



Michigan Technological University  
*Create the Future* Digital Commons @ Michigan Tech

---

Dissertations, Master's Theses and Master's  
Reports - Open

Dissertations, Master's Theses and Master's  
Reports

---

2009

## Estimation of the degree of polarization through computational sensing

Wei Wang  
*Michigan Technological University*

Follow this and additional works at: <https://digitalcommons.mtu.edu/etds>



Part of the [Electrical and Computer Engineering Commons](#)

Copyright 2009 Wei Wang

---

### Recommended Citation

Wang, Wei, "Estimation of the degree of polarization through computational sensing", Dissertation, Michigan Technological University, 2009.  
<https://doi.org/10.37099/mtu.dc.etds/59>

Follow this and additional works at: <https://digitalcommons.mtu.edu/etds>



Part of the [Electrical and Computer Engineering Commons](#)

ESTIMATION OF THE DEGREE OF POLARIZATION THROUGH  
COMPUTATIONAL SENSING

By

WEI WANG

A DISSERTATION

Submitted in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

(Electrical Engineering)

MICHIGAN TECHNOLOGICAL UNIVERSITY

2009

© 2009 Wei Wang



This dissertation, "Estimation of the degree of polarization through computational sensing", is hereby approved in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY in the field of Electrical Engineering.

DEPARTMENT:  
Electrical and Computer Engineering

Signatures:

Dissertation Advisor \_\_\_\_\_  
Dr. Timothy Schulz

Department Chair \_\_\_\_\_  
Dr. Daniel Fuhrmann

Date \_\_\_\_\_



## **Dedication**

To My Parents, Family and Friends.



# Contents

<b>List of Figures</b>	<b>xii</b>
<b>Acknowledgments</b>	<b>xiii</b>
<b>Abstract</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Background</b>	<b>5</b>
2.1 The Speckle Effect in Coherent Imaging . . . . .	5
2.1.1 Mathematical Model . . . . .	7
2.1.2 Measurement Statistical Models . . . . .	9



2.1.2.1	Four-channel Polarimeter Intensities . . . . .	9
2.1.2.2	Orthogonal Intensities . . . . .	10
2.1.2.3	Total Intensity . . . . .	10
2.1.3	Circular Complex Gaussian Random Variables . . . . .	11
2.2	Estimation Theory . . . . .	12
2.2.1	The Maximum-likelihood principle . . . . .	12
2.2.2	Expectation-Maximization (EM) Algorithm . . . . .	14
2.2.3	The Cramér-Rao Lower Bound . . . . .	15
2.3	Detector Noise Model . . . . .	16
<b>3</b>	<b>Sensing Modalities</b>	<b>19</b>
3.1	Fully Polarimetric Sensing . . . . .	20
3.1.1	Stokes' Procedure . . . . .	21
3.1.2	Collett's Procedure . . . . .	22

3.1.3	Division-of-Amplitude Photopolarimetry . . . . .	24
3.1.4	Comments . . . . .	25
3.2	Orthogonal Intensities Sensing . . . . .	26
3.3	Total Intensity Sensing . . . . .	27
<b>4</b>	<b>Cramer-Rao Lower Bounds</b>	<b>29</b>
4.1	Four-Channel Polarimetric Sensing . . . . .	30
4.1.1	Speckle Only . . . . .	34
4.1.2	Gaussian Noise . . . . .	35
4.1.3	Poisson Noise . . . . .	36
4.1.4	Combined Gaussian and Poisson Noise . . . . .	38
4.1.5	Plot of Bounds . . . . .	40
4.2	Orthogonal Intensities Sensing . . . . .	41
4.2.1	Speckle Only . . . . .	42

4.2.2	Gaussian Noise . . . . .	43
4.2.3	Poisson Noise . . . . .	44
4.2.4	Combined Gaussian and Poisson Noise . . . . .	45
4.2.5	Plot of Bounds . . . . .	47
4.3	Total Intensity Sensing . . . . .	48
4.3.1	Speckle Only . . . . .	49
4.3.2	Gaussian Noise . . . . .	50
4.3.3	Poisson Noise . . . . .	51
4.3.4	Combined Gaussian and Poisson Noise . . . . .	53
4.3.5	Plot of Bounds . . . . .	55
4.4	Discussion . . . . .	56
4.4.1	Comparison Of Bounds . . . . .	56
4.4.2	Adjusting Cramer-Rao bound for Bias . . . . .	59

4.4.3	Signal-to-noise Ratio . . . . .	60
4.4.4	Summary . . . . .	62
<b>5</b>	<b>Maximum-Likelihood Estimation from Total Intensity Data</b>	<b>65</b>
5.1	EM algorithms for Total Intensity Sensing . . . . .	67
5.1.1	Laser Speckle Only . . . . .	69
5.1.2	Laser Speckle with Gaussian Noise . . . . .	70
5.1.3	Laser Speckle with Poisson Noise . . . . .	71
5.1.4	Laser Speckle with Combined Gaussian and Poisson Noise . . . . .	74
5.1.5	Summary . . . . .	75
5.2	Simulation . . . . .	76
<b>6</b>	<b>Estimation from Orthogonal Intensity Data</b>	<b>79</b>
6.1	Algebraic Estimator . . . . .	80
6.2	Maximum-Likelihood Estimator . . . . .	81

6.2.1	Expectation-Maximization Algorithm . . . . .	83
6.2.1.1	Speckle Only . . . . .	85
6.2.1.2	Speckle with Gaussian Noise . . . . .	85
6.2.1.3	Speckle with Poisson Noise . . . . .	87
6.3	Simulation . . . . .	87
<b>7</b>	<b>Conclusion</b>	<b>91</b>
<b>A</b>	<b>Joint PDF of the instantaneous Stokes parameters</b>	<b>95</b>
<b>B</b>	<b>Derivation of EM Algorithms</b>	<b>99</b>
B.1	Estimation from Orthogonal Intensities . . . . .	99
B.1.1	Laser Speckle Only . . . . .	99
B.1.2	Gaussian Noise . . . . .	105
B.1.3	Poisson Noise . . . . .	111
B.2	Estimation from Total Intensity . . . . .	114

B.2.1	Laser Speckle Only . . . . .	114
B.2.2	Laser Speckle and Gaussian Noise . . . . .	116
B.2.3	Laser Speckle and Poisson Noise . . . . .	119
B.2.4	Laser Speckle Combined with Gaussian Noise and Poisson Noise .	122

<b>Bibliography</b>	<b>125</b>
---------------------	------------



# List of Figures

1.1	Degree or polarization image. . . . .	2
1.2	Determining surface orientations of transparent objects based on polarization degrees . . . . .	3
2.1	Laser speckle on a digital camera image from a green laser pointer . . . . .	6
3.1	General Sensing Modality . . . . .	20
3.2	Schematic diagram of Stokes's procedure . . . . .	21
3.3	Schematic diagram of Collett's procedure . . . . .	23
3.4	Division-of-amplitude photopolarimeter (DOAP) . . . . .	25



4.1	Diagram of the four-channel polarimeter. The incident beam is split by a 80/20 polarizing splitter into two beams of orthogonal polarization, s-polarized and p-polarized. 80% p-polarized and 20% s-polarized go to one direction while the rest go to the other direction, followed by a $1/4$ wave plate at $45^\circ$ and a $1/2$ wave plate at $22.5^\circ$ in each leg. The two beams are split again by two polarizing splitters into 4 beams, which go to 4 intensity detectors to register the intensity data. . . . .	30
4.2	Cramer-Rao bounds on the root-mean-square (RMS) estimation error for unbiased estimators of the degree of polarization from four-channel polarimeter intensity measurements with various noise model; for Gaussian noise case $\bar{I} = 100$ , $\sigma = 30$ ; for Poisson noise case, $\bar{I} = 15/\text{photon counts}$ , $a = 1$ , $b = 0$ ; for Gaussian and Poisson noise case, $\bar{I} = 50/\text{photon counts}$ , $\sigma = 10$ , $a = 1$ , $b = 0$ . . . . .	40
4.3	Cramer-Rao bounds on the root-mean-square (RMS) estimation error for unbiased estimators of the degree of polarization from two orthogonal intensities measurements with various noise model; for Gaussian noise case $\bar{I} = 100$ , $\sigma = 30$ ; for Poisson noise case, $\bar{I} = 15/\text{photon counts}$ , $a = 1$ , $b = 0$ ; for Gaussian and Poisson noise case, $\bar{I} = 50/\text{photon counts}$ , $\sigma = 10$ , $a = 1$ , $b = 0$ . . . . .	48

