Computational Polarimetry: A Bayesian Framework for Polarimetric System Design

by

Shahid Abbas Haider

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2019

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Tim Lee Associate Professor, Dept. of Medicine, Dermatology and Skin Science, University of British Columbia
Supervisor(s):	Alexander Wong Associate Professor, Dept. of Systems Design Engineering, University of Waterloo
Internal Member:	Paul Fieguth Professor, Dept. of Systems Design Engineering, University of Waterloo
Internal-External Member:	Vassilli Karinossios Professor, Dept. of Chemistry, University of Waterloo
Internal Member:	Maud Gorbet Associate Professor, Dept. of Systems Design Engineering, University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In this thesis, we propose a novel polarimetric system design framework that computationally evaluates a design to solve an optical problem. It does this by explicitly formulating the logical connections and dependencies in the design of the components in a polarimetric system (i.e the components chosen affect the choice of their parameters; the parameters of each component affects the measurements, and the system design requirements and constraints affect them all) in a Bayesian network. With this Bayesian network formulation, for the first time, prior knowledge of components, system component parameters, and measurement processes can be explicitly modelled in conjunction with system design requirements and constraints in a unified way for the design of polarimetric systems. With this prior knowledge of system processes, component behaviour, and design requirements and constraints, we can design polarimetric systems to achieve design objectives while reducing the effects of stochastic and deterministic error.

We demonstrate these capabilities in this thesis by first designing a single linear polariser, polarisation angle estimation system to produce a desired and measurable angular sensitivity given prior knowledge of component behaviour and prior knowledge of stochastic and deterministic error sources. Using the Computational Polarimetry Framework, we were able to estimate ideal linear polariser orientations under various orthodox and unorthodox design conditions to achieve minimal and desired levels of measurable angular sensitivities. An unintentional consequence of this system was producing stable parameter configurations where the system performance is optimal over tens of degrees.

Next, we used the Computational Polarimetry Framework to estimate the optical activity of a sample using spirally polarised beams and a spatial detector array. The computational optical rotary dispersion (CORD) system incorporated prior knowledge of the beam polarisation distribution, the measurement system, and the measurement process to arrive at an inference model to estimate a sample's optical activity. This system was able to estimate accurate optical activities under synthetic conditions with varying amounts of stochastic error to a lower detectable limit of two millidegrees. The system was able to estimate more accurate angular changes caused by a polarisation rotator in comparison to the state of art linear polarisation orientation (LPO) scanning systems with only a single measurement. Finally, it was demonstrated to provide an accurate estimate of Sucrose optical activity over the LPO system to a tenth of a degree.

Due to the probabilistic and Bayesian foundation of this framework, it is flexible enough to accommodate a range of system prior models for expected system behaviour. However, there can be cases where the distribution of a particular parameter in the framework is

unknown or needs to be known. Future directions for this framework can be to estimate the distribution of a system component parameter for a known component. Given prior information of all the the immediate parent and dependant parameters of that system component, we can use their logical dependencies and the properties of Bayesian networks to infer their distributions. This framework has the additional benefits of being generalisable to model the behaviour of other optical processes.

Acknowledgements

I would like to thank my supervisor, Alexander Wong, for encouraging me to make my crazy ideas, that merge all my diverse interests and nowledge, a reality.

I would also like to thank Farnoud Kazemzadeh. Our chats about optics, life, and everything were always enjoyable and worthwhile.

Dedication

This thesis is dedicated to my Father and Mother. Throughout my high and lows, whether they be personal or academic, they have been supportive of my decisions and respectful of my independence.

Thanks, Ami and Abu.

Table of Contents

List of Tables

List of Figures

Nomenclature

α	Linear polariser angle
$\cdot$	Dot product between two vectors
δ	Phase delay difference between the orthogonal wave components, $\delta_y - \delta_x$
δ_i	Phase delay in the wave propagation in direction $i = \{x, y\}$
λ	The wavelength of a wave
$\mathcal{C}$	Set of System Components
$\mathcal{D}$	Set of System Design requirements and constraints
$\mathcal{D}_C$	Set of Design Constraints
$\mathcal{D}_R$	Set of Design requirements
$\mathcal{E}$	Set of edges in directed acyclic graph D
$\mathcal{I}$	Set of Measurements
$\mathcal{Q}$	Set of System Component Parameters
$\mathcal{U}$	Set of nodes in directed acyclic graph D
ω	The angular frequency of a wave
$\|\| \cdot \|\|$	Magnitude of a vector or complex number
ϕ	Angle of linear polarisation or optical activity
ρ	Radial distance

σ	Uncertainty or stochastic error parameter
τ	Single variable representation of the wave propagator in the z direction, $\omega t - kz$
$\mathbf{M}$	Mueller matrix
$\mathbf{W}$	Measurement matrix of a Stokes or Mueller Matrix estimation system
$\underline{k}$	The wave vector of a wave
$\underline{M}$	Vectorised Mueller Matrix
$\underline{r}$	Vector representation of some point in 3-space
$\underline{S}$	Stokes vector
c_j	Component j; member in Set $\mathcal{C}$
E_i	The wave amplitude in the $i = \{x, y, z\}$ orthogonal directions
$E_{o\{i\}}$	The maximum wave amplitude in the $i = \{x, y, z\}$ orthogonal directions
f_i	Measurement i; member of Set $\mathcal{I}$
i	Measurement indices in $\mathcal{I}$
J	Number of Components
j	Component indices in $\mathcal{C}$
N	Number of measurements
q_j	System component parameter j; member in Set $\mathcal{Q}$
t	Variable representation of time
x, y, z	Co-ordinates to represent spatial location in Cartesian three-space
Z_{nm}	Zernike Moment
D	Directed acyclic graph

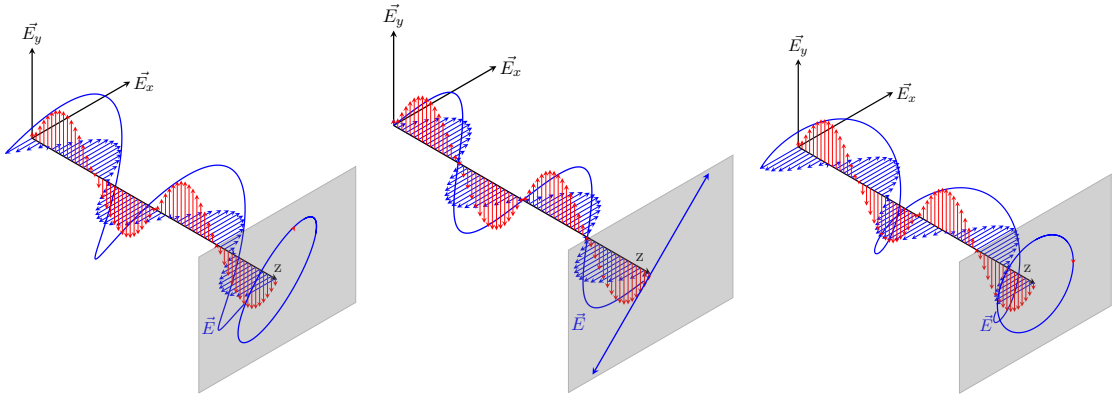
Chapter 1

Introduction

Polarisation is the vectorial description of how light is vibrating at any given point in space and moment in time. Examples of the ways light can vibrate are given in Figure 1.1. Light polarisation can change through the anisotropic absorption or retardation of the electromagnetic wave's orthogonal components when interacting with anisotropic materials or materials with inhomogeneous distributions of matter [52]. By keeping track of how light changes as it passes through these materials, scientists can infer material structure and composition. Using light polarisation for material analysis has proven effective for probing the structural makeup of cells [5, 38], determining skin cellular activity and structure [19, 20, 28], characterising aerosols [31], assessing food quality [49], and performing thin film analysis [47]

Devices that measure the change in light polarisation are called polarimeters and can be classified into two types of devices: 1) those that estimate the polarisation of light, and 2) those that estimate the polarising properties of a material. These polarimeters each arrange their components in a particular way, take an intensity measurement, and go on to another configuration to take another measurement. After the measurements have been captured, optical polarisation or polarisation properties are estimated.

The choice of component parameters in the design of these systems is an active field of study and is focused on choosing component parameters that minimise the influence of error on the estimates of light polarisation or polarisation properties. This error can be stochastic or deterministic. Stochastic error is typically the result of noise from the measuring device while converting light intensity to machine readable values [15], while deterministic error comes from predictable deviations in measured intensity, like beam wander [27, 46] or parameter temperature dependence [53]. Parameter choices to minimise



(a) Elliptical polarisation (b) Linear polarisation (c) Circular polarisation

Figure 1.1: Examples of light polarisations. $\vec{E}$ is the travelling electromagnetic wave. $\vec{E}_x$ is the amplitude of the wave in the horizontal direction, represented by the blue arrows and $\vec{E}_y$ is the amplitude of the electromagnetic wave in the vertical direction, represented by the red arrows. The grey panel contains a trace of the path of the electromagnetic wave projected on the $\vec{E}_x$ and $\vec{E}_y$ plane. z is the direction of wave propagation in space or time.

the effects of stochastic error sources are the most widely discussed and optimised for [2, 8, 11, 15, 32, 40, 41, 54, 53, 57]. An additional benefit, that is not as coveted as it should be, in performing component optimisation is that the number of measurements needed to estimate the light polarisation or polarisation properties are kept to a minimum. As shown by Tyo *et al.* [57], Layden *et al.* [26], and Goudail *et al.* [15], increasing the number of measurements will make the measurements more resistant to stochastic error, however, it is not desirable due to the fact that increasing the number of measurements increases the time for acquisition, thereby limiting any applications in monitoring dynamic systems, like in monitoring birefringence changes of bio-materials undergoing varying levels of mechanical stress [60]. An alternative to performing component optimisation is to separate the intensity signal from the noise through temporal modulation [7, 25]. This method does increase the overall cost and complexity of the system and is therefore not desirable either if designing systems in low-resource settings [13].

The optimisation of component configuration parameters take into account prior knowledge of the stochastic error sources [8], however, design requirements and constraints and deterministic error from experimental uncertainty are unaccounted for in design metrics. Therefore to perform full system optimisation, it is unclear as to how to incorporate all this prior knowledge and additional project information into a single unifying metric to design optimal systems. From a system design perspective, having a unified design framework to infer and model polarimetric system configurations and parameters based on prior knowledge of model performance, of stochastic and deterministic error sources, and of project requirements would be ideal as a single optimisation procedure can occur to arrive at a final design that accounts for all the trade-offs between the aforementioned parameters.

Utilising prior knowledge is a necessary tool in Bayesian techniques for parameter estimation or general inference. A Bayesian network is a formalism that applies a sense of logical dependence between parameters and state in a directed graphical structure and a set of rules to model the random variables in the graph [36, 37]. These networks have been applied to system design frameworks in the past for inference of component health and their monitoring [30], and electrical hardware design [50]. The logical dependence of Bayesian networks is what allows us to incorporate the design requirements and constraints into the framework providing a single unifying metric that contains prior information of component behaviour and error.

Therefore, to create a unified design framework for polarimetric systems, we propose a Computational Polarimetry Framework that presents a joint probability distribution over a set of parameters representing the measurements, system component parameters, components, and system requirements and constraints. The logical dependencies between them are modelled using a Bayesian network. With this network model, given prior knowledge,

the joint distribution can be a metric that is optimised over any set of parameters to achieve the best system for a given purpose. The proposed framework can be written with the joint probability function in Eq. 1.1.

$$P\left(\mathcal{I}, \mathcal{Q}, \mathcal{C}, \mathcal{D}\right) \tag{1.1}$$

where $\mathcal{I}$ is the set of measurements, $\mathcal{Q}$ is the set of vectors describing the system component parameters in $\mathcal{C}$, $\mathcal{C}$ is the set of variables representing the system components, and $\mathcal{D}$ is the set of project requirements and constraints. A brief description of each parameter is given below

Measurements ($\mathcal{I}$) represents the set of measurements received from the photodetector or spatial detector array. These measurements can be random variables, random vectors, or be sets of random variables and random vectors.

Set of Component Parameters ($\mathcal{Q}$) represents the set of parameters that describe each component in the system, including any samples being measured. These parameters are random vectors.

Set of Components ($\mathcal{C}$) represents the set of components in the system.

Project Requirements and Constraints ($\mathcal{D}$) represents the set of hyper-parameters quantitatively describing the requirements and constraints placed on the design of the polarimetric system.

These parameters are modelled through a Bayesian network and utilise its properties to produce a metric on which full systems design can occur. In addition, the metric can be used to produce inference models for sample parameter estimation by exploiting the graphical structure of the network and the inclusion of prior knowledge of component behaviour and error sources.

1.1 Contributions

This PhD thesis contributes a novel unifying framework for the design of polarimetric systems. The framework utilises Bayesian networks to mimic the logical design process of polarimetric optical systems resulting in a singular metric that can be optimised for full system design that includes prior knowledge of component behaviour, error sources, as well

as design requirements and constraints. In addition, through the use of this framework, inference models can be set up that utilise prior knowledge of components behaviour and prior knowledge of stochastic error sources for sample parameter estimation. This is the first unified polarimetric system design framework and the first polarimetric design tool that incorporates design requirements and constraints.

The framework will be used in the design of two systems to replicate two design scenarios:

Design Problem 1: Designing a single linear polariser optical system to estimate the angle of linear polarisation for material differentiation tasks. The design requirement is that the system will detect the smallest angular difference possible given a detector contaminated with stochastic error and a linear polariser holder with spaced graduations, representing a source of deterministic error.

Design Problem 2: Designing an optical polarimetry system that uses heterogeneously polarised beams to infer the optical activity of a sample using a spatial detector array. The spatial detector array is contaminated with stochastic error and measures a discrete set of polarisation states simultaneously.

1.2 Outline

This thesis is outlined as such:

- Chapter 2 provides background information about polarisation, sample modelling, polarisation measurement systems, sources of error, heterogeneously polarised beams, and Bayesian networks.
- Chapter 3 reviews work relating to the design of optimal polarimetric systems.
- Chapter 4 introduces the Computational Polarimetry Framework.
- Chapter 5 demonstrates the Computational Polarimetry Framework in the design of a single linear polariser optical system to estimate the angle of linear polarisation.
- Chapter 6 demonstrates the use of the Computational Polarimetry Framework in the design of a system to infer optical activity.
- Chapter 7 discusses future work to expand upon this framework and concludes this thesis.

Chapter 2

Theoretical Background

Polarised light has been well modelled and studied through the 20^{th} and 21^{st} century. In this chapter, we will present the tools and models designed to study the polarisation of light and how it interacts with materials. In addition, we will introduce how polarimetric systems are modelled along with special types of heterogeneously polarised beams. Finally, we will introduce Bayesian networks, outline and demonstrate two of their properties for modelling logical dependencies.

2.1 Polarised Light

In this section, we will introduce the mathematical formulation of polarised light and the generalised Polarisation ellipse along with its degenerate states. After, a more in-depth introduction to Stokes Parameters will be given.

2.1.1 Polarisation Ellipse

Ignoring diffraction, light can be described as an oscillating, monochromatic electromagnetic (EM) planar wave emitting from a point source with some wavelength λ and angular frequency ω. For the scope of this thesis, the polarisation of non-monochromatic waves will not be described.

If the EM wave is imagined to be travelling in some arbitrary direction through a homogeneous, isotropic medium, $\underline{r} = \underline{r}(x, y, z)$, then the electric wave's amplitude can be described as

$$E(\underline{r},t) = \mathbf{E_o}\cos\left(\omega t - \underline{k}\cdot\underline{r} + \delta_i\right) \tag{2.1}$$

where $\mathbf{E_o} = \mathbf{E_o}(x,y,z)$ is the maximum amplitude, δ_i is some phase delay in the wave, and $\underline{k}$ is the wave vector, where $\parallel \underline{k} \parallel = \frac{2\pi}{\lambda}$. The wave is assumed to be travelling in three-space (x,y,z), allowing for eq. 2.1 to be decomposed into a scalar representation in three orthogonal directions

$$E_x(\underline{r},t) = E_{ox}\cos\left(\omega t - \underline{k}\cdot\underline{r} + \delta_x\right) \tag{2.2}$$

$$E_y(\underline{r},t) = E_{oy}\cos\left(\omega t - \underline{k}\cdot\underline{r} + \delta_y\right) \tag{2.3}$$

$$E_z(\underline{r},t) = E_{oz}\cos\left(\omega t - \underline{k}\cdot\underline{r} + \delta_z\right) \tag{2.4}$$

where E_{ox}, E_{oy}, and E_{oz} are the maximum amplitudes in each of the three orthogonal directions. The term polarisation is actually this vectorial nature of the wave and polarimetry is the measurement of this vectorial nature. For the rest of this thesis, it is assumed that there will be no change in the wave's propagation direction. The propagation direction will be along the z axis and since EM waves are defined as transverse waves in free space, there will only be vibration in the x and y directions for the electric waves, thereby reducing the amplitude equations from eq. 2.2, 2.3, 2.4 to eq. 2.5 and eq. 2.6.

$$E_x(z,t) = E_{ox}\cos\left(\omega t - kz + \delta_x\right) \tag{2.5}$$

$$E_y(z,t) = E_{oy}\cos\left(\omega t - kz + \delta_y\right) \tag{2.6}$$

The term $\omega t - kz$ can be called the propagator and be represented by τ. Using τ, eq. 2.5 and 2.6 can be rephrased as eq. 2.7 and 2.8 to make the proceeding formulation easier.

$$E_x(z,t) = E_{ox}\cos\left(\tau + \delta_x\right) \tag{2.7}$$

$$E_y(z,t) = E_{oy}\cos\left(\tau + \delta_y\right) \tag{2.8}$$

These two equations can be combined to form the equation of an ellipse. The term (z,t) will be dropped for the following equations in this section.

$$\frac{E_x^2}{E_{ox}^2} + \frac{E_y^2}{E_{oy}^2} - 2\frac{E_x}{E_{ox}}\frac{E_y}{E_{oy}}\cos\delta = \sin^2\delta \tag{2.9}$$

where the phase delay difference, δ, is defined as $\delta = \delta_y - \delta_x$. This derivation demonstrates that the wave generally propagates as an ellipse. This behaviour is coined the optical ellipse while eq. 2.9 is the polarisation ellipse.

Special (Degenerate) forms of the polarisation ellipse can occur, leading to these polarisations: i) E_{oy} or E_{ox} is 0, ii) δ is 0 or π, iii) δ is $\frac{\pi}{2}$ or $\frac{3\pi}{2}$, and iv) $E_{oy} = E_{ox} = E_o$ and $\delta = \frac{\pi}{2}$ or $\delta = \frac{3\pi}{2}$. These special cases are important in describing polarimetric states. Polarisations that occur that do not fall in these cases are referred to generally as elliptical polarisations.

1. In eq. 2.7 and eq. 2.8, when E_{oy} is zero, E_y is also zero leaving only the E_x component. The resulting vectorial description describes a horizontally polarised wave. Similarly, if E_{ox} is zero, this describes a vertically polarised wave.

2. With a phase delay difference of 0 or π, the polarisation ellipse in eq. 2.9 reduces to

$$\begin{aligned} \frac{E_x^2}{E_{ox}^2} + \frac{E_y^2}{E_{oy}^2} \pm 2\frac{E_x}{E_{ox}}\frac{E_y}{E_{oy}} &= 0 \\ \left(\frac{E_x}{E_{ox}} \pm \frac{E_y}{E_{oy}}\right) &= 0 \end{aligned} \tag{2.10}$$

   which has a solution of

$$E_x = \pm\left(\frac{E_{ox}}{E_{oy}}\right) E_y \tag{2.11}$$

   The solution in eq. 2.11 shows that there is a linear relation between the orthogonal components. This describes a linearly polarised wave. In the case where $\delta = 0$, there is a negative linear relation and conversely when $\delta = \pi$, the linear relation is positive. If the maximum amplitudes are equal, i.e., $E_{ox} = E_{oy}$, then this describes a $\pm 45^\circ$ linearly polarised wave.

3. For the third case, when there is a phase delay difference of $\delta = \frac{\pi}{2}$ or $\frac{3\pi}{2}$, the polarisation ellipse reduces to being the general description of the non-rotated ellipse in eq. 2.12. In this case, unlike case two, it is not apparent whether the phase delay difference is $\delta = \frac{\pi}{2}$ or $\frac{3\pi}{2}$. This is still termed as a elliptical polarisation.

$$\frac{E_x^2}{E_{ox}^2} + \frac{E_y^2}{E_{oy}^2} = 1 \tag{2.12}$$

4. In the final case, when the maximum amplitudes are the same and the phase delay differences are the same as case three, the polarisation ellipse and eq. 2.12 reduces to

$$E_x^2 + E_y^2 = E_o^2 \tag{2.13}$$

Depending on the phase delay difference, the resulting waves can be either right-handed circularly polarised (RHCP) or left-handed circularly polarised (LHCP). When $\delta = \frac{\pi}{2}$, the wave is RHCP and when $\delta = \frac{3\pi}{2}$, the wave is LHCP. The polarisations are named as such because if an observer places their thumb in the direction of $\underline{k}$, the light wave will travel radially around their thumb. If the wave travels in the direction of the fingers on their right hand, then the wave is right-hand circularly polarised and conversely named if the wave follows the fingers on their left hand. Unfortunately, again, like in case three, from the equation alone, the phase difference is ambiguous. In all the special cases, the phase difference is only clear in case two.

In this section, polarisation was formulated mathematically along with the resulting polarisation ellipse. This ellipse will degenerate to some special cases under certain circumstances.

The polarisation ellipse described here has some limitations. This ellipse will be traced out during such a short duration that it cannot be easily observed thereby making the polarisation of the light wave difficult to determine. In addition, this method assumes the light waves are entirely polarised, which could not be the case. This motivates the search for a formulation of the polarisation ellipse in terms of observables. This was a task taken on by Gabriel Stokes in 1852 [13].

