parameters to be optimized. In this case, the differential evolution [195] method of **NMaximize** in Mathematica 14 was used to maximize the similarity $\mathcal{S}(U, \tilde{U}) = |\text{tr}(U^\dagger \tilde{U})|/2$. The best performance was obtained for the σ_z gate, yielding:

$$\tilde{U} = \begin{bmatrix} 0.981 & -0.019 + 0.012i \\ 0.019 - 0.012i & -0.981 \end{bmatrix} \tag{5.40}$$

for $A_2 = \pi/2$, $L = 16.54$ cm, and $K = 17\,266.6$ m⁻¹. The average gate fidelity $\bar{F}_g(\sigma_z, \tilde{U}) \approx 97.5\%$.

Sequences of square-wave gratings with more than two layers were also explored for achieving traceless unitaries. Numerical investigation suggests the addition of more layers in this case does not improve the performance.

5.5 Summary

We numerically demonstrated that a sequence of thin sinusoidal phase gratings with varying propagation distances can approximate two-beam and three-beam unitaries — such as Pauli gates and the discrete Fourier transform — with $> 99\%$ average gate fidelity given sufficient layers. Increasing the number of layers tends to improve the performance, with a sharp rise at three or four layers. Smaller improvements may be observed for more than five layers. Fixing the propagation distance did not significantly affect the performance of the explored unitaries. The good performance with equidistant gratings is particularly encouraging, as its physical implementation would be compatible with standard MPLC systems. We also considered two spaced thin square-wave gratings for achieving unitary operations, and found traceless operations to be most accessible in this case.

Achieving the desired transformations relies on the interference of the various diffraction orders as they propagate between the phase gratings. The relative phases that contribute to this interference effect are the propagation phases $e^{i\sqrt{k^2-(mK)^2}L}$ and the phases $e^{im\theta_0}$ associated with the grating diffraction orders. As our treatment of the problem makes use of plane waves, it is important to ensure that the grating wavenumber and the propagation length(s) in experimental implementations are chosen to ensure sufficient spatial overlap between the propagating Gaussian beams. Along the same lines, the beam diameter should be several times larger than the grating pitch (e.g., ~ 3 mm for ~ 300 μm) to properly sample the sinusoid. Moreover, the grating pitch should be large enough relative to the SLM pixels (≥ 100 μm) to avoid discretization errors. It is also important that the beam remains within the SLM aperture over all reflections to prevent clipping and phase distortions.

The numerical results for two and three-mode unitary transformations are very promising. They suggest that a significantly smaller number of parameters, simpler phase patterns, and a less computationally intensive approach, compared to general MPLC and wave-front matching, can achieve the desired performance levels.

Chapter 6

Conclusion

In this thesis, we closely investigated efficient state characterization and unitary transformation based on several figures of merit including average fidelity, asymptotic infidelity scaling, speed, and gate fidelity.

Following a review of optimal state estimation, we sought to clarify the effect of the pure-state assumption for the unknown state due to prior information. We carried out Monte Carlo simulations of state tomography, where the MLE constraints were chosen to reflect the presence or absence of the pure-state assumption. Standard Pauli tomography and the two-stage adaptive scheme of Ref. [46] were then compared. Our simulations highlight the fact that in the presence of the *a priori* pure state assumption, Pauli tomography achieves the optimal scaling, consistent with previous asymptotic infidelity calculations [109]. We showed that once the pure-state assumption is removed, the Pauli tomography infidelity scaling deteriorates to $O(\sim 1/\sqrt{N})$. Moreover, the simulation results validated that the two-stage adaptive tomography scheme introduced in Ref. [46] reaches the optimal infidelity scaling $O(1/N)$ — improving quadratically upon Pauli tomography — when the unknown input states cannot be assumed as pure. We further characterized this method by comparing its performance to the Gill-Massar bounds for pure and mixed states.

Furthermore, an experimental demonstration of the Massar and Popescu optimal collective measurement for state estimation was performed for a two-particle ensemble. The results showed that the collective measuring strategy performs at least as well as the optimal LOCC strategy and can surpass it with further reduction of systematic errors. The entangling component of the measurement was found to play a crucial role in achieving near-optimal infidelity scaling.

In addition, a fast, automated, and direct polarization compensation method was pro-

posed and experimentally demonstrated, achieving over 99% fidelity with an average of fewer than six compensation iterations. This scheme has the potential to significantly improve the stability and efficiency of quantum information processing experiments.

The use of a sequence of thin phase gratings for implementing unitaries on two or more diffraction orders was also explored. The results showed that this approach can approximate useful unitaries with $> 99\%$ gate fidelity, with the addition of more layers generally improving performance. This suggests positive prospects both from the perspective of applicability in standard MPLC setups, and further reducing the number of system parameters. We are hopeful that further study of this model can bring insights into achieving unitary transformations with much simpler phase planes and acquire insights into any underlying decomposition.

This thesis can serve as a resource for the application of efficient characterization and manipulation of states of light in quantum information processing. The methods and results presented can be used to inform the design of future experiments and protocols.

6.1 Future directions

Having discussed optimal state estimation in terms of scaling with the ensemble size N , it is also important to identify the application circumstances before choosing a state characterization scheme. For example, the system size d , the size N of the available ensemble, the level of characterization required (e.g., complete state estimate versus certification), the desired level of accuracy, the imperfections and losses associated with experimental implementation of a scheme, as well as any prior information on the distribution of the ensembles in the state-space, should be taken into account. These considerations will help determine, for instance, the cost versus benefit of using a collective, as opposed to an adaptive or fixed separable scheme.

Collective measurements tend to appear as the optimal solution in a range of estimation and discrimination problems: multi-parameter estimation ultimately limited by the Holevo-Cramér-Rao bound [49], N -copy Nagaoka bounds [52], the Massar and Popescu result for multi-copy qubit estimation, and subsequent extensions to mixed qubit and qudit states. Notably, optimal collective measurements for state estimation have been theoretically formulated when using: more than two particles [127, 128], mixed qubit states [51], as well as qudit states [123, 132]. For collective measurements on tripartite systems, restricted measurements, such as biseparable collective measurements have been considered [45], with performance bounds being derived to complement the known bounds for separable and adaptive cases.

In principle, general collective measurements could be one day realized with universal quantum computers. However, an experimental implementation scheme of general N -copy collective measurements is an open question. A central problem is whether the use of only linear optics, post-selection, and polynomial overhead in resources would be sufficient to implement N -copy collective measurements. Practical challenges such as independent polarization control of multiple photons and reduced interference visibility in multi-port beamsplitters would need to be carefully considered.

In parallel to these efforts, it is worthwhile to study imperfect collective measurements, both theoretically and experimentally. In particular, we need to determine the robustness and noise-tolerance of these measurements, and perhaps derive bounds on the level of tolerance for imperfections before their advantage fades, e.g., the minimum required efficiency of a successful measurement.

Optimal quantum state tomography of entangled states is also a significant topic of interest. There is some evidence that Bayesian adaptive tomography can improve the infidelity scaling compared to product POVMs [48, 114] for two-qubit states. Such approaches are challenged by significant computational cost. Others have applied schemes involving artificial neural-networks for computational speed-up [140] or to improve performance compared to fixed tomography [139].

Meanwhile, the expectation that state estimates will continue to improve significantly by measuring more quantum systems (a larger identical ensemble) is challenged by experimental sources of error, such as imperfect measurement settings or unreliable reference input states (see for example the Supplementary Material of Ref. [46]). While discussion in this thesis focused on reducing statistical uncertainty, an important direction is the development of adaptive tomography schemes for minimizing systematic sources of error encountered in experiments.

The related problem of optimal state discrimination among an arbitrary (yet finite) number of states remains unsolved, despite recent efforts to provide tighter upper bounds [147]. Although efficiently solvable numerically, the known nonlinear and globally coupled optimality conditions admit no general constructive solution. An analytical formula may be out of reach, but classifications based on the type of ensemble by studying the symmetry could provide conceptual understanding.

For the realization of unitaries discussed in Chapter 5, future work will involve the experimental implementation of the proposed systems using SLM or liquid crystal platforms. Extensions to composite unitary operations, as well as phase masks across both transverse dimensions, may be explored.

Additionally, analytical results and bounds on the range of unitary transformations

achievable by a sequence of thin phase gratings would provide further insight into the capabilities and limitations of the approach discussed in Chapter 5. Such insight can inform how arbitrary unitaries can be constructed with the minimum number of phase planes in MPLC schemes. With regards to this challenge, recent work using a DFT and phase masks [196, 197] — or with coupled waveguides in integrated optics [198] — attempt to minimize the length of the device while providing analytical guarantees for the decomposition. Even in the absence of an analytical decomposition for the spaced thin sine gratings system, establishing the scaling of the required number of layers with the dimension d for achieving given performance thresholds could be very useful.

References

- [1] C. H. Bennett and G. Brassard, “Quantum cryptography: Public key distribution and coin tossing,” [Theoretical Computer Science](#), vol. 560, pp. 7–11, 2014. Theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84.
- [2] P. Shor, “Algorithms for quantum computation: discrete logarithms and factoring,” in [Proceedings 35th Annual Symposium on Foundations of Computer Science](#), pp. 124–134, 1994.
- [3] L. K. Grover, “Quantum mechanics helps in searching for a needle in a haystack,” [Phys. Rev. Lett.](#), vol. 79, pp. 325–328, Jul 1997.
- [4] J. L. Park and W. Band, “A general theory of empirical state determination in quantum physics: Part I,” [Foundations of Physics](#), vol. 1, pp. 211–226, 1971. Publisher: Springer Science and Business Media LLC.
- [5] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, “Measurement of qubits,” [Physical Review A](#), vol. 64, p. 052312, Oct. 2001.
- [6] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.
- [7] J. B. Altepeter, D. F. James, and P. G. Kwiat, “4 Qubit Quantum State Tomography,” in *Quantum State Estimation* (M. Paris and J. Řeháček, eds.), vol. 649, pp. 113–145, Berlin, Heidelberg: Springer Berlin Heidelberg, Aug. 2004. Series Title: Lecture Notes in Physics.
- [8] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*. Cambridge University Press, 3 ed., 2020.
- [9] P. Kaye, R. Laflamme, and M. Mosca, *An Introduction to Quantum Computing*. Oxford University Press, 11 2006.