4.4	Cramer-Rao bounds on the root-mean-square (RMS) estimation error for unbiased estimators of the degree of polarization from total intensity measurements with various noise model; for Gaussian noise case $\bar{I} = 100$ , $\sigma = 30$ ; for Poisson noise case, $\bar{I} = 15$ /photon counts, $a = 1$ , $b = 0$ ; for Gaussian and Poisson noise case, $\bar{I} = 50$ /photon counts, $\sigma = 10$ , $a = 1$ , $b = 0$ . . . . .	56
4.5	Comparison of normalized Cramer-Rao bounds on RMS estimation error for unbiased estimators of the DOP. (a) detector noise free; (b) Gaussian read out noise. $\bar{I} = 100, \sigma = 30$ ; (c) Poisson shot noise. $\bar{I} = 15$ ; (d) combined Gaussian and Poisson noise. $\bar{I} = 50, \sigma = 10$ . . . . .	57
4.6	CRLBs for both unbiased and biased estimators for total intensity data with Poisson noise. Average Intensity $\bar{I} = 2$ , $K = 500$ speckle realizations. . . . .	59
4.7	Normalized Cramer-Rao bounds on RMS error for estimators of the degree of polarization from measurements corrupted by Gaussian and Poisson noise with the same SNRs; $\bar{I} = 15$ and $\sigma^2 = 15$ . . . . .	62
4.8	Performance bounds with different SNRs for total intensity measurements with Gaussian noise. . . . .	62
5.1	Diagram of a computational sensor for total intensity data . . . . .	66

5.2	Comparison of the performances of ML estimators and algebraic estimators with corresponding modified Cramer-Rao bounds on RMS estimation error for biased estimators of the DOP. (a) detector noise free. 100 speckle realizations; (b) Gaussian read out noise. $\bar{I} = 100$ , $\sigma = 30$ , 100 speckle realizations; (c) Poisson shot noise. $\bar{I} = 15$ , 100 speckle realizations; (d) combined Gaussian and Poisson noise. $\bar{I} = 50$ , $\sigma = 10$ , 100 speckle realizations. . . . .	78
6.1	Schematic diagram of orthogonal intensities sensing . . . . .	80
6.2	Comparison of the performances of ML estimator and algebraic estimators with corresponding modified Cramer-Rao bounds on RMS estimation error for biased estimators of the DOP. (a) detector noise free. 100 speckle realizations; (b) Gaussian read out noise. $\bar{I} = 100$ , $\sigma = 30$ , 100 speckle realizations; (c) Poisson shot noise. $\bar{I} = 15$ , 100 speckle realizations; (d) combined Gaussian and Poisson noise. $\bar{I} = 50$ , $\sigma = 10$ , 100 speckle realizations. . . . .	88

## Acknowledgments

It has been an exceptional journey at Michigan Tech since August, 2005. My research career began with Dr. Timothy Schulz. I am deeply indebted to Dr. Schulz for his advice, mentorship, research support, and encouragement throughout my master's and doctoral studies. I also greatly appreciate his patience and tolerance of my slowness and numerous mistakes.

I am very fortunate to have the opportunity to work with a group of exceptional people in the department. I have enjoyed every course I took, and every moment I had worked with them. I appreciate all the friendships and their collaborations to finish this dissertation. I want especially to thank Kerkil Choi for sharing his wisdom, and to thank Dr. Michael Roggeman, Dr. David Volz, Dr. Warren Perger, Dr. Allen Struthers and Dr. Zhi Tian for their research ideas and advice.

Finally, it is impossible to finish my doctoral degree without my parents' love and support. This dissertation is dedicated to them.

And to my parents, family, friends and everybody else, thank you!



## Abstract

The degree of polarization of a reflected field from active laser illumination can be used for object identification and classification. The goal of this study is to investigate methods for estimating the degree of polarization for reflected fields with active laser illumination, which involves the measurement and processing of two orthogonal field components (complex amplitudes), two orthogonal intensity components, and the total field intensity. We propose to replace interferometric optical apparatuses with a computational approach for estimating the degree of polarization from two orthogonal intensity data and total intensity data. Cramer-Rao bounds for each of the three sensing modalities with various noise models are computed. Algebraic estimators and maximum-likelihood (ML) estimators are proposed. Active-set algorithm and expectation-maximization (EM) algorithm are used to compute ML estimates. The performances of the estimators are compared with each other and with their corresponding Cramer-Rao bounds. Estimators for four-channel polarimeter (intensity interferometer) sensing have a better performance than orthogonal intensities estimators and total intensity estimators. Processing the four intensities data from polarimeter, however, requires complicated optical devices, alignment, and four CCD detectors. It only requires one or two detectors and a computer to process orthogonal intensities data and total intensity data, and the bounds and estimator performances demonstrate that reasonable estimates may still be obtained from orthogonal intensities or total intensity data. Computational sensing is a promising way to estimate the degree of polarization.

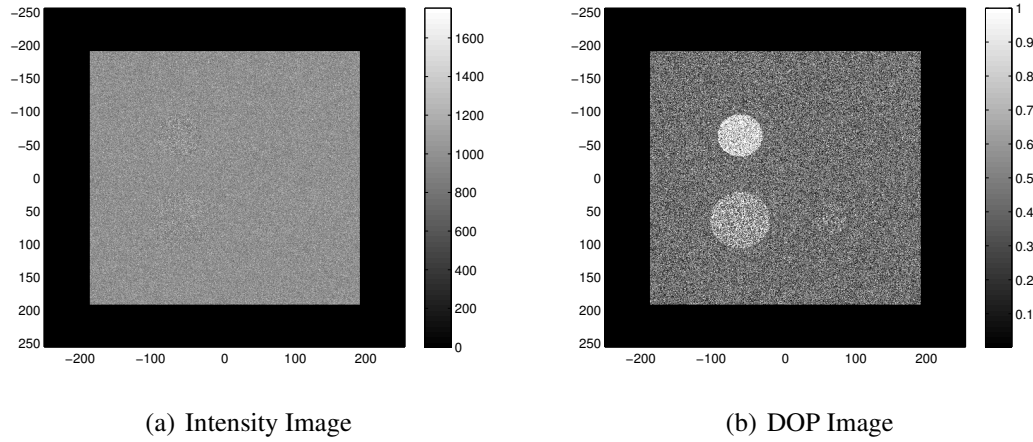


# Chapter 1

## Introduction

The degree of polarization (DOP) is a quantity that characterizes the degree to which an electromagnetic wave is polarized[1]. A linearly polarized wave has a DOP of 1, whereas an unpolarized wave (polarization direction is totally random and unpredictable, such as natural light) has a DOP of 0. The degree of polarization of the back scattered field has been used for target detection and object identification by several authors [2, 3, 4]. The concept is to measure the speckle intensities and use the intensity statistics to estimate the degree of polarization, which has been shown to discriminate between different target materials and find the surface orientations. A high degree of polarization (DOP) typically implies a single reflection from a surface (for example, metal target), whereas a low DOP indicates multiple scattering or surface penetration process (such as dielectric target) [5, 6]. Measurements at multiple angles (for example, binocular stereo and motion analysis techniques) or at



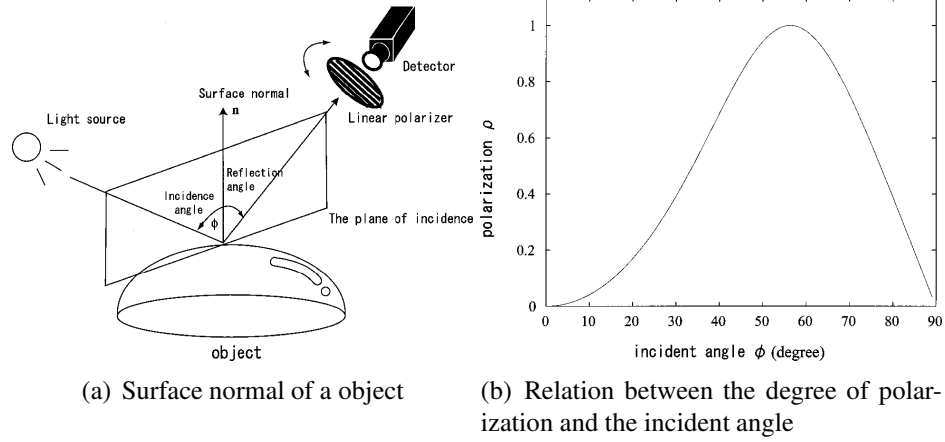


**Figure 1.1:** Degree or polarization image.

the same position under different illumination conditions (photometric stereo method) can also provide surface orientation information [7, 8, 9]. A further advantage of using these techniques is that they provide night vision capability and improve image resolution for a given aperture because laser wavelengths are shorter than infrared (IR) [10].