To formulate the polarisation in terms of observables, both of the orthogonal waves in eq. 2.5 and 2.6 can be imagined to be at the spatial position $z = 0$, leaving only time as the wave propagator.

$$E_x(t) = E_{ox}\cos(\omega t + \delta_x) \tag{2.14}$$

$$E_y(t) = E_{oy}\cos(\omega t + \delta_y) \tag{2.15}$$

This leads to eq. 2.9, but now with only a time dependence.

$$\frac{E_x^2(t)}{E_{ox}^2} + \frac{E_y^2(t)}{E_{oy}^2} - 2\frac{E_x(t)}{E_{ox}}\frac{E_y(t)}{E_{oy}}\cos\delta = \sin^2\delta \tag{2.16}$$

To make this observable, the time average ($\langle \cdots \rangle$) of the instantaneous amplitude term can be taken.

$$\frac{\langle E_x^2(t)\rangle}{E_{ox}^2} + \frac{\langle E_y^2(t)\rangle}{E_{oy}^2} - 2\frac{\langle E_x(t)}{E_{ox}}\frac{E_y(t)\rangle}{E_{oy}}\cos\delta = \sin^2\delta \tag{2.17}$$

where the time average is

$$\langle E_i(t)E_j(t)\rangle = \lim_{T\to\infty}\frac{1}{T}\int_0^T E_i(t)E_j(t)dt \tag{2.18}$$

Each time average is evaluated, using eq. 2.14 and 2.15 in eq. 2.18. Due to the periodicity of the monochromatic radiation, only the average over a single oscillation needs to be taken

$$\langle E_x^2(t)\rangle = \frac{1}{2}E_{ox}^2 \tag{2.19}$$

$$\langle E_y^2(t)\rangle = \frac{1}{2}E_{ox}^2 \tag{2.20}$$

$$\langle E_x(t)E_y(t)\rangle = \frac{1}{2}E_{ox}E_{oy}\cos\delta \tag{2.21}$$

and to make the waves observable, eq. 2.17 is multiplied by by $4E_{ox}^2E_{oy}^2$ and $E_{ox}^2 + E_{oy}^2$ is added to both sides of the equation to produce

$$\left(E_{ox}^2 + E_{oy}^2\right)^2 - \left(E_{ox}^2 - E_{oy}^2\right)^2 - \left(2E_{ox}^2E_{oy}^2\cos\delta\right)^2 = \left(2E_{ox}^2E_{oy}^2\sin\delta\right)^2 \tag{2.22}$$

Each of the square terms is recognized as the Stokes parameters of the light wave. The parameters are all observable intensities.

$$S_0 = E_{ox}^2 + E_{oy}^2 \tag{2.23}$$

$$S_1 = E_{ox}^2 - E_{oy}^2 \tag{2.24}$$

$$S_2 = 2E_{ox}E_{oy}\cos\delta \tag{2.25}$$

$$S_3 = 2E_{ox}E_{oy}\sin\delta \tag{2.26}$$

And if the derivation is done using the phasor representation,

$$E_x = E_{ox}e^{i\omega t}e^{i\delta_x} \tag{2.27}$$

$$E_y = E_{oy} e^{i\omega t} e^{i\delta_y} \tag{2.28}$$

the Stokes parameters take the form of

$$S_0 = E_x E_x^* + E_y E_y^* \tag{2.29}$$
$$S_1 = E_x E_x^* - E_y E_y^* \tag{2.30}$$
$$S_2 = E_x E_y^* + E_y E_x^* \tag{2.31}$$
$$S_3 = i(E_x E_y^* - E_y E_x^*) \tag{2.32}$$

Eq. 2.23, eq. 2.29 and the first term in eq. 2.22 are recognized as the full intensity of the light wave. Rearranging eq. 2.22 shows that the squared sum of S_{1-3} is equal to the square of the full intensity of the light wave, S_0, shown in eq 2.33.

$$S_0^2 = S_1^2 + S_2^2 + S_3^2 \tag{2.33}$$

The sum in eq 2.33 is true under the assumption that the wave is perfectly polarised. In the case of partial polarisation, it can be shown in eq. 2.34 that the relation in eq. 2.34 holds.

$$S_0^2 \geq S_1^2 + S_2^2 + S_3^2 \tag{2.34}$$

A metric can then be defined to describe the degree of polarisation, P, using eq. 2.33 and eq. 2.34

$$P = \frac{I_{pol}}{I_{tot}} = \frac{(S_1^2 + S_2^2 + S_3^2)^{\frac{1}{2}}}{S_0} \qquad 0 \leq P \leq 1 \tag{2.35}$$

2.1.2 Stokes Vectors

The four Stokes parameters can be expressed in terms of a vector with each element being a Stokes parameter. The use of a vector representation allows other mathematical tools to be used later on.

$$\underline{S} = \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} E_{ox}^2 + E_{oy}^2 \\ E_{ox}^2 - E_{oy}^2 \\ 2E_{ox}E_{oy}\cos\delta \\ 2E_{ox}E_{oy}\sin\delta \end{bmatrix} \tag{2.36}$$

Each of the special (degenerate) cases of the polarisation ellipse can be represented using this observable representation from eq. 2.36. Each representation is scaled by the intensity of the light wave, F_o.

Horizontally Polarised Light

In horizontally polarised light, $E_{oy} = 0$, eq. 2.36 reduces to

$$\underline{S} = F_o \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{2.37}$$

where $F_o = E_{ox}^2$.

Vertically Polarised Light

In vertically polarised light, $E_{ox} = 0$, eq. 2.36 reduces to

$$\underline{S} = F_o \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \tag{2.38}$$

where $F_o = E_{oy}^2$.

$+45°$ Polarised Light

In $+45°$ polarised light, $E_{oy} = E_{ox} = E_o$ and $\delta = 0$, so eq. 2.36 reduces to

$$\underline{S} = F_o \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \tag{2.39}$$

where $F_o = 2E_o^2$

$-45°$ Polarised Light

In $-45°$ polarised light, $E_{oy} = E_{ox} = E_o$ and $\delta = \pi$, so eq. 2.36 reduces to

$$\underline{S} = F_o \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} \tag{2.40}$$

where $F_o = 2E_o^2$.

Right-hand Circularly Polarised Light

In RHCP light, $E_{oy} = E_{ox} = E_o$ and $\delta = \frac{\pi}{2}$, so eq. 2.36 reduces to

$$\underline{S} = F_o \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{2.41}$$

where $F_o = 2E_o^2$.

Left-hand Circularly Polarised Light

In LHCP light, $E_{oy} = E_{ox} = E_o$ and $\delta = \frac{3\pi}{2}$, so eq. 2.36 reduces to

$$\underline{S} = F_o \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \tag{2.42}$$

where $F_o = 2E_o^2$.

2.1.3 Mueller Matrices

Polarimetry is the measurement of the polarisation state of light or how the polarisation state of light changes as it interacts with matter. The mathematical characterisation of this matter and how it changes a wave's polarisation is described using Mueller matrices.

Mueller matrices are four-by-four real matrices (i.e. $\mathbf{M} \in \mathbb{R}^{4\times4}$) that map an incoming light beam's Stokes vector to an outgoing light beam's Stokes vector, demonstrated in eq. 2.43, mimicking the interaction of the polarized light wave with matter: i) as it changes the amplitudes of components of the light wave, ii) as the phases change between the orthogonal components, iii) as it changes the direction of the orthogonal field components (i.e., rotates the Stokes Vector), and iv) as energy transfers from polarised to unpolarised states.This tool offers the ability, under certain constraints [18], to theoretically observe the change in light polarisation for a wide range of possible input Stokes vectors.

$$\underline{S'} = \mathbf{M}\underline{S} \tag{2.43}$$

$$\begin{bmatrix} S'_0 \\ S'_1 \\ S'_2 \\ S'_3 \end{bmatrix} = \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} \tag{2.44}$$

The size of the Mueller matrix requires there to be 16 measurements taken with different input and output Stokes vectors to determine the effect of the material on the polarisation of a light wave. There is no guarantee that it will result in a well posed system [12]. However, work has been done to determine the best input Stokes vectors for the optimal estimation of a Mueller matrix [1, 26].

In the modelling of polarimetric systems, components such as polarisers, which force the light polarisation into a particular type of polarisation, retarders, which cause a phase delay between two orthogonal components, and rotators, which cause a transfer the energy from one orthogonal field component to the other, have idealised Mueller matrices. Examples of these components are shown in Eq. 2.45, 2.46, and 2.47.

Horizontal Linear Polariser

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.45}$$

Rotator (for some angle ϕ)

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\phi & \sin 2\phi & 0 \\ 0 & -\sin 2\phi & \cos 2\phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.46}$$

Retarder (for some trasmission axis θ and some retardance δ)

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2\theta + \sin^2 2\theta \cos\delta & \sin 2\theta \cos 2\theta(1-\cos\delta) & -\sin 2\theta \sin\delta \\ 0 & \sin 2\theta \cos 2\theta(1-\cos\delta) & \sin^2 2\theta + \cos^2 2\theta \cos\delta & \cos 2\theta \sin\delta \\ 0 & \sin 2\theta \sin\delta & -\cos 2\theta \sin\delta & \cos\delta \end{bmatrix} \tag{2.47}$$

2.2 Polarimetric Systems

In this section, we outline the general components in a polarimetric system and their associated mathematical system models.

2.2.1 System Components

Generally, a polarimetric system can be composed of five parts. These five parts are:

Light source (LS) The Light Source (LS) produces a temporally incoherent or coherent light beam. This light beam can be conditioned to have any spatial intensity distribution and can be diverging, converging, or collimated. Unless stated otherwise, we will be using coherent, collimated beams in our systems in this thesis.

Polarisation State Generator (PSG) The Polarisation State Generator (PSG) provides the light beam from the LS with a polarisation profile that can be homogeneous across the entire beam or be heterogeneous across the beam. In the former case, the Stokes vector across the beam profile can be constant, while in the latter, the Stokes vector is a function of space across the beam profile.

Sample (S) The Sample (S) changes the polarisation of the beam in some way.

Polarisation State Analyser (PSA) The Polarisation State Analyser (PSA) polarises the beam from S in such a way that the resulting Stokes vector contains information representative of the way that S changed the beam polarisation that was set by the PSG.

Detector (D) The Detector (D) measures the optical energy (typically the S_0 Stokes parameter) and converts the energy into machine and human readable quantities.

For a Stokes polarimetry system, there is no PSG, however for a Mueller matrix polarimetry system, there is a PSG. The components mentioned in this section, along with their optomechanics, are the components described in $\mathcal{C}$.

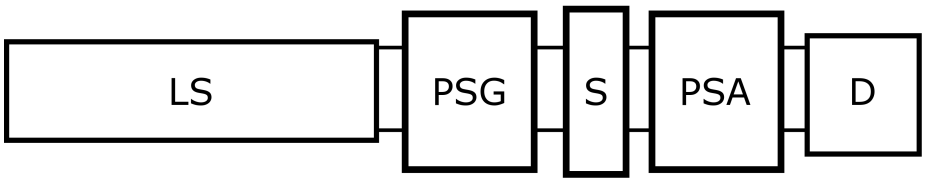


Figure 2.1: General Polarimetric System design consisting of an Light source (LS), Polarisation state generator (PSG), Sample (S), Polarisation state analyser (PSA), and the Detector (D).

2.2.2 System Model

The mathematical modelling of Polarimeters is done in two ways depending on what kind of system it is: a Stokes polarimetry system [41] or a Mueller matrix polarimetry system [53]. For either case, multiple system configurations are necessary. The following models apply only to a system with a polarisation insensitive (measures S_0 only) detector.

For a Stokes polarimetry system, each configuration has a PSA whose Stokes vector is $\underline{A} = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 \end{bmatrix}$. The $\underline{A}$ vectors can be stacked row-wise into a matrix, $\mathbf{W}$, called the measurement matrix. The intensity measured on the detector from all configurations can be neatly described then as

$$\underline{F} = \mathbf{W}\underline{S}F_0 \tag{2.48}$$

where $\underline{F}$ is a vector of N measurements, $\mathbf{W}$ is the $N \times 4$ measurement matrix, $\underline{S}$ is the Stokes vector being measured, and F_0 is the beam intensity from the light source. From here, if $N = 4$ and W is a non-singular matrix, $\underline{S}$ can be estimated by

$$\underline{S} = \frac{1}{F_0}\mathbf{W}^{-1}\underline{F} \tag{2.49}$$

however, if $N \neq 4$, then

$$\underline{S} = \frac{1}{F_0}\mathbf{W}^{+}\underline{F} \tag{2.50}$$

where $\mathbf{W}^+$ is the Moore-Penrose inverse of the measurement matrix.

For a Mueller matrix polarimetry system, along with $\underline{A}$, there is also the Stokes vector for the PSG, $\underline{G} = \begin{bmatrix} g_0 & g_1 & g_2 & g_3 \end{bmatrix}$. This system also has a measurement matrix as well, however, it is described differently. For the N measurements being taken, let $\underline{A}_i = \begin{bmatrix} a_{i0} & a_{i1} & a_{i2} & a_{i3} \end{bmatrix}$ and $\underline{G}_i = \begin{bmatrix} g_{i0} & g_{i1} & g_{i2} & g_{i3} \end{bmatrix}$ be the analyser and generator for the i^{th} measurement where $i = 0, 1, \cdots, N-1$, respectively. The $\mathbf{W}_i$ row in the measurement matrix is

$$\mathbf{W}_i = \begin{bmatrix} a_{i0}\underline{G}_i & a_{i1}\underline{G}_i & a_{i2}\underline{G}_i & a_{i3}\underline{G}_i \end{bmatrix} \tag{2.51}$$

with this, the system can be modelled as

$$\underline{F} = \mathbf{W}\underline{M}F_0 \tag{2.52}$$

where $\underline{F}$ is a vector of N measurements, $\mathbf{W}$ is a $N \times 16$ measurement matrix, $\underline{M}$ is a vectorised Mueller matrix, and F_0 is the beam intensity from the light source. To estimate $\underline{M}$, we can follow the same procedure as the Stokes polarimetry system where if $N = 16$ and $\mathbf{W}$ is non-singular, then

$$\underline{M} = \frac{1}{F_0}\mathbf{W}^{-1}\underline{F} \tag{2.53}$$

however, if $N \neq 16$, then

$$\underline{M} = \frac{1}{F_0}\mathbf{W}^{+}\underline{F} \tag{2.54}$$

where $\mathbf{W}^+$ is the Moore-Penrose inverse of the measurement matrix.

2.3 Error

2.3.1 Stochastic Error

The detector plays an important role in the detection of polarimetric information as methods for Stokes vector estimation and Mueller matrix estimation relying heavily upon intensity quantities. Goudail and Tyo have published several papers on system designs to

minimise the effect of additive Gaussian noise, signal-dependant Poisson shot noise, and mixed Poisson-Gaussian noise on polarimetry [2, 8, 40, 54, 57]. Each of these noise models comes from fundamental properties of the measurement systems: additive Gaussian is from thermal noise in the semiconducting material and other circuit parameters and Poisson shot noise is from the combination of the stochastic arrival of photons and the stochastic conversion of photons into current [43].

We will mainly be concerned with additive Gaussian stochastic error in this thesis and model it as

$$f = F + \eta \tag{2.55}$$

where F is the ideal measure EM wave intensity received from the light source on the detector, and η is stochastic noise sampled from a zero-mean Gaussian with standard deviation σ_f ($\eta \sim \mathcal{N}(0, \sigma_f^2)$) under the assumption that enough photons are incident on the detector that we can assume the error to be Gaussian distributed [43].

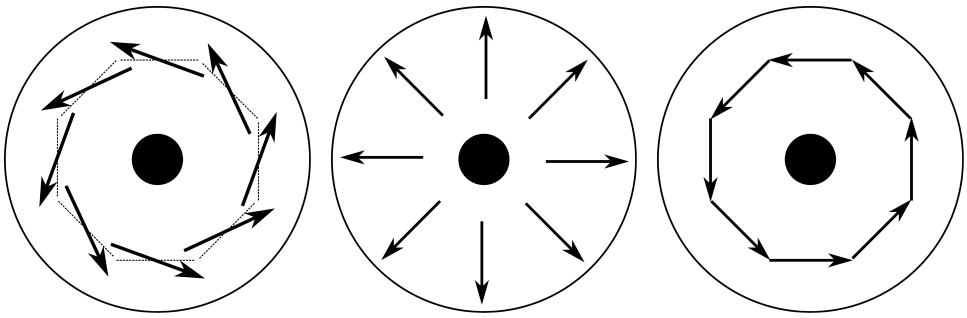
2.3.2 Deterministic Error

Polarimetry systems rely upon setting up PSGs and PSAs in different configurations to detect the Stokes vectors of light [13]. However, these systems work optimally when the components are arranged perfectly relative to each other, and do not affect the direction of light propagation.

Research has been done to determine the optimal combinations of PSG and PSAs such that error in the arrangement do not have a significant affect on the estimation of the Stokes parameters [41, 53, 57]. However, there are some effects that cannot be corrected for through arrangement because they are a result of the components themselves. For example, beam wander is a result of rotating retarding elements [46]. Additional computational work has to be done to centre the beam [27].

2.4 Heterogeneously Polarised Beams

Homogeneously polarised light maintains a constant polarisation across the entire beam profile. In contrast, Heterogeneously polarised light has many polarisations across the entire beam profile.



(a) Spirally Polarised Beam (b) Radially polarised SPB (c) Azimuthally polarised SPB

Figure 2.2: Examples of Radial and Azimuthal SPB polarisation profiles. The centre of the beam profile is marked with a black dot and the extent of its detectable intensity is the larger circle. The arrows represent the polarisations across the beam profile.

Relative to homogeneously polarised beams, heterogeneously polarised beams offer the ability to send multiple light polarisations into a sample to determine its polarising properties. If a homogeneously polarised beam is used for this purpose, it would have to modulate to several different type of polarisations to provide the same amount of information. Heterogeneously polarised beams have been utilised for this exact purpose recently [9, 10, 44]. In this section, I will review a type of heterogeneously polarised beam: Spirally polarised beams [4, 9, 10, 24, 34, 35, 39, 62, 44, 51].

2.4.1 Spirally Polarised beams

Heterogeneously polarised beams provide a spatial distribution of polarisation states over the profile of the beam. Spirally polarised beams (SPB) are a particular subset of heterogeneous polarised beams where there is an axially symmetric linear polarisation profile. The appeal of using SPBs is that they provide a constant polarisation profile as they travel through space and all linear polarisation states are present across the beam [10]. Two commonly seen versions of SPBs are shown in Figure 2.2: azimuthally and radially polarised beams. Creative methods for generating SPBs have been theorised and demonstrated in the past and will be reviewed here.

Nguyen *et al.* [34] utilised q-plates to generate SPBs. Briefly, q-plates are liquid crystal molecules imprinted between two glass plates and the orientation of the molecules is controlled with a variable voltage supply. A q-plate can be tuned to convert any input

polarisation into a radially and azimuthally polarised beam. Passily *et al.* [35] used an interferometric methods to produce SPBs. They combined beams of two orthogonal transverse electromagnetic modes (TEM), the TEM_{01} and TEM_{10} Hermite-Gaussian Modes, to produce a radially polarised beam. This can later be translated to an azimuthal polarisation through the use of a half-wave retarder. Brown and Beckley [4] discuss the use of stress engineered optics to produce SPBs. Applying forces to the edges of an optically transmissive element will produce a smooth birefringence pattern. The centre of the stressed element attains a azimuthally-independant half-wave retardation profile. This element can be rotated to produce an azimuthally or radially polarised SPBs. Lai *et al.* [24] were able to create SPBs by etching an α-BBO crystal to depths in a stepwise fashion. This etched crystal is then placed between two quarter-wave retarders which are then followed by two half-wave retarders to produce radially or azimuthally polarised beams. Radwell *et al.* [39] utilised Fresnel principles to generate a spirally polarised beam through the back reflections of light against a cone. This method exploits the geometric phase shift that occurs in beams during total internal refraction. It so happens that using a glass cone surrounded by air acts as a quarter-wave retarder with a spatially varying fast-axis. This allows the conversion of a circularly polarised beam into a SPB. McEldowney *et al.* [29] use a liquid polymer that has been baked and photoaligned with UV light. The photoalignment is azimuthally varying about the centre of the polymer plate that is thick enough to act as a half-wave retarder. The spatially varying fast axis of a half-wave retarder acts as a spatially varying rotator. With the input of a linearly polarised light at any arbitrary angle, a SPB is produced. This liquid polymer plate is sold by Thorlabs as a Vortex half-wave retarder (Thorlabs WPV10L-633) and what is used in this thesis to generate SPBs.

2.5 Bayesian Networks

Bayesian networks were conceived by Judea Pearl in 1985 [36] to formalise the objectives of connected models of knowledge and then solidified them as a field of study in his book *Probabilistic Reasoning in Intelligent Systems* in 1988 [37].

Bayesian networks are defined as being directed acyclic graphs (DAG) with directed edge structure. A DAG $D = (\mathcal{U}, \mathcal{E})$ is defined by a set of nodes $\mathcal{U} = \{X_1, \cdots, X_n\}$ and a set of directed edges $\mathcal{E}$. A property of DAGs is if you start at any node, you cannot trace a path along the directed edges back to that same node. An example of a DAG is in Figure 2.3.

Bayesian networks use the directed edge structure of DAGs to signify logical dependencies and causalities in evidential updating of information meant to model the joint

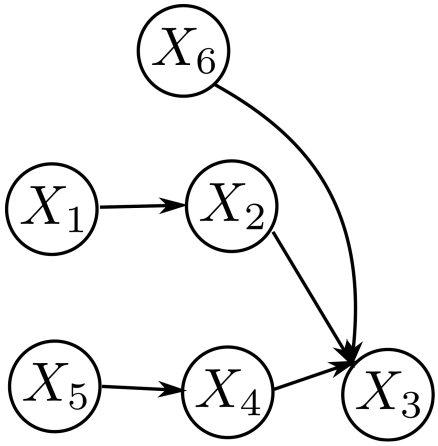


Figure 2.3: An example of a directed acyclic graph ($D = (\mathcal{U}, \mathcal{E})$) representing a Bayesian network mapping the of the joint probability on the set of nodes $\mathcal{U}$. The set of directed edges, $\mathcal{E}$, dictate the logical causality of evidential information.

probability of all the nodes in the DAG. The acyclic structure of D is important here because the update in knowledge of a node in D should not be then used to update itself. A DAG also has an ordering d to the nodes in $\mathcal{U}$, which we will use to reflect the ordering of the polarimetric system design.