- [10] E. Schrödinger, “Probability relations between separated systems,” [Mathematical Proceedings of the Cambridge Philosophical Society](#), vol. 32, p. 446–452, 1936.
- [11] L. P. Hughston, R. Jozsa, and W. K. Wootters, “A complete classification of quantum ensembles having a given density matrix,” [Physics Letters A](#), vol. 183, pp. 14–18, 1993.
- [12] N. Gisin, “Stochastic quantum dynamics and relativity,” [Helv. Phys. Acta](#), vol. 62, pp. 363–371, 1989.
- [13] N. Hadjisavvas, “Properties of mixtures on non-orthogonal states,” [Letters in Mathematical Physics](#), vol. 5, pp. 327–332, 1981.
- [14] E. T. Jaynes, “Information theory and statistical mechanics. II,” [Phys. Rev.](#), vol. 108, pp. 171–190, Oct 1957.
- [15] C. A. Fuchs, *Coming of Age With Quantum Information: Notes on a Paulian Idea*. Cambridge University Press, 2011.
- [16] N. D. Mermin, “What do these correlations know about reality? Nonlocality and the absurd,” [Foundations of Physics](#), vol. 29, pp. 571–587, 1999.
- [17] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, “Quantum entanglement,” [Rev. Mod. Phys.](#), vol. 81, pp. 865–942, Jun 2009.
- [18] R. F. Werner, “Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model,” [Phys. Rev. A](#), vol. 40, pp. 4277–4281, Oct 1989.
- [19] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 2nd ed., 2012.
- [20] D. J. Griffiths and D. F. Schroeter, *Introduction to Quantum Mechanics*. Cambridge University Press, 3 ed., 2018.
- [21] D. M. Appleby, “Symmetric informationally complete–positive operator valued measures and the extended Clifford group,” [Journal of Mathematical Physics](#), vol. 46, p. 052107, 04 2005.
- [22] A. Y. Kitaev, A. Shen, and M. N. Vyalıy, *Classical and quantum computation*. No. 47, American Mathematical Soc., 2002.

- [23] L. Yeh and J. van de Wetering, “Constructing all qutrit controlled Clifford+T gates in Clifford+T,” in *International Conference on Reversible Computation*, pp. 28–50, Springer, 2022.
- [24] E. Hecht, *Optics*. Pearson, 2012.
- [25] D. Goldstein, *Polarized Light*. CRC Press, 2011.
- [26] R. C. Jones, “A new calculus for the treatment of optical systems. Description and discussion of the calculus,” *J. Opt. Soc. Am.*, vol. 31, pp. 488–493, Jul 1941.
- [27] A. Yariv and P. Yeh, *Photonics: optical electronics in modern communications*, vol. 6. Oxford University Press New York, 2007.
- [28] P. Soleillet, “Sur les paramètres caractérisant la polarisation partielle de la lumière dans les phénomènes de fluorescence,” *Ann. Phys. (Paris)*, vol. 10, pp. 23–97, 1929.
- [29] N. G. Parke III, “Optical algebra,” *Journal of Mathematics and Physics*, vol. 28, pp. 131–139, 1949.
- [30] Gil, J. J., “Polarimetric characterization of light and media - physical quantities involved in polarimetric phenomena,” *Eur. Phys. J. Appl. Phys.*, vol. 40, pp. 1–47, 2007.
- [31] J. J. Gil Pérez and R. Ossikovski, *Polarized light and the Mueller matrix approach*. Series in optics and optoelectronics, CRC Press, second ed., 2022.
- [32] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, “Experimental realization of any discrete unitary operator,” *Physical Review Letters*, vol. 73, pp. 58–61, July 1994.
- [33] S. Danakas and P. K. Aravind, “Analogies between two optical systems (photon beam splitters and laser beams) and two quantum systems (the two-dimensional oscillator and the two-dimensional hydrogen atom),” *Phys. Rev. A*, vol. 45, pp. 1973–1977, Feb 1992.
- [34] H. J. Bernstein, “Must quantum theory assume unrestricted superposition?,” *Journal of Mathematical Physics*, vol. 15, pp. 1677–1679, 10 1974.
- [35] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, “Optimal design for universal multiport interferometers,” *Optica*, vol. 3, pp. 1460–1465, Dec 2016.

- [36] J. Carolan, C. Harrold, C. Sparrow, E. Martín-López, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh, G. D. Marshall, M. G. Thompson, J. C. F. Matthews, T. Hashimoto, J. L. O’Brien, and A. Laing, “Universal linear optics,” [Science](#), vol. 349, pp. 711–716, 2015.
- [37] Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljačić, “Deep learning with coherent nanophotonic circuits,” [Nature Photonics](#), vol. 11, pp. 441–446, Jul 2017.
- [38] N. K. Fontaine, R. Ryf, H. Chen, D. Neilson, and J. Carpenter, “Design of High Order Mode-Multiplexers using Multiplane Light Conversion,” in [2017 European Conference on Optical Communication \(ECOC\)](#), (Gothenburg), pp. 1–3, IEEE, Sept. 2017.
- [39] N. K. Fontaine, R. Ryf, H. Chen, D. T. Neilson, K. Kim, and J. Carpenter, “Laguerre-Gaussian mode sorter,” [Nature Communications](#), vol. 10, Apr. 2019. Publisher: Springer Science and Business Media LLC.
- [40] F. Brandt, M. Hiekkamäki, F. Bouchard, M. Huber, and R. Fickler, “High-dimensional quantum gates using full-field spatial modes of photons,” [Optica](#), vol. 7, p. 98, Feb. 2020.
- [41] A. C. Martinez-Becerril, S. Luo, L. Li, J. T. R. Pagé, L. Giner, R. A. Abrahao, and J. S. Lundeen, “Reconfigurable unitary transformations of optical beam arrays,” [Optics Express](#), vol. 32, p. 41111, Nov. 2024.
- [42] V. López Pastor, J. Lundeen, and F. Marquardt, “Arbitrary optical wave evolution with Fourier transforms and phase masks,” [Optics Express](#), vol. 29, p. 38441, Nov. 2021. Publisher: Optica Publishing Group.
- [43] H. M. Wiseman and G. J. Milburn, *Quantum measurement theory*, p. 1–50. Cambridge University Press, 2009.
- [44] I. M. Gelfand and M. A. Neumark, “On the imbedding of normed rings into the ring of operators in Hilbert space,” *Rec. Math. [Mat. Sbornik]*, vol. 54, pp. 197–217, 1943.
- [45] C. Yi, K. Zhou, Z. Hou, G.-Y. Xiang, and H. Zhu, “Optimal estimation of three parallel spins with genuine and restricted collective measurements,” [Physical Review A](#), vol. 111, p. 062409, June 2025.
- [46] D. H. Mahler, L. A. Rozema, A. Darabi, C. Ferrie, R. Blume-Kohout, and A. M. Steinberg, “Adaptive Quantum State Tomography Improves Accuracy Quadratically,” [Phys. Rev. Lett.](#), vol. 111, p. 183601, Oct. 2013.

- [47] J. Liu, H. Yuan, X.-M. Lu, and X. Wang, “Quantum Fisher information matrix and multiparameter estimation,” [Journal of Physics A: Mathematical and Theoretical](#), vol. 53, p. 023001, dec 2019.
- [48] F. Huszár and N. Houlsby, “Adaptive Bayesian quantum tomography,” [Physical Review A](#), vol. 85, p. 052120, May 2012.
- [49] A. Holevo, “Statistical decision theory for quantum systems,” [Journal of Multivariate Analysis](#), vol. 3, pp. 337–394, 1973.
- [50] S. Massar and S. Popescu, “Optimal extraction of information from finite quantum ensembles,” [Phys. Rev. Lett.](#), vol. 74, pp. 1259–1263, Feb 1995.
- [51] G. Vidal, J. I. Latorre, P. Pascual, and R. Tarrach, “Optimal minimal measurements of mixed states,” [Phys. Rev. A](#), vol. 60, pp. 126–135, July 1999.
- [52] H. Nagaoka, *Asymptotic Theory of Quantum Statistical Inference*. World Scientific, Hayashi, M. ed., 2005.
- [53] Y.-C. Liang and A. C. Doherty, “Better Bell-inequality violation by collective measurements,” [Phys. Rev. A](#), vol. 73, p. 052116, May 2006.
- [54] E. Bagan, M. A. Ballester, R. D. Gill, R. Muñoz Tapia, and O. Romero-Isart, “Separable measurement estimation of density matrices and its fidelity gap with collective protocols,” [Phys. Rev. Lett.](#), vol. 97, p. 130501, Sep 2006.
- [55] M. Tsang, “Quantum nonlocality in weak-thermal-light interferometry,” [Phys. Rev. Lett.](#), vol. 107, p. 270402, Dec 2011.
- [56] M. D. Vidrighin, G. Donati, M. G. Genoni, X.-M. Jin, W. S. Kolthammer, M. Kim, A. Datta, M. Barbieri, and I. A. Walmsley, “Joint estimation of phase and phase diffusion for quantum metrology,” [Nature communications](#), vol. 5, p. 3532, 2014.
- [57] A. Muñoz, A. B. Klimov, M. Grassl, and L. L. Sánchez-Soto, “Tomography from collective measurements,” [Quantum Information Processing](#), vol. 17, p. 286, Sep 2018.
- [58] K.-D. Wu, E. Bäumer, J.-F. Tang, K. V. Hovhannisyan, M. Perarnau-Llobet, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, “Minimizing backaction through entangled measurements,” [Phys. Rev. Lett.](#), vol. 125, p. 210401, Nov 2020.

- [59] L. O. Conlon, F. Eilenberger, P. K. Lam, and S. M. Assad, “Discriminating mixed qubit states with collective measurements,” [Commun Phys](#), vol. 6, p. 337, Nov. 2023.
- [60] J. Crossman, S. Dimitroff, L. Cincio, and M. Sarovar, “Quantum computer-enabled receivers for optical communication,” [Quantum Science and Technology](#), vol. 9, p. 045005, Jul 2024.
- [61] Z. Hou, J.-F. Tang, J. Shang, H. Zhu, J. Li, Y. Yuan, K.-D. Wu, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, “Deterministic realization of collective measurements via photonic quantum walks,” [Nature Communications](#), vol. 9, p. 1414, Apr 2018.
- [62] K. Zhou, C. Yi, W.-Z. Yan, Z. Hou, H. Zhu, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, “Experimental Realization of Genuine Three-Copy Collective Measurements for Optimal Information Extraction,” [Phys. Rev. Lett.](#), vol. 134, p. 210201, May 2025.
- [63] M. A. C. Rossi, F. Albarelli, D. Tamascelli, and M. G. Genoni, “Noisy quantum metrology enhanced by continuous nondemolition measurement,” [Phys. Rev. Lett.](#), vol. 125, p. 200505, Nov 2020.
- [64] L. O. Conlon, T. Vogl, C. D. Marciniak, I. Pogorelov, S. K. Yung, F. Eilenberger, D. W. Berry, F. S. Santana, R. Blatt, T. Monz, P. K. Lam, and S. M. Assad, “Approaching optimal entangling collective measurements on quantum computing platforms,” [Nat. Phys.](#), vol. 19, pp. 351–357, Mar. 2023.
- [65] L. O. Conlon, J. M. Koh, B. Shajilal, J. Sidhu, P. K. Lam, and S. M. Assad, “Attainability of quantum state discrimination bounds with collective measurements on finite copies,” [Phys. Rev. A](#), vol. 111, p. 022438, Feb 2025.
- [66] C. K. Hong, Z. Y. Ou, and L. Mandel, “Measurement of subpicosecond time intervals between two photons by interference,” [Phys. Rev. Lett.](#), vol. 59, pp. 2044–2046, Nov 1987.
- [67] H. Fearn and R. Loudon, “Quantum theory of the lossless beam splitter,” [Optics Communications](#), vol. 64, pp. 485–490, 1987.
- [68] J. G. Rarity and P. R. Tapster, “Fourth-order interference in parametric downconversion,” [J. Opt. Soc. Am. B](#), vol. 6, pp. 1221–1226, Jun 1989.
- [69] J. Rarity and P. Tapster, “Nonclassical effects in parametric downconversion,” in *Photons and Quantum Fluctuations*, pp. 122–150, CRC Press, 2020.