Fig. 1.1 shows a simple example. The intensity image, which suggests a homogeneous scene with no features, is the averaged total intensity of a series of independent speckle realizations. However, if we look at the degree of polarization image, we can see the features of the scene. Fig.1.2 gives an example for application on the determination of surface orientation [9]. The degree of polarization is a function of the refractive index  $n$  and the incident angle  $\phi$ . Thus by obtaining the degree of polarization from the data, we can determine the incident angle  $\phi$  given the refractive index  $n$

Speckle is a random intensity pattern caused by the mutual interference of many back-scattered waves [1]. When a rough surface is illuminated by quasimonochromatic light,



**Figure 1.2:** Determining surface orientations of transparent objects based on polarization degrees

and the surface roughness varies on the scale of the illumination wavelength, the reflected path-length differences induce random variation in the phases of the scattered waves. As a result, the intensity of the scattered light varies in a random way. If the source-scene-sensor geometry is perturbed slightly as data are acquired, the relative microscopic structure will be changed and independent speckle realizations can be acquired. If the scene is homogenous, independent speckle realizations can also be obtained by spatial sampling instead of temporal sampling. By using different sensing modalities, a sequence of data from independent speckle realizations can be collected to form an estimate of the degree of polarization at each spatial location.

We studied three common sensing modalities: 1) full polarimetric (Stokes parameters) sensing, 2) orthogonal intensities sensing, and 3) total intensity sensing. Chapter 2 gives a background on mathematical models and theories we use. Chapter 3 gives a review of some common techniques for the three sensing modalities. For each sensing modality, per-

formance bounds (Cramer-Rao bounds on root-mean-square error) for different detector noise models are provided in Chapter 4. Chapter 5 and Chapter 6 give a detailed presentation of the proposed computational methods to estimate degree of polarization from total intensity and orthogonal intensities.

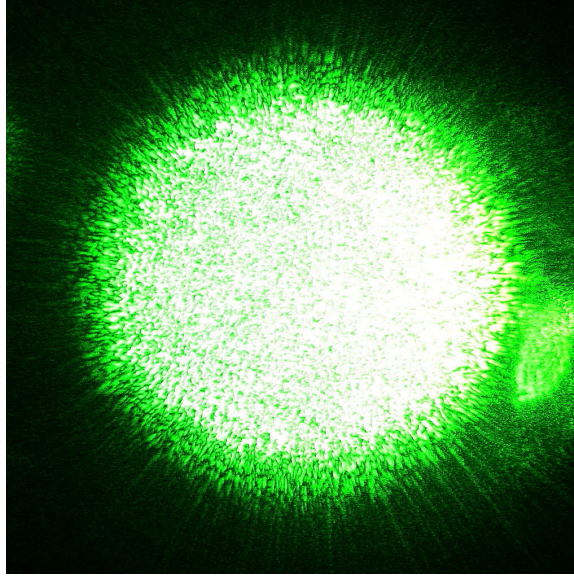
# **Chapter 2**

## **Background**

### **2.1 The Speckle Effect in Coherent Imaging**

When images of complex objects are formed by use of the highly coherent light produced by a laser, an important image characteristic soon becomes apparent. If the object consists of surfaces that are rough on the scale of an optical wavelength (as most objects are), the image is found to be grainy, with a multitude of bright and dark spots that bear no apparent relationship to the macroscopic scattering properties of the object. These random patterns are known as "speckle". A typical example is the random pattern created when a laser beam is scattered off a rough surface. See Fig. 2.1, an image of laser speckle [11].

The speckle effect is a result of the interference of many waves having different phases



**Figure 2.1:** Laser speckle on a digital camera image from a green laser pointer

that add together to give a resultant wave whose amplitude and intensity vary randomly. If the complex amplitudes of each wave are modeled by a vector, then it can be seen that if a number of vectors with random phases are added together, then the length of the resulting vector can vary from zero to the sum of the individual vector lengths - this effect is sometimes known as a random walk. According to the Central Limit Theorem, if the phases of the individual scattered contributions from the object are approximately uniformly distributed over  $(-\pi, \pi)$ , the field associated with any single linear polarization component of the image must be a circular complex Gaussian random variable [1].

### 2.1.1 Mathematical Model

The field at a particular spatial location in the focal plane of an imaging lens is an image of the field that is reflected from a corresponding spatial location on the object or scene. We denote the complex amplitude of this field as

$$U = \begin{bmatrix} U_x \\ U_y \end{bmatrix}, \quad (2.1)$$

where  $U_x$  and  $U_y$  represent the complex amplitudes for two orthogonal components of the field. Whereas the amplitude and phase of this field is influenced in a deterministic manner by the macroscopic characteristics of the scene, the field will also contain a random component that is influenced by the scene's microscopic structure (roughness variations on the scale of the illumination wavelength). This randomness in the field is commonly referred to as coherent laser speckle [1]. It is common practice to model the random fluctuations due to the coherent laser speckle effect as a Gaussian random process [1].

The coherency matrix is defined as:

$$J = E \left[ UU^\dagger \right] = \begin{bmatrix} J_{xx} & J_{xy} \\ J_{xy}^* & J_{yy} \end{bmatrix}, \quad (2.2)$$

where, for our applications, the expectation  $E[*]$  refers to the ensemble of speckle realizations, and  $^\dagger$  is the complex conjugate transpose operator. Because  $J$  is a Hermitian matrix,

there exists a unitary transformation  $P$  such that

$$PJP^\dagger = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad (2.3)$$

where  $\lambda_1$  and  $\lambda_2$  are the nonnegative eigenvalues of  $J$ . The unitary transforming matrix can be written as:

$$P = \begin{bmatrix} 1 & 0 \\ 0 & \exp(-j\phi_{xy}) \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}, \quad (2.4)$$

where  $\theta$  represents a coordinate rotation and  $\phi_{xy}$  represents a relative phase shift between the orthogonal field components.

The degree of polarization (DOP) is defined as

$$\text{DOP} = \sqrt{1 - 4 \frac{\det(J)}{[\text{tr}(J)]^2}} = \frac{|\lambda_1 - \lambda_2|}{\lambda_1 + \lambda_2}, \quad (2.5)$$

where  $\det(J) = J_{xx}J_{yy} - |J_{xy}|^2$  and  $\text{tr}(J) = J_{xx} + J_{yy}$  are the determinant and trace of the coherency matrix, respectively.

## 2.1.2 Measurement Statistical Models

### 2.1.2.1 Four-channel Polarimeter Intensities

Based on the discussion in Chapter 1, if the phases of the individual scattered contributions from the object are approximately uniformly distributed over  $[-\pi, \pi]$  (i.e. if the object is rough on the scale of a wavelength), the field associated with any single linear polarization component is a circular complex Gaussian random variable [1]; therefore, the probability density function (PDF) for the field's complex amplitudes is

$$p_U(u|J) = \frac{1}{\pi^2 \det(J)} \exp(-u^\dagger J^{-1} u), \quad (2.6)$$

where  $^\dagger$  denotes complex conjugate transpose,  $\det(J) = J_{xx}J_{yy} - |J_{xy}|^2$  is the determinant of the coherency matrix, and  $J^{-1}$  is its inverse. This PDF can be also written as

$$p_U(u|J) = \frac{1}{\pi^2 \det(J)} \exp \left[ -\text{tr} \left( J^{-1} u u^\dagger \right) \right]. \quad (2.7)$$

In practice, we can not measure the field directly; instead, optical interferometric methods are used to sense field data through intensity measurements. In this paper, we use a four-channel polarimeter as an example of an interferometric method. The statistical model and more details are discussed in section 4.1



### 2.1.2.2 Orthogonal Intensities

The instantaneous intensities for the two orthogonal components are defined as

$$I_x = |U_x|^2 \quad (2.8)$$

$$I_y = |U_y|^2, \quad (2.9)$$

and the joint PDF for the orthogonal intensities is [12]

$$p_{I_x, I_y}(i_x, i_y) = \frac{1}{\det(J)} \exp\left(-\frac{J_{yy}i_x + J_{xx}i_y}{\det(J)}\right) I_0\left(2\frac{|J_{xy}|}{\det(J)}\sqrt{i_x i_y}\right), \quad (2.10)$$

where  $I_0$  is the zero-order modified Bessel function of the first kind. Note that the joint PDF for orthogonal intensities depends only on  $J_{xx}$ ,  $J_{yy}$  and the magnitude of  $J_{xy}$ . Consequently, the joint PDF does not depend on the relative phase shift  $\phi_{xy}$ .