Highlighted below are two properties of Bayesian networks:

d-separation

This definition is quoted from Pearl [37]. If $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ are three disjoint (non-intersecting) subsets of nodes in D, then $\mathcal{Z}$ is said to **d-separate** $\mathcal{X}$ from $\mathcal{Y}$ if there is no path from a node in $\mathcal{X}$ and a node in $\mathcal{Y}$ along which the following two conditions hold

1. every node with converging arrows is in $\mathcal{Z}$ or has descendants in $\mathcal{Z}$
2. every other node is outside $\mathcal{Z}$

The implications here are if you have a path from $\mathcal{X}$ to $\mathcal{Y}$ converging through $\mathcal{Z}$ (i.e. $\mathcal{X} \rightarrow \mathcal{Z} \leftarrow \mathcal{Y}$), $\mathcal{Z}$ d-separates the two sets. The path from $\mathcal{X}$ to $\mathcal{Y}$ is blocked, until information is learned about $\mathcal{Z}$. Another way to look at it is that until $\mathcal{Z}$ is learned, $\mathcal{X}$ to $\mathcal{Y}$ are independent, but when $\mathcal{Z}$ is learned, they become mutually dependant.

Markov Property

Given a DAG D and a joint probability distribution of all the nodes in D, any node

in D is conditionally dependant on only its direct parents (i.e. for any node X_i, its probability is $P(X_i = x_i|\mathbf{\Pi}_{X_i})$ where the set $\mathbf{\Pi}_{X_i}$ is the minimal set of connected parents of X_i).

The combination of these two properties allows us to structure a Bayesian network and use Bayesian techniques for polarising parameter inference and system design evaluation. To demonstrate the use of a Bayesian network for modelling joint probabilities let's model the Bayesian network in Figure 2.3. This DAG $D = (\mathcal{U}, \mathcal{E})$, defines the joint probability of $P(X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4, X_5 = x_5, X_6 = x_6)$. Using the Markov property we can define the joint probability as

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4, X_5 = x_5, X_6 = x_6) = \prod_i P(X_i = x_i|\mathbf{\Pi}_{X_i}) \quad (2.56)$$

$$\begin{aligned} P(X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4, X_5 = x_5, X_6 = x_6) = \\ P(X_1 = x_1)P(X_2 = x_2|X_1 = x_1) \\ P(X_3 = x_3|X_2 = x_2, X_4 = x_4, X_6 = x_6)P(X_4 = x_4|X_5 = x_5)P(X_5 = x_5)P(X_6 = x_6) \end{aligned} \quad (2.57)$$

From this graph,X_2 d-separates X_1 and X_6 and X_3 blocks X_4 and X_5 from X_1, X_2, and X_6. Taking X_2 for example, if we learn its outcome, then we can infer X_1 and X_6.

2.6 Summary

In this chapter, we have reviewed the theory for modelling polarised light, modelling the polarisation effects of components, models of stochastic error, heterogeneously polarised beams, and Bayesian networks. With this information, the reader will now know enough to understand the theory used in this thesis to describe the novel Computational Polarimetry Framework and using it in the following two chapters.

Chapter 3

Literature Review

In this chapter, we will introduce the state of art methods for polarimetric system design. These methods focus primarily on optimising the choice of Stoke vector to reduce the variance in intensity measurements received from the detector. We review them in the context of full polarimetric system design.

3.1 Polarimetry System Optimisation

The typical design methodologies for polarimetry systems revolve around choosing the best PSA configurations or the best PSG/PSA configurations to estimate accurate Stokes vector of a beam or a sample's Mueller matrices. For Stokes polarimetry, PSA designs are chosen that reduce the effects of stochastic and deterministic error on the estimate of the Stokes vector. For the Mueller matrix polarimetry, the PSA designs are chosen that not only reduce the effects of stochastic and deterministic error on the estimate of the Mueller matrix, but also on the choice of Stokes vectors from the PSG to optimally estimate the matrix.

For Mueller matrix polarimetry, the optimal choice of Stokes vector to estimate the Mueller matrix has been shown to be those that create Platonic solids in the Poincaré sphere [3, 26]. Platonic solids can be phrased simply as solids that sit inside a sphere with the maximum volume possible. Azzam was the first to show that the optimal estimation of a Mueller matrix with four Stokes vectors is done with any four linearly independent Stokes vectors that form the largest volume tetrahedron possible in the Poincaré Sphere [3]. Layden *et al.* went on to show that the optimal estimate of a Mueller matrix, such that

there is the smallest possible amount of error from random sources, can be done with any number of Stokes vectors, so long as they form Platonic solids inside the Poincaré Sphere [26]. Anna and Goudail went on to show that the increase in the number of measurements decreases the variance in the estimate of the Mueller matrix [2]. In addition to choosing the optimal probing Stokes vectors, the choice of the PSA can be used to reduce the effects of stochastic and deterministic error.

For a system to be resilient to stochastic error, the sensitivity of $\mathbf{W}^+$ to small changes in measurement intensity that can be attributed to the stochastic error sources in Section 2.3.1 must be minimised. There have been several metrics devised on how to choose the optimal configurations of the PSA and PSG such that the inverse of the measurement matrix is not sensitive to minor stochastic variations. For Stokes polarimeters with 4 measurements, the optimisation procedures focus on minimising the condition number of the measurement matrix [1, 56]. The condition number is calculated as

$$\mathrm{CN}(\mathbf{W}) = ||\mathbf{W}||||\mathbf{W}^{-1}|| \tag{3.1}$$

where $\mathrm{CN}(\mathbf{W})$ is the condition number of $\mathbf{W}$, and $||\cdot||$ is either the $L1$ or L_∞ matrix norm. For PSAs with more than four measurements, the optimisation of the measurement matrix revolves around the reduction of figures of merit that are functions of the measurement matrix's singular values [41, 42]. Sabatke *et al.* proposed two figures of merit: The reciprocal absolute determinant (RAD) and the equally weighted variance (EWV).

$$\mathrm{RAD} = \prod_j \frac{1}{\mu_j} \tag{3.2}$$

$$\mathrm{EWV} = \sum_j \frac{1}{\mu_j^2} \tag{3.3}$$

where μ_j is the j^{th} singular value of the measurement matrix. Zallat took a different approach to determining the optimal measurement matrix configuration by recognising that the detector cannot detect small changes in intensity, therefore, instrument components should be chosen that produce large intensity changes to reflect small polarisation changes [61]. The modulation efficiency factor metric was introduced for each row of the measurement matrix. If we define $\mathbf{Q} = \mathbf{W}^+$, then the modulation efficiency factor is

$$\zeta_i = \left(\sum_j q_{ij}^2 \right)^{-\frac{1}{2}} \tag{3.4}$$

where q_{ij} is the matrix element of $\mathbf{Q}$ in the i^{th} row and j^{th} column. The mean system efficiency is the average of all the modulation efficiencies of $\mathbf{W}^{+}$.

Deterministic error sources can affect the measurement matrix and cause some error in the estimate of the Stokes vector or Mueller matrix. Ambirajan and Look were wary of deterministic error from experimental component setup and opted to choose a logic-based route to reduce error by proposing an optimal configuration of components with whole angles rather than half angles [1]. Sabatke *et al.* address deterministic error briefly as a nice consequence of using their figures of merit for determining a rotating-retarder fixed-polariser (RRFP) configuration where the chosen parameters are stables to within 2.5 degrees for the orientation angle and 4 degrees of the retarder rotation [41]. Tyo did an analysis of deterministic error on the estimation of Stokes vectors and Mueller matrices with a rotating retarder. He noticed that there is a trade-off in the reduction of deterministic error in Mueller matrices and the reduction of the condition number of measurement matrices in rotating retarder systems. The trade-off exists due to the fact that the trajectory of the polarisation around the Poincaré sphere is greater for larger retardances, which leads to a smaller condition number, however a variation from that retardance can therefore lead to a larger estimation error [55, 57]. Tyo also discovered that, by increasing the number of measurements, the error in the produced Stokes vector from deterministic error in the measurement matrix can be reduced [57]. Twietmeyer and Chipman devised an error metric on the trace of the covariance matrix of the Mueller matrix error from known error sources [53].

The metrics and methods mentioned have proven effective enough to produce accurate polarimetry systems based off of them, but these methods have several disadvantages:

1. Primarily, optimisation methods weigh more towards the reduction of stochastic error.

2. Save for Zallat [61], the methods do not take into account the detectability of polarisation changes.

3. The methods focus on the reduction of error through the design of the system components and do not use any no prior models to enhance system estimates after the measurements have been taken.

4. Optimisation methods focus on the optimisations of null intensity systems only.

5. These methods only focus on system parameter optimisation to reduce noise, but do not provide further methods to optimise for parameter estimation or optimise computational algorithms for parameter estimation.

Building on the last point, these methods provide methods to optimise the component parameters for reducing the variance in the estimates, but that is only a small part of the full design of application based polarimetric systems. Design requirements and constraints need to be considered for component and component parameter choice. As well, inference models need to be developed and evaluated with potential design trade-offs given prior knowledge of component behaviour and error. These methods cannot provide a metric that offers a complete picture of polarimetric system design.

3.2 Summary

In this chapter, we review polarimetric design metrics for system optimisation and identify the shortcomings of their methods. Primarily, these metrics focus on the reduction on the effects of stochastic error on measurements and do not take into account overall system requirements and constraints, prior knowledge of component behaviour to enhance the measurements, and are not flexible to optimise any system other than a null intensity polarimetry system. The result of this is that while they are useful tools, they do not provide a guiding framework or metric for full polarimetric system design.

Chapter 4

Computational Polarimetry

In this chapter, we will describe the formulation of the Computational Polarimetry Framework and relate it to the design of polarimetric systems from a systems design perspective. We will first explore the meaning of each of the sets in Eq. 1.1 and their place in polarimetric system design, then we will formulate the Computational Polarimetry Framework as a Bayesian network, and finally outline the design procedure.

4.1 Framework Parameters

4.1.1 Design Parameters ($\mathcal{D}$)

In the design of systems to achieve a particular objective, there are a subset of objectives that need to be reached during the design process. These sub-objectives are split into two groups: requirements and constraints. Requirements are design sub-objectives that can be adhered to within a certain margin, while Constraints have to be adhered to. The set $\mathcal{D}$ is defined for each system design at the outset. It can be split into two subsets of random variables, such that $\mathcal{D} = \{\mathcal{D}_R, \mathcal{D}_C\}$ where $\mathcal{D}_R$ are the random variables for the requirements and $\mathcal{D}_C$ are the random variables for the constraints.

The design requirements and constraints can affect all or some of the other parameters in the Computational Polarimetry Framework.

4.1.2 System Components ($\mathcal{C}$)

The set of system components ($\mathcal{C}$) is a set of random components such that

$$\mathcal{C} = \{c_0, c_1, \cdots, c_j, \cdots, c_{J-2}, c_{J-1}\} \tag{4.1}$$

where c_j is the j^{th} random system component. These components include all the components outlined in Section 2.2.1 as well as their optomechanical housings. The space of possible components is a discrete space given the physical realisability of each component.

4.1.3 System Component Parameters ($\mathcal{Q}$)

The set of system component parameters ($\mathcal{Q}$) is a set of random vectors describing the parameters for each component in the set of system components, $\mathcal{C}$. This implies that if there are J system component, then there are J system component parameters such that:

$$\mathcal{Q} = \left\{\underline{q}_0, \underline{q}_1, \cdots, \underline{q}_j, \cdots, \underline{q}_{J-2}, \underline{q}_{J-1}\right\} \tag{4.2}$$

where $\underline{q}_j$ is a random vector of component parameters for the j^{th} random system component. The components with the corresponding component parameters can be combined to describe Mueller matrices.

Most polarimetric systems are reference based systems and require at least one configuration to measure the reference beam or sample and at least one more to measure with either the beam of interest or with a sample of interest. In that case there can be many possible configurations of certain system parameters, and $\mathcal{Q}$ is the complete set of unique system parameters. This implies that the $\underline{q}_j$ random vectors can all be of different sizes.

4.1.4 Measurements ($\mathcal{I}$)

The measurements achieved by photodetectors or spatial detector arrays in this framework are represented by the set $\mathcal{I}$. The set is composed of N random variables such that

$$\mathcal{I} = \{f_0, f_1, \cdots, f_i, \cdots, f_{N-2}, f_{N-1}\} \tag{4.3}$$

where f_i is a random variable representing the measurement from the i^{th} component configuration.

Table 4.1: Polarimetric system logical design parameter dependancies

Parameter	Dependant upon
$\mathcal{I}$	$\mathcal{Q}$,$\mathcal{D}$
$\mathcal{Q}$	$\mathcal{C}$,$\mathcal{D}$
$\mathcal{C}$	$\mathcal{D}$

4.2 Bayesian Network

The Computational Polarimetry Framework is a joint probability between all the parameters involved in the design of a polarimetric systems. Each of the parameters in the joint probability are directly influenced by a subset of the other parameters; this is summarised in Table 4.1. It can be seen from the dependencies in Table 4.1, that the parameters are sequential and dependant upon only the previous parameter (save for the $\mathcal{D}$ parameter, which can influence them all, but is not influenced by any of them). With this knowledge of dependencies, we can formulate the joint probability as a Bayesian network. This network is defined as a directed acyclic graph, $D = (\mathcal{U}, \mathcal{E})$, where $\mathcal{E}$ is the set of edges and $\mathcal{U}$ is a set of nodes. Each node is a set of random variables or vectors representing the parameters of the joint probability, defined as

$$\mathcal{U} = \{\mathcal{D}, \mathcal{C}, \mathcal{Q}, \mathcal{I}\} \tag{4.4}$$

The graph is visualised in Figure 4.1. The arrows represent the conditional dependence of the random variables at the tip of the arrow to the ones at the base of the arrow. However, it should be noted that not all elements in the the set $\mathcal{D}$ directly influence all the components, component parameters, and measurements.

With the graph formulated and the dependencies visualised, we can use the properties of the Bayesian network, outlined by Pearl [37], to formulate the joint probability in Eq. 1.1 as shown in Figure 4.1.

$$P\left(\mathcal{I}, \mathcal{Q}, \mathcal{C}, \mathcal{D}\right) = P(\mathcal{D})P(\mathcal{C}|\mathcal{D})P(\mathcal{Q}|\mathcal{C}, \mathcal{D})P(\mathcal{I}|\mathcal{Q}, \mathcal{D}) \tag{4.5}$$

Each term in the framework can be qualitatively explained as:

Term 1: $P(\mathcal{D})$

The prior probability on the design requirements and constraints limit the search

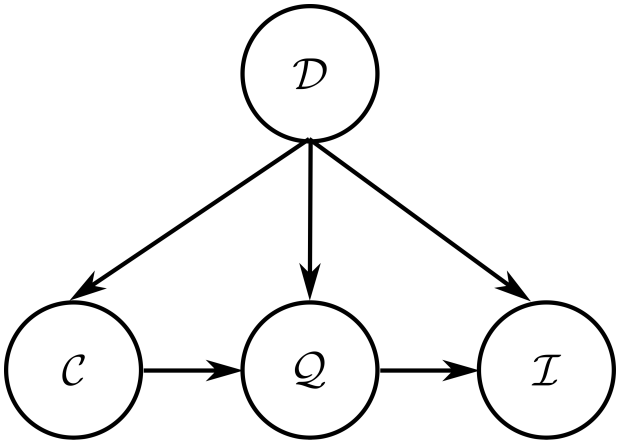


Figure 4.1: Graphical conceptualisation of the conditional dependencies in the Computational Polarimetry Framework. The set of components ($\mathcal{C}$), the set of system component parameters, ($\mathcal{Q}$), the measurement set ($\mathcal{I}$), and the set of design requirements and constraints ($\mathcal{D}$) have the dependencies of $\mathcal{C}$ on $\mathcal{D}$, $\mathcal{Q}$ on $\mathcal{C}$ and $\mathcal{D}$, $\mathcal{I}$ on $\mathcal{Q}$ and $\mathcal{D}$.

space for the random variables being inferred in the Computational Polarimetry Framework.

Term 2: $P(\mathcal{C}|\mathcal{D})$

The likelihood of component choice given the design requirements and constraints. An example of requirements and constraints that can affect choice of components can be size constraints, or budgetary requirements.

Term 3: $P(\mathcal{Q}|\mathcal{C},\mathcal{D})$

The probability of system component parameters occurring given the set of components and design requirements and constraints. An example can be the probability of angular orientation of a polariser given an optomechanical device that has uncertainty in its orientation due to the graduation of angles. Another example can the probability of retardance given a retarder that has uncertainty based on the environmental temperature with the environment temperature constrained to a range of operation.

Term 4: $P(\mathcal{I}|\mathcal{Q},\mathcal{D})$

The measurements produced from the components of the system can occur with a measure of uncertainty from the detector component parameters and other system component parameters in the system. As well, the measurements themselves

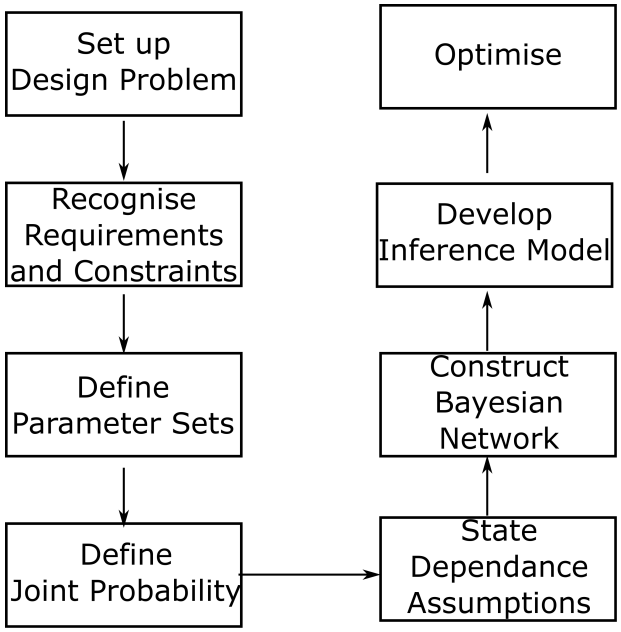


Figure 4.2: The design procedure utilising the Computational Polarimetry Framework. Prior knowledge of component behaviour and error sources will be incorporated in to the framework used throughout the design procedure.

are produced from the underlying physics parametrised by the polarimetry system component parameters in $\mathcal{Q}$.

4.2.1 Formulation for Polarimetry Systems

For a polarimetry system design problem, we will formulate the Bayesian network following the procedure shown in Figure 4.2.

First the design problem will be stated and requirements and constraints will be recognised. Then the parameters outlined in this chapter will be defined. The joint probability with those parameters will be defined right after. Once that is done so, logical independence and dependence assumptions are performed between parameter sets or within parameters themselves. Once the logical structure has been assumed, the Bayesian network can be constructed as a DAG $D = (\mathcal{U}, \mathcal{E})$ with all the parameters, defined or otherwise, as nodes in $\mathcal{U}$ connected by directed edges $\mathcal{E}$. Properties of Bayesian networks can now be utilised to develop the inference model of parameters in the network given knowledge of the other

nodes. The measurement set, in particular, is defined using knowledge of all the other nodes and the underlying physics of how optical energy travels through polarimetry systems. This is also where probabilistic models can be formulated or probabilistic tables can be defined. With the inference model complete, optimisation algorithms can be used to estimate system parameters.

4.3 Summary

In this chapter, we have outlined the formulation of the Computational Polarimetry Framework. The framework was formulated as a Bayesian network in a directed acyclic graph that followed the logical design structure of polarimetric systems incorporating a set of design requirements and constraints in addition to the set of components, their component parameters, and the resulting measurements.

With the Computational Polarimetry Framework, we can create an inference model to estimate different types of parameters in a polarimetric system for specific purposes. The remainder of this thesis will be using this framework to infer parameters of a polarimetric system. The first system will be a demonstration of using the framework to infer a system component parameter given other system information and design requirements, and the second system will demonstrate the use of the framework to computationally estimate a sample parameter from a novel system configuration.

Chapter 5

Computational Polarimetry in Action

Objects can be differentiated based on their dielectric properties by the angle of linear polarisation of natural illumination reflected from their surfaces for the purposes of saliency analysis in robotic vision [16], remote sensing [48], and in material identification tasks [59]. In this chapter, we demonstrate the Computational Polarimetry Framework to optimise the design of a system to estimate the angle of linear polarisation for this purpose.

The Computational Polarimetry Framework designs a joint probability that not only incorporates problem design requirements, but also prior information of component behaviour and prior information of both stochastic and deterministic error sources using a Bayesian network. We optimise the system under different experimental conditions in simulation against the design requirements.

5.1 Background

The angle of light polarisation can be measured simply using a linear polariser. The angle of the linear polariser affects the measurement of the beam following Malus' Law outlined in Eq. 5.1.

$$F_1 = F_o \cos^2(\alpha) \tag{5.1}$$

where α is the angle of the linear polariser's transmission axis relative to the horizontal. If a beam with a polarisation angle of ϕ encounters the polariser, the subsequent measurement will be described by Eq. 5.2.

$$F_2 = F_o \cos^2(\alpha + \phi) \tag{5.2}$$

The smallest angle to be detected is theoretically an infinitesimally small amount, however, with commercially available detectors, there is stochastic error (noise) from the measurement process of the detector. This noise defines a difference in measurements that can be confidently said to come from two different measurements. As a result, there is subsequently a lower limit placed on the minimum possible angular sensitivity from a linear polariser placed at some angle α. In addition, since the second derivative exists for $\cos^2(\alpha+\phi)$ with respect to some α, a difference in measurement can correspond to different ϕ values.