- [70] F. Bouchard, A. Sit, Y. Zhang, R. Fickler, F. M. Miatto, Y. Yao, F. Sciarrino, and E. Karimi, “Two-photon interference: the hong-ou-mandel effect,” [Reports on Progress in Physics](#), vol. 84, p. 012402, dec 2020.
- [71] G. S. Thekkadath, L. Giner, X. Ma, J. Flórez, and J. S. Lundeen, “Projecting onto any two-photon polarization state using linear optics,” [New Journal of Physics](#), vol. 20, p. 083033, Aug 2018.
- [72] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, “Experimental quantum teleportation,” [Nature](#), vol. 390, pp. 575–579, Dec 1997.
- [73] Z.-Y. J. Ou, *Multi-photon Quantum interference*. New York: Springer, 2007.
- [74] W. K. Wootters, “Entanglement of formation and concurrence,” [Quantum Inf. Comput.](#), vol. 1, pp. 27–44, 2001.
- [75] W. Heisenberg, *The physical principle of Quantum Theory*. Univ. Chicago Press, Dover, NY, 1930.
- [76] S. Rajasekar and R. Velusamy, *Quantum Mechanics II: Advanced Topics*. Boca Raton: CRC Press, 2 ed., Oct. 2022.
- [77] W. K. Wootters and W. H. Zurek, “A single quantum cannot be cloned,” [Nature](#), vol. 299, pp. 802–803, Oct 1982.
- [78] D. Dieks, “Communication by EPR devices,” [Physics Letters A](#), vol. 92, pp. 271–272, 1982.
- [79] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher, “Noncommuting mixed states cannot be broadcast,” [Phys. Rev. Lett.](#), vol. 76, pp. 2818–2821, Apr 1996.
- [80] T. Mor, “No cloning of orthogonal states in composite systems,” [Phys. Rev. Lett.](#), vol. 80, pp. 3137–3140, Apr 1998.
- [81] U. Fano, “Description of states in quantum mechanics by density matrix and operator techniques,” [Rev. Mod. Phys.](#), vol. 29, pp. 74–93, Jan 1957.
- [82] A. Royer, “Measurement of quantum states and the Wigner function,” [Foundations of Physics](#), vol. 19, pp. 3–32, Jan. 1989. Publisher: Springer Science and Business Media LLC.

- [83] G. M. D'Ariano, M. G. Paris, and M. F. Sacchi, "Quantum tomography," *Advances in imaging and electron physics*, vol. 128, pp. S1076–5670, 2003.
- [84] K. Vogel and H. Risken, "Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase," *Physical Review A*, vol. 40, pp. 2847–2849, Sept. 1989. Publisher: American Physical Society (APS).
- [85] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, "Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum," *Phys. Rev. Lett.*, vol. 70, pp. 1244–1247, Mar 1993.
- [86] J. R. Ashburn, R. A. Cline, P. J. M. van der Burgt, W. B. Westerveld, and J. S. Risley, "Experimentally determined density matrices for $H(n=3)$ formed in H^+ -He collisions from 20 to 100 keV," *Phys. Rev. A*, vol. 41, pp. 2407–2421, Mar 1990.
- [87] T. J. Dunn, I. A. Walmsley, and S. Mukamel, "Experimental determination of the quantum-mechanical state of a molecular vibrational mode using fluorescence tomography," *Phys. Rev. Lett.*, vol. 74, pp. 884–887, Feb 1995.
- [88] D. Leibfried, D. M. Meekhof, B. E. King, C. Monroe, W. M. Itano, and D. J. Wineland, "Experimental determination of the motional quantum state of a trapped atom," *Phys. Rev. Lett.*, vol. 77, pp. 4281–4285, Nov 1996.
- [89] G. Klose, G. Smith, and P. S. Jessen, "Measuring the quantum state of a large angular momentum," *Phys. Rev. Lett.*, vol. 86, pp. 4721–4724, May 2001.
- [90] I. L. Chuang, N. Gershenfeld, and M. Kubinec, "Experimental implementation of fast quantum searching," *Phys. Rev. Lett.*, vol. 80, pp. 3408–3411, Apr 1998.
- [91] U. Fano, "A Stokes-Parameter Technique for the Treatment of Polarization in Quantum Mechanics," *Physical Review*, vol. 93, pp. 121–123, Jan. 1954.
- [92] P. G. Kwiat and B.-G. Englert, "Quantum erasing the nature of reality: or, perhaps, the reality of nature?," in *Science and Ultimate Reality: Quantum Theory, Cosmology, and Complexity* (J. D. Barrow, P. C. W. Davies, and C. L. Harper, Jr, eds.), p. 306–328, Cambridge University Press, 2004.
- [93] Y. S. Teo, *Introduction to quantum-state estimation*. Singapore ; Hackensack, NJ: World Scientific, 2016.

- [94] J. Altepeter, E. Jeffrey, and P. Kwiat, “Photonic State Tomography,” in *Advances In Atomic, Molecular, and Optical Physics*, vol. 52, pp. 105–159, Elsevier, 2005.
- [95] G. M. D’Ariano, M. G. Paris, and M. F. Sacchi, “2 Quantum Tomographic Methods,” in *Quantum State Estimation* (M. Paris and J. Řeháček, eds.), vol. 649, pp. 7–58, Berlin, Heidelberg: Springer Berlin Heidelberg, Aug. 2004. Series Title: Lecture Notes in Physics.
- [96] G. M. D’Ariano, P. Perinotti, and M. F. Sacchi, “Informationally complete measurements and group representation,” *Journal of Optics B: Quantum and Semiclassical Optics*, vol. 6, p. S487, may 2004.
- [97] Z. Hradil, “Quantum-state estimation,” *Phys. Rev. A*, vol. 55, pp. R1561–R1564, Mar 1997.
- [98] K. Banaszek, G. M. D’Ariano, M. G. A. Paris, and M. F. Sacchi, “Maximum-likelihood estimation of the density matrix,” *Phys. Rev. A*, vol. 61, p. 010304, Dec 1999.
- [99] M. Guță, J. Kahn, R. Kueng, and J. A. Tropp, “Fast state tomography with optimal error bounds,” *Journal of Physics A: Mathematical and Theoretical*, vol. 53, p. 204001, apr 2020.
- [100] J. Watrous, *The Theory of Quantum Information*. Cambridge University Press, 2018.
- [101] A. A. Mele, “Introduction to Haar measure tools in quantum information: A beginner’s tutorial,” *Quantum*, vol. 8, p. 1340, 2024.
- [102] M. J. Hall, “Random quantum correlations and density operator distributions,” *Physics Letters A*, vol. 242, pp. 123–129, 1998.
- [103] M. D. de Burgh, N. K. Langford, A. C. Doherty, and A. Gilchrist, “Choice of measurement sets in qubit tomography,” *Phys. Rev. A*, vol. 78, p. 052122, Nov 2008.
- [104] R. Jozsa, “Fidelity for mixed quantum states,” *Journal of Modern Optics*, vol. 41, pp. 2315–2323, 1994.
- [105] D. Walls, G. J. Milburn, and G. J. Milburn, *Quantum Optics*. Berlin, Heidelberg: Springer Nature, 2nd ed. ed., 2008.
- [106] A. Gilchrist, N. K. Langford, and M. A. Nielsen, “Distance measures to compare real and ideal quantum processes,” *Physical Review A*, vol. 71, p. 062310, June 2005.

- [107] N. Gisin and S. Popescu, “Spin flips and quantum information for antiparallel spins,” [Phys. Rev. Lett.](#), vol. 83, pp. 432–435, Jul 1999.
- [108] R. D. Gill and S. Massar, “State estimation for large ensembles,” [Phys. Rev. A](#), vol. 61, p. 042312, Mar 2000.
- [109] E. Bagan, M. Baig, and R. Muñoz Tapia, “Optimal scheme for estimating a pure qubit state via local measurements,” [Phys. Rev. Lett.](#), vol. 89, p. 277904, Dec 2002.
- [110] E. Bagan, A. Monras, and R. Muñoz-Tapia, “Comprehensive analysis of quantum pure-state estimation for two-level systems,” [Phys. Rev. A](#), vol. 71, p. 062318, June 2005.
- [111] S. Popescu, “Bell’s inequalities versus teleportation: What is nonlocality?,” [Physical Review Letters](#), vol. 72, pp. 797–799, Feb. 1994.
- [112] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, “Proposed Experiment to Test Local Hidden-Variable Theories,” [Physical Review Letters](#), vol. 23, pp. 880–884, Oct. 1969.
- [113] K. S. Kravtsov, S. S. Straupe, I. V. Radchenko, N. M. T. Houlsby, F. Huszár, and S. P. Kulik, “Experimental adaptive Bayesian tomography,” [Physical Review A](#), vol. 87, June 2013. Publisher: American Physical Society (APS).
- [114] G. I. Struchalin, I. A. Pogorelov, S. S. Straupe, K. S. Kravtsov, I. V. Radchenko, and S. P. Kulik, “Experimental adaptive quantum tomography of two-qubit states,” [Phys. Rev. A](#), vol. 93, p. 012103, Jan. 2016.
- [115] F. A. An, A. Ransford, A. Schaffer, L. R. Sletten, J. Gaebler, J. Hostetter, and G. Vittorini, “High Fidelity State Preparation and Measurement of Ion Hyperfine Qubits with $I > 1/2$,” [Physical Review Letters](#), vol. 129, p. 130501, Sept. 2022.
- [116] X. Wu, H. Yan, G. Andersson, A. Anferov, M.-H. Chou, C. R. Conner, J. Grebel, Y. J. Joshi, S. Li, J. M. Miller, R. G. Povey, H. Qiao, and A. N. Cleland, “Modular Quantum Processor with an All-to-All Reconfigurable Router,” [Physical Review X](#), vol. 14, p. 041030, Nov. 2024.
- [117] S. J. Evered, D. Bluvstein, M. Kalinowski, S. Ebadi, T. Manovitz, H. Zhou, S. H. Li, A. A. Geim, T. T. Wang, N. Maskara, H. Levine, G. Semeghini, M. Greiner, V. Vuletić, and M. D. Lukin, “High-fidelity parallel entangling gates on a neutral-atom quantum computer,” [Nature](#), vol. 622, pp. 268–272, Oct. 2023.