### 2.1.2.3 Total Intensity

The instantaneous total intensity of the field is defined as:

$$I = \|U\|^2 = I_x + I_y, \quad (2.11)$$

and the PDF for the total intensity is [1]

$$p_I(i) = \frac{1}{\lambda_1 - \lambda_2} \left[ \exp\left(-\frac{i}{\lambda_1}\right) - \exp\left(-\frac{i}{\lambda_2}\right) \right]. \quad (2.12)$$

where[13]

$$\lambda_1 = \frac{J_{xx} + J_{yy}}{2} \left[ 1 + \sqrt{1 - 4 \frac{\det(J)}{[\text{tr}(J)]^2}} \right], \quad (2.13a)$$

$$\lambda_2 = \frac{J_{xx} + J_{yy}}{2} \left[ 1 - \sqrt{1 - 4 \frac{\det(J)}{[\text{tr}(J)]^2}} \right]. \quad (2.13b)$$

Therefore, the PDF for the total intensity depends only on the eigenvalues for the coherency matrix, and does not depend on the coordinate rotation angle  $\theta$  or the phase shift  $\phi_{xy}$ .

### 2.1.3 Circular Complex Gaussian Random Variables

An important property of circular complex Gaussian random variables is the *complex Gaussian moment theorem*. Let  $U_1, U_2, \dots, U_{2k}$  be zero-mean jointly circular complex Gaussian random variables. Then

$$\overline{u_1^* \cdots u_k^* u_{k+1} \cdots u_{2k}} = \sum_{\pi} \overline{u_1^* u_p} \overline{u_2^* u_q} \cdots \overline{u_k^* u_r} \quad (2.14)$$

where  $\sum_{\pi}$  denotes a summation over the  $k!$  possible permutations  $(p, q, \dots, r)$  of  $(1, 2, \dots, k)$ .

For the simplest case of  $k = 2$ , we have

$$\overline{u_1^* u_2^* u_3 \cdots u_4} = \overline{u_1^* u_3} \overline{u_2^* u_4} + \overline{u_1^* u_4} \overline{u_2^* u_3} \quad (2.15)$$

## 2.2 Estimation Theory

Estimation is the process of making decisions over a set of parameters. There are two major types of estimators: the Neyman-Pearson approach, in which no prior probability is assumed on the parameters; and the Bayes approach, in which a prior probability is assumed. When no prior distribution is assumed, the estimation is commonly based upon the principle of *maximum likelihood (ML)*. On the other hand, when a prior distribution for the parameter is assumed, a *maximum a posteriori (MAP)* estimate is formed

### 2.2.1 The Maximum-likelihood principle

The essential principle of maximum likelihood (ML) is that it requires one to choose a value, as an estimate of a parameter, for which the probability of obtaining the observed sample is as large as possible. That is, having obtained the observations, one looks back and computes the probability from the point of view of one about to perform the experi-

ment from which the given sample values will be observed. This probability will in general depend on the parameter, which is then given the value for which this probability is maximized.

Suppose that the random variable  $X$  has a probability distribution that depends on a parameter  $\theta \in \Theta$ . Let  $f_X(x|\theta)$  denote either the probability mass function (PMF) or the probability density function (PDF) of  $X$ . We assume that the form of  $f_X$  is known, but not the value of the parameter  $\theta$ . The Joint PMF of  $m$  independent random variables evaluated at the sample points  $x_1, \dots, x_m$  is

$$\ell(\theta, x_1, \dots, x_m) = \ell(\theta, \mathbf{x}) = f_{\mathbf{X}}(\mathbf{x}|\theta) = \prod_{i=1}^m f_X(x_i|\theta) \quad (2.16)$$

This function is also known as the *likelihood function* of the sample. The principle of maximum likelihood requires us to choose an estimate of the unknown parameter that maximizes the likelihood function.

$$\hat{\theta}_{ML} = \arg \max_{\theta} \ell(\theta, \mathbf{x}) \quad (2.17)$$

It is actually more convenient, for many applications, to consider the logarithm of the likelihood function, which is defined as

$$\Lambda(\theta, \mathbf{x}) = \log f_{\mathbf{X}}(\mathbf{x}|\theta) \quad (2.18)$$

and called the *log-likelihood function*. Then, the maximum-likelihood estimate is

$$\left. \frac{\partial \Lambda(\theta, \mathbf{x})}{\partial \theta} \right|_{\theta = \hat{\theta}_{ML}} = 0 \quad (2.19)$$

### 2.2.2 Expectation-Maximization (EM) Algorithm

In many cases, it is impossible to directly access to the data necessary to make the maximum-likelihood (ML) estimation of parameters, or sometimes, some of the data are missing. The expectation-maximization (EM) algorithm was discovered to solve problems like this. It produces ML estimates of parameters when there is a many-to-one mapping from a distribution to the underlying distribution governing the observation. The observation is usually referred as "incomplete data" and the underlying data as the "complete data". The EM algorithm consists of two primary steps: an expectation step and a maximization step. The expectation is obtained with respect to the unknown underlying variables, using the current estimates of the parameters and conditioned upon the observations. The maximization step then provides a new estimate of the parameters. The two steps are iterated until convergence. At every iteration of the algorithm, a value of parameter is computed so that the likelihood function increase until a local maximum is achieved [14].

### 2.2.3 The Cramér-Rao Lower Bound

The Cramer-Rao bound (CRB) or Cramer-Rao lower bound (CRLB) expresses a lower bound on the variance of estimators of a deterministic parameter[15]. In its simplest form, the bound states that the value of any unbiased estimator is at least as high as the inverse of the Fisher information. An unbiased estimator that achieves this lower bound is said to be efficient. The Cramer-Rao bound can also be used to bound the variance of biased estimators. In some cases, a biased estimator can have both its variance and mean squared error below the unbiased Cramer-Rao lower bound.

Suppose  $\theta$  is an unknown deterministic parameter that is to be estimated from measurements  $x$ , distributed according to some probability density function  $f(x|\theta)$ . The variance of any unbiased estimator  $\hat{\theta}$  of  $\theta$  is then bounded by the inverse of the Fisher information  $F(\theta)$ :

$$\text{var}(\hat{\theta}) \geq \frac{1}{F(\theta)} \quad (2.20)$$

where the Fisher information is defined by

$$F(\theta) = \text{E} \left[ \left( \frac{\partial \ell(x|\theta)}{\partial \theta} \right)^2 \right] = \text{E} \left[ \frac{\partial^2 \ell(x|\theta)}{\partial^2 \theta} \right] \quad (2.21)$$

For multivariate case, define a parameter column vector

$$\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_d]^T \quad (2.22)$$

with PDF  $f(x|\boldsymbol{\theta})$ . The Fisher information matrix is a  $d \times d$  matrix with element  $F_{m,k}$  defined as

$$F_{m,k} = \mathbb{E} \left[ \frac{\partial}{\partial \theta_m} \log f(x|\boldsymbol{\theta}) \frac{\partial}{\partial \theta_k} \log f(x|\boldsymbol{\theta}) \right]. \quad (2.23)$$

Let  $p$  be a function of the parameters,  $p = \psi(\boldsymbol{\theta})$ . The Cramer-Rao bound for estimating  $p$  becomes:

$$\text{var}(\hat{p}) \geq \frac{\partial \psi(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}^T [F(\boldsymbol{\theta})]^{-1} \frac{\partial \psi(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \quad (2.24)$$

where  $\partial \psi(\boldsymbol{\theta})/\partial \boldsymbol{\theta}$  is a vector whose element is  $\partial \psi(\boldsymbol{\theta})/\partial \theta_j$ .