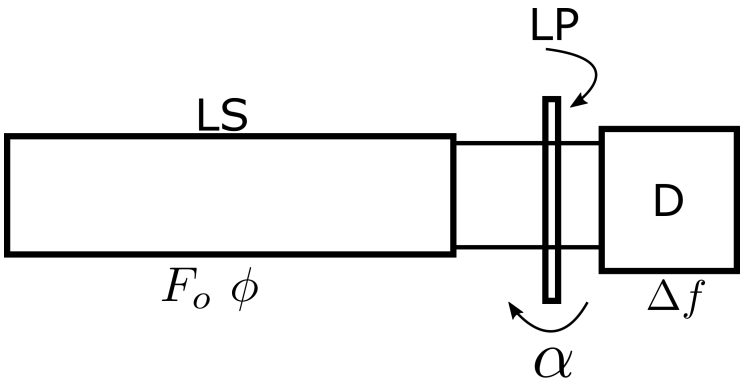
To combat the influence of noise, prior knowledge of the noise statistics can be used to identify the minimal detectable measurement difference (Δf) by defining a level of confidence and using a two-tailed z-table

$$\Delta f = 2(1.96)\sigma_f \tag{5.3}$$

where σ_f is the standard deviation of the detector converting measurements to machine readable values. We are assuming that there are enough photons incident upon the detector that we can assume an additive Gaussian error distribution [43]. With this we can find the linear polariser orientation that gives the smallest angular sensitivity for the smallest detectable measurement amount. While seemingly straight forward, the experimental implementation and construction of the polarimetric system can have uncertainties that inhibit the perfect construction. For example, the linear polariser could be in an optomechanical housing that can only achieve a certain angle with a certain precision due to project cost constraints that make purchasing a motorised mount prohibitive. In addition, there is a trade-off in the design between the measurement difference and angular sensitivity which can be relaxed depending on the required angular sensitivity of the system.

5.2 Design Problem

We will define our design objective as designing a polarimetric system that is capable of estimating the polarisation angle from two photodetector measurements. This system will be composed of a light source that emits light at some measurement value (F_o) and polarisation angle (ϕ), a linear polariser at some angle (α), and a photodetector. The system should be able to estimate small polarisation changes, therefore having a high angular sensitivity.



	F_o	ϕ	α	Δf
f_1	F_o	ϕ_1	α	Δf
f_2	F_o	ϕ_2	α	Δf

Figure 5.1: The system design of the single linear polariser angle estimator system. The components are the light source (LS) with component parameters F_o and ϕ, a rotated linear polariser (LP) rotated by α from the horizontal with component parameter α, and a detector (D) with component parameter Δf. A summary of the parameters for each measurement are summarised in the table under the setup. The parameter that changes between each measurement is the angle of linear polarisation ϕ.

5.2.1 System Design

The system is shown in Figure 5.1. This system is similar to a Stokes polarimetry system with a light source, a PSA, and a detector. In this system, what we are measuring is the angle of light polarisation from the light source.

The light source's polarisation $\underline{S}_{LS}$ can be modelled as having a rotated linear polarisation of

$$\underline{S}_{LS} = \begin{bmatrix} 1 \\ \cos 2\phi \\ -\sin 2\phi \\ 0 \end{bmatrix} F_o \tag{5.4}$$

The measurement matrix of the rotated linear polariser, the PSA, is modelled as

$$\mathbf{W} = \underline{W} = \begin{bmatrix} 1 & \cos 2\alpha & -\sin 2\alpha & 0 \end{bmatrix} \tag{5.5}$$

resulting in the detector measurement being

$$F = F_o \cos^2(\alpha + \phi) \tag{5.6}$$

The detector is modelled as producing continuous measurements, f, with infinite well capacity. The measurements are modelled as

$$f = F + \eta \tag{5.7}$$

where F is the ideal measure on the detector, and η is stochastic noise sampled from a zero-mean Gaussian with standard deviation σ_f ($\eta \sim \mathcal{N}(0, \sigma_f^2)$).

5.2.2 Requirements and Constraints

The only requirement of this system is that the angular sensitivity be as high as possible. We can quantify this requirement as $\Delta\phi_R$ and set it equal to zero. We will incorporate this into the set of design requirements ($\mathcal{D}$).

5.2.3 Computational Polarimetry Framework

With the design problem outlined above, we can start building the graphical model with parameters. First we will outline the model parameters and then provide models for each node. Then we will build an inference model of the linear polariser angle.

Table 5.1: System components and their parameters.

Component	Parameters
Light Source	Beam measurement (F_0)
	Polarisation angle (ϕ_1)
	Polarisation angle (ϕ_2)
Linear Polariser	Polariser Angle (α)
Photodetector	Detectable measurement difference (Δf) (Eq. 5.3)

Model Parameters

The set of design requirements is defined as

$$\mathcal{D} = \{\mathcal{D}_R = \{\Delta\phi_R = 0\}\} \tag{5.8}$$

The components of the system are chosen already and are defined as

$$\mathcal{C} = \{c_0 = \text{Light Source}, c_1 = \text{Linear Polariser}, c_2 = \text{Photodetector}\} \tag{5.9}$$

The components will need to perform two measurements under two configurations. The first measurement will be used as a reference to determine the beam measurement when the beam is completely horizontal, while the second will be a beam with a polarisation angle. We define the parameters for each component and the configurations in Table 5.1 where ϕ_1, is the angle of polarisation for the first measurement and ϕ_2 is the angle of polarisation for the second measurement. These parameters are random variables, however, some will be fixed to values for us to infer α and $\Delta\phi$. The fixed variables are defined in Eq. 5.10.

$$\mathcal{Q} = \left\{\underline{q}_0 = [F_o = F_o, \phi_1 = 0, \phi_2 = \Delta\phi], \underline{q}_1 = [\alpha], \underline{q}_2 = [\Delta f = \Delta f]\right\} \tag{5.10}$$

In this system, there will be two measurements f_1 and f_2, where f_1 is the reference beam and f_2 is the rotated beam. The reference beam can be modelled ideally as using Eq. 5.1 and f_2 they will be defined using Eq. 5.2.

$$\mathcal{I} = \left\{f_1 = F_o \cos^2(\alpha), f_2 = F_o \cos^2(\alpha + \phi)\right\} \tag{5.11}$$

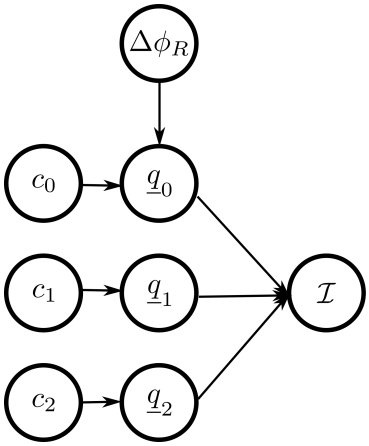


Figure 5.2: Bayesian network for the parameters and assumptions outlined in Section 5.2.3 and 5.2.3.

Assumptions

We will include the following assumptions before we model the joint probability as a Bayesian network following the procedure for the Computational Polarimetry Framework. The assumptions are as follows:

1. The $\Delta\phi_R$ design requirement effects the component parameters of the light source only, since it is the only random vector of component parameters with the angle of polarisation, ϕ.
2. The system component parameters are conditionally dependant on their associated components only.

With these assumptions and the parameters defined above, we can construct a DAG.

Bayesian Network

Using the assumptions and parameters defined above, we can now construct a Bayesian network as a DAG $D = (\mathcal{U}, \mathcal{E})$ where $\mathcal{U}$ is the set of parameters defined in this system, and $\mathcal{E}$ are the edges. The graph has some ordering d such that it can be constructed in the form of Figure 5.2.

With the network defined, we can formulate the joint probability in Eq. 1.1 as

$$P(\mathcal{I}=\mathcal{I},\mathcal{Q}=\mathcal{Q},\mathcal{C}=\mathcal{C},\mathcal{D}=\mathcal{D}) = \\ P(\Delta\phi_R=0)P(c_0=c_0)P(c_1=c_1)P(c_2=c_2) \\ P(F_o=F_o,\phi_1=0,\phi_2=\Delta\phi|c_0=c_0,\Delta\phi_R=0)P(\alpha|c_1=c_1)P(\Delta f=\Delta f|c_2=c_2) \\ P(f_1=F_1,f_2=F_2|F_o=F_o,\phi_1=0,\phi_2=\Delta\phi,\alpha,\Delta f=\Delta f) \tag{5.12}$$

where F_1 is defined by Eq. 5.1, and F_2 is defined by Eq. 5.2.

Probabilistic Models

With the joint probability being modelled as the product of prior probabilities and likelihoods, we will now define the probabilistic models we will use to define some of the terms. Terms that have fully defined random variables for this system will not be discussed and their product will be later interpreted as a constant during the inference process. We will now define each term in Eq. 5.12 to then use in a computational algorithm to infer α and $\Delta\phi$.

Term 1: $P(F_o=F_o,\phi_1=0,\phi_2=\Delta\phi|c_0=c_0,\Delta\phi_R=0)$

We will assume that the component parameters for c_0 and its configurations are all uncorrelated and independent. And we will assume that only the ϕ_2 parameter is dependant on $\Delta\phi_R=0$. With these assumptions the term becomes

$$P(F_o=F_o,\phi_1=0,\phi_2=\Delta\phi|c_0=c_0,\Delta\phi_R=0) = \\ P(F_o=F_o|c_0=c_0)P(\phi_1=0|c_0=c_0)P(\phi_2=\Delta\phi|c_0=c_0,\Delta\phi_R=0) \tag{5.13}$$

With the first two terms being constants with respect to the random variables α and $\Delta\phi$, we need to define a model for $P(\phi_2=\Delta\phi|c_0=c_0,\Delta\phi_R=0)$ as part of the inference framework. To do this, we will also assume that all $\Delta\phi$ values are possible given the light source, c_0, however, with the design requirement, the highest chance should be $\Delta\phi=\Delta\phi_R$. With these assumptions, we can assert that the likelihood of the parameter $\Delta\phi$ given the component and the design requirement can be modelled as a Gaussian distribution with mean at $\Delta\phi_R$ and a tolerance factor $\sigma_{\Delta\phi}$.

$$P(\phi_2 = \Delta\phi | c_0 = c_0, \Delta\phi_R = 0) = A_1 \exp -\frac{1}{2}\left(\frac{\Delta\phi - \Delta\phi_R}{\sigma_{\Delta\phi}}\right)^2 \tag{5.14}$$

$$= A_1 \exp -\frac{1}{2}\left(\frac{\Delta\phi}{\sigma_{\Delta\phi}}\right)^2 \tag{5.15}$$

where A_1 is a normalisation constant.

Term 2: $P(\alpha|c_1 = c_1)$

Given the component c_1, the choice of α could be restricted to a specific set of α values for a defined precision. We can assert the likelihood of α given the component c_1 as

$$P(\alpha|c_1 = c_1) = \sum_{\text{all k}} \frac{g_k(\alpha)}{\mathbb{Z}} \tag{5.16}$$

where $\mathbb{Z}$ is a normalisation factor, and the function $g_k(\cdot)$ is defined as

$$g_k(\alpha) = \exp -\frac{1}{2}\left(\frac{\alpha - \alpha_k}{\sigma_\alpha}\right)^2 \tag{5.17}$$

where α_k is a possible angular value (graduation) given the component c_1, and σ_α represents the precision of the component. We can rationalise this assertion by assuming our confidence in the angle of the polarisation holder is symmetric and we are the most confident when the polariser angle is on the angle graduation of the polarisation holder, therefore modelling the likelihood as a Gaussian distribution supports those assumptions.

Term 3: $P(f_1 = F_1, f_2 = F_2 | F_o = F_o, \phi_1 = 0, \phi_2 = \Delta\phi, \alpha, \Delta f = \Delta f)$

This term defines the likelihood of the measurements occurring given the defined system parameters. With the given definitions of F_1 and F_2 in Eq. 5.1 and 5.2, respectively, and the purpose of our system, we will assert the model for this term as a Gaussian distribution:

$$P(f_1 = F_1, f_2 = F_2 | F_o = F_o, \phi_1 = 0, \phi_2 = \Delta\phi, \alpha, \Delta f = \Delta f) = A_2 \exp -\frac{1}{2}\left(\frac{(|F_0 \cos^2(\alpha + \Delta\phi) - F_0 \cos^2(\alpha)|) - \Delta f}{\sigma_{\Delta f}}\right)^2 \tag{5.18}$$

where A_2 is a normalisation constant and $\sigma_{\Delta f}$ is a tolerance factor. We assume a Gaussian distribution because we want the difference between the measurements to be as close to the smallest recognisable difference possible, but since the possible set of angles from the polarisation holder is limited (resulting in a source of deterministic error), we need to be able to relax the adherence to this difference to accommodate for that with $\sigma_{\Delta f}$.

With the assumed and asserted models, the joint probability in Eq. 5.12 becomes

$$P(\mathcal{I}=\mathcal{I},\mathcal{Q}=\mathcal{Q},\mathcal{C}=\mathcal{C},\mathcal{D}=\mathcal{D}) = \beta\left[A_1\exp-\frac{1}{2}\left(\frac{\Delta\phi}{\sigma_{\Delta\phi}}\right)^2\right]\left[\sum_{\text{all k}}\frac{1}{\mathbb{Z}}\exp-\frac{1}{2}\left(\frac{\alpha-\alpha_k}{\sigma_\alpha}\right)^2\right] \left[A_2\exp-\frac{1}{2}\left(\frac{(|F_0\cos^2(\alpha+\Delta\phi)-F_0\cos^2(\alpha)|)-\Delta f}{\sigma_f}\right)^2\right] \tag{5.19}$$

where β is the product of all the probabilities that are constant and do not contain the random variables α and $\Delta\phi$. With Eq. 5.19, we can formulate an optimisation function to determine the optimal α and $\Delta\phi$ values as then

$$\hat{\Delta\phi},\hat{\alpha} = \underset{\alpha,\Delta f}{\arg\max}\Gamma(\alpha,\Delta\phi) \tag{5.20}$$

where

$$\Gamma(\alpha,\Delta\phi) = A_1A_2\left[\exp-\frac{1}{2}\left(\frac{\Delta\phi}{\sigma_{\Delta\phi}}\right)^2\right]\left[\sum_{\text{all k}}\frac{1}{\mathbb{Z}}\exp-\frac{1}{2}\left(\frac{\alpha-\alpha_k}{\sigma_\alpha}\right)^2\right] \left[\exp-\frac{1}{2}\left(\frac{(|F_0\cos^2(\alpha+\Delta\phi)-F_0\cos^2(\alpha)|)-\Delta f}{\sigma_f}\right)^2\right] \tag{5.21}$$

5.3 Experimental Setup

To demonstrate the Computational Polarimetry Framework in this design problem, we will infer the α and $\Delta\phi$ parameters that maximise the joint probability in Eq. 5.19 and the optimisation problem in Eq. 5.21.

We will perform the optimisation using a global basinhopping approach [58] to maximise the function in Eq. 5.21. Basinhopping is a stochastic global optimisation algorithm that performs local optimisation at every stochastic jump. In these experiments we use the Broyden, Fletcher, Goldfarb, and Shanno bounded local optimisation method that uses first derivatives [63]. We choose a stochastic global approach rather than a gradient based approach for the fact that the deterministic error from our system can produce multiple local extreme to occur in our $\Gamma(\alpha, \Delta\phi)$ function (This is seen to occur in Experiment 4 in Figure 5.4). Gradient based approaches would be caught in the local extrama for each graduation. This is implemented and done through the Scipy package in Python [21]. We will define the domain of the two random variables as

$$\Delta\phi \in [0, 180)\ ^\circ \tag{5.22}$$

$$\alpha \in [0, 180)\ ^\circ \tag{5.23}$$

and let α_k be defined as

$$\alpha_k = k\Delta\alpha \tag{5.24}$$

where $k = 0, 1, \cdots, K-2, K-1$ where $K = \text{floor}\left[\frac{2\pi}{\Delta\alpha}\right]$.

To demonstrate the effectiveness of this framework in designing polarimetric systems under different conditions, we will estimate the two parameters, α and $\Delta\phi$ under eight different experimental conditions. Each experiment will change the beam measurement (i.e. F_o), noise (i.e. σ_f), precision (i.e. σ_α), and graduations (i.e. $\Delta\alpha$) conditions. The parameter combinations are in Table 5.2. In addition, the ideal component case will be simulated (i.e. $\Delta\alpha \to 0$) and then the case with no noise will be simulated (i.e. $\sigma_f \to 0$, therefore $\Delta f \to 0$). Each experiment is described and given purpose below:

Experiment 1: The noise free ($\sigma_f \to 0$), infinite angle case ($\Delta\alpha \to 0$), it is expected that $\alpha = \frac{\pi}{4}$ and $\Delta\phi \to 0$ [6]. This experiment will determine if the Computational Polarimetry Framework can perform similar system optimisation to the prior art.

Experiment 2 and 3: Observes the change in system parameters to a change in the light source parameters. The expectation is that the design and sample parameters will change to estimate a larger $\Delta\phi$ value from Experiment 2 to 3.

Experiment 4: Observe the system parameter response to a change in the angular step size of the linear polariser from Experiment 2. The expectation is that by increasing

Table 5.2: Parameter sets for System Design Experiments. The − symbol represents that parameter being either excluded or moot during the experiment. The → symbol represents there only being an infinitesimally small amount of deterministic or stochastic error.

Experiment No.	F_o	σ_f	$\sigma_{\Delta f}$	$\sigma_\alpha(^\circ)$	$\Delta\alpha(^\circ)$	$\Delta\phi_R(^\circ)$
1	2^{10}	$\to 0$	1	–	$\to 0$	-
2	2^{10}	1	1	1	2	-
3	2^8	1	1	1	2	-
4	2^{10}	1	1	1	5	-
5	2^{10}	2	1	1	2	-
6	2^{12}	1	1	–	$\to 0$	0.01
7	2^{12}	1	1	1	2	1
8	2^{12}	2	1	–	$\to 0$	1

the step size of the linear polariser, we change the α parameter to be a multiple of 5 rather than 2 in comparison to Experiment 2.

Experiment 5: Observes the change in system parameters when the detector noise increases. The resulting change in the effective signal to noise ratio should decrease the $\Delta\phi$ valuefrom Experiment 2.

Experiment 6,7, and 8: Observes the change in system parameter performance when design requirements change. The resulting systems should all produce parameters that adhere as close to the design requirements as possible.

We can assess the accuracy of the Computational Polarimetry Framework for each experiment by generating 2^{16} samples for a reference beam (f_1) and 2^{16} for a beam rotated at the estimated angle $\Delta\phi$ (f_2) and plotting the histogram (probability distribution function) of each of those sets. We will define the histogram from f_1 as $H_{f_1}[f_b]$ and the histogram from f_2 as $H_{f_2}[f_b]$ where f_b is the histogram bin centre. We will be assuming this is all captured on a continuous detector. The performance of the framework's estimates will be quantitatively assessed by calculating the probability of error using Eq. 5.25. The higher the error, the less distinguishable the two measurement distributions are implying that the system is not capable of producing distinguishable measurements from the detector and the less desirable of a system it will be.

$$P(\epsilon) = \sum_{\text{all bins}} 1_{f_1} H_{f_1}[f_b] + 1_{f_2} H_{f_2}[f_b] \tag{5.25}$$

Table 5.3: Results of the Computational Polarimetry Framework. High $P(\epsilon)$ implies minimal separation between the measurement sets. Experiment 1 has a high $P(\epsilon)$ value implying large overlap between the measurement sets. Experiment 7 has the lowest $P(\epsilon)$ value implying that the two measurement sets are the most separable.

Experiment No.	$\hat{\alpha}(^\circ)$	$\hat{\Delta\phi}(^\circ)$	$P(\epsilon)$
1	45.0	0	1.0
2	134.00	0.22	0.0510
3	134.00	0.88	0.0682
4	135.00	0.22	0.0506
5	112.00	0.62	0.0545
6	135.33	0.05	0.0628
7	178.00	1.07	0.1611
8	176.62	1.12	0.0527

where

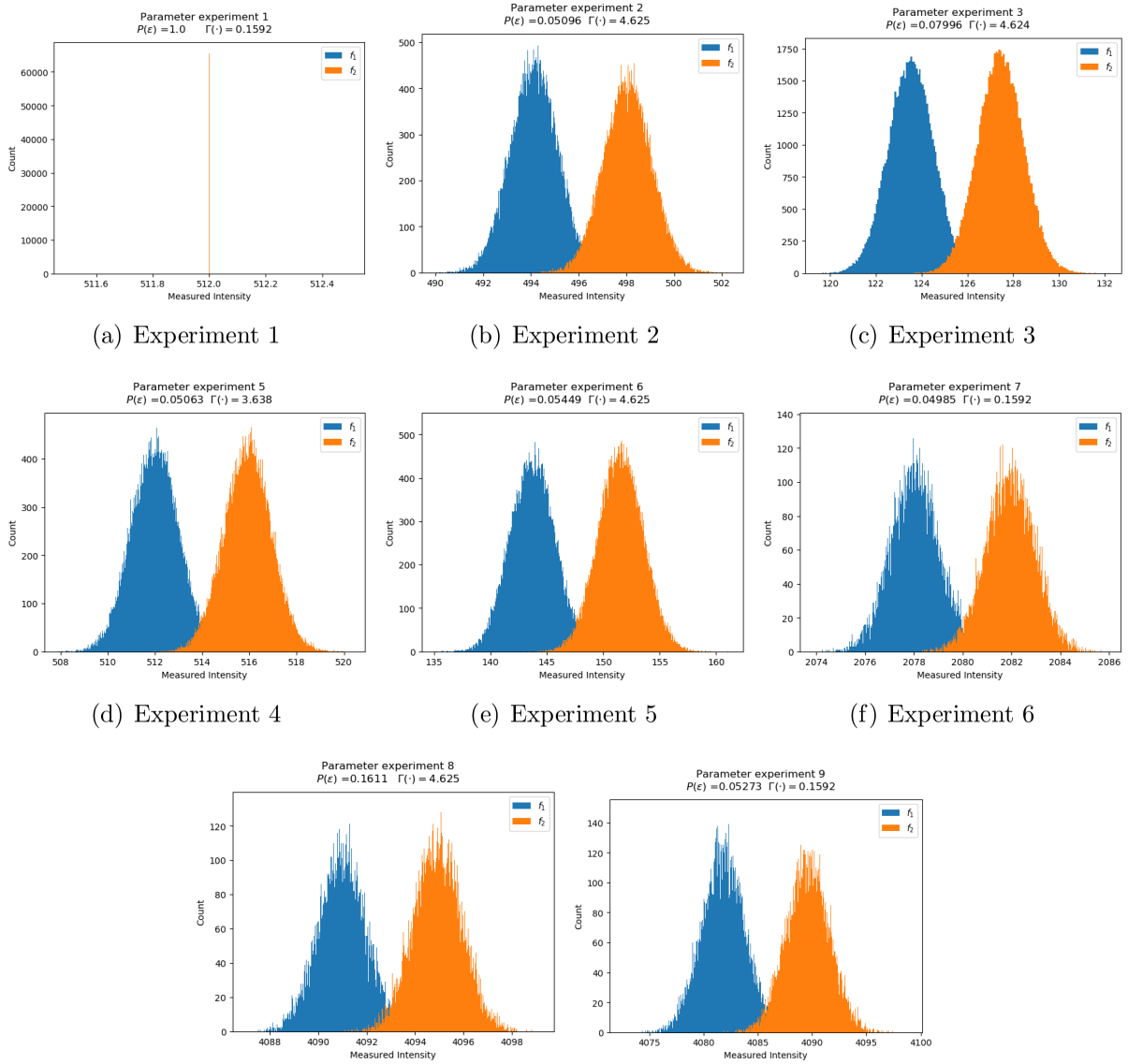
$$1_{f_1} = \begin{cases} 1, & H_{f_1}[f_b] \leq H_{f_2}[f_b] \\ 0, & \text{otherwise} \end{cases} \tag{5.26}$$

$$1_{f_2} = \begin{cases} 1, & H_{f_2}[f_b] \leq H_{f_1}[f_b] \\ 0, & \text{otherwise} \end{cases} \tag{5.27}$$

5.4 Results and Analysis

Figure 5.3 shows the histograms for all eight experiments along with the $P(\epsilon)$ and the estimated parameter are tabulated in Table 5.3. Observations from each experiment are discussed below.