- [118] M. A. Nielsen, “A simple formula for the average gate fidelity of a quantum dynamical operation,” [Physics Letters A](#), vol. 303, pp. 249–252, 2002.
- [119] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, “Surface codes: Towards practical large-scale quantum computation,” [Physical Review A](#), vol. 86, p. 032324, Sept. 2012.
- [120] D. Aharonov, A. Kitaev, and J. Preskill, “Fault-tolerant quantum computation with long-range correlated noise,” [Phys. Rev. Lett.](#), vol. 96, p. 050504, Feb 2006.
- [121] D. Aharonov and M. Ben-Or, “Fault-tolerant quantum computation with constant error rate,” [SIAM Journal on Computing](#), vol. 38, pp. 1207–1282, 2008.
- [122] A. Katabarwa, K. Gratsea, A. Caesura, and P. D. Johnson, “Early fault-tolerant quantum computing,” [PRX Quantum](#), vol. 5, p. 020101, Jun 2024.
- [123] D. Bruß and C. Macchiavello, “Optimal state estimation for d -dimensional quantum systems,” [Physics Letters A](#), vol. 253, pp. 249–251, Mar. 1999.
- [124] J. Bae and A. Acín, “Asymptotic quantum cloning is state estimation,” [Phys. Rev. Lett.](#), vol. 97, p. 030402, Jul 2006.
- [125] S. Croke, E. Andersson, and S. M. Barnett, “No-signaling bound on quantum state discrimination,” [Phys. Rev. A](#), vol. 77, p. 012113, Jan 2008.
- [126] A. Peres, *Quantum Theory: Concepts and Methods*. Fundamental Theories of Physics, Kluwer Academic Publishers, 1995.
- [127] R. Derka, V. Buzek, and A. K. Ekert, “Universal Algorithm for Optimal Estimation of Quantum States from Finite Ensembles via Realizable Generalized Measurement,” [Physical Review Letters](#), vol. 80, pp. 1571–1575, Feb. 1998.
- [128] J. I. Latorre, P. Pascual, and R. Tarrach, “Minimal Optimal Generalized Quantum Measurements,” [Physical Review Letters](#), vol. 81, pp. 1351–1354, Aug. 1998.
- [129] G. Vidal, J. I. Latorre, P. Pascual, and R. Tarrach, “Optimal minimal measurements of mixed states,” [Physical Review A](#), vol. 60, pp. 126–135, July 1999.
- [130] V. Bužek and M. Hillery, “Quantum copying: Beyond the no-cloning theorem,” [Physical Review A](#), vol. 54, pp. 1844–1852, Sept. 1996.

- [131] N. Gisin and S. Massar, “Optimal Quantum Cloning Machines,” [Physical Review Letters](#), vol. 79, pp. 2153–2156, Sept. 1997.
- [132] A. Hayashi, T. Hashimoto, and M. Horibe, “Reexamination of optimal quantum state estimation of pure states,” [Physical Review A](#), vol. 72, p. 032325, Sept. 2005.
- [133] L. Zambrano, L. Pereira, S. Niklitschek, and A. Delgado, “Estimation of pure quantum states in high dimension at the limit of quantum accuracy through complex optimization and statistical inference,” [Sci Rep](#), vol. 10, p. 12781, July 2020.
- [134] D. G. Fischer, S. H. Kienle, and M. Freyberger, “Quantum-state estimation by self-learning measurements,” [Phys. Rev. A](#), vol. 61, p. 032306, Feb. 2000.
- [135] T. Hannemann, D. Reiss, C. Balzer, W. Neuhauser, P. E. Toschek, and C. Wunderlich, “Self-learning estimation of quantum states,” [Phys. Rev. A](#), vol. 65, p. 050303, May 2002.
- [136] J. Carrasquilla and R. G. Melko, “Machine learning phases of matter,” [Nature Physics](#), vol. 13, pp. 431–434, May 2017.
- [137] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, “Neural-network quantum state tomography,” [Nature Physics](#), vol. 14, pp. 447–450, May 2018.
- [138] A. M. Palmieri, E. Kovlakov, F. Bianchi, D. Yudin, S. Straupe, J. D. Biamonte, and S. Kulik, “Experimental neural network enhanced quantum tomography,” [npj Quantum Information](#), vol. 6, p. 20, Feb 2020.
- [139] H. Lange, M. Kebrič, M. Buser, U. Schollwöck, F. Grusdt, and A. Bohrdt, “Adaptive Quantum State Tomography with Active Learning,” [Quantum](#), vol. 7, p. 1129, Oct. 2023. arXiv:2203.15719 [cond-mat, physics:quant-ph].
- [140] Y. Quek, S. Fort, and H. K. Ng, “Adaptive quantum state tomography with neural networks,” [npj Quantum Inf](#), vol. 7, p. 105, June 2021.
- [141] C. Ferrie, “Self-Guided Quantum Tomography,” [Phys. Rev. Lett.](#), vol. 113, p. 190404, Nov. 2014.
- [142] R. J. Chapman, C. Ferrie, and A. Peruzzo, “Experimental demonstration of self-guided quantum tomography,” [Phys. Rev. Lett.](#), vol. 117, p. 040402, Jul 2016.

- [143] M. Rambach, M. Qaryan, M. Kewming, C. Ferrie, A. G. White, and J. Romero, “Robust and efficient high-dimensional quantum state tomography,” [Phys. Rev. Lett.](#), vol. 126, p. 100402, Mar 2021.
- [144] C. Granade, C. Ferrie, and S. T. Flammia, “Practical adaptive quantum tomography,” [New J. Phys.](#), vol. 19, p. 113017, Nov. 2017.
- [145] S. Chen, B. Huang, J. Li, A. Liu, and M. Sellke, “When Does Adaptivity Help for Quantum State Learning?,” May 2023. arXiv:2206.05265 [quant-ph].
- [146] S. Massar, “Collective versus local measurements on two parallel or antiparallel spins,” [Phys. Rev. A](#), vol. 62, p. 040101, Sep 2000.
- [147] E. R. Loubenets, “General lower and upper bounds under minimum-error quantum state discrimination,” [Physical Review A](#), vol. 105, p. 032410, Mar. 2022.
- [148] V. P. Belavkin, “Optimal multiple quantum statistical hypothesis testing,” *Stochastics: an international journal of probability and stochastic processes*, vol. 1, pp. 315–345, 1975.
- [149] P. Hausladen and W. K. Wootters, “A ‘pretty good’ measurement for distinguishing quantum states,” *Journal of Modern Optics*, vol. 41, pp. 2385–2390, 1994.
- [150] J. Shang, Z. Zhang, and H. K. Ng, “Superfast maximum-likelihood reconstruction for quantum tomography,” [Phys. Rev. A](#), vol. 95, p. 062336, Jun 2017.
- [151] A. Mansouri, K. M. Jordan, R. A. Abrahao, and J. S. Lundeen, “Experimental demonstration of a multi-particle collective measurement for optimal quantum state estimation,” May 2025. arXiv:2505.04607 [quant-ph].
- [152] T. Kitagawa, M. A. Broome, A. Fedrizzi, M. S. Rudner, E. Berg, I. Kassal, A. Aspuru-Guzik, E. Demler, and A. G. White, “Observation of topologically protected bound states in photonic quantum walks,” *Nature communications*, vol. 3, p. 882, 2012.
- [153] X. Qiang, Y. Wang, S. Xue, R. Ge, L. Chen, Y. Liu, A. Huang, X. Fu, P. Xu, T. Yi, *et al.*, “Implementing graph-theoretic quantum algorithms on a silicon photonic quantum walk processor,” *Science Advances*, vol. 7, p. eabb8375, 2021.
- [154] K. M. Jordan, R. A. Abrahao, and J. S. Lundeen, “Quantum metrology timing limits of the Hong-Ou-Mandel interferometer and of general two-photon measurements,” [Phys. Rev. A](#), vol. 106, p. 063715, Dec 2022.

- [155] D. W. Berry, H. Jeong, M. Stobińska, and T. C. Ralph, “Fair-sampling assumption is not necessary for testing local realism,” [Physical Review A](#), vol. 81, p. 012109, Jan. 2010.
- [156] M. Ringbauer, A. Fedrizzi, D. W. Berry, and A. G. White, “Information Causality in the Quantum and Post-Quantum Regime,” [Scientific Reports](#), vol. 4, p. 6955, Nov. 2014.
- [157] E. Roccia, I. Gianani, L. Mancino, M. Sbroscia, F. Somma, M. G. Genoni, and M. Barbieri, “Entangling measurements for multiparameter estimation with two qubits,” [Quantum Science and Technology](#), vol. 3, p. 01LT01, oct 2017.
- [158] M. Parniak, S. Borówka, K. Boroszko, W. Wasilewski, K. Banaszek, and R. Demkowicz-Dobrzański, “Beating the Rayleigh limit using two-photon interference,” [Phys. Rev. Lett.](#), vol. 121, p. 250503, Dec 2018.
- [159] R. W. Boyd, A. L. Gaeta, and E. Giese, *Nonlinear optics*. Springer, 2008.
- [160] M. O. Scully and M. S. Zubairy, *Quantum optics*. Cambridge University Press, 1997.
- [161] S. K. Liao, W. Q. Cai, W. Y. Liu, L. Zhang, Y. Li, J. G. Ren, J. Yin, Q. Shen, Y. Cao, Z. P. Li, F. Z. Li, X. W. Chen, L. H. Sun, J. J. Jia, J. C. Wu, X. J. Jiang, J. F. Wang, Y. M. Huang, Q. Wang, Y. L. Zhou, L. Deng, T. Xi, L. Ma, T. Hu, Q. Zhang, Y. A. Chen, N. L. Liu, X. B. Wang, Z. C. Zhu, C. Y. Lu, R. Shu, C. Z. Peng, J. Y. Wang, and J. W. Pan, “Satellite-to-ground quantum key distribution,” [Nature](#), vol. 549, pp. 43–47, 2017.
- [162] C. He, H. He, J. Chang, B. Chen, H. Ma, and M. J. Booth, “Polarisation optics for biomedical and clinical applications: a review,” [Light: Science and Applications](#), vol. 10, 2021.
- [163] Y. Intaravanne, R. Wang, H. Ahmed, Y. Ming, Y. Zheng, Z. K. Zhou, Z. Li, S. Chen, S. Zhang, and X. Chen, “Color-selective three-dimensional polarization structures,” [Light: Science and Applications](#), vol. 11, 2022.
- [164] C. B. Czegledi, M. Karlsson, E. Agrell, and P. Johansson, “Polarization Drift Channel Model for Coherent Fibre-Optic Systems,” [Sci Rep](#), vol. 6, p. 21217, Feb. 2016.
- [165] S. P. Neumann, A. Buchner, L. Bulla, M. Bohmann, and R. Ursin, “Continuous entanglement distribution over a transnational 248 km fiber link,” [Nature Communications](#), vol. 13, 2022.