## 2.3 Detector Noise Model

Optical detectors convert radiative energy into an electrical signal [16]. The incident photons interact with the sensor material and are converted into an electrical charge or current. Data are read out of the device by converting the charge packets into a voltage. Such transformation, the read-out process, is corrupted by zero mean Gaussian read-out noise. Because photon-electron generation itself is a random process with a Poisson statistic, Poisson

noise, or shot noise, also exists in the device. Meanwhile, a detector can generate a signal even if the incident radiative flux is zero, and this signal is called *dark current*, denoted by  $b$ , which brings a bias to the image. If we denote the combined *gain* of the output amplifier and the camera circuitry by  $a$ , then the output (intensity measurement) of each collection site is given by [17]

$$D = \text{Poisson}(aI + b) + N \quad (2.25)$$

where  $N \sim N(0, \sigma^2)$  is a Gaussian random variable with mean zero and variance  $\sigma^2$ . For a more detailed overview on noise in CCD camera, please refer to [16, 17]





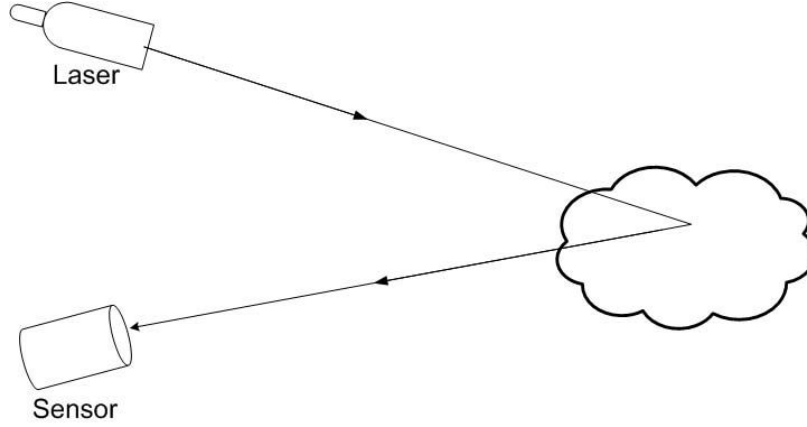
# Chapter 3

## Sensing Modalities

The general sensing scenario is depicted in Fig.3.1. We illuminate the scene by a monochromatic laser, and measure the degree of polarization of the backscattered field with some types of sensors. The degree of polarization (DOP) of a quasi-monochromatic light beam at a point is the ratio of the average intensity of the completely polarized portion of the radiation to the average total intensity. It has the following relation with Stokes parameters [18]:

$$\text{DOP} = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0}. \quad (3.1)$$

Therefore, one way is to measure the Stokes parameters first and use the Eq.3.1 to calculate the DOP, which is usually done by fully polarimetric sensing methods. Full polarimetric sensing needs complicated optical apparatus, and any misalignment, bad calibration or



**Figure 3.1:** General Sensing Modality

defect of the optic components will bring errors. Instead, Emil Wolf [19] proposed the possibility of estimating the degree of polarization from the correlation of the orthogonal intensities, which is similar to intensity interferometry. Alternatively, these two modalities can be replaced by simple total intensity sensing, which requires only one camera and a computer to process the data. However, the formation of an estimate of the degree of polarization from the total intensity depends on the fluctuation of the total intensity of a series of independent speckle frames.

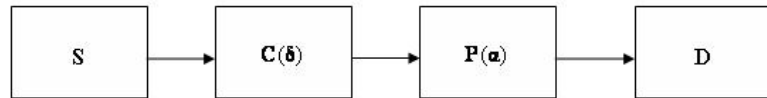
### 3.1 Fully Polarimetric Sensing

In general, optical setups to measure the state of polarization (Stokes parameters) of light can be divided into roughly three groups. The first group is based on the use of two optical elements, such as linear polarizer and a quarter-wave plate, successively placed in four different positions. In the second group, the polarization of light is modulated by rotating

elements [18, 20]. The third group uses parallel treatment of the incoming light. The Following are three examples in detail.

### 3.1.1 Stokes' Procedure

In the Stokes scheme, Fig. 3.2 [18], the measurement of the state of polarization of light is performed by rotating a retarding wave plate and a polarizer to four pairs of angular orientations and recording the four transmittances through the polarizer, as shown in Fig.3.2.  $\mathbf{S}$  is the input light source,  $\mathbf{C}(\delta)$  is the compensator,  $\mathbf{P}(\alpha)$  is the linear polarizer, and  $D$  is the detector.



**Figure 3.2:** Schematic diagram of Stokes's procedure

We denote by  $S_j(\alpha, \delta)$  the averaged intensity of light, at the output of the linear polarizer with its electric field vector in a direction at an angle  $\alpha$  to the x-axis. The Stokes parameters

are given by Eq(3.2)

$$\langle S_0 \rangle = \langle S_0(0,0) \rangle + \left\langle S_0 \left( \frac{\pi}{2}, 0 \right) \right\rangle \quad (3.2a)$$

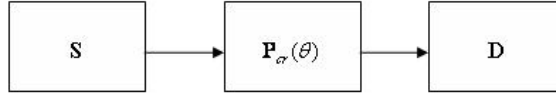
$$\langle S_1 \rangle = \langle S_0(0,0) \rangle - \left\langle S_0 \left( \frac{\pi}{2}, 0 \right) \right\rangle \quad (3.2b)$$

$$\langle S_2 \rangle = \left\langle S_0 \left( \frac{\pi}{4}, 0 \right) \right\rangle - \left\langle S_0 \left( \frac{3\pi}{4}, 0 \right) \right\rangle \quad (3.2c)$$

$$\langle S_3 \rangle = \left\langle S_0 \left( \frac{\pi}{4}, \frac{\pi}{2} \right) \right\rangle - \left\langle S_0 \left( \frac{3\pi}{4}, \frac{\pi}{2} \right) \right\rangle \quad (3.2d)$$

### 3.1.2 Collett's Procedure

Collett's Procedure, Fig.3.3 [18] simplifies the Stokes procedure by use only a single circular polarizer, constructed from a quarter-wave retardation plate and a linear polarizer whose transmission axis is at  $45^\circ$  with respect to the horizontal axis, placed in a rotatable mechanical mount. The Stokes parameters are measured at selected values of the rotation angle. The circular polarizer is then flipped  $180^\circ$ , and the Stokes parameters are again measured at selected values of the rotation angle, as shown in Fig 3.3. **S** is the input light source, **P<sub>cr</sub>** is the rotatable right circular polarizer, and **D** is the detector.



**Figure 3.3:** Schematic diagram of Collett's procedure

The Mueller matrix for the rotating right circular polarizer is expressed as

$$\mathbf{P}_{\text{cr}}(\theta) = \mathbf{R}(-\theta)\mathbf{C}(\mathbf{0})\mathbf{P}\left(\frac{\pi}{4}\right)\mathbf{R}(\theta)$$

where  $\mathbf{R}(\theta)$  is the Mueller matrix of a rotator corresponding to a rotation of an angle  $\theta$ .

The output-averaged intensity emerging from the circular polarizer is

$$\langle S_0(\theta) \rangle_c = \frac{1}{2} (\langle S_0 \rangle - \sin(2\theta) \langle S_1 \rangle + \langle S_2 \rangle \cos(2\theta))$$

Similarly, the output averaged intensity emerging from the linear polarizer (flipped to its linear side) is

$$\langle S_0(\theta) \rangle_l = \frac{1}{2} (\langle S_0 \rangle + \langle S_3 \rangle)$$

Therefore, the formulas for the Stokes parameters are

$$\langle S_0 \rangle = \langle S_0(0) \rangle_c + \left\langle S_0\left(\frac{\pi}{2}\right) \right\rangle_c \quad (3.3a)$$

$$\langle S_1 \rangle = \langle S_0 \rangle - 2 \left\langle S_0\left(\frac{\pi}{4}\right) \right\rangle_c \quad (3.3b)$$

$$\langle S_2 \rangle = \langle S_0(0) \rangle_c - \left\langle S_0\left(\frac{\pi}{2}\right) \right\rangle_c \quad (3.3c)$$

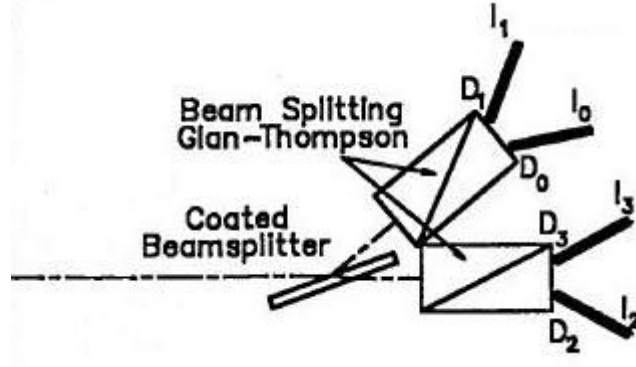
$$\langle S_3 \rangle = 2 \langle S_0(0) \rangle_l + \langle S_0 \rangle \quad (3.3d)$$

### 3.1.3 Division-of-Amplitude Photopolarimetry

Fig.3.4 is a schematic diagram of the Division-of-Amplitude Photopolarimetry [21, 22].