Experiment 1: This experiments validates that the Computational Polarimetry Framework mimics the parameter choice in literature for a similar system [6]. It is of note here that the metrics in Table 5.3 show this to be very bad configuration, and so does the corresponding histogram. However, since there is no noise and the detector is continuous, there is no need for the two measurements to be any further apart to be detectable.



(a) Experiment 1 (b) Experiment 2 (c) Experiment 3

(d) Experiment 4 (e) Experiment 5 (f) Experiment 6

(g) Experiment 7 (h) Experiment 8

Figure 5.3: The resulting histogram from each of the 8 parameter sets. Included in each plot is the probability of error $P(\epsilon)$ for that parameter set and also the value of the $\Gamma(\cdot)$ function for the chosen parameters from the optimisation of the joint probability. Aside from Experiment 1, we see good separation between the measurement groups f_1 and f_2 and low $P(\epsilon)$ values. In Experiment 7 however, the probability of error is higher than Experiment 8 due to its adherence to the design requirement and adherence to the polariser graduations.

Experiment 2 and 3: We see that by decreasing the measurement of the beam from the light source, there is an increase in the smallest detectable $\Delta\phi$ value, as expected. This demonstrates that the Computational Polarimetry Framework is an adaptive model that performs what Zallat's metric was intended to assess [61].

Experiment 4: By increasing the angular step, we see there has been a change to the $\hat{\alpha}$ design parameter to accommodate for the uncertainty in the polariser holder graduations relative to Experiment 2.

Experiment 5: With the increase in detector noise we see an in crease in the $\Delta\phi$ value in comparison to Experiment 2. This is due to the fact that a two-fold increase in possible fluctuation does not create a 4 fold decrease in signal. The Computational Polarimetry Framework as able to optimise the system to arrive at this value, but it is also able to simultaneously choose the appropriate α design parameter to not only achieve the smallest $\Delta\phi$ value, but also adhere to the prior knowledge of the detector measurement process to produce confidently different measurements.

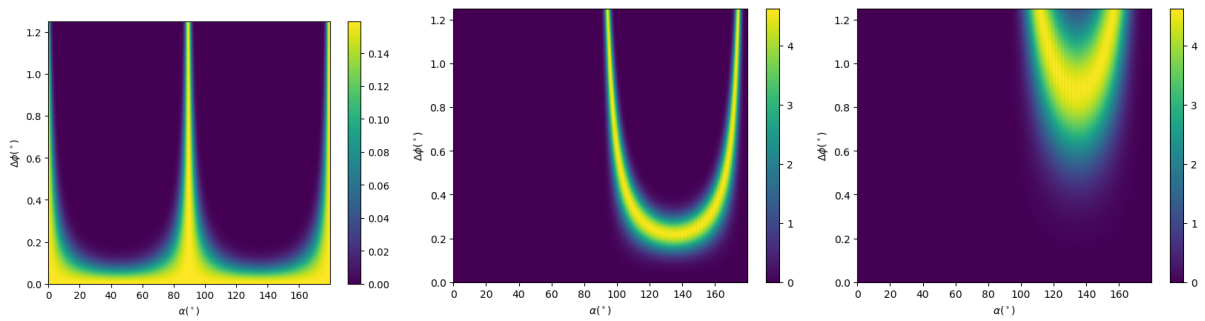
Experiment 6,7, and 8: With a design requirement for angular sensitivity specified, the inference model was able to optimise the system parameters to adhere to it. In experiment 7 however, the design requirements were met, but when uncertainty in the angular graduation was introduced, the probability of error increased.

5.5 Discussions and Future Work

In this chapter, we presented a framework to infer system parameters that adhere to design requirements to design a single linear polariser angle estimation system. In this section, we discuss the stability of the estimated parameters and discuss the uniqueness of the results.

5.5.1 System Stability

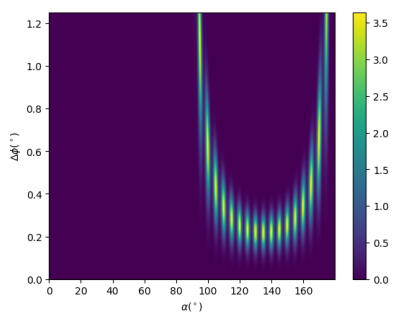
The stability of a measurement to system error has been explored by Sabatke *et al.* [41], Tyo [55, 57] and Twietmeyer [53]. Analysis is done by observing the change in the condition number of the measurement matrix to small changes in parameters and by observing Stokes or Mueller reconstruction error with small changes. We can look at the stability of every configuration by observing the $\Gamma(\cdot)$ function for each of our configurations. These function plots are shown in Figure 5.4 with $\alpha \in [0, 180)^\circ]$ and $\Delta\phi \in [0, 1.5)^\circ]$. The higher the value on the plot, the better that parameter is for maximising the joint probability function for



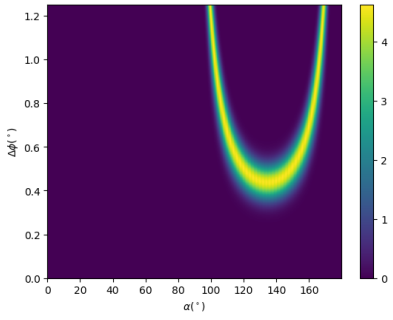
(a) Experiment 1

(b) Experiment 2

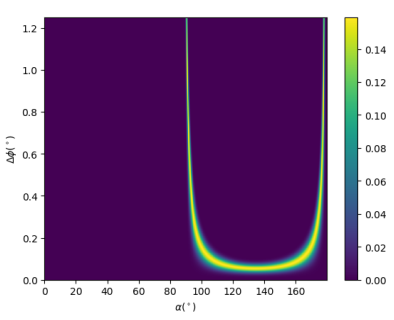
(c) Experiment 3



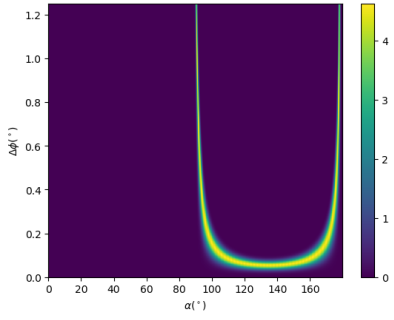
(d) Experiment 4



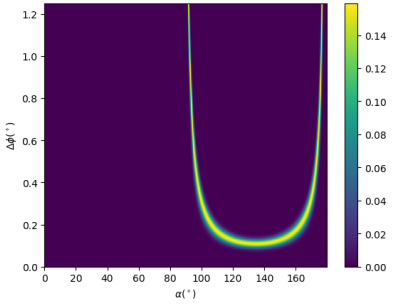
(e) Experiment 5



(f) Experiment 6



(g) Experiment 7



(h) Experiment 8

Figure 5.4: The resulting $\Gamma(\alpha, \Delta\phi)$ function from each of the 8 parameter sets. The higher the value of $\Gamma(\alpha, \Delta\phi)$, the more optimal those parameters are for a linear polarisation estimation system under different light source parameters, detector parameters, and the $\Delta\phi_R$ requirement. These plots demonstrate that there is some stability in the α position to achieve a minimum $\Delta\phi$.

the system. In the proposed framework, there is some stability in the α position to achieve a minimum $\Delta\phi$. In all the experiments, stability in the desired sensitivity is in the tens of degrees. Further study should be done to explicitly incorporate system stability into the Computational Polarimetry Framework to design stable polarimetric systems.

In Experiment 7 and 8, system stability actually decreases when the linear polariser graduation is removed ($\Delta\alpha \to 0$) while there is little change between experiment 2 and 5. This is due to the fact that 7 and 8 need to adhere to the sensitivity requirement and there are fewer possible values for that requirement to be met at.

5.5.2 Uniqueness

From the same plots in Figure 5.4, each the functions contains several optimum values along the α axis. This reflects the observation by Tyo [55] and Ambirajan and Look [1] where there can be many optimal configurations to minimise the system optimisation metrics in polarimetry. However, in the computational optical polarimetry framework, we explicitly adhere to the design requirement and find an optimum value that minimises the $\Delta\phi$ value. This is related to the intrinsic logical connections made in the Bayesian network formulation to mimic the logical choice that Ambirajan and Look took in their system component parameter choice [1].

5.6 Summary

In this chapter, we demonstrated the capability of the framework to be used in the design of a known system under various component conditions and design requirements. It was capable of achieving known system parameter configurations with no intentional direct intervention from the designer and design systems under stochastic error from the measurement process and deterministic error from spaced graduations on linear polariser optomechanics. In addition, due to the explicit incorporation of design requirement, unique optimal configurations can be achieved.

Chapter 6

Computational Optical Rotary Dispersion

In this chapter, we use the Computational Polarimetry Framework to develop an inference model to estimate the optical activity of a sample given an novel system that utilises SPBs and spatial detector arrays. The Computational Polarimetry Framework utilises prior knowledge of the system in a Bayesian network to develop the inference model. We simulate the performance of the inference model and validate the performance in the lab.

6.1 Background

Optical rotary dispersion (ORD) measures the amount of linear polarisation rotation over a range of wavelengths. This is utilised for concentration estimation, and material identification [6, 52].

Current methods for optical rotary dispersion are based off of null intensity methods where either two orthogonal polarisation states are measured or where all possible polarisations are scanned through. With only two orthogonally polarised measurements, noise can have a large impact on the estimation of optical activity and significantly reduce the accuracy of ORD systems demonstrated in Castaglioni *et al.* [6]. Sweeping through the possible orientations of linear polarisation provides prior information about the optical system to produce more accurate ORD estimates, however, the system requires additional complex components like electro-mechanical optical choppers, and lock-in amplifiers to decouple the signal from any contamination, requiring not only more measurements, but in

increasing system complexity and cost [7, 25]. Reduction in system complexity and the number of measurements is a design goal for polarimetric systems so that faster and more portable systems can be developed for looking at dynamic systems [60] and in low-resource settings [17].

A way to reduce system complexity is through the use of heterogeneously polarised beams, like SPBs. SPBs are capable of probing a sample with multiple linear polarisation states simultaneously. This is similar to scanning through all linear polarisation states, like in the prior art, however, it happens as a function of space rather than time. This results in the removal of electromechanical components that reduce system complexity and cost.

Sande *et al.* have demonstrated the use of SPBs for sample parameter estimation [9, 10, 44]. They chose to use three measurements from the beam corresponding to three Stokes vectors whose measurement matrix condition number was minimal. This was to reduce the impact of stochastic noise variation in the estimation of sample properties. While successful, this did not take full advantage of all the measurements captured and also did not take into account prior information of the detector characteristic to computationally estimate the sample parameters or apply it the field of ORD estimation.

Using the Computational Polarimetry Framework, we propose and demonstrate the novel Computational Optical Rotary Dispersion (CORD) system. The CORD system utilises an SPB in combination with an inference model developed using the Computational Polarimetry Framework for ORD estimation. With the Framework the inference model utilises prior knowledge of component behaviour and spatial relations of the measurements. This results in a system that is not only compact, and less complex than scanning polarimetry systems, but is as accurate as commercial lab based optical activity systems while being a fraction of the size and cost.

6.2 Design Problem

We will define our design objective as designing a polarimetric system that is capable of inferring the optical activity of a sample using a SPB. This system will be composed of a light source, a linear polariser and vortex half-wave retarder (VHWP) PSG (this will be used to generate the SPB), an optically active sample, a fixed linear polariser PSA, and a spatial detector array.

6.2.1 System Design

The arrangement of components in the CORD system is illustrated in Figure 6.1. The PSG of the system is composed of a horizontal (or vertical) linear polariser and a half-wave retarder with a continuous fast axis as a function of the azimuth. This produces a beam with a heterogeneously polarised beam profile composed of spatially distributed linear polarisations. The Stokes vector incident on a single spatial detector array can be imagined as a rotated linear polarisation, $\underline{S}'(\alpha)$

$$\underline{S}'(\alpha(x,y)) = \mathbf{M}_{rot}(\alpha(x,y))\underline{S} \tag{6.1}$$

where $\alpha(x,y)$ is the rotation amount, as a function of the spatial detector arrays Cartesian position from the centre of the beam, and $\mathbf{M}_{rot}$ is described by a Mueller matrix

$$\mathbf{M}_{rot}(\alpha(x,y)) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\alpha(x,y) & \sin 2\alpha(x,y) & 0 \\ 0 & -\sin 2\alpha(x,y) & \cos 2\alpha(x,y) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6.2}$$

If we assume that the sample being measured is only optically active, then the sample can be modelled as a rotation matrix as well, where the beam polarisation is rotated by ϕ. The output Stokes parameter $\underline{S}_o(\alpha(x,y),\phi)$ is then modelled as

$$\underline{S}_o(\alpha(x,y),\phi) = \mathbf{M}_s(\phi)\underline{S}'(\alpha(x,y)) \tag{6.3}$$

where $\mathbf{M}_s(\phi)$ is the Mueller matrix of the optically active sample.

With a PSA of a vertical linear polariser crossed relative to the linear polariser in the PSG [1], the resulting system model is

$$\mathbf{M}_{sys}(\alpha(x,y),\phi) = M_{sys}(\alpha(x,y),\phi) = \begin{bmatrix} 1 & -1 & 0 & 0 \end{bmatrix} \mathbf{M}_s(\phi)\underline{S}'(\alpha(x,y)) \tag{6.4}$$

The intensity sampled by a detector at a certain position from the centre of the beam can be modelled then as

$$F(\alpha(x,y),\phi) = \cos^2\left(\alpha(x,y)+\phi\right) F_o \tag{6.5}$$

[1] The system can work in the parallel configuration as well.

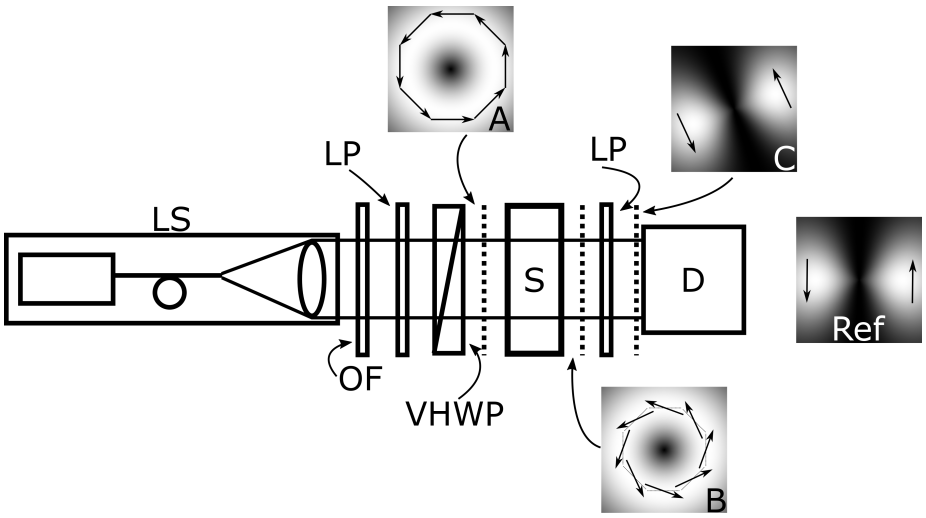


Figure 6.1: Schematic of optical system for the CORD system. LS is the light source, OF is a laser-line optical filter, LP is a linear polariser, VHWP is vortex half-wave plate, S is the sample, and D is a spatial detector array. Inset are pictures of the beam profile after each component past the VHWP: (A) the intensity profile and the polarisation distribution after the VHWP; (B) the intensity and rotated polarisation distribution after S; (C) the polarisation and intensity distribution after the LP; (Ref) the intensity and polarisation distribution for no optically active sample.

where F_o is the intensity of the light source.

The function in Eq. 6.5 is minimised when $\alpha = k\pi - \phi$ and maximised when $\alpha = k\frac{\pi}{2} - \phi$, where k is the number of 180 degree rotations the SPB undergoes as a function of its azimuth. [2]

After scanning through all linear polarisation orientations, finding the maxima and minima is the principle that linear orientation scanning ORD systems typically operate to determine the optical activity of the sample. While this technique sequentially scans through polarisation as a function of time, the CORD system scans the linear polarisations as a function of space. With this prior knowledge of the CORD system, for any given optical activity, we can predict the locations of the maxima and minima. We show the calculation of the location of the minima below

$$\begin{aligned} x' &= \rho\cos(k\pi - \phi) \\ y' &= \rho\sin(k\pi - \phi) \end{aligned} \tag{6.6}$$

where x' and y' are the Cartesian locations of the minima caused by the angular rotation, and ρ is some radius. Conversely, given the location of the minima, we can determine the optical activity, ϕ as

$$\phi = k\pi - \arctan\left(\frac{y'}{x'}\right) \tag{6.7}$$

However, due to noise, finding the minimum is not a trivial task. While prior art dealt with stochastic error through optimal choice of measurements, or the addition of complex timing mechanics, we will demonstrate that this can be done computationally instead with a single measurement. In the next section we will define the CORD system in the Computational Polarimetry Framework and utilise its properties to develop an inference model to estimate the optical activity.

6.2.2 Requirements and Constraints

We want to have a novel system that is less complex and costly in comparison to the existing state of art for ORD. With the market variability of component cost, we will focus

[2] For this thesis, we will assume $k = 1$ in formulation to match the WPV10L-633 Thorlabs Optic used in the Real Experiments in Section 6.3.3.

on component complexity as being a constraint in the design of the novel CORD system. While complexity is a qualitative descriptor, in the ORD application we can give it a quantitative value by defining it as the number of electro-mechanical and electro-optical components in our system design. We will describe the constraint with the random variable C_T. Since we want our system to not be complex at all, we will say that $C_T = 0$.

To calculate the complexity of our CORD system, we will define a function $f_C(\mathcal{C})$ that, given a set of components $\mathcal{C}$, produces a scalar value representing the complexity of the set. Since the function takes a set of random variables as its parameter, the output of the function will also be a random variable. The range will be assumed to be $f(\mathcal{C}) \in [0, \infty)$. It is defined as

$$f_C(\mathcal{C}) = \sum_{j=0}^{J-1} \mathbb{1}(c_j) \tag{6.8}$$

where

$$\mathbb{1}(c_j) = \begin{cases} 1 & \text{if } c_j \text{ is electro-mechanical or electro-optical} \\ 0 & \text{otherwise} \end{cases} \tag{6.9}$$

for the simplicity, we will assume the output of the function $f_C(\mathcal{C})$ is continuous.

In addition, we want the system to be constrained in its size and be smaller than the state of art. We will define the random variable L_T for this constraint. We can compare this to the JASCO 815 spectropolarimeter with its 1150 mm system width. Since this might be exaggerated due to its additional features, we will instead use its significantly smaller depth dimension as our size constraint. The depth of the system is 576 mm [**?**]. Similar to the previous complexity constraint, we will define a function that will calculate the length of the system based off the system component. We will define this function as $f_L(\mathcal{C})$

$$f_L(\mathcal{C}) = \sum_{j=0}^{J-1} \ell(c_j) \tag{6.10}$$

where

$$\ell(c_j) = \text{Width of the component} + \text{Minimum point-to-point distance between } c_j \text{ to } c_{j+1} \quad (6.11)$$

with the range being defined to $f_L(\mathcal{C}) \in [0, \infty)$. We will set $L_T = 576$mm for our system design.

With the constraints and the way to extract them from the set defined, we will incorporate them into the design constraint set $\mathcal{D}_C$ in the computational polarimetry framework.

6.2.3 Computational Polarimetry Framework

With the design outlined above, we can start to build the inference model with parameters. First we will define the parameters and define the fixed values, and then we will provide models that we can use in the inference of the sample optical activity.