- [166] Y. Tan, J. Wang, J. Wu, and Z. He, “Real-time polarization compensation method in quantum communication based on channel Muller parameters detection,” [Communications Engineering](#), vol. 3, pp. 1–10, 2024.
- [167] S. Li, L. Hu, and J. Cao, “High-Sensitivity Polarization Sensitive Optical Coherence Tomography Based on Numerical Correction for Perfect Circularly Polarized Light,” [Applied Sciences](#), vol. 14, p. 2525, Mar. 2024.
- [168] C. He, J. Antonello, and M. J. Booth, “Vectorial adaptive optics,” [eLight](#), vol. 3, p. 23, Nov. 2023.
- [169] Z. Zhuang, S.-W. Suh, and J. S. Patel, “Polarization controller using nematic liquid crystals,” [Optics Letters](#), vol. 24, p. 694, 1999.
- [170] Y. Shi, H. S. Poh, A. Ling, and C. Kurtsiefer, “Fibre polarisation state compensation in entanglement-based quantum key distribution,” [Optics Express](#), vol. 29, p. 37075, 2021.
- [171] T. Strömberg, P. Schiansky, and P. Walther, “Prescriptive method for fiber polarization compensation in two bases,” [Appl. Opt.](#), vol. 63, pp. 1822–1827, Mar 2024.
- [172] M. F. Ramos, N. A. Silva, N. J. Muga, and A. N. Pinto, “Full polarization random drift compensation method for quantum communication,” [Opt. Express](#), vol. 30, p. 6907, Feb. 2022.
- [173] M. Peranić, M. Clark, R. Wang, S. Bahrani, O. Alia, S. Wengerowsky, A. Radman, M. Lončarić, M. Stipčević, J. Rarity, R. Nejabati, and S. K. Joshi, “A study of polarization compensation for quantum networks,” [EPJ quantum technology](#), vol. 10, pp. 30–, 2023.
- [174] O. Bedroya, C. Li, W. Wang, J. Hu, H.-K. Lo, and L. Qian, “Resource-Efficient Real-Time Polarization Compensation for MDI-QKD with Rejected Data,” [Quantum](#), vol. 8, p. 1435, Aug. 2024.
- [175] Y. Dai, C. He, J. Wang, R. Turcotte, L. Fish, M. Wincott, Q. Hu, and M. J. Booth, “Active compensation of extrinsic polarization errors using adaptive optics,” [Opt. Express](#), vol. 27, pp. 35797–35810, Nov 2019.
- [176] B. I. Gramatikov, “A method of calculating compensators in polarization-sensitive optical systems,” [Optik](#), vol. 201, p. 163474, 2020.

- [177] K. Kikuchi, “Fundamentals of Coherent Optical Fiber Communications,” [Journal of Lightwave Technology](#), vol. 34, pp. 157–179, Jan. 2016.
- [178] A. Laing, V. Scarani, J. G. Rarity, and J. L. O’Brien, “Reference-frame-independent quantum key distribution,” [Physical Review A](#), vol. 82, p. 012304, July 2010.
- [179] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python,” [Nature Methods](#), vol. 17, pp. 261–272, 2020.
- [180] S. E. Gilman, T. G. Baur, D. J. Gallagher, and N. K. Shankar, “Properties Of Tunable Nematic Liquid Crystal Retarders,” in [Polarization Considerations for Optical Systems II](#) (R. A. Chipman, ed.), vol. 1166, pp. 461 – 471, International Society for Optics and Photonics, SPIE, 1990.
- [181] N. P. Schnoor, R. C. Niemeier, A. L. Woods, and J. D. Rogers, “Calibration of liquid crystal variable retarders using a common-path interferometer and fit of a closed-form expression for the retardance curve,” [Appl. Opt.](#), vol. 59, pp. 10673–10679, 2020.
- [182] J. M. López-Téllez and N. C. Bruce, “Stokes polarimetry using analysis of the non-linear voltage-retardance relationship for liquid-crystal variable retarders,” [Review of Scientific Instruments](#), vol. 85, p. 033104, 03 2014.
- [183] J. C. Gladish and D. D. Duncan, “LCVR-based polarimetry: calibration issues,” in [Optics in Tissue Engineering and Regenerative Medicine IV](#) (S. J. Kirkpatrick and R. Wang, eds.), vol. 7566, p. 756609, International Society for Optics and Photonics, SPIE, 2010.
- [184] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, “Quantum cryptography,” [Reviews of Modern Physics](#), vol. 74, pp. 145–195, 2002.
- [185] A. Anwar, C. Perumangatt, F. Steinlechner, T. Jennewein, and A. Ling, “Entangled photon-pair sources based on three-wave mixing in bulk crystals,” [Review of Scientific Instruments](#), vol. 92, pp. 1–29, 2021.

- [186] S. Bobach, A. Hidic, J. J. Arlt, and A. J. Hilliard, “Note: A portable rotating waveplate polarimeter,” [Review of Scientific Instruments](#), vol. 88, p. 036101, 03 2017.
- [187] T. Gaylord and M. Moharam, “Analysis and applications of optical diffraction by gratings,” [Proceedings of the IEEE](#), vol. 73, pp. 894–937, 1985.
- [188] R. Magnusson and T. K. Gaylord, “Diffraction efficiencies of thin phase gratings with arbitrary grating shape,” [Journal of the Optical Society of America](#), vol. 68, p. 806, June 1978.
- [189] R. Alferness, “Analysis of optical propagation in thick holographic gratings,” [Applied Physics](#), vol. 7, pp. 29–33, May 1975.
- [190] C. Neipp, A. Márquez, A. Hernandez, M. Alvarez, and A. Beléndez, “Thin and thick diffraction gratings: Thin matrix decomposition method,” [Optik](#), vol. 115, pp. 385–392, 2004.
- [191] J. Goodman, *Introduction to Fourier Optics*. Electrical Engineering Series, McGraw-Hill, 1996.
- [192] T. K. Moon and W. C. Stirling, *Mathematical methods and algorithms for signal processing*. Prentice Hall, 2000.
- [193] F. Gori, “Diffractive optics: an introduction,” in *Diffractive Optics and Optical Microsystems*, pp. 3–22, Springer, 1997.
- [194] A. F. Gad, “Pygad: An intuitive genetic algorithm python library,” [Multimedia tools and applications](#), vol. 83, pp. 58029–58042, 2024.
- [195] R. Storn and K. Price, “A simple and efficient heuristic for global optimization over continuous spaces,” [J. Glob. Optim.](#), vol. 11, pp. 341–359, 1997.
- [196] V. Girouard and N. Quesada, “Near-optimal decomposition of unitary matrices using phase masks and the discrete fourier transform,” [J. Opt. Soc. Am. B](#), vol. 43, pp. A66–A73, Mar 2026.
- [197] J. Álvarez Vizoso and D. Barral, “Universality and Optimal Architectures for Layered Programmable Unitary Decompositions,” Oct. 2025. arXiv:2510.19397 [quant-ph].
- [198] M. Markowitz, M.-A. Miri, A. Ovchinnikov, and K. Zelaya, “Photonic interlacing architectures for learning discrete linear unitaries,” [Phys. Rev. Res.](#), vol. 7, p. 033205, Sep 2025.

- [199] D. L. Colton, R. Kress, and R. Kress, *Inverse acoustic and electromagnetic scattering theory*, vol. 93. Springer, 1998.

APPENDICES

Appendix A

Abstract of “Boson Sampling with Gaussian Input States: Toward Efficient Scaling and Certification”

A universal quantum computer of large scale is not available yet, however, intermediate models of quantum computation would still permit demonstrations of a quantum computational advantage over classical computing and could challenge the extended Church-Turing thesis. One of these models based on single photons interacting via linear optics is called boson sampling. Although boson sampling was demonstrated and the threshold to claim quantum computational advantage was achieved, the question of how to scale up boson sampling experiments remains. To make progress with this problem, here we present a practically achievable pathway to scale boson sampling experiments by combining continuous-variable quantum information and temporal encoding. We propose the combination of switchable dual-homodyne and single-photon detections, the temporal loop technique, and scattershot-based boson sampling. We detail the required assumptions for concluding computational hardness for this configuration. Furthermore, this particular combination of techniques moves towards an efficient scaling and certification of boson sampling, all in a single experimental setup.

R. A. Abrahao, A. Mansouri, and A. P. Lund, “Boson sampling with Gaussian input states: Toward efficient scaling and certification,” [Phys. Rev. A](#), vol. 110, p. 052437, Nov 2024. © 2024 American Physical Society. Reproduced with permission.

[APS copyright policies state](#): “Yes, the author has the right to use the article or a portion of the article in a thesis or dissertation without requesting permission from APS, provided

the bibliographic citation and the APS copyright credit line are given on the appropriate pages."

Appendix B

Separability of a Two-Qubit State

Recall the two-qubit polarization state:

$$\begin{aligned} |\psi_{2Q}\rangle &= c_{HH} |H_a\rangle \otimes |H_b\rangle + c_{HV} |H_a\rangle \otimes |V_b\rangle + c_{VH} |V_a\rangle \otimes |H_b\rangle + c_{VV} |V_a\rangle \otimes |V_b\rangle \\ &\equiv c_{HH} |H_a H_b\rangle + c_{HV} |H_a V_b\rangle + c_{VH} |V_a H_b\rangle + c_{VV} |V_a V_b\rangle. \end{aligned} \quad (\text{B.1})$$

We will show that $|\psi_{2Q}\rangle$ is separable if and only if $c_{HH}c_{VV} = c_{HV}c_{VH}$.

Proof.

($\Rightarrow$) Suppose $|\psi_{2Q}\rangle$ is separable. Then there exist one qubit states $|\psi^{(1)}\rangle = x_0 |H\rangle + x_1 |V\rangle$ and $|\psi^{(2)}\rangle = y_0 |H\rangle + y_1 |V\rangle$ ($|x_0|^2 + |x_1|^2 = 1$ and $|y_0|^2 + |y_1|^2 = 1$) such that

$$\begin{aligned} |\psi_{2Q}\rangle &= |\psi^{(1)}\rangle \otimes |\psi^{(2)}\rangle \\ &= (x_0 |H\rangle + x_1 |V\rangle) \otimes (y_0 |H\rangle + y_1 |V\rangle) \\ &= x_0 y_0 |HH\rangle + x_0 y_1 |HV\rangle + x_1 y_0 |VH\rangle + x_1 y_1 |VV\rangle \end{aligned} \quad (\text{B.2})$$

Comparing Eq. (B.2) with Eq. (B.1), we see that

$$c_{HH}c_{VV} = (x_0 y_0)(x_1 y_1) = (x_0 y_1)(x_1 y_0) = c_{HV}c_{VH}. \quad (\text{B.3})$$

($\Leftarrow$) There are two possible cases: $c_{HH} = 0$ or $c_{HH} \neq 0$.

Suppose $c_{HH} = 0$. Then at least one of c_{HV} or c_{VH} must be zero (so that $c_{HH}c_{VV} = c_{HV}c_{VH}$). Without loss of generality, suppose $c_{VH} = 0$. Then we have:

$$|\psi_{2Q}\rangle = c_{HV} |HV\rangle + c_{VV} |VV\rangle = (c_{HV} |H\rangle + c_{VV} |V\rangle) \otimes |V\rangle, \quad (\text{B.4})$$

meaning $|\psi_{2Q}\rangle$ is separable.