It is an arrangement of four detectors,  $D_0, D_1, D_2$  and  $D_3$ . The key element is a beam-splitter that divides an incident beam into reflected beam  $r$  and a transmitted beam  $t$ . The intensities in two orthogonal transverse directions of the reflected and transmitted beams are measured. The unknown Stokes parameters under measurement have the following relation to the intensities of the four beams

$$\mathbf{S} = \mathbf{F}^{-1} \mathbf{I}, \quad (3.4)$$



**Figure 3.4:** Division-of-amplitude photopolarimeter (DOAP)

The elements  $f_{ij}$  of  $\mathbf{F}$  are given by:

$$f_{1j} = r_{1j} + r_{2j} \cos 2A_1 + r_{3j} \sin 2A_1$$

$$f_{2j} = r_{1j} - r_{2j} \cos 2A_1 - r_{3j} \sin 2A_1$$

$$f_{3j} = t_{1j} + t_{2j} \cos 2A_2 + t_{3j} \sin 2A_2$$

$$f_{4j} = t_{1j} - t_{2j} \cos 2A_2 - t_{3j} \sin 2A_2$$

where  $\mathbf{R} = (r_{ij})$  is the reflection Mueller matrix and  $\mathbf{T} = (t_{ij})$  is the transmission Mueller matrix.  $A_1$  and  $A_2$  are the azimuth angles of the last two BS.

### 3.1.4 Comments

The errors associated with the Stokes' Procedure and Collett's Procedure may be due to the misalignment of the compensator, the polarizer, the rotating device and the internal prop-



erties of the birefringent material. The arrangements of both methods make it impossible to measure all four Stokes parameters simultaneously. The advantages of DOAP include the fact that DOAP is capable of time-resolved simultaneous measurements of the four stokes parameters, no moving parts and no modulation. Disadvantages of DOAP are the calibration procedure which may be somewhat cumbersome and the 4 by 4 matrix may be singular. (Azzam and his coworkers have shown that the condition that the matrix is nonsingular requires the plane of incidence to be rotated between successive reflections by other than  $\frac{\pi}{2}$ , where  $\pm\frac{\pi}{4}$  and  $\pm\frac{3\pi}{4}$  are the optimum rotations [23])

## 3.2 Orthogonal Intensities Sensing

To replace the polarizing or compensating devices that conventional optical fully polarimetric sensing methods require, Emil Wolf [19] investigated the theoretical possibility that the degree of polarization can be determined from the measurements of the correlation in the output of two photoelectric detectors illuminated by two orthogonal portions of the beam.

Instead of estimating the degree of polarization, investigators sometimes measure a similar quantity, which is defined as [10, 24, 25]

$$R = \frac{|I_x - I_y|}{I_x + I_y}. \quad (3.5)$$

The scene is illuminated with a single elliptical polarization state laser light. The light backscattered by the scene is analyzed in the polarization states parallel and orthogonal to the incident one. Two images thus obtained:  $I_x$  for the intensity of the light in the parallel state and  $I_y$  for the intensity of the light in the orthogonal state. The total intensity backscattered by the scene is equal to the sum of these two images. However, this  $R$  is *not* the degree of polarization unless the coherency matrix is diagonal [26].

Based on Wolf's work, we propose a computational way that algebraic and maximum-likelihood estimators are used. The details are given in Chapter 6.

### 3.3 Total Intensity Sensing

The total intensity of a polarized random electromagnetic field is generally fluctuating. Having both the speckle model and detector noise model, we find the first and second moments of the total intensity data. The mean of the data is

$$E[D] = a(\lambda_1 + \lambda_2) + b = a\bar{I} + b, \quad (3.6)$$

and the variance of the data is

$$\text{VAR}[D] = a^2(\lambda_1^2 + \lambda_2^2) + a(\lambda_1 + \lambda_2) + b + \sigma^2 \quad (3.7)$$

$$= \frac{a^2 \bar{I}^2}{2}(1 + \text{DOP}^2) + a\bar{I} + b + \sigma^2. \quad (3.8)$$

Then, we have the relation between the DOP and the mean and variance of the data

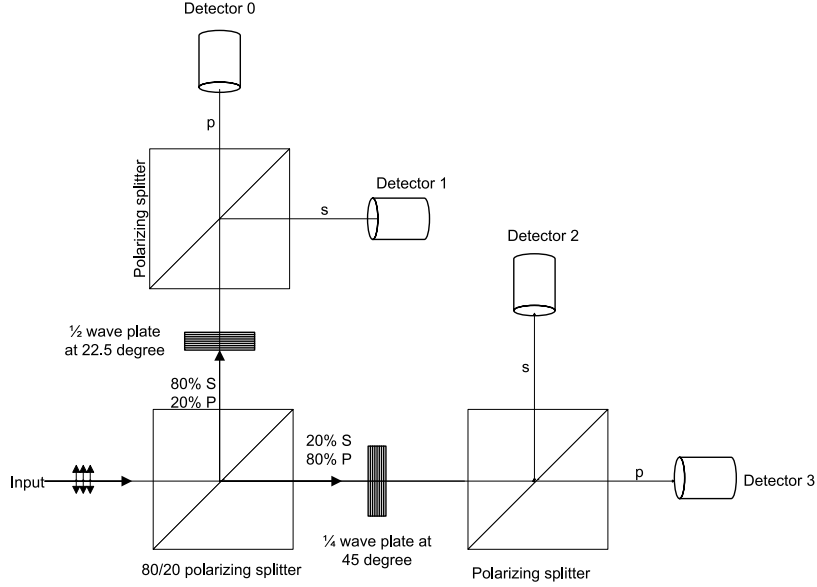
$$\text{DOP} = \sqrt{\frac{2\text{VAR}[D]}{a^2 \bar{I}^2} - \frac{2}{a\bar{I}} - \frac{2(b + \sigma^2)}{a^2 \bar{I}^2}} - 1. \quad (3.9)$$

Therefore, we have a simple *algebraic estimator*, which use the sample mean and sample variance to estimate the degree of polarization [27]. Alternatively, one can find the estimate using maximum-likelihood method, which is investigated in Chapter 5.

# Chapter 4

## Cramer-Rao Lower Bounds

The problem of estimating the degree of polarization from active laser illumination is subject to both speckle randomness and measurement noise. Because the practical estimation algorithms for these tasks are rarely optimal, a real concern is determining "the best that can be done", and comparing a given algorithm's performance with that estimation error lower bound to see if the estimation algorithm is adequate. The Cramer-Rao lower bound (CRLB) is powerful in such performance assessment. In this chapter, we compute the Cramer-Rao bounds on the estimation of the degree of polarization for each sensing modality.



**Figure 4.1:** Diagram of the four-channel polarimeter. The incident beam is split by a 80/20 polarizing splitter into two beams of orthogonal polarization, s-polarized and p-polarized. 80% p-polarized and 20% s-polarized go to one direction while the rest go to the other direction, followed by a  $1/4$  wave plate at  $45^\circ$  and a  $1/2$  wave plate at  $22.5^\circ$  in each leg. The two beams are split again by two polarizing splitters into 4 beams, which go to 4 intensity detectors to register the intensity data.