Model Parameters

Our set of design constraints is defined as

$$\mathcal{D} = \{\mathcal{D}_C = \{C_T = 0, L_T = 576\text{mm}\}\} \quad (6.12)$$

The components in the system are defined as

$$\mathcal{C} = \{c_0 = \text{Light Source}, c_1 = \text{Horizontal Linear Polariser}, c_2 = \text{Vortex Half-wave Retarder}, c_3 = \text{Sample}, c_4 = \text{ Vertical Linear Polariser}, c_5 = \text{Spatial Detector Array}\} \quad (6.13)$$

The definition of the set of system component parameters is a little bit more difficult than in the design problem of Chapter 5. While the design parameters of components $c_{0-1,3-5}$ are not random variables and in a single configuration, the component parameters for the VHWP vary across the entire spatial domain of measurements. We will define the parameter set for c_2 as

$$(\underline{q}_2)_i = [\alpha(x, y) = \alpha(x, y)] \quad (6.14)$$

where $\alpha(x, y)$ is the defined rotation at some index (x, y), (x, y) are members of a set that describe the spatial locations of all N measurements relative to the centre of the beam, and j is the i^{th} row of the system component parameter random vector.

$$\mathcal{Q} = \left\{ \underline{q}_0 = [F_o = F_o], \underline{q}_1 = [], \underline{q}_2 = [\underline{q}_2], \underline{q}_3 = [\phi], \underline{q}_4 = [], \underline{q}_5 = [\sigma_f = \sigma_f] \right\} \tag{6.15}$$

where $\underline{q}_2 = [\underline{q}_2]$ is defined in Eq. 6.14, and σ_f is the standard deviation of the noise from the detector [3]. The system component parameter random vectors $\underline{q}_1$ and $\underline{q}_4$ are undefined because we are assuming that their orientation is perfectly defined and are not random.

These system parameters will produce a set of measurements $\mathcal{I}$ defined as

$$\mathcal{I} = \{f_0, f_1, f_2, \cdots, f_i, \cdots, f_{N-2}, f_{N-1}\} \tag{6.16}$$

Each measurement is produced independently by the varying rotation angle of the VHWP centred about the beam centre and since we are using a VHWP element, there is some spatial relation in the measurement intensities being measured on the spatial detector array. We will take advantage of that spatial relation and perform a transform of the set of measurements into the Zernike domain. This is done through the calculation of the Zernike moments for all the radial and azimuthal modes. The procedure for calculating the Zernike moments is in the next section.

Zernike Moments

The measured intensities are indexed on a Cartesian grid, however, they can be decomposed as the sum of weighted Zernike polynomials that form a complete orthogonal basis over a circle of unit radius [23, 33]. A Zernike polynomial of order (n, m) is composed of a radial polynomial $(R_{nm}(\rho))$ and a complex phase term such that it is defined as

$$V_{nm}(\rho, \theta) = R_{nm}(\rho) \exp(jm\theta), \quad \rho \leq 1 \tag{6.17}$$

where $R_{nm}(\rho)$ is defined as

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^2 \frac{(n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} \rho^{n-2s} \tag{6.18}$$

[3] We are assuming the same noise model as in Eq. 5.7

where n is a positive integer or zero, m is an integer subject to the constraints, $n - |m|$ is even, and $|m| \leq n$.

For a spatial distribution of measurements defined as $F(\rho(x,y), \alpha(x,y))$, the two dimensional Zernike moment is defined as

$$Z_{nm} = \frac{n+1}{\pi} \sum_{\text{all } \rho} \sum_{\text{all } \theta} V^*_{nm}(\rho, \theta) F(\rho(x,y), \alpha(x,y)) \tag{6.19}$$

where * symbol represents the complex conjugate. An interesting property of Zernike moments is if we have a some intensity distribution in polar coordinates, $F(\rho, \alpha)$, and it is rotated from a reference, $F^r(\rho, \alpha)$, by ϕ the Zernike moments of F will be the the Zernike moments of F^r but with a complex phase factor. This is shown in Eq. 6.20 and 6.21

$$F(\rho, \theta) = F^r(\rho, \theta + \phi) \tag{6.20}$$

$$Z_{nm} = Z^r_{nm} \exp(jm\phi) \tag{6.21}$$

With the Zernike moments for an intensity distribution defined, we can redefine $\mathcal{I}$ as a set of random Zernike moments

$$\mathcal{I} = \{Z_{00}, Z_{11}, Z_{1-1} \cdots, Z_{nm}, \cdots\} \tag{6.22}$$

Assumptions

With the parameters defined, we can now define our assumptions that will assist in modelling the joint probability for this polarimetric system as a Bayesian network. In the CORD system there is only one assumption:

1. The system component parameters are conditionally dependant on their associated components only.
2. The set of components are conditionally dependant on the C_T and L_T constraints.

With this assumption, we can now construct the DAG representing the Bayesian network.

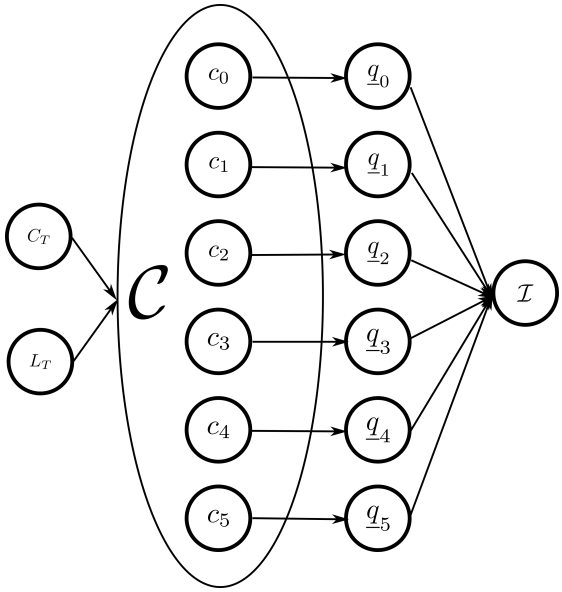


Figure 6.2: Bayesian network for the CORD system using the parameters and assumptions outlined in Section 6.2.3 and 6.2.3.

Bayesian Network

Using the assumptions and parameters defined above, we can now construct a Bayesian network as a DAG $D = (\mathcal{U}, \mathcal{E})$ where $\mathcal{U}$ is the set of parameters defined in this system, and $\mathcal{E}$ are the edges. The graph has some ordering d such that it can be constructed in the form of Figure 6.2.

With the network defined, we can formulate the joint probability in Eq. 1.1 as

$$\begin{aligned} P(\mathcal{I} = \mathcal{I}, \mathcal{Q} = \mathcal{Q}, \mathcal{C} = \mathcal{C}, \mathcal{D} = \mathcal{D}) = \\ P(\mathcal{C} = \mathcal{C} | C_T = 0, L_T = 587\text{mm}) \\ P(F_o = F_o | c_0 = c_0) P(\underline{q}_2 = \underline{q}_2 | c_2 = c_2) P(\phi | c_3 = c_3) P(\sigma_f = \sigma_f | c_5 = c_5) \\ P(\mathcal{I} = \mathcal{I} | F_o = F_o, \underline{q}_2 = \underline{q}_2, \phi, \sigma_f = \sigma_f) \end{aligned} \quad (6.23)$$

Probabilistic Models

With the joint probability being modelled as the product of prior probabilities and likelihoods, we will now define the probabilistic models we will use for some of the terms. Terms that have fully defined random variables for this system will not be discussed and their product will be later interpreted as a constant during the inference process. With this system being the first to introduce constraints, we will define $P(\mathcal{C} = \mathcal{C}|C_T = 0, L_T = 587\text{mm})$, as well. We will now define each term in Eq. 6.23 to then use in a computational algorithm to infer the system parameters.

Term 1: $P(\mathcal{C} = \mathcal{C}|C_T = 0, L_T = 587\mathbf{mm})$ This likelihood term defines the probability of the components given the constraints. In the set of components we defined, we will be able to satisfy the complexity and size constraints (and will touch upon this again in the lab experiments in Section 6.3), however, if the set of components are unknown, we can assert the distribution of $P(\mathcal{C}|C_T, L_T)$ to be uniformly distributed and with the functions $f_C(\cdot)$ and $f_L(\cdot)$, we can assume conditional independence of the cost of the system on the size constraint, and the size of the system on the cost constraint. As a result, we can simplify the likelihood and therefore $P(\mathcal{C}|C_T, L_T) = P(f_C(\mathcal{C})|C_T)P(f_L(\mathcal{C})|L_T)$. We will assert the distributions to be defined as follows:

$$P(f_C(\mathcal{C})|C_T) = \begin{cases} \frac{1}{C_T} & \text{if } f_C(\mathcal{C}) \leq C_T \\ 0 & \text{otherwise} \end{cases} \tag{6.24}$$

$$P(f_L(\mathcal{C})|L_T) = \begin{cases} \frac{1}{L_T} & \text{if } f_L(\mathcal{C}) \leq L_T \\ 0 & \text{otherwise} \end{cases} \tag{6.25}$$

And with these asserted distributions, should we wish to change the defined set of components and define them as random design variables, we can assess our choices and incorporate them in the design process.

Term 2: $P(\phi|c_3 = c_3)$ This prior can provide limits on the possible values of ϕ to be inferred. We will assert that there are no limits on ϕ and it is uniformly distributed over the range of $[0, 2\pi)$.

Term 3: $P(\mathcal{I} = \mathcal{I}|F_o = F_o, \underline{q}_2 = \underline{q}_2, \phi, \sigma_f = \sigma_f)$ This likelihood term dictates the probability of the set of measurements occurring given the system configuration. Make note that from Figure 6.2, we know the outcomes of all nodes in the network, save for $\underline{q}_3 = \phi$, which is the parameter we are designing this system to infer. With this,

we can use the d-separation property of Bayesian networks and can infer the probability of $\underline{q}_3 = \phi$ occurring, given the information of all the other nodes. We can simplify this likelihood to $P(\mathcal{I}|\phi)$ to make the next steps in the formulation easier to see and logically know the component parameters that produced the measurements. Using d-separation, we can use Baye's theorem to model the posterior as

$$P(\mathcal{I}|\phi) \propto P(\phi|\mathcal{I})P(\mathcal{I}) \tag{6.26}$$

We will dedicate the remainder of this section to the definition of the likelihood term $P(\phi|\mathcal{I})$ and prior term $P(\mathcal{I})$.

The likelihood and prior term defined above logically state that an estimate of the optical activity can be obtained from the set of Zernike moments that inherently describe the intensity distribution of the measurements and their spatial relation. To formulate the two terms, we will follow the method of Kim and Kim [23] for angular estimation using Zernike moments to come up with an estimate of the optical activity.

Referring back to Eq. 6.21, the magnitudes of the rotated Zernike moments ($||Z_{nm}||$) and the reference Zerike moments ($||Z^r_{nm}||$) are the same. Therefore to estimate the angle of rotation from a pair of Z^r_{nm} and Z_{nm} moments:

$$\Phi_{nm} \equiv \arg\left(\frac{Z_{nm}}{Z^r_{nm}}\right) = m\phi \tag{6.27}$$

where $0 \leq \Phi_{nm} \leq 2m\pi$. It should be noticed that there is going to be ambiguity in the angles calculated due to the range of Φ_{nm}. Similar to Kim and Kim [23], we will incorporate a correction factor such that

$$\Phi_{nm} = \Omega_{nm} + 2\pi K_m, \quad K_m \in \{0, 1, 2, \cdots, m-1\} \tag{6.28}$$

where Ω_{nm} is the true rotation and $2\pi K_m$ is a correction factor. With this ambiguity correction, there will be m solutions for Φ_{nm}. However, due to stochastic noise from a spatial detector array, the calculation of Φ_{nm} is contaminated. Kim and Kim assumed the estimate of Φ_{nm} is contaminated with additive zero-mean Gaussian noise due to the measurements also being contaminated with additive zero-mean Gaussian noise and modelled it as

$$\Phi_{nm} = m\phi + \eta \tag{6.29}$$

where $\eta \sim \mathcal{N}(0, \sigma^2)$. An estimate of ϕ can be found as

$$\hat{\phi}_{nm} = \frac{\Omega_{nm}}{m} + \frac{2\pi}{m} K_m \tag{6.30}$$

Since we are estimating an optical activity from each Zernike moment using the method of Kim and Kim [23], we can model the posterior instead as

$$P(\mathcal{I}|\phi) = \sum_{\text{all } Z_{nm}} P(\phi|Z_{nm}) P(Z_{nm}) \tag{6.31}$$

with this new formulation of the $P(\phi|Z_{nm})$ likelihood and the assumed additive noise model on the angular estimated, it can be modelled as

$$P(\phi|Z_{nm}) = A_1 \exp\left(-\frac{1}{2}\left(\frac{\phi - \hat{\phi}_{nm}}{\sigma}\right)^2\right) \tag{6.32}$$

where A_1 is a normalisation constant, σ is representative of the noise from the detector harming the angular estimate. The prior term can be defined as the normalised Zernike moments implying that the stronger the moment, the more likely the occurrence of the estimate. The prior term, $P(Z_{nm})$, can be asserted then

$$P(Z_{nm}) = \frac{||Z_{nm}||}{\sum_{\text{all } m,n} ||Z_{nm}||} \tag{6.33}$$

With this, we can define the joint probability as

$$P(\mathcal{I} = \mathcal{I}, \mathcal{Q} = \mathcal{Q}, \mathcal{C} = \mathcal{C}, \mathcal{D} = \mathcal{D}) = \beta \sum_{\text{all } Z_{nm}} \left(\left[A_1 \exp\left(-\frac{1}{2}\left(\frac{\phi - \hat{\phi}_{nm}}{\sigma}\right)^2\right)\right] \left[\frac{||Z_{nm}||}{\sum_{\text{all } Z_{nm}} ||Z_{nm}||}\right] \right) \tag{6.34}$$

where β is a constant representing the product of all the probabilities in the model. With the final framework of the inference determined in Eq. 6.34, we can formulate the estimation of the optical activity, ϕ as the value that maximises the joint probability. The inference of ϕ is formulated then as the following optimisation problem

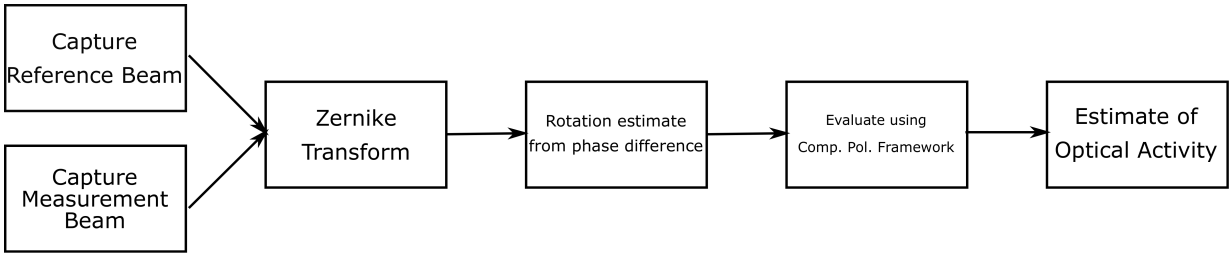


Figure 6.3: The measurement procedure using the CORD system. First a reference beams is captured without an optically active sample. Then a measurement beam is captured with an optically active sample. The set of Zernike moments for each of the beams is produced and used to produce a set of angular estimates. The angular estimates are evaluated using the Computational Polarimetry Framework and an accurate estimate is inferred given prior knowledge of the measurement system.

$$\hat{\phi} = \arg\max_{\phi} \Gamma(\phi) \tag{6.35}$$

$$= \arg\max_{\phi} \sum_{\text{all } Z_{nm}} \left(\left[A_1 \exp\left(-\left(\frac{\phi - \hat{\phi}_{nm}}{\sigma} \right)^2 \right) \right] \left[\frac{||Z_{nm}||}{\sum_{\text{all } Z_{nm}} ||Z_{nm}||} \right] \right) \tag{6.36}$$

6.2.4 Measurement Procedure

With the CORD system components defined and the inference model derived for the estimating the sample optical activity, we can now define a procedure for the system. A flowchart of the procedure is shown in Figure 6.3. To use the CORD system for inferring sample optical activity, first a reference measurement set must be captured. This reference is the intensity distribution $F^r(\alpha(x,y),\phi)$ in Eq. 6.5, except where $\phi = 0$. The reference beam has travelled through all the components in the CORD system, except there is no optically active sample. This intensity distribution is saved for later use. Next, the optically active sample is placed in the system and another intensity distribution is captured. This is the measured intensity distribution $F(\alpha(x,y),\phi)$ for some value ϕ.

Once the reference and the measured intensity distribution have been captured, their Zernike moments are calculated, and optical activity estimates from the ratio of the Zernike moments are produced for each (n,m) index. These are put into the joint probability model and evaluated. The best angular estimate will produce the highest probability.

6.3 Experimental Setup

To demonstrate the effectiveness of this system for inferring the optical activity of a sample, we will perform simulated and real experiments. The simulated experiments will empirically demonstrate the performance of the CORD system to:

1. Infer the optical activity of a sample to limits similar to commercial ORD devices.
2. Infer the optical activity under varying levels of noise.

The real experiments will estimate the rotation in the intensity distribution from a polarisation rotator and from a sucrose solution. We will first explain the procedure for parameter inference, then setup the simulated and real experiments.

6.3.1 Inference Procedure

The inference of ϕ is formulated as the optimisation problem in Eq. 6.36 and done using the procedure outlined in Section 6.2.4. We practically implement the inference procedure by using the first 35 Zernike moments to model the intensity distributions, similar to Kim and Kim [23]. The σ parameter was set to $0.001°$. A grid search optimisation was done to select $\hat{\phi}$ over 1 million points spanning $[0, 2\pi)$

6.3.2 Simulated Experiments

In the simulated experiments, a reference intensity distribution is generated using the intensity model proposed by Gori [14]

$$F_o(\rho) = AJ_1(\beta\rho)\exp\left(-\frac{\rho^2}{w_0^2}\right) \tag{6.37}$$

where A is a constant, $J_1(\cdot)$ is a first order Bessel function, w_0 is width of a Gaussian envelope, and β is a transverse scaling of the Bessel function. For the purposes of this experiment, we will use the following parameters presented in Table 6.1. An example of the spatial intensity distribution is in inset A in Figure 6.1.

The rotated intensity distribution is determined by the angular scaling factor in Eq. 6.5 such that

$$F(\rho(x,y), \alpha(x,y), \phi) = \cos^2(\alpha(x,y) + \phi)\, F_o(\rho(x,y)) \tag{6.38}$$

Table 6.1: Parameter values for CORD Synthetic Experiments

Parameter	Value
A	1 Watt
w_0	$M/6$
β	1

Optical Activity Estimation

ORD systems typically measure the optical activity of samples to a millidegree [6]. We will investigate the lower limit of angular estimation of the CORD system by simulating the following rotations (ϕ) and inferring them:

1. 0.001°
2. 0.002°
3. 0.003°
4. 0.004°
5. 0.005°
6. 0.006°
7. 0.007°
8. 0.008°

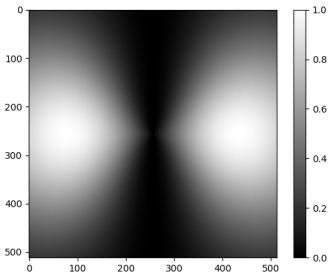
The intensity distribution from each of the rotations will be sampled on a 512×512 Cartesian detector array (i.e $M = 512$). Zero-mean additive Gaussian noise (i.e. $v \sim \mathcal{N}(0, \sigma^2)$) will be added for each intensity distribution to mimic various Signal-to-Noise Ratio (SNR) values (SNR $=20 \log\left(\frac{1}{\sigma}\right)$). The following SNRs will be used:

1. 80 dB
2. 40 dB
3. 30 dB
4. 20 dB

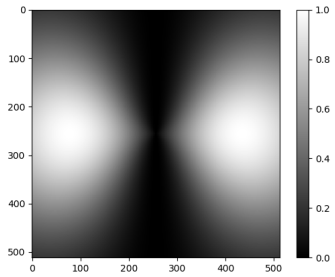
The level of noise contamination can be visually observed in Figure 6.4 with the reference beam.

The measurement procedure for the CORD system relies upon the measurement of a reference beam. With the reference beam being captured by the same spatial detector array as the measured beam, it will be contaminated with the same amount of noise and have a similar SNR. We will explore the impact of this on the inference of the optical activity by using a noise-free reference and a noise-contaminated reference for each of the angular rotations.

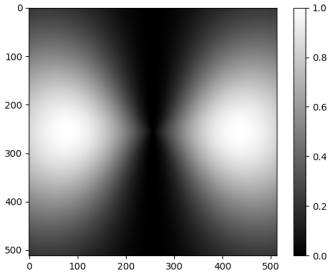
Each estimate of the angle from 0.001-0.008 degrees in the simulations under each of the noise conditions will be performed twenty times with different noise patterns. The results presented will be the mean and standard deviation of the angular estimates. Statistical



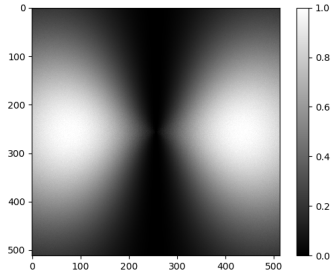
(a) Reference Intensity Distribution



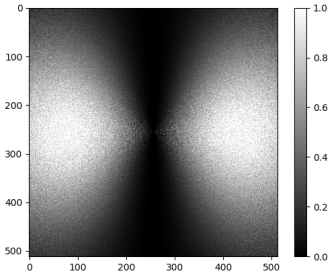
(b) SNR 80 dB



(c) SNR 60 dB



(d) SNR 40 dB



(e) SNR 20 dB

Figure 6.4: Examples of the intensity distribution after having gone through the CORD system and being contaminated with various noise amounts. The intenstiy distribution is measured in watts, showing values from 0 watts (black) to 1 watt (white) with the axes representing the locations of the spatial detector array measuring the intensities. (a) shows the reference intensity distribution without rotation (referred to here on as ref). (b) Ref with a SNR of 80 dB. (c) Ref with 60 dB. (d) Ref with 40 dB. (e) Ref with 20 dB.

significance tests will be performed for each angle at each SNR value for each reference type. Statistical significance will be tested with a two-tailed student t-test of unequal variance looking for a $p < 0.05$ value. The test is done using the Scipy package in Python [21].

To evaluate the accuracy of the angular estimates in the simulation, we will use a relative error metric between the average angle estimated and the actual angle:

$$e = \frac{|\hat{\phi} - \phi|}{\phi} \tag{6.39}$$

6.3.3 Real Experiments

In these experiments, we used a Avantes HAL light source, collimated into a 632.8 nm laserline filter (Thorlabs FL632.8-3), with a 633nm vortex half-wave retarder (Thorlabs WPV10L-633) to produce a SPB. The beam will pass through a sample and then a static Glan-Taylor (Thorlabs GL10) polarising prism (acting as the analyser). The intensity distribution is measured on a FLIR GS3-U3-23S6M-C CMOS Camera. The measurements will be reported in electrons (e^-) and the detector has a maximum well depth of 34078 e^-. The reference and sample measurement are captured with the same set of camera parameters in the FLIR FlyCap2 Application. The CORD and LPO method will each perform 10 estimates of optical activity and average the results. The reported values with be the average and the standard deviation.