We now consider the case where $c_{HH} \neq 0$. Then we can write:

$$\begin{aligned}
|\psi_{2Q}\rangle &= |H\rangle \otimes (c_{HH}|H\rangle + c_{HV}|V\rangle) + |V\rangle \otimes (c_{VH}|H\rangle + c_{VV}|V\rangle) \\
&= |H\rangle \otimes (c_{HH}|H\rangle + c_{HV}|V\rangle) + \frac{1}{c_{HH}}|V\rangle \otimes (c_{HH}c_{VH}|H\rangle + c_{HH}c_{VV}|V\rangle) \\
&= |H\rangle \otimes (c_{HH}|H\rangle + c_{HV}|V\rangle) + \frac{1}{c_{HH}}|V\rangle \otimes (c_{HH}c_{VH}|H\rangle + c_{HV}c_{VH}|V\rangle) \\
&= |H\rangle \otimes (c_{HH}|H\rangle + c_{HV}|V\rangle) + \frac{c_{VH}}{c_{HH}}|V\rangle \otimes (c_{HH}|H\rangle + c_{HV}|V\rangle) \\
&= \left(|H\rangle + \frac{c_{VH}}{c_{HH}}|V\rangle \right) \otimes (c_{HH}|H\rangle + c_{HV}|V\rangle) \tag{B.5}
\end{aligned}$$

To ensure the tensor product is between two normalized one qubit states, we rewrite Eq. (B.5) as

$$|\psi_{2Q}\rangle = \left(\frac{|H\rangle + \frac{c_{VH}}{c_{HH}}|V\rangle}{\sqrt{1 + \frac{|c_{VH}|^2}{|c_{HH}|^2}}} \right) \otimes \left(\frac{c_{HH}|H\rangle + c_{HV}|V\rangle}{\sqrt{|c_{HH}|^2 + |c_{HV}|^2}} \right) \equiv |\psi_0\rangle \otimes |\psi_1\rangle, \tag{B.6}$$

concluding that $|\psi_{2Q}\rangle$ is separable, with the one qubit states given by $|\psi_0\rangle$ and $|\psi_1\rangle$.

Appendix C

Group Axioms

The elements of $\mathbf{U}(d)$ satisfy the group axioms regarding:

- Closure: The product of any two unitary matrices is also a unitary matrix.
- Associativity: $(U_1 U_2) U_3 = U_1 (U_2 U_3)$ for any unitaries $U_1, U_2, U_3 \in \mathbf{U}(d)$.
- Existence of identity element: The identity matrix $\mathbb{I}_d \in \mathbf{U}(d)$ acts as $\mathbb{I}_d U = U \mathbb{I}_d = U$ for any $U \in \mathbf{U}(d)$.
- Existence of inverses: For any unitary $U \in \mathbf{U}(d)$, there exists an inverse element $U^{-1} = U^\dagger \in \mathbf{U}(d)$ such that $U U^\dagger = U^\dagger U = \mathbb{I}$.

Appendix D

Multi-particle Collective Measurements for Quantum State Estimation

Reproduced from A. Mansouri, K. M. Jordan, R. A. Abrahao, and J. S. Lundeen, “Experimental demonstration of a multi-particle collective measurement for optimal quantum state estimation,” [Optics Express](#), vol. 34, p. 5718, Feb. 2026.

[Optica Publishing Group policies state](#): “Authors who wish to use all or part of papers they previously authored and published in Optica Publishing Group Journals within a thesis or dissertation may do so according to the license and reuse terms outlined [here](#).”

D.1 Rotated Tetrahedron States

In the implementation of the collective measurement, we used projections onto rotated versions of the Massar and Popescu states, which were more convenient to realize using the imperfect reflection port of a polarizing beamsplitter. This is due to the fact that a ratio $|t_H|^2 : |t_V|^2 \approx 62 : 1$ is needed to project onto the Massar and Popescu states, while the imperfect reflection port of the polarizing beamsplitter provides the reciprocal ratio $|t_V|^2 : |t_H|^2 \approx 62 : 1$. This reciprocal ratio results in projection onto a modified set of Massar and Popescu states, which are rotated by 180° (up to a global phase), shown in Eq. (D.1).

One can directly verify (using Eq. (3.1)) that the modified set of Massar and Popescu and tetrahedron states result in the same theoretical optimal average fidelity of $\bar{\mathcal{F}}_{N=2}^{\text{Coll}} = 3/4 = 0.75$ and $\bar{\mathcal{F}}_{\text{Tetra}}^{\text{Coll}} = 5/6 \approx 0.8333$.

$$\begin{aligned}
|\uparrow_{n_1}\rangle &= |H\rangle \\
\rightarrow |\uparrow_{n_1}\rangle' &= \hat{R}_y(\pi) |\uparrow_{n_1}\rangle = |V\rangle \\
|\uparrow_{n_2}\rangle &= \frac{1}{\sqrt{3}} |H\rangle + \sqrt{\frac{2}{3}} |V\rangle \\
\rightarrow |\uparrow_{n_2}\rangle' &= \hat{R}_y(\pi) |\uparrow_{n_2}\rangle = \sqrt{\frac{2}{3}} |H\rangle - \frac{1}{\sqrt{3}} |V\rangle \\
|\uparrow_{n_3}\rangle &= \frac{1}{\sqrt{3}} |H\rangle + \sqrt{\frac{2}{3}} e^{2i\pi/3} \\
\rightarrow |\uparrow_{n_3}\rangle' &= \hat{R}_y(\pi) |\uparrow_{n_3}\rangle = \sqrt{\frac{2}{3}} |H\rangle + \frac{e^{i\pi/3}}{\sqrt{3}} |V\rangle \\
|\uparrow_{n_4}\rangle &= \frac{1}{\sqrt{3}} |H\rangle + \sqrt{\frac{2}{3}} e^{4i\pi/3} \\
\rightarrow |\uparrow_{n_4}\rangle' &= \hat{R}_y(\pi) |\uparrow_{n_4}\rangle = \sqrt{\frac{2}{3}} |H\rangle + \frac{e^{5i\pi/3}}{\sqrt{3}} |V\rangle, \\
\text{with } \hat{R}_y(\pi) &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.
\end{aligned} \tag{D.1}$$

D.2 Data Collection

The coincidence detection data was collected from time histograms with a ~ 2 ns coincidence window and a 3 s integration time. To increase the number of samples in the estimate, corresponding 3 s histograms were combined together. The number of experimental trials is based on the number of available histograms once they were combined, respectively $\{150, 75, 50, 37, 30, 25, 21, 18, 16, 15, 13, 12, 11, 10, 10, 9, 8, 8, 7, 7, 7, 6, 6, 6\}$ for the data points in each subfigure of Fig. 3.9-3.10.

D.3 Reference states: Mitigating bias

Each reference state in Chapter 3, Tab. 3.1 corresponds to the reconstruction using the largest number of samples available ($N_{\text{ens}} \sim 15000$) for the state. This is somewhat of a necessity, because the state produced will differ from the nominal state due to experimental imperfections. The reconstruction corresponding to the N_{ens} values shown in Fig. 3.9 of Chapter 3 come from smaller subsets of the data available for each state. When the subsets grow large enough, each has significant overlap with the reference state data, and the bias becomes apparent.

Given this issue, we can examine the variation in the slope of a log-log infidelity plot (the scaling exponent) if we progressively included more data points to perform the fit. Looking at a moving average of the slopes (Fig. D.1) for two states where the issue was prominent, it appears that the value of the slope is relatively stable when including the first 7-10 of the points from the infidelity plots, before there is a significant drop.

To mitigate bias, the power-law fits are thus performed by including only the first eight data points in Fig. 3.9 of Chapter 3. The same reasoning applies to the case of nearly-pure reference states (Fig. 3.10), where power-law fits are also performed using the first eight data points.

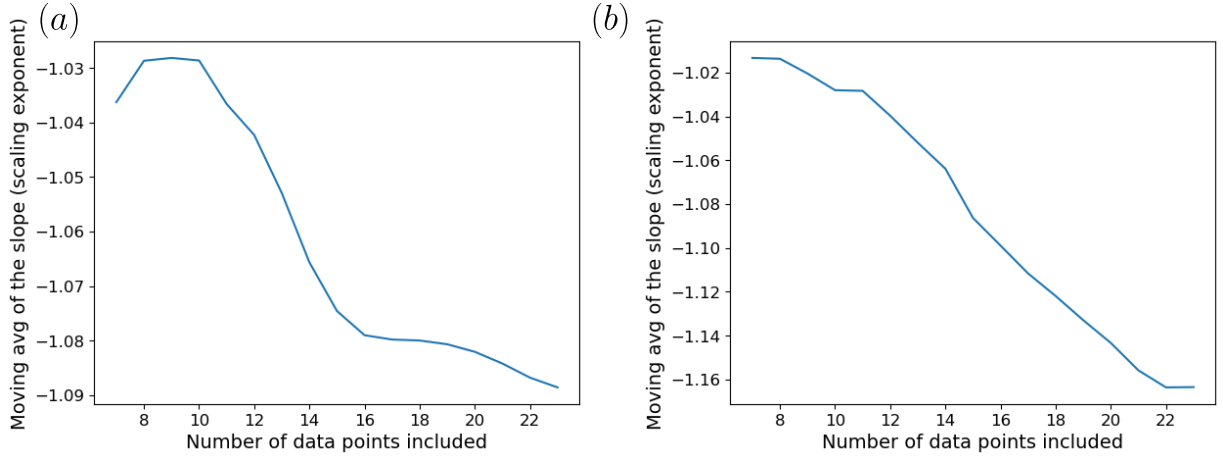


Figure D.1: Moving average (window size of four) for the slope of the infidelity log-log plot of two states as a function of the number of data points included in the fit. Part (a) shows the slope's variation for $0.308|H\rangle - (0.916 + 0.256i)|V\rangle$, while (b) shows the slope's variation for $0.549|H\rangle - (0.777 + 0.309i)|V\rangle$. The slope appears relatively stable with 7-10 data points included, before there is a significant drop.

Appendix E

Fast and Automated Compensation of Fiber Unitaries

Reproduced from N. Braband, A. Mansouri, R. Fazili, S. Czischek, and J. S. Lundeen, “Fast and automated optical polarization compensation of fiber unitaries,” [Review of Scientific Instruments](#), vol. 96, p. 063101, June 2025, with the permission of AIP Publishing.

[AIP copyright policies state](#): “Authors do not need permission from AIP Publishing to: [...] reuse your own AIP Publishing article in your thesis or dissertation (please format your credit line: “Reproduced from [FULL CITATION], with the permission of AIP Publishing”).”

E.1 Polarization Tomography

In this subsection of the appendix, we discuss the experimental details of the tomography intensity measurements with a PD and the angle measurement from the rotation mount. The output voltage of the PD is linearly proportional to the incoming intensity $I = k \cdot V + b$. Before starting the measurement, it is important to first measure the background noise V_{back} without the laser turned on and subtract it from all following measurements:

$$V_n = V_n^{\text{raw}} - V_{\text{back}}. \quad (\text{E.1})$$

Here, the index n refers to the sample number in the tomography dataset, while the superscript ‘raw’ represents the raw voltage measurement recorded by the photodetector during each polarization measurement. Since the PD intensity is proportional to the voltage, the

calculated parameters are now A', B', C' and D' . These parameters are linearly proportional to A_0, B_0, C_0 , and D_0 , respectively, with coefficient k . This issue is addressed by the normalization of the Stokes parameters (see Eq. (E.2)-Eq. (E.6)). Our polarization compensation is a unitary transformation which cannot compensate for loss of degree of polarization (mixed states). This is not a significant issue, since single-mode optical fibers do not generally affect the degree of polarization, as was confirmed by our polarization tomography. Therefore, we base the compensation on the fidelity (Eq. (4.12)), which depends only on the Stokes parameters S_1 to S_3 . We can normalize the Stokes parameters by $\sqrt{S_1^2 + S_2^2 + S_3^2}$, where we have assumed a pure polarization state.