## 4.1 Four-Channel Polarimetric Sensing

Consider a four-channel polarimeter as illustrated in Fig. 4.1. The incident beam is split by a 80/20 polarizing splitter into two beams of orthogonal polarization, s-polarized and p-polarized. 80% p-polarized and 20% s-polarized go to one direction while the rest go to the other direction, followed by a  $1/4$  wave plate at  $45^\circ$  and a  $1/2$  wave plate at  $22.5^\circ$  in each leg. The two beams are split again by two polarizing splitters into 4 beams, which go to 4 intensity detectors to register the intensity data. The polarimeter is calibrated to be optimized according to Azzmam [28]. The intensity measurements  $I = \begin{bmatrix} I_0 & I_1 & I_2 & I_3 \end{bmatrix}^T$

are linearly related to the four *instantaneous* Stokes parameters,  $S = \begin{bmatrix} S_0 & S_1 & S_2 & S_3 \end{bmatrix}^T$  by a  $4 \times 4$  instrument matrix

$$I = AS \tag{4.1}$$

$$\begin{bmatrix} I_0 \\ I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} 0.25 & 0.15 & -0.2 & 0 \\ 0.25 & 0.15 & 0.2 & 0 \\ 0.25 & -0.15 & 0 & -0.2 \\ 0.25 & -0.15 & 0 & 0.2 \end{bmatrix} \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix}.$$

The instantaneous Stokes parameters are related to the field according to the following equations [18]:

$$S_0 = |U_x|^2 + |U_y|^2 = a_x^2 + a_y^2, \tag{4.2a}$$

$$S_1 = |U_x|^2 - |U_y|^2 = a_x^2 - a_y^2, \tag{4.2b}$$

$$S_2 = U_x U_y^* + U_x^* U_y = 2a_x a_y \cos(\Delta), \tag{4.2c}$$

$$S_3 = j(U_x U_y^* - U_x^* U_y) = 2a_x a_y \sin(\Delta), \tag{4.2d}$$

where  $a_x$  and  $a_y$  are the real-valued amplitudes,  $\theta_x$  and  $\theta_y$  are the real-valued phases, and  $\Delta = \theta_y - \theta_x$  is the relative phase. Therefore,  $S$  is obtained by

$$S = A^{-1}I \quad (4.3)$$

$$\begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 5/3 & 5/3 & -5/3 & -5/3 \\ -2.5 & 2.5 & 0 & 0 \\ 0 & -0 & -2.5 & 2.5 \end{bmatrix} \begin{bmatrix} I_0 \\ I_1 \\ I_2 \\ I_3 \end{bmatrix}.$$

The four-element instantaneous Stokes vector  $S$  contains all the information on the polarization state except the absolute phase of the field, which is seldom of interest. The usual Stokes parameters or Stokes vector are the ensemble average, or the time average if the field is stationary and ergodic. We denote the usual Stokes vector as  $\langle S \rangle = [\langle S_0 \rangle \ \langle S_1 \rangle \ \langle S_2 \rangle \ \langle S_3 \rangle]^T$ , which has the following relation with the coherency matrix:

$$\langle S_0 \rangle = J_{xx} + J_{yy}, \quad (4.4a)$$

$$\langle S_1 \rangle = J_{xx} - J_{yy}, \quad (4.4b)$$

$$\langle S_2 \rangle = 2\text{Re}(J_{xy}), \quad (4.4c)$$

$$\langle S_3 \rangle = -2\text{Im}(J_{xy}). \quad (4.4d)$$

We show in Appendix A that the joint PDF for the four intensity data is

$$p_{I_0, I_1, I_2, I_3}(i_0, i_1, i_2, i_3) = \frac{125\delta\left(s_0 - \sqrt{s_1^2 + s_2^2 + s_3^2}\right)}{12\pi d s_0} \times \exp\left[-\frac{1}{2d}(\langle S_0 \rangle s_0 - \langle S_1 \rangle s_1 - \langle S_2 \rangle s_2 - \langle S_3 \rangle s_3)\right], \quad (4.5)$$

where  $d$ , is the determinant of the coherency matrix, and the relation between  $s$  and  $i$  is given in Eq. (4.3), i.e.

$$s_0 = i_0 + i_1 + i_2 + i_3, \quad (4.6a)$$

$$s_1 = \frac{5}{3}(i_0 + i_1 - i_2 - i_3), \quad (4.6b)$$

$$s_2 = -2.5(i_0 - i_1), \quad (4.6c)$$

$$s_3 = -2.5(i_2 - i_3). \quad (4.6d)$$

The joint PDF for the four intensities is parameterized by the Stokes vector

$$\langle S \rangle = \begin{bmatrix} \langle S_0 \rangle & \langle S_1 \rangle & \langle S_2 \rangle & \langle S_3 \rangle \end{bmatrix}^T. \quad (4.7)$$

Using Eq. (2.23) the Fisher information matrix for these parameters is defined as

$$F_{m,k} = \mathbb{E} \left[ \frac{\partial}{\partial \langle S_m \rangle} \ln p_{D_0, D_1, D_2, D_3}(d_0, d_1, d_2, d_3) \frac{\partial}{\partial \langle S_k \rangle} \ln p_{D_0, D_1, D_2, D_3}(d_0, d_1, d_2, d_3) \right], \quad (4.8)$$



where  $D = \begin{bmatrix} D_0 & D_1 & D_2 & D_3 \end{bmatrix}^T$  are the four intensity measurements with or without noise. We notice that in terms of the Stokes vector, the degree of polarization is given by

$$\text{DOP} = \psi(\langle S \rangle) = \frac{\sqrt{\langle S_1 \rangle^2 + \langle S_2 \rangle^2 + \langle S_3 \rangle^2}}{\langle S_0 \rangle}. \quad (4.9)$$

To bound the unbiased estimator variance for the estimation of degree of polarization, using the transformation property Eq. (2.24), we have

$$\text{var}(\widehat{\text{DOP}}) \geq \frac{\partial \psi(\langle S \rangle)}{\partial \langle S \rangle} F^{-1} \left( \frac{\partial \psi(\langle S \rangle)}{\partial \langle S \rangle} \right)^T, \quad (4.10)$$

where

$$\frac{\partial \psi(\langle S \rangle)}{\partial \langle S \rangle} = \begin{bmatrix} \frac{\partial \psi(\langle S \rangle)}{\partial \langle S_0 \rangle} & \frac{\partial \psi(\langle S \rangle)}{\partial \langle S_1 \rangle} & \frac{\partial \psi(\langle S \rangle)}{\partial \langle S_2 \rangle} & \frac{\partial \psi(\langle S \rangle)}{\partial \langle S_3 \rangle} \end{bmatrix}. \quad (4.11)$$

### 4.1.1 Speckle Only

When there is no detection noise,  $D = I$ , and the joint PDF of the measurements is

$$p_{D_0, D_1, D_2, D_3}(d_0, d_1, d_2, d_3) = p_{I_0, I_1, I_2, I_3}(d_0, d_1, d_2, d_3), \quad (4.12)$$

where  $p_{I_0, I_1, I_2, I_3}(i_0, i_1, i_2, i_3)$  is given by Eq. (4.5). By substituting Eq. (4.12) into Eq. (4.8), we can obtain the Fisher information matrix. Although analytic expressions are available for the derivatives, analytic expressions are not available for the expectations. However,

we can evaluate the expectations by simulating data and computing the expectations with statistical averages. After we obtain the Fisher information matrix, the Cramer-Rao bound can be computed by substituting the Monte Carlo result into Eq. (4.10). The result is shown in Fig. 4.2 for the case when there is no detector noise.

### 4.1.2 Gaussian Noise

If read-out noise corrupts the measurements, then

$$D = \begin{bmatrix} D_0 \\ D_1 \\ D_2 \\ D_3 \end{bmatrix} = \begin{bmatrix} I_0 + N_0 \\ I_1 + N_1 \\ I_2 + N_2 \\ I_3 + N_3 \end{bmatrix}, \quad (4.13)$$

where  $N_j \sim N(0, \sigma^2)$ ,  $j = 0, 1, 2, 3$ . We assume that each detector has the same noise level, and that the noise is mutually independent and independent of the input incident light. The joint PDF of the four intensities measurements is then given by

$$\begin{aligned} p_{D_0, D_1, D_2, D_3}(d_0, d_1, d_2, d_3) &= \int_{-\infty}^{d_3} \int_{-\infty}^{d_2} \int_{-\infty}^{d_1} \int_{-\infty}^{d_0} p_{I_0, I_1, I_2, I_3}(d_0 - n_0, d_1 - n_1, d_2 - n_2, d_3 - n_3) \\ &\quad \times p_{N_0}(n_0) p_{N_1}(n_1) p_{N_2}(n_2) p_{N_3}(n_3) dn_0 dn_1 dn_2 dn_3, \end{aligned} \quad (4.14)$$

where  $p_{I_0, I_1, I_2, I_3}(d_0 - n_0, d_1 - n_1, d_2 - n_2, d_3 - n_3)$  is given by Eq. (4.5) and  $p_{N_k}(n_k) = 1/(\sqrt{2\pi}\sigma) \exp[-n_k^2/(2\sigma^2)]$ ,  $k = 0, 1, 2, 3$ . The integral can be evaluated by numerical integration, and by substituting Eq. (4.14) into Eq. (4.8), we can obtain the Fisher information matrix. In this case, analytic expressions for either derivatives or expectations are not easily available; but we can approximate the derivatives by the local slope of the secant line:  $m = [f(a+h) - f(a)]/h$ , where  $h$  is a small number close to zero, and evaluate the expectations by simulating data and computing the statistical average. After we obtain the Fisher information matrix, the Cramer-Rao bound can be computed by substituting the Monte Carlo result into Eq. (4.10). The result is shown in Fig. 4.2 for the case of Gaussian noise.