As mentioned previously, we have two constraints for the complexity and size of our system. The system needed to have a complexity of zero ($C_T = 0$) and a size less than 587mm ($L_T = 587$mm). With the system shown in Figure 6.5, we do adhere to our design constraints and achieve a complexity of zero and a system size of of less than 587mm at roughly 300mm in length.

Experiment 1: Half-wave retarder plate

The first experiment will involve estimating the amount of angular rotation caused by an angled half-wave retarder plate (Thorlabs WPH10E-633). This experiment will determine the CORD systems capabilities of estimating the angular rotation under purposeful conditions. A reference beam will first be measured with the half-wave retarder plate fast axis aligned at zero degrees. Beams will then be measured with the half-axes aligned with the following angles (the resulting rotation from the rotator will be in brackets):

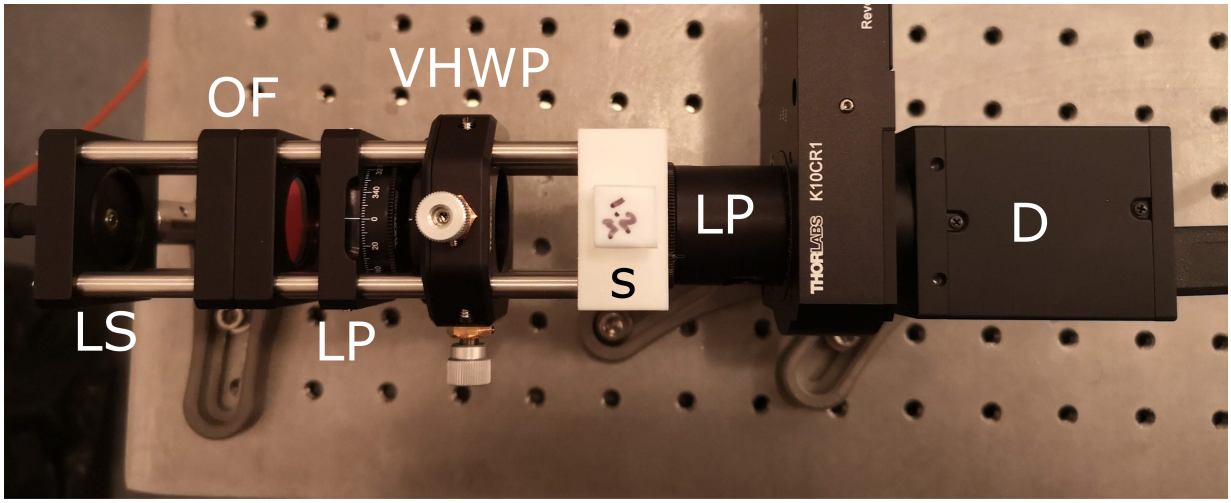


Figure 6.5: The laboratory setup for the CORD system. LS: Avantes HAL light source, OF: 632.8 nm laserline filter (Thorlabs FL632.8-3), VHWP: 633nm vortex half-wave retarder (Thorlabs WPV10L-633), LP: Glan-Taylor (Thorlabs GL10), and D: FLIR GS3-U3-23S6M-C. The Sample (S) is the 1.2M Sucrose solution. This system is approximately 300mm (one foot) in length.

1. 10° (20°)
2. 20° (40°)
3. 30° (60°)
4. 40° (80°)

Experiment 2: Sucrose Solution

The second experiment will be done with a cuvette of 1.2 M sucrose solution (Alfa Aesar, MA) in a quartz cuvette (Thorlabs CV10Q3500) with a 1 cm path length. The reference beam for this experiment will be a measured with a sample of deionized water in a similar quartz cuvette. The expected optical activity from this solution is 2.3433° [45].

The CORDS system will be compared against the linear polarisation orientation (LPO) scanning method for measuring ORD which is done by using a rotating analyser and determining the ORD by estimating the angle where peak intensity occurs. The LPO system will be mimicked using a Glan-Taylor polarising prism attached to a Thorlabs

K10CR1 motorised rotator. The prism will be rotated at a speed of $0.5°/s$ with the camera capturing at 30 frames per second thereby mimicking an angular resolution of 1 arc minute ($1/60^{\text{th}}$ of a degree) per frame. A single detector pixel location will be chosen from the frame and the intensity of that location through time will produce the measured signal for the LPO system. The pixel location for all frames was chosen to be the location with the highest value in the first frame captured.

6.4 Results and Analysis

6.4.1 Simulated Experiments

The results from the simulated experiments are summarised in Tables 6.2, 6.3, 6.4, 6.5, 6.6, 6.7,.

From that results in Table 6.2, 6.3, it can be seen that with a noise-free reference, the CORD system is capable of estimating the angle from optical activity to within a millidegree at an SNR of 80 dB, with similar performance at 60 and 40 dB. However, at 20 dB, the system loses accuracy. In addition, at no SNR is the system capable of estimating 0.001° accurately, having relative errors from 1677% to 23%. From the results in Table 6.4, 6.5, it can be seen that even with a reference image captured at the same SNR as the rotated intensity, the CORD system can estimate angles down to the millidegree for the SNR of 80, 60, and 40 dB. Similar to the noise-free reference case, the CORD algorithm does not perform well at 20 dB.

The results for both reference types show statistical significance for all the results at 80 and 60 dB and no statistical significance at 20 dB. This, combined with the results from the angular SNR tests show that at 80 and 60 dB, a 0.001° angular resolution is achievable. However, shown in Tables 6.6 and 6.7, statistical significance is not achieved between some angles and the one adjacent. In Table 6.6 statistical significance is maintained for all angles below 0.002 degrees, however, from 0.003 degrees onward, statistical significance is not maintained for all angles between them and the adjacent angle. This occurs with 0.003, 0.004, 0.006, and 0.007 degrees. And in Table 6.7 statistical significance does not occur between an angle and either of its adjacent angles for any of the angles simulated, save for 0.003 degrees. This leads us to conclude that at 40 dB SNR, the angular resolution drops from one to two millidegrees regardless of the type of reference captured. However, with a two millidegree level of resolution, we can detect differences of roughly 3 $\frac{\text{mmol}}{\text{L}}$. In comparison to systems that measure glucose in blood for diabetics, which report at 0.1 $\frac{\text{mmol}}{\text{L}}$ resolution in a range of 4 to 11 $\frac{\text{mmol}}{\text{L}}$, this system shows potential for that application.

From the results of this experiment, we have demonstrated that the CORD system, based off the Computational Polarimetry Framework, is capable of inferring the sample activity accurately to a at least two millidegrees with both an ideal reference intensity distribution and a reference distribution captured by the same detector under the same noise conditions. However, when the standard deviation of the noise is one-tenth of the signal, the system looses accuracy. Given the simulated performance, this system is capable of performing ORD on the level of accuracy seen in JASCO J-700/800 instruments (JASCO, OK)(surveyed in Castiglioni *et al.* [6]) of $0.002^{(\circ)}$.

In addition, it should be pointed out that for a circle inscribed in a 512×512 square, the highest angular resolution between two adjacent squares is $\tan^{-1}(1/256) = 0.2238^{\circ}$. With the results shown from the simulated experiments, the CORD system is capable of estimating angles to $1/100^{th}$ the accuracy of a strictly geometric method.

Table 6.2: Angular Estimates and Error for simulated experiments with a noise free reference. $\sigma_{\hat{\phi}}$ is the standard deviation of the angular estimations.

SNR (dB)	0.001°			0.002°			0.003°			0.004°		
	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$
80	1.3500e-1	1.6957e-1	86.500e1	1.9800	0.0000	9.9990e-1	2.9880	8.8182e-2	3.9990e-1	3.9600	7.7650e-16	9.9990e-1
60	3.2400e-1	3.8857e-1	6.7600e1	1.9980	1.9636e-1	9.9900e-2	2.9880	1.7452e-1	3.9990e-1	3.9960	1.6693e-1	9.9900e-2
40	1.2330	9.8306e-1	2.3300e1	2.1330	1.1862	6.6501	2.8170	1.3759	6.0999	4.0590	1.3452	1.4751
20	1.2303e1	8.1467	1.1303e3	1.2096e1	8.3872	5.0480e2	1.1853e1	8.6481	2.9510e2	1.1826e1	8.9312	1.9565e2

Table 6.3: Angular Estimates and Error for simulated experiments with a noise free reference. $\sigma_{\hat{\phi}}$ is the standard deviation of the angular estimations.

SNR (dB)	0.005°			0.006°			0.007°			0.008°		
	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$
80	5.0400	0.0000	8.0010e-1	5.9850	7.7942e-2	2.4990e-1	7.0110	3.9230e-2	1.5724e-1	7.9380	5.4000e-2	7.7490e-1
60	4.9860	1.7171e-1	2.7990e-1	5.9760	1.6693e-1	3.9990e-1	6.9840	1.6693e-1	2.2847e-1	7.9740	1.5167e-1	3.2490e-1
40	5.0400	1.5231	8.0010e-1	5.9580	1.6798	6.9990e-1	6.8580	1.8575	2.0285	8.2890	9.7012e-1	3.6126
20	1.1700e1	9.2526	1.3400e2	1.0989e1	9.8339	8.3150e1	1.1979e1	9.7489	7.1129e1	1.2312e1	9.8584	5.3900e1

Table 6.4: Angular Estimates and Error for simulated experiments with a noisy reference. $\sigma_{\hat{\phi}}$ is the standard deviation of the angular estimations.

SNR(dB)	0.001°			0.002°			0.003°			0.004°		
	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$
80	1.2600e-1	1.7171e-1	8.7400e1	1.9800	0.0000	9.9990e-1	2.9790	8.9549e-2	6.9990e-1	3.9600	7.7650e-16	9.9990e-1
60	2.5200e-1	4.3870e-1	7.4800e1	1.9620	2.1222e-1	1.8999	2.9970	1.9113e-1	9.9900e-2	3.9870	2.0739e-1	3.2490e-1
40	1.2600	1.4196	2.6000e1	2.1690	1.6155	8.4501	2.6280	1.5962	1.2400e1	4.0950	1.8566	2.3751
20	1.7775e1	1.2462e1	1.6775e3	1.7406e1	1.2523e1	7.7030e2	1.6965e1	1.2763e1	4.6550e2	1.6605e1	1.2920e1	3.1513e2

Table 6.5: Angular Estimates and Error for simulated experiments with a noisy reference. $\sigma_{\hat{\phi}}$ is the standard deviation of the angular estimations.

SNR (dB)	0.005°			0.006°			0.007°			0.008°		
	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$	$\hat{\phi} \times 10^{-3}$	$\sigma_{\hat{\phi}} \times 10^{-3}$	$e(\%)$
80	5.0400	0.0000	8.0010e-1	5.9760	7.2000e-2	3.9990e-1	7.0110	3.9230e-2	1.5724e-1	7.9470	6.4273e-2	6.6240e-1
60	4.9950	1.8771e-1	9.9900e-2	6.0030	2.1506e-1	5.0100e-2	6.9750	1.8771e-1	3.5704e-1	8.0100	1.8445e-1	1.2510e-1
40	5.2200	1.7618	4.4001	5.8320	2.2757	2.7999	7.2090	2.0762	2.9858e	8.1720	2.2463	2.1501
20	1.6380e1	1.3044e1	2.2760e2	1.5741e1	1.3484e1	1.6235e2	1.6137e1	1.3157e1	1.3053e2	1.6074e1	1.3267e1	1.0093e2

Table 6.6: Table of Statistical Significance for the noise-free reference case at 40 dB. . This table compares the distribution of the angular estimate of the rows by the distribution of the angular estimates of the columns and shows whether the distributions are statistically significant according to the two tailed student t-test of unequal variance.. $\times$ represents a comparisons of an angular distributions with its own distribution. * represents a statistical significance with $p < 0.05$. - represents $p \geq 0.05$. Statistical significance is maintained for all angles below 0.002 degrees, however, from 0.003 degrees onward, statistical significance is not maintained for all angles between them and the adjacent angle. This occurs with 0.003, 0.004, 0.006, and 0.007 degrees.

Angles (°)	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008
0.001	$\times$	*	*	*	*	*	*	*
0.002	*	$\times$	*	*	*	*	*	*
0.003	*	*	$\times$	-	*	*	*	*
0.004	*	*	-	$\times$	*	*	*	*
0.005	*	*	*	*	$\times$	*	*	*
0.006	*	*	*	*	*	$\times$	-	*
0.007	*	*	*	*	*	-	$\times$	-
0.008	*	*	*	*	*	*	-	$\times$

6.4.2 Real Experiments

The results from the real experiments are summarised in Tables 6.8, and 6.9.

Experiment 1: Half-wave retarder plate

Table 6.8 shows that, in comparison to the LPO method for ORD, the CORD system provides a more accurate measurement of the rotation caused by half-wave retarder with very small amounts of standard deviation around each estimate. However, while better in comparison to the LPO method, it does not compete with the accuracy of commercially available ORD systems like during the simulations.

Experiment 2: Sucrose Solution

In Figure 6.6, we show the measured and reference beams for this experiment. It is not readily visible the rotation caused by the Sucrose solution however, using the procedure

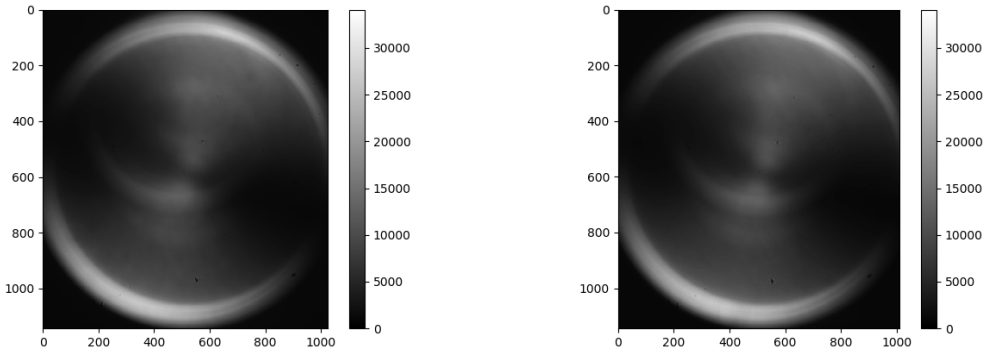
Table 6.7: Table of Statistical Significance for the noisy reference case at 40 dB. This table compares the distribution of the angular estimate of the rows by the distribution of the angular estimates of the columns and shows whether the distributions are statistically significant according to the two tailed student t-test of unequal variance. $\times$ represents a comparisons of an angular distributions with its own distribution. * represents a statistical significance with $p < 0.05$. - represents $p \geq 0.05$. While using a noisy reference intensity distribution, statistical significance does not occur between an angle and either of its adjacent angles for any of the angles simulated, save for 0.003 degrees.

Angles (°)	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008
0.001	$\times$	-	*	*	*	*	*	*
0.002	-	$\times$	-	*	*	*	*	*
0.003	*	-	$\times$	*	*	*	*	*
0.004	*	*	*	$\times$	-	*	*	*
0.005	*	*	*	-	$\times$	-	*	*
0.006	*	*	*	*	-	$\times$	-	*
0.007	*	*	*	*	*	-	$\times$	-
0.008	*	*	*	*	*	*	-	$\times$

outlined in Section 6.3.1, we were able to estimate a rotation from these two intensity distributions.

In Table 6.9, the CORD system estimates the optical activity as accurately of LPO, however the CORD system produces more consistent angular estimates of optical activity in comparison over the LPO system. Similar to the previous experiment, the CORD system is not as accurate as in the simulations, however, it is very stable estimating the optical activity. Both estimates of the sample result in a 0.1 degree error in the optical activity estimate. If we converted this back to Sucrose concentrations, this would result in a roughly 600 $\frac{\text{mmol}}{\text{L}}$ concentration error.

As a sanity check to see that the algorithm is estimating the angular rotation correctly, we show the angular estimation procedure using measurements from the 1.2M Sucrose solution using the measured and reference beams shown in Figure 6.6. In the formulation of our inference model, we assumed that the Zernike moments of the reference and rotated measurement distributions are the same, however in Figure 6.7 there is some difference. While the system is quite confident in its estimation of the optical activity from Table 6.9, it appears that the changes in the Zernike moment significantly affect the accuracy of the sucrose optical activity estimate.



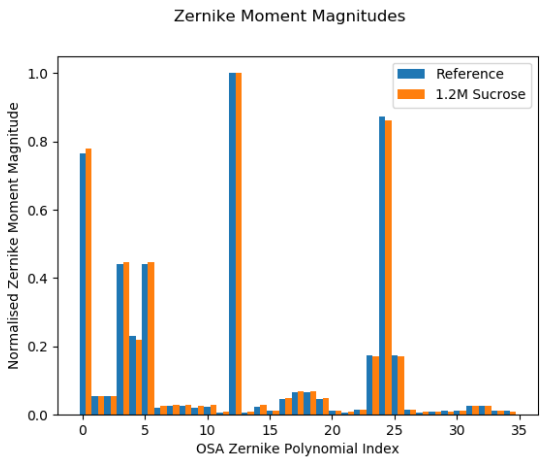
(a) Reference Intensity distribution captured with a cuvette of Deionised Water

(b) Intensity distribution captured with a cuvette of 1.2M Sucrose solution

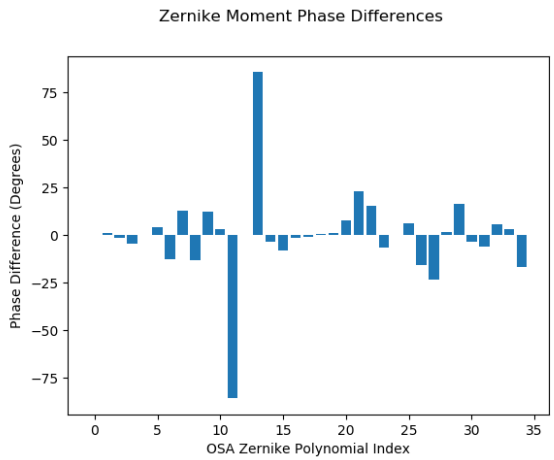
Figure 6.6: Intensity distributions captured reported in e^- for the purpose of estimating the optical activity of 1.2M Sucrose Solution. White represents the full well capacity of a detector in the array with black representing no electrons being measured. The axes represent the locations of the detectors on the array. It is not readily visible the rotation caused by the Sucrose. However, using the procedure outlined in Section 6.3.1, we were able to estimate the rotation from these two intensity distributions.

Table 6.8: Results of the CORD system in estimating the rotation caused by the half-wave rotator. Std is the standard deviation of that estimation process. Bold numbering belongs to the system that produced the better estimate. The CORD system produce a closer estimates of the polarisation in comparison to the LPO across all angles.

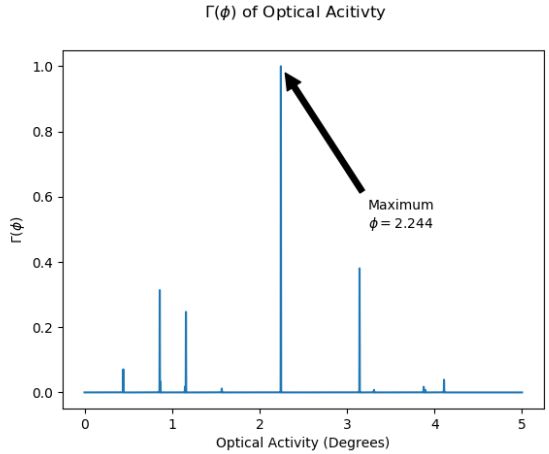
Angle(°)	CORD(°)	Std. of CORD(°)	LPO(°)	Std. of LPO(°)
20	**20.5468**	7.1910e-4	19.0617	7.2284e-2
40	**39.9623**	1.4962e-3	38.7867	5.2068e-2
60	**59.8619**	8.8932e-4	61.4950	1.0463e-1
80	**79.8080**	2.0425e-3	80.7817	1.6707e-1



(a) Zernike Moment Magnitude



(b) Zernike Moment Phase Difference



(c) $\Gamma(\phi)$ result

Figure 6.7: The Zernike moment magnitude and phase difference from the intensity distributions measured during the Sucrose experiment. (a) and(b) validate relation in Eq. 6.21 where the magnitude of the moments of two rotated intensity distributions are the same, while their phases are different. (c) shows the resulting $\Gamma(\phi)$ using Eq. 6.34 and with Eq. 6.36, we find the maximum being the rotation amount. The OSA Zernike polynomial index $= \frac{n(n+2)+m}{2}$.

Table 6.9: Results for the CORD system to estimate the optical activity of a sucrose solution. Std is the standard deviation of that estimation process. Bold numbering belongs to the system that produced the better estimate. Unlike the previous experiment, the CORD and LPO systems produce similar estimates, however, the CORD system produces more consistent estimates of the optical activity.

Solution	Optical Activity(°)	CORD(°)	Std of CORD(°)	LPO(°)	Std of LPO(°)
1.2 M Sucrose Solution	2.3433	**2.4416**	5.9181e-4	2.2417	1.4850

6.5 Discussion

In the CORD system, we demonstrated the use of the Computational Polarimetry Framework to infer the derived parameters of optical activity accurately through computational means, to a degree of accuracy better than the LPO ORD systems with only a single measurement. The framework accounts for prior knowledge of measurement characteristics, system characteristics, and was able to incorporate spatial relations between pixel intensities from the spatial detector array. This system was able to, during simulation, provide accurate estimates of optical activity on the same level of accuracy as commercial lab devices, however, in estimating the optical activity of the sucrose solution, it was not capable of the same level of accuracy.

Some general considerations with the CORD system and the chosen component configuration for the Computational Polarimetry Framework are discussed here in relation to optical activity estimation.