There is an offset α between the QWP fast axis and the 0° on the rotation mount. The angle α needs to be subtracted from the measured angle ϕ from the motorized rotation mount as $2n\Delta\phi = \phi(2n) - \alpha$, and respectively $4n\Delta\phi = \phi(4n) - \alpha$, resulting in:

$$\begin{aligned} A_0 &= \frac{2}{N} \sum_{n=1}^N (I_n^{\text{raw}} - I_{\text{back}}) \\ &= k \frac{2}{N} \sum_{n=1}^N (V_n^{\text{raw}} - V_{\text{back}}) = kA', \end{aligned} \quad (\text{E.2})$$

$$\begin{aligned} B_0 &= \frac{4}{N} \sum_{n=1}^N (I_n^{\text{raw}} - I_{\text{back}}) \sin(\phi(2n) - \alpha) \\ &= k \frac{4}{N} \sum_{n=1}^N (V_n^{\text{raw}} - V_{\text{back}}) \sin(\phi(2n) - \alpha) = kB', \end{aligned} \quad (\text{E.3})$$

$$\begin{aligned} C_0 &= \frac{4}{N} \sum_{n=1}^N (I_n^{\text{raw}} - I_{\text{back}}) \cos(\phi(4n) - \alpha) \\ &= k \frac{4}{N} \sum_{n=1}^N (V_n^{\text{raw}} - V_{\text{back}}) \cos(\phi(4n) - \alpha) = kC', \end{aligned} \quad (\text{E.4})$$

$$\begin{aligned} D_0 &= \frac{4}{N} \sum_{n=1}^N (I_n^{\text{raw}} - I_{\text{back}}) \sin(\phi(4n) - \alpha) \\ &= k \frac{4}{N} \sum_{n=1}^N (V_n^{\text{raw}} - V_{\text{back}}) \sin(\phi(4n) - \alpha) = kD', \end{aligned} \quad (\text{E.5})$$

$$S'_i = \frac{k \cdot S'_i}{k \cdot \sqrt{S_1'^2 + S_2'^2 + S_3'^2}} = \frac{S_i}{\sqrt{S_1^2 + S_2^2 + S_3^2}} = S_i. \quad (\text{E.6})$$

Here, S'_i are the Stokes parameters calculated from A', B', C' and D' . S'_i are the normalized S'_i parameters and S_i are the correctly normalized Stokes parameters.

E.2 Polarization Transformation

In this subsection of the Appendix, we discuss the experimental details for the LCVR characterization and the matrix for three LCVRs. As in case of the polarization tomography, the background needs to be subtracted from the measurement of the LCVR characterization. For each applied voltage to the LCVR, ten samples were measured by the PD. The value of one measurement V_{meas} is the mean value of the ten samples. The measured voltage of the PD is proportional to the input optical power: $I = k \cdot V + b$. Therefore, the retardance is given by:

$$\delta = \arccos \left(1 - \frac{2 \cdot (V_{\text{meas}} - V_{\text{back}})}{V_{\text{max}} - V_{\text{back}}} \right), \quad (\text{E.7})$$

which is equal to the retardance equation in Eq. (4.20) as k and b respectively cancel each other out. The uncertainty for the LCVR characterization is calculated through Gaussian error propagation and results in

$$\Delta\delta = \sqrt{\left(\frac{\partial\delta}{\partial V_{\text{meas}}} \Delta V_{\text{meas}} \right)^2 + \left(\frac{\partial\delta}{\partial V_{\text{back}}} \Delta V_{\text{back}} \right)^2} \quad (\text{E.8})$$

$$= \sqrt{\frac{\Delta V_{\text{meas}}^2 + \Delta V_{\text{back}}^2}{(V_{\text{meas}} - V_{\text{back}})(V_{\text{max}} - V_{\text{meas}})}}. \quad (\text{E.9})$$

Here, ΔV_{meas} and ΔV_{back} represent the errors in the measured voltages and the background measurement, respectively, and are estimated using the standard error of the mean for each case. The error associated with V_{max} is excluded from the error propagation analysis.

The matrix describing the polarization transformation caused by the LCVRs is

$$\begin{aligned} & \mathbf{M}_{\text{LCVRs}}(\delta_1, \delta_2, \delta_3) \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \delta_2 & \sin \delta_1 \sin \delta_2 & -\cos \delta_1 \sin \delta_2 \\ 0 & \sin \delta_2 \sin \delta_3 & \cos \delta_1 \cos \delta_3 - \cos \delta_2 \sin \delta_1 \sin \delta_3 & \cos \delta_3 \sin \delta_1 + \cos \delta_1 \cos \delta_2 \sin \delta_3 \\ 0 & \cos \delta_3 \sin \delta_2 & -\cos \delta_2 \cos \delta_3 \sin \delta_1 - \cos \delta_1 \sin \delta_3 & -\sin \delta_1 \sin \delta_3 + \cos \delta_1 \cos \delta_2 \cos \delta_3 \end{bmatrix}, \quad (\text{E.10}) \end{aligned}$$

where δ_i is the retardance of the i^{th} LCVR.

E.3 Apparatus Specification

For the polarization tomography, we used a Thorlabs LP808-SA60 808 nm CW laser with a ~ 1 nm bandwidth, a Thorlabs PDA20CS photo detector that was read out using a National Instruments 782258-01 USB-6361 DAQ, and a Thorlabs PRM1Z8 motorized rotation stage. The Thorlabs LCC1413-B compensated full-wave LCVRs were driven by JDS6600 two-channel signal generators. All components used in this work are commercially available.

Appendix F

Spatial Unitaries with a Sequence of Thin Phase Gratings

F.1 Fourier Coefficient Calculation

F.1.1 Sinusoidal Phase Grating

Recall that the amplitude transmittance of a sinusoidal phase grating is given by

$$t_A^{\text{sine}}(x) = e^{iA \sin(Kx + \theta_0)}. \quad (\text{F.1})$$

We can use the Jacobi-Anger expansion [199],

$$e^{iz \sin \alpha} \equiv \sum_{m=-\infty}^{\infty} J_m(z) e^{im\alpha}, \quad (\text{F.2})$$

where $J_m(\cdot)$ is the m^{th} order Bessel function of the first kind to rewrite $t_A^{\text{sine}}(x)$ as

$$e^{iA \sin(Kx + \theta_0)} = \sum_{m=-\infty}^{\infty} J_m(A) e^{im(Kx + \theta_0)} \quad (\text{F.3})$$

$$= \sum_{m=-\infty}^{\infty} e^{im\theta_0} J_m(A) e^{imKx}. \quad (\text{F.4})$$

Eq. (F.4) is simply the complex Fourier series expansion of $t_A^{\text{sine}}(x)$, meaning the Fourier coefficients are given by:

$$c_m^{\text{sine}} = e^{im\theta_0} J_m(A). \quad (\text{F.5})$$

F.1.2 Square-Wave Phase Grating

Recall that the amplitude transmittance of a square-wave phase grating is given by

$$t_A^{\text{square}}(x) = e^{iA \text{sgn}(\sin(Kx + \theta_0))}. \quad (\text{F.6})$$

Using Eq. (5.2),

$$\begin{aligned} c_m^{\text{square}} &= \frac{1}{\Lambda} \int_{-\Lambda/2}^{\Lambda/2} t_A^{\text{square}}(x) e^{-imKx} dx \\ &= \frac{1}{\Lambda} \int_{-\Lambda/2}^{-\theta_0/K} e^{-iA} e^{-imKx} dx + \frac{1}{\Lambda} \int_{-\theta_0/K}^{-\theta_0/K + \Lambda/2} e^{iA} e^{-imKx} dx \\ &\quad + \frac{1}{\Lambda} \int_{-\theta_0/K + \Lambda/2}^{\Lambda/2} e^{-iA} e^{-imKx} dx \quad (\text{Assumption: } \theta_0 \in [0, \pi]) \\ &= \begin{cases} -\frac{i}{2m\pi} e^{-i(A+m\pi)} (-1 + e^{im\pi}) (1 + e^{im\pi} - e^{im\theta_0} + e^{i(2\pi+m\theta_0)}), & m \neq 0 \\ \frac{e^{-iA}}{\Lambda} \left(-\frac{\theta_0}{K} + \frac{\Lambda}{2} + \frac{\Lambda}{2} + \frac{\theta_0}{K} - \frac{\Lambda}{2}\right) + \frac{e^{iA}}{\Lambda} \left(-\frac{\theta_0}{K} + \frac{\Lambda}{2} + \frac{\theta_0}{K}\right), & m = 0 \end{cases} \\ &= \begin{cases} 0 & \text{if } m \text{ is even} \\ \frac{2}{m\pi} e^{im\theta_0} \sin(A) & \text{if } m \text{ is odd} \\ \cos(A) & \text{if } m = 0 \end{cases} \quad (\text{F.7}) \end{aligned}$$

F.2 Optimized Parameter Values

	A_1	θ_1	L_1 (m)	A_2	θ_2	L_2 (m)	A_3	θ_3	L_3 (m)	A_4	θ_4	L_4 (m)	A_5	θ_5	K (1/m)
H	0.3826	1.125	0.1379	1.050	0.1753	0.07603	1.875	2.431	0.1319	1.631	0.8423	0.02113	1.051	6.082	171516.89
SBS	1.138	1.551	0.1349	0.3925	4.720	0.0812	0.5273	1.684	0.1699	1.348	1.512	0.1738	0.9853	1.526	129349.91
σ_x	0.6650	3.131	0.04729	1.3904	3.141	0.09367	1.506	6.278	0.04420	1.330	6.283	0.09057	0.6379	3.136	16969.01

Table F.1: Parameter values for a sequence of five thin sinusoidal gratings optimized for the Hadamard gate, symmetric beamsplitter (SBS) transformation, and σ_x (NOT) gate.

The genetic algorithm optimization code for achieving unitaries using sequences of spaced sinusoidal gratings is available at:

<https://github.com/uOttawaQuantumPhotonics/MultiPlaneLightConverter>

F.3 Two Square-Wave Gratings: Zero Trace

We recall the matrix for two spaced square-wave gratings is given by $\tilde{U} = T_{A_2, \theta_2}^{\text{square}} D_L T_{A_1, \theta_1}^{\text{square}}$. In this Appendix, we first show that if $A_1 = \pi/2$, $\theta_1 = 0$, and $\theta_2 = \pi/2$, the trace $\tilde{U}_{-1, -1} + \tilde{U}_{+1, +1} = 0$. We recall the expressions for these two matrix elements:

$$\begin{aligned} \tilde{U}_{-1, -1} &= -\frac{4}{\pi^2} e^{ikL} \sin(A_1) \sin(A_2) e^{i(\theta_1 - \theta_2)} + \cos(A_1) \cos(A_2) e^{iL\sqrt{k^2 - K^2}} \end{aligned} \quad (\text{F.8})$$

$$\begin{aligned} &- \frac{4}{\pi^2} \sin(A_1) \sin(A_2) \sum_{n=1}^{n_{\max}} \left(\frac{e^{-i(2n+1)(\theta_2 - \theta_1)}}{(2n+1)^2} + \frac{e^{i(2n-1)(\theta_2 - \theta_1)}}{(2n-1)^2} \right) e^{iL\sqrt{k^2 - 4n^2 K^2}} \\ &\equiv \alpha(k, L, A_1, A_2) e^{i(\theta_1 - \theta_2)} + \beta(k, L, K, A_1, A_2) + \gamma(A_1, A_2) \sum_{n=1}^{n_{\max}} b_n(\theta_1, \theta_2) \delta_n(k, L, K) \end{aligned} \quad (\text{F.9})$$

$$\begin{aligned} \tilde{U}_{+1, +1} &= -\frac{4}{\pi^2} e^{ikL} \sin(A_1) \sin(A_2) e^{i(\theta_2 - \theta_1)} + \cos(A_1) \cos(A_2) e^{iL\sqrt{k^2 - K^2}} \end{aligned} \quad (\text{F.10})$$

$$\begin{aligned} &- \frac{4}{\pi^2} \sin(A_1) \sin(A_2) \sum_{n=1}^{n_{\max}} \left(\frac{e^{i(2n+1)(\theta_2 - \theta_1)}}{(2n+1)^2} + \frac{e^{-i(2n-1)(\theta_2 - \theta_1)}}{(2n-1)^2} \right) e^{iL\sqrt{k^2 - 4n^2 K^2}} \\ &\equiv \alpha(k, L, A_1, A_2) e^{-i(\theta_1 - \theta_2)} + \beta(k, L, K, A_1, A_2) + \gamma(A_1, A_2) \sum_{n=1}^{n_{\max}} b_n^*(\theta_1, \theta_2) \delta_n(k, L, K). \end{aligned} \quad (\text{F.11})$$

In Eq. (F.9) and Eq. (F.11), we have introduced the definitions

$$\alpha(k, L, A_1, A_2) \equiv -\frac{4}{\pi^2} e^{ikL} \sin(A_1) \sin(A_2) \quad (\text{F.12})$$

$$\beta(k, L, K, A_1, A_2) \equiv \cos(A_1) \cos(A_2) e^{iL\sqrt{k^2 - K^2}} \quad (\text{F.13})$$

$$\gamma(A_1, A_2) \equiv -\frac{4}{\pi^2} \sin(A_1) \sin(A_2) \quad (\text{F.14})$$

$$b_n(\theta_1, \theta_2) \equiv \frac{e^{-i(2n+1)(\theta_2 - \theta_1)}}{(2n+1)^2} + \frac{e^{i(2n-1)(\theta_2 - \theta_1)}}{(2n-1)^2} \quad (\text{F.15})$$

$$\delta_n(k, L, K) \equiv e^{iL\sqrt{k^2 - 4n^2 K^2}}. \quad (\text{F.16})$$

From Eq. (F.9) and Eq. (F.11), it follows that the trace:

$$\begin{aligned}\tilde{U}_{-1,-1} + \tilde{U}_{+1,+1} &= 2\alpha(k, L, A_1, A_2) \cos(\theta_1 - \theta_2) + 2\beta(k, L, K, A_1, A_2) \\ &\quad + \gamma(A_1, A_2) \sum_{n=1}^{n_{\max}} 2 \operatorname{Re}[b_n(\theta_1, \theta_2)] \delta_n(k, L, K).\end{aligned}\quad (\text{F.17})$$

We will now evaluate each term in Eq. (F.17) for $A_1 = \pi/2$, $\theta_1 = 0$, and $\theta_2 = \pi/2$. The first term vanishes, as $\cos(\theta_1 - \theta_2) = \cos(-\pi/2) = 0$. For the second term:

$$2\beta(k, L, K, A_1 = \pi/2, A_2) = 2e^{iL\sqrt{k^2 - K^2}} \cos(\pi/2) \cos(A_2) = 0 \quad (\text{F.18})$$

The sum in Eq. (F.17) also vanishes, as

$$2 \operatorname{Re}[b_n(\theta_1 = 0, \theta_2 = \pi/2)] = 2 \operatorname{Re} \left[\frac{e^{-i(2n+1)(\theta_2 - \theta_1)}}{(2n+1)^2} + \frac{e^{i(2n-1)(\theta_2 - \theta_1)}}{(2n-1)^2} \right] \quad (\text{F.19})$$

$$= \frac{2 \cos[(2n+1)(\theta_2 - \theta_1)]}{(2n+1)^2} + \frac{2 \cos[(2n-1)(\theta_2 - \theta_1)]}{(2n-1)^2} \quad (\text{F.20})$$

$$= \frac{2 \cos[(2n+1)\pi/2]}{(2n+1)^2} + \frac{2 \cos[(2n-1)\pi/2]}{(2n-1)^2} \quad (\text{F.21})$$

$$= 0. \quad (\text{F.22})$$

Therefore, the trace $\tilde{U}_{-1,-1} + \tilde{U}_{+1,+1} = 0$ when $A_1 = \pi/2$, $\theta_1 = 0$, and $\theta_2 = \pi/2$.

F.4 Equidistant Sinusoidal Gratings: Performance for Other Target Unitaries

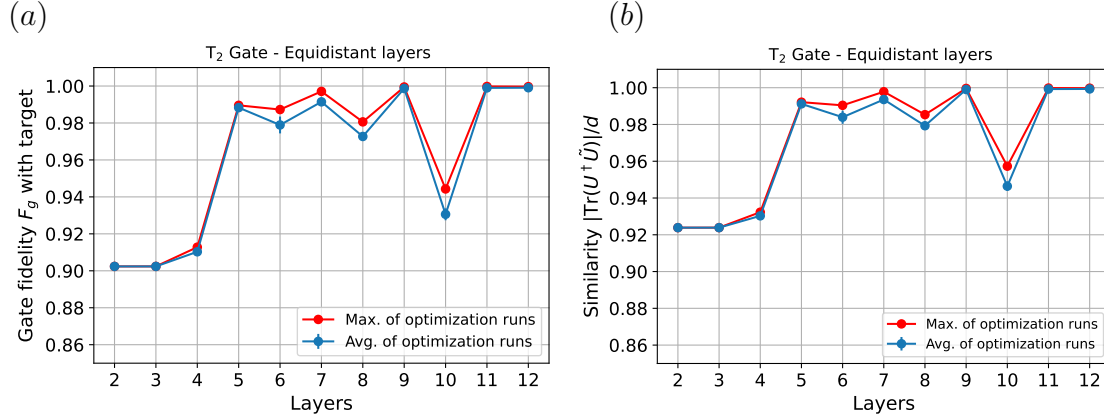


Figure F.1: (a) Gate fidelity and (b) similarity with the $\mathbf{T}_2$ gate from parameter optimization of a sequence of thin equidistant sinusoidal gratings. The average gate fidelity based on 20 genetic algorithm optimization runs is shown in blue (error bars correspond to the standard error of the mean). The maximum gate fidelity among the optimization runs is shown in red.

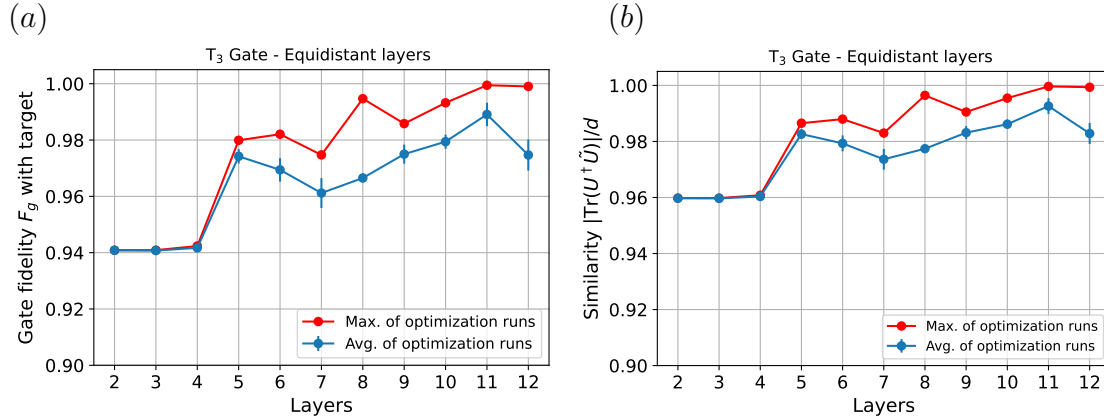


Figure F.2: (a) Gate fidelity and (b) similarity with the $\mathbf{T}_3$ gate from parameter optimization of a sequence of thin equidistant sinusoidal gratings. The average gate fidelity based on 20 genetic algorithm optimization runs is shown in blue (error bars correspond to the standard error of the mean). The maximum gate fidelity among the optimization runs is shown in red.

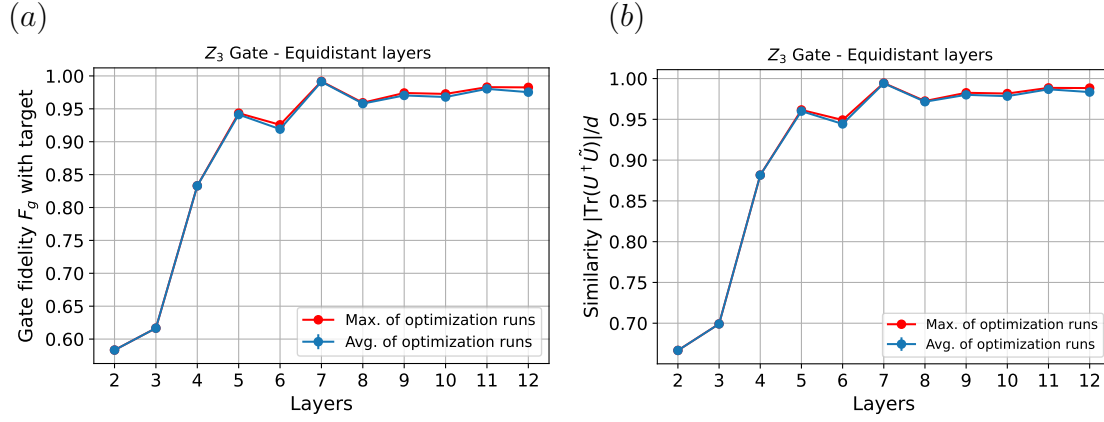


Figure F.3: (a) Gate fidelity and (b) similarity with the $\mathbf{Z}_3$ gate from parameter optimization of a sequence of thin equidistant sinusoidal gratings. The average gate fidelity based on 20 genetic algorithm optimization runs is shown in blue (error bars correspond to the standard error of the mean). The maximum gate fidelity among the optimization runs is shown in red.