6.5.1 Reference Beam

The CORD system relies upon the prior knowledge of a continuously varying polarisation orientation over the azimuth of the beam to estimate the rotation caused by an optically active sample. This does not imply that the system relies upon the beam intensity profile matching that of Gori [14]. The choice of beam intensity profile is simply limited to having a an axially-symmetric non-zero value around a centroid. This implies that flat top intensities, Gaussian modes, or Bessel modes can be used in this system.

6.5.2 Dispersion

While demonstrating the ability for the system to estimate the optical activity from a transparent solution, it is assumed that the system will decompose the beam into its

spectral parts in order to create a full dispersion plot in a single capture. Using a diffraction grating, this is possible, however, in contrast to a typical diffractive spectroscopy setup, the CORD system requires a spatial array of detectors to correspond to each wavelength. To map the dispersion of the sample using this setup, such a system would have to:

1. Provide enough separation between two wavelengths such that each intensity distribution can be isolated and decomposed
2. Provide the same spectral spread for each intensity distribution
3. Be able to provide that separation over a reasonable detector array width while still accommodating the desired spectral range

The first and second consideration can be addressed through the choice of illumination source. The illumination source must have spectral lines with a full width half max that produces a spread of less than a detector on the detector array, or a consistent full-width half max across all spectral peaks. The third consideration can be addressed by splitting the beam into various diffraction gratings with the detectors at different positions in the 1st diffracted mode.

The design of systems to disperse the beam in accordance with the considerations mentioned can be formulated in the Computational Polarimetry Framework as requirements or constraints on the components or their parameters.

6.5.3 Heterogeneously Polarised Beams

For the purposes of the CORD system, we only discuss the use of a vortex half-wave retarder that is effective at 632.8 nm. This highlights one of the limitations of using this optic since it only works as designed at a particular wavelength. Generating broadband heterogeneously polarised beams is an active area of research demonstrating positive results [39]. The CORD system is a wavelength agnostic method for estimating the optical activity, allowing it to be used in conjunction with broadband heterogeneous beams.

In addition, the heterogeneously polarised beam discussed in this system is an azimuthally polarised SPB. It is worth noting that the polarisations in the heterogeneous beam need not be exclusively linear and trace a complete path around the equator of the Poincaré Sphere. This can be shown using the model described by Eq. 6.4 where instead of $\underline{S}$ being a linear polarisation, it is an arbitrary elliptical polarisation, the intensity distribution in Eq. 6.5 becomes

$$F = \left(\frac{1}{2} - \frac{1}{2}\left(S_1 \cos 2\left(\alpha + \phi\right) + S_2 \sin 2\left(\alpha + \phi\right)\right)\right) F_o \tag{6.40}$$

where S_1 and S_2 are the second and third Stokes parameters. Then differentiating with respect to θ, we can find that an extrema exists when

$$\theta = \frac{\arctan\left(\frac{S_2}{S_1}\right)}{2} - \phi \tag{6.41}$$

This will simplify to the case outlined in the Section 6.2.1 if $S_2 = 0$.

6.6 Summary

In this chapter, we have introduced, discussed, and demonstrated the use of the Computational Polarimetry Framework for the computational inference of optical activity in the novel CORD system. In simulation, the system produces estimates of optical activity as accurate as lab-based ORD estimation systems with potential for blood glucose monitoring. This system demonstrated being able to produce estimates of polarisation rotation caused by a half-wave retarder and a sample of sucrose that are similar to, if not better than, the LPO method for ORD using only a less complex, less costly and smaller system using a single measurement.

Chapter 7

Conclusion

With the Computational Polarimetry Framework presented in this thesis, we were capable of designing polarimetric systems for the inference of system component parameters and sample parameters under design requirements and constraints. In this chapter, we will discuss future directions and applications for this work.

7.1 Future Framework Developments

7.1.1 Unknown Parameter Models

In the two design problems we have discussed, we had all the terms in the inference model defined, however, in some cases, the distribution for an node in the network maybe be unknown. We can utilise the inherent Markov Blanket property of Bayesian networks (this is a result of d-separation and the local Markov property) [36], to estimate the distribution of that node. The Markov blanket property states that for any node X_i there will be some subset $\mathcal{S}$ that shields X_i from the influence of all the other nodes in the graph.

We can show this with an example where if we have a DAG $D = (\mathcal{U}, \mathcal{E})$ where $U = \{X_1, X_2, X_3, X_4, X_5\}$, shown in Figure 7.1, we can state that for node X_3, the Markov blanket is the entire set $\mathcal{U}$. With this, and the outlined dependencies in the graph, we can formulate the joint probability of D as

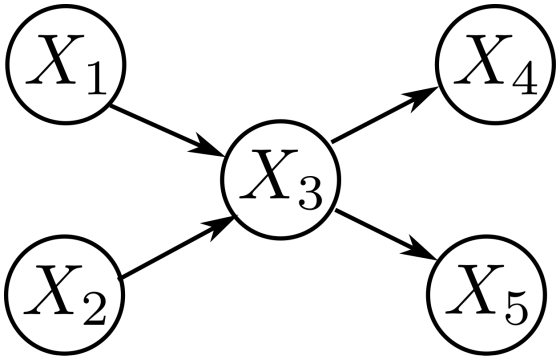


Figure 7.1: Bayesian graph of a DAG in a Markov Blanket.

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4, X_5 = x_5) = \\ P(x_1)P(x_2)P(x_3|x_2, x_1)P(x_4|x_3)P(x_5|x_3) \quad (7.1)$$

the distribution of $P(x_3|x_2, x_1)$ is unknown, however, should the rest of the terms be known, the distribution can be inferred [22]. This would be a likely scenario in situations where black-box polarimetry or imaging components are used in the full polarimetry system where the distribution of parameters in those component are unknown but need to be estimated for sample parameter inference.

7.1.2 Component Choice

During the design problems proposed, we have had the basic structure of the polarimetric system given with only the parameters of the components needing to be estimated. This is common practice with other polarimetric system optimisation [1, 11, 15, 40, 41, 53, 56, 57]. In some design scenarios however, there are requirements and constraints on the choice of components and these components are chosen from a discrete set, rather than being pre-chosen.

With there being flexibility in the component choice, the component choices can change the possible system parameter configurations. For example, a PSA configuration consisting of a dual rotating retarder has two sets of system parameters that concern the angular orientation and retardance of the components, while a PSA configuration of a fixed polariser and rotating retarder has two different sets of system parameters. The demonstration of

this framework to be used with different component configurations with design requirements and constraints will be an important next step in demonstrating this frameworks effectiveness as a complete and unifying tool for polarimetric system design and polarimetric system parameter inference.

7.2 Future Applications of the Framework

7.2.1 Whole Sample Parameter Inference

In this thesis, we proposed two design problems that focused on the estimation of polarisation angle and sample optical acitivity. However, samples can have many polarising properties that can affect the polarisation of light in more ways than rotating its orientation [13]. These sample properties are summarised in a Mueller matrix of that sample making the estimation of a Mueller matrix important.

From Section 2.2.2, we know that to estimate a Mueller matrix, several measurements are needed to estimate the Mueller matrix of a sample and the alignment of those measurements are important for estimation accuracy [46]. Using the Computational Polarimetry Framework, a measurement alignment constraint can be put into place and components and their parameters can be chosen to ensure alignment. In addition, the Computational Polarimetry Framework can be also used to evaluate computational algorithms for beam alignment given a particular component setup.

7.2.2 General Optical System Design

The Computational Polarimetry Framework is designed for the system design of polarimetric systems, however, there are possibilities for this framework to be used imaging, and interferometric applications. The components in the system and their parameters can be generalised into using optical diffraction or ray tracing models and have them be included into the model's $\mathcal{C}$ and $\mathcal{Q}$ set parameters. With this change and prior models of a sample's optical properties, inference models can be set up to infer sample properties and design optical systems under certain requirements and constraints across different system types.

7.3 Conclusion

In conclusion, we have demonstrated a computational polarimetry system design framework that utilises Bayesian networks to infer system properties for optimal system design and create inference model for sample properties. The framework is the first of its kind to formulate polarimetry system design as an all encompassing probabilistic framework that incorporate system design requirements and constraints.

The framework's effectiveness was demonstrated in inferring the ideal system parameter to design a polarimetric system to determine the angle of light polarisation that satisfy design requirements and address component limitations. In this case, the framework had to optimise system parameters for not only stochastic error from the detector, but also from deterministic error from the optomechanical components. In addition, it ensured that the linear polariser would produce a detectable change in intensity for a small angular change. This was all due to the model being able to incorporate prior system knowledge about component behaviour into the inference of system parameters. It was also able to arrive at expected component configurations under ideal conditions.

We then demonstrated the framework's ability to infer sample optical activity with a spatial detector array and an SPB. In this case, we demonstrated that the unifying system design framework is flexible enough to compute system parameters utilising novel polarimetric measurement methods and still produce accurate estimations of sample parameters over the state of art methods using fewer measurements and exploiting prior knowledge of system components.

The Framework has the potential to be improved upon and extended for whole sample inference applications, and general optical system designs and with it change the way in which scientists design polarimetry systems.

References

[1] Amrit Ambirajan. Optimum angles for a polarimeter: part II. *Optical Engineering*, 34(6):1656, jun 1995.

[2] Guillaume Anna and François Goudail. Optimal mueller matrix estimation in the presence of poisson shot noise. *Optics Express*, 20(19):21331, sep 2012.

[3] RMA Azzam, IM Elminyawi, and AM El-Saba. General analysis and optimization of the four-detector photopolarimeter. *JOSA A*, 5(5):681–689, 1988.

[4] Thomas G. Brown and Amber M. Beckley. Stress engineering and the applications of inhomogeneously polarized optical fields. *Frontiers of Optoelectronics*, 6(1):89–96, jan 2013.

[5] Brent D. Cameron, Harshal W. Gorde, Bhavana Satheesan, and Gerard L. Cote. The use of polarized laser light through the eye for noninvasive glucose monitoring. *Diabetes Technology & Therapeutics*, 1(2):135–143, jun 1999.

[6] Ettore Castiglioni, Sergio Abbate, and Giovanna Longhi. Experimental methods for measuring optical rotatory dispersion: Survey and outlook. *Chirality*, 23(9):711–716, aug 2011.

[7] Gerard L. Cote. Multispectral polarimetric glucose detection using a single pockels cell. *Optical Engineering*, 33(8):2746, aug 1994.

[8] Jun Dai, François Goudail, Matthieu Boffety, and Jun Gao. Estimation precision of full polarimetric parameters in the presence of additive and poisson noise. *Optics Express*, 26(26):34081, dec 2018.

[9] Juan Carlos González de Sande, Gemma Piquero, and Massimo Santarsiero. Polarimetry with azimuthally polarized light. *Optics Communications*, 410:961–965, mar 2018.

[10] Juan Carlos González de Sande, Massimo Santarsiero, and Gemma Piquero. Spirally polarized beams for polarimetry measurements of deterministic and homogeneous samples. *Optics and Lasers in Engineering*, 91:97–105, apr 2017.

[11] Hui Dong, Ming Tang, and Yandong Gong. Noise properties of uniformly-rotating RRFP stokes polarimeters. *Optics Express*, 21(8):9674, apr 2013.

[12] Paul Fieguth. *Statistical Image Processing and Multidimensional Modeling*. Springer New York, 2011.

[13] Dennis H Goldstein. *Polarized Light, revised and expanded*. CRC Press, third edition, 2003.

[14] Franco Gori. Polarization basis for vortex beams. *Journal of the Optical Society of America A*, 18(7):1612, jul 2001.

[15] François Goudail. Equalized estimation of stokes parameters in the presence of poisson noise for any number of polarization analysis states. *Optics Letters*, 41(24):5772, dec 2016.

[16] S. A. Haider, C. Scharfenberger, F. Kazemzadeh, A. Wong, and D. A. Clausi. Multi-polarimetric textural distinctiveness for outdoor robotic saliency detection. In Juha Röning and David Casasent, editors, *Intelligent Robots and Computer Vision XXXII: Algorithms and Techniques*. SPIE, feb 2015.

[17] Shahid A. Haider, Megan Y. Tran, and Alexander Wong. Computational circular dichroism estimation for point-of-care diagnostics via vortex half-wave retarders. In Gerard L. Coté, editor, *Optical Diagnostics and Sensing XVIII: Toward Point-of-Care Diagnostics*. SPIE, feb 2018.

[18] J. W. Hovenier. Structure of a general pure mueller matrix. *Applied Optics*, 33(36):8318, dec 1994.

[19] Steven L. Jacques, Jessica C. Ramella-Roman, and Ken Lee. Imaging skin pathology with polarized light. *Journal of Biomedical Optics*, 7(3):329, 2002.

[20] Steven L. Jacques, Ravikant Samatham, Scott Isenhath, and Ken Lee. Polarized light camera to guide surgical excision of skin cancers. In Nikiforos Kollias, Bernard Choi, Haishan Zeng, Reza S. Malek, Brian J. Wong, Justus F. R. Ilgner, Kenton W. Gregory, Guillermo J. Tearney, Henry Hirschberg, and Steen J. Madsen, editors, *Photonic Therapeutics and Diagnostics IV*. SPIE, feb 2008.

[21] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–. [Online; accessed April 2019].

[22] Dinu Kaufmann, Sonali Parbhoo, Aleksander Wieczorek, Sebastian Keller, David Adametz, and Volker Roth. Bayesian markov blanket estimation. In *Artificial Intelligence and Statistics*, pages 333–341, 2016.

[23] Whoi-Yul Kim and Young-Sung Kim. Robust rotation angle estimator. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8):768–773, 1999.

[24] W. J. Lai, B. C. Lim, P. B. Phua, K. S. Tiaw, H. H. Teo, and M. H. Hong. Generation of radially polarized beam with a segmented spiral varying retarder. *Optics Express*, 16(20):15694, sep 2008.

[25] Steven J. Lascos and Daniel T. Cassidy. Optical phase and intensity modulation from a rotating optical flat: effect on noise in degree of polarization measurements. *Applied Optics*, 48(9):1697, mar 2009.

[26] D. Layden, M. F. G. Wood, and I. A. Vitkin. Optimum selection of input polarization states in determining the sample mueller matrix: a dual photoelastic polarimeter approach. *Optics Express*, 20(18):20466, aug 2012.

[27] A. Lizana, I. Estévez, A. Turpin, C. Ramirez, A. Peinado, and J. Campos. Implementation and performance of an in-line incomplete stokes polarimeter based on a single biaxial crystal. *Applied Optics*, 54(29):8758, oct 2015.

[28] Daniel C. Louie, Lioudmila Tchvialeva, Sunil Kalia, Harvey Lui, and Tim K. Lee. One-shot stokes polarimetry for low-cost skin cancer detection. In David Levitz and Aydogan Ozcan, editors, *Optics and Biophotonics in Low-Resource Settings V*. SPIE, feb 2019.

[29] Scott C. McEldowney, David M. Shemo, and Russell A. Chipman. Vortex retarders produced from photo-aligned liquid crystal polymers. *Optics Express*, 16(10):7295, may 2008.

[30] Ole J Mengshoel. Designing resource-bounded reasoners using bayesian networks: System health monitoring and diagnosis. In *Proc. of the 18th International Workshop on Principles of Diagnosis (DX-07)*, page 330, May 2007.

[31] Michael I Mishchenko and Larry D Travis. Satellite retrieval of aerosol properties over the ocean using polarization as well as intensity of reflected sunlight. *Journal of Geophysical Research: Atmospheres*, 102(D14):16989–17013, 1997.

[32] Tingkui Mu, Donghao Bao, Chunmin Zhang, Zeyu Chen, and Jionghui Song. Optimal reference polarization states for the calibration of general stokes polarimeters in the presence of noise. *Optics Communications*, 418:120–128, jul 2018.

[33] Rafael Navarro, Justo Arines, and Ricardo Rivera. Direct and inverse discrete zernike transform. *Optics Express*, 17(26):24269, dec 2009.

[34] Thien An Nguyen, Yisa Rumala, Giovanni Milione, Daniel A. Nolan, Ebrahim Karimi, Sergei Slussarenko, Lorenzo Marrucci, and Robert R. Alfano. Incoherent polarized white-light vecctor vortex from a q-plate. In David L. Andrews, Enrique J. Galvez, and Jesper Glückstad, editors, *Complex Light and Optical Forces VIII*. SPIE, feb 2014.

[35] Nicolas Passilly, François Treussart, Rolland Hierle, Renaud de Saint Denis, Kamel Aït-Ameur, and Jean-François Roch. Simple interferometric technique for generation of a radially polarized light beam. *Journal of the Optical Society of America A*, 22(5):984, may 2005.

[36] Judea Pearl. Bayesian networks: A model of self-activated memory for evidential reasoning. In *Proceedings of the 7th conference of the cognitive science society*, 1985.

[37] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference (Representation and Reasoning)*. Morgan Kaufmann Pub, 1988.

[38] Angelo Pierangelo, Abdelali Benali, Maria-Rosaria Antonelli, Tatiana Novikova, Pierre Validire, Brice Gayet, and Antonello De Martino. Ex-vivo characterization of human colon cancer by mueller polarimetric imaging. *Optics Express*, 19(2):1582, jan 2011.

[39] N. Radwell, R. D. Hawley, J. B. Götte, and S. Franke-Arnold. Achromatic vector vortex beams from a glass cone. *Nature Communications*, 7(1), feb 2016.

[40] Stéphane Roussel, Matthieu Boffety, and François Goudail. Polarimetric precision of micropolarizer grid-based camera in the presence of additive and poisson shot noise. *Optics Express*, 26(23):29968, oct 2018.

[41] D. S. Sabatke, M. R. Descour, E. L. Dereniak, W. C. Sweatt, S. A. Kemme, and G. S. Phipps. Optimization of retardance for a complete stokes polarimeter. *Optics Letters*, 25(11):802, jun 2000.

[42] Derek S. Sabatke, Ann M. Locke, Michael R. Descour, William C. Sweatt, John P. Garcia, Eustace L. Dereniak, Shanalyn A. Kemme, and Gary S. Phipps. Figures of

merit for complete stokes polarimeter optimization. In David B. Chenault, Michael J. Duggin, Walter G. Egan, and Dennis H. Goldstein, editors, *Polarization Analysis, Measurement, and Remote Sensing III*. SPIE, nov 2000.

[43] Bahaa E. A. Saleh and Malvin Carl Teich. *Fundamentals of Photonics (Wiley Series in Pure and Applied Optics)*. John Wiley & Sons, 1991.

[44] J. C. G. De Sande, G. Piquero, M. Santarsiero, Roma Tre, and Via V. Volterra. Polarimetry with non uniformly polarized beams. In *Trends in Electromagnetic Coherence, 2nd Joensuu Conference on Coherence and Random Polarization*, 2018.

[45] Sigma-Aldrich. Sucrose. Online; accessed April 2019.

[46] Matthew H. Smith, Jacob B. Woodruff, and James D. Howe. Beam wander considerations in imaging polarimetry. In Dennis H. Goldstein and David B. Chenault, editors, *Polarization: Measurement, Analysis, and Remote Sensing II*. SPIE, oct 1999.

[47] JB Theeten and DE Aspnes. Ellipsometry in thin film analysis. *Annual Review of Materials Science*, 11(1):97–122, 1981.

[48] Vimal Thilak, Charles D. Creusere, and David G. Voelz. Passive polarimetric imagery based material classification for remote sensing applications. In *2008 IEEE Southwest Symposium on Image Analysis and Interpretation*. IEEE, mar 2008.

[49] Pavel Tománek, Jan Mikláš, Hamed Mohamed Abubaker, and Lubomír Grmela. Optical sensing of polarization states changes in meat due to the ageing. In *AIP Conference Proceedings*, volume 1288, pages 127–131. AIP, 2010.

[50] Hakki Mert Torun, Madhavan Swaminathan, Anto Kavungal Davis, and Mohamed Lamine Faycal Bellaredj. A global bayesian optimization algorithm and its application to integrated system design. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 26(4):792–802, apr 2018.

[51] Santosh Tripathi and Kimani C. Toussaint. Rapid mueller matrix polarimetry based on parallelized polarization state generation and detection. *Optics Express*, 17(24):21396, nov 2009.

[52] Valery V. Tuchin. Polarized light interaction with tissues. *Journal of Biomedical Optics*, 21(7):071114, apr 2016.

[53] K. M. Twietmeyer and R. A. Chipman. Optimization of mueller matrix polarimeters in the presence of error sources. *Optics Express*, 16(15):11589, jul 2008.

[54] J. S. Tyo. Optimum linear combination strategy for an n-channel polarization-sensitive imaging or vision system. *Journal of the Optical Society of America A*, 15(2):359, feb 1998.

[55] J. Scott Tyo. Considerations in polarimeter design. In David B. Chenault, Michael J. Duggin, Walter G. Egan, and Dennis H. Goldstein, editors, *Polarization Analysis, Measurement, and Remote Sensing III*. SPIE, nov 2000.

[56] J. Scott Tyo. Noise equalization in stokes parameter images obtained by use of variable-retardance polarimeters. *Optics Letters*, 25(16):1198, aug 2000.

[57] J. Scott Tyo. Design of optimal polarimeters: maximization of signal-to-noise ratio and minimization of systematic error. *Applied Optics*, 41(4):619, feb 2002.

[58] David J. Wales and Jonathan P. K. Doye. Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms. *The Journal of Physical Chemistry A*, 101(28):5111–5116, jul 1997.

[59] L.B. Wolff. Polarization-based material classification from specular reflection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(11):1059–1071, 1990.

[60] Xianyu Wu, Mark Pankow, Hsiao-Ying Shadow Huang, and Kara Peters. High-speed polarized light microscopy for in situ, dynamic measurement of birefringence properties. *Measurement Science and Technology*, 29(1):015203, dec 2017.

[61] J Zallat, Aïnouz S, and M Ph Stoll. Optimal configurations for imaging polarimeters: impact of image noise and systematic errors. *Journal of Optics A: Pure and Applied Optics*, 8(9):807–814, jul 2006.

[62] Qiwen Zhan. Cylindrical vector beams: from mathematical concepts to applications. *Advances in Optics and Photonics*, 1(1):1, jan 2009.

[63] Ciyou Zhu, Richard H. Byrd, Peihuang Lu, and Jorge Nocedal. Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization. *ACM Trans. Math. Softw.*, 23(4):550–560, December 1997.