### 4.1.3 Poisson Noise

Consider the Poisson noise case. Poisson noise, or shot noise exists because light and electrical charge are quantized and modeled as a random process with a Poisson statistic:

$$D = \begin{bmatrix} D_0 \\ D_1 \\ D_2 \\ D_3 \end{bmatrix} = \begin{bmatrix} \text{Poisson}(aI_0 + b) \\ \text{Poisson}(aI_1 + b) \\ \text{Poisson}(aI_2 + b) \\ \text{Poisson}(aI_3 + b) \end{bmatrix}, \quad (4.15)$$

where  $a$  is the gain of the amplifier and  $b$  is the dark current. The conditional probability mass function of the data (number of photons collected) is given by

$$p_{D|I}(d_0, d_1, d_2, d_3 | i_0, i_1, i_2, i_3) = \prod_{k=0}^3 \frac{\exp[-(ai_k + b)](ai_k + b)^{d_k}}{d_k!}. \quad (4.16)$$

Because the joint PDF of  $I$  is known, Eq. (4.5), we have

$$p_{D,I}(d_0, d_1, d_2, d_3, i_0, i_1, i_2, i_3) = p_{D|I}(d_0, d_1, d_2, d_3 | i_0, i_1, i_2, i_3) p_I(i_0, i_1, i_2, i_3). \quad (4.17)$$

Therefore, by computing the marginal PDF, the joint probability mass function of  $D$  is given by

$$\begin{aligned} p_{D_0, D_1, D_2, D_3}(d_0, d_1, d_2, d_3) &= \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty p_{D,I}(d_0, d_1, d_2, d_3, i_0, i_1, i_2, i_3) di_0 di_1 di_2 di_3 \\ &= \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \prod_{k=0}^3 \frac{\exp[-(ai_k + b)](ai_k + b)^{d_k}}{d_k!} P_{I_0, I_1, I_2, I_3}(i_0, i_1, i_2, i_3) di_0 di_1 di_2 di_3, \end{aligned} \quad (4.18)$$

where  $P_{I_0, I_1, I_2, I_3}(i_0, i_1, i_2, i_3)$  is given by Eq. (4.5)

After we have the joint probability density function of the four measured intensity data, we can compute the Cramer-Rao bound by the same Monte Carlo method used for the Gaussian noise case. The result is shown in Fig. 4.2 for the case of Poisson noise.

#### 4.1.4 Combined Gaussian and Poisson Noise

In the presence of both Gaussian and Poisson noise, the four intensities data from the polarimeter are

$$D = \begin{bmatrix} D_0 \\ D_1 \\ D_2 \\ D_3 \end{bmatrix} = \begin{bmatrix} \text{Poisson}(aI_0 + b) + N_0 \\ \text{Poisson}(aI_1 + b) + N_1 \\ \text{Poisson}(aI_2 + b) + N_2 \\ \text{Poisson}(aI_3 + b) + N_3 \end{bmatrix} = \begin{bmatrix} L_0 + N_0 \\ L_1 + N_1 \\ L_2 + N_2 \\ L_3 + N_3 \end{bmatrix}, \quad (4.19)$$

where  $N_j \sim N(0, \sigma^2)$ ,  $j = 0, 1, 2, 3$ . We assume each detector has the same noise level,  $a$  is the gain of the amplifier, and  $b$  is the dark current. Similar to the conditional probability mass function for total intensity data, the conditional probability mass function is

$$p_{L_0, L_1, L_2, L_3 | I_0, I_1, I_2, I_3}(l_0, l_1, l_2, l_3 | i_0, i_1, i_2, i_3) = \sum_{m_0=0}^{\infty} \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \sum_{m_3=0}^{\infty} \prod_{j=0}^3 \frac{\exp[-(ai_j + b)](ai_j + b)^{l_j}}{l_j!} \times \delta(l_0 - m_0, l_1 - m_1, l_2 - m_2, l_3 - m_3). \quad (4.20)$$

Because Gaussian noise in the four channels are mutually independent, and also independent on the four intensities, we have that

$$p_{N_0, N_1, N_2, N_3 | I_0, I_1, I_2, I_3}(n_0, n_1, n_2, n_3 | i_0, i_1, i_2, i_3) = \prod_{j=0}^3 p_{N_j}(n_j) = \prod_{j=0}^3 \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n_j^2}{2\sigma^2}\right). \quad (4.21)$$

Because  $D_j = L_j + N_j$ ,  $j = 0, 1, 2, 3$ , we have

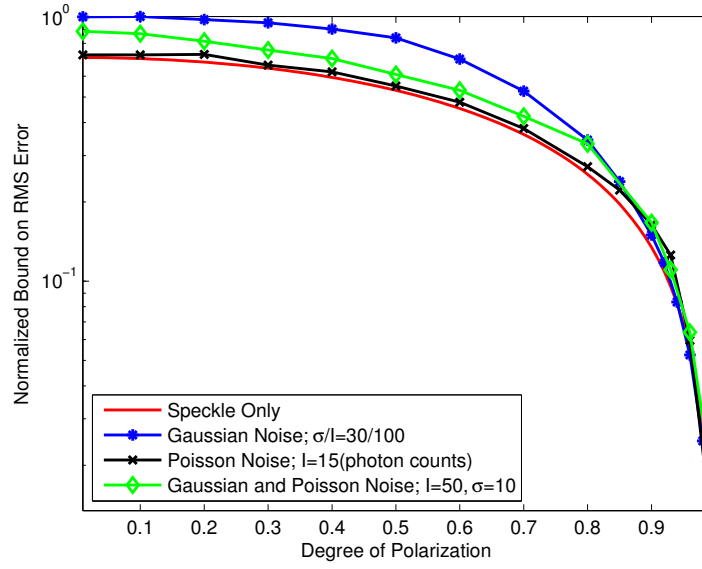
$$\begin{aligned}
& p_{D_0, D_1, D_2, D_3 | I_0, I_1, I_2, I_3}(d_0, d_1, d_2, d_3 | i_0, i_1, i_2, i_3) \\
&= \sum_{m_0=0}^{\infty} \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \sum_{m_3=0}^{\infty} P \left( \begin{array}{c} l_0 = m_0, l_1 = m_1, l_2 = m_2, l_3 = m_3 \\ n_0 = d_0 - m_0, n_1 = d_1 - m_1, n_2 = d_2 - m_2, n_3 = d_3 - m_3 \end{array} \middle| I \right) \\
&= \sum_{m_0=0}^{\infty} \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \sum_{m_3=0}^{\infty} \prod_{j=0}^3 \frac{\exp[-(ai_j + b)](ai_j + b)^{m_j}}{m_j!} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d_j - m_j)^2}{2\sigma^2}\right).
\end{aligned} \tag{4.22}$$

The joint PDF of the measurements  $D$  is then given by the integral

$$\begin{aligned}
p_{D_0, D_1, D_2, D_3}(d_0, d_1, d_2, d_3) &= \int_0^{\infty} \int_0^{\infty} \int_0^{\infty} \int_0^{\infty} p_{D|I}(d_0, d_1, d_2, d_3 | i_0, i_1, i_2, i_3) \\
&\quad \times p_I(i_0, i_1, i_2, i_3) di_0 di_1 di_2 di_3, \tag{4.23}
\end{aligned}$$

where the conditional probability density function  $p_{D|I}(d_0, d_1, d_2, d_3 | i_0, i_1, i_2, i_3)$  and the joint PDF of four intensities  $p_I(i_0, i_1, i_2, i_3)$  is given by Eq. (4.22) and Eq. (4.5) respectively.

The above integral is computed by numerical integration, and the Cramer-Rao bound is computed by the same approach used in the previous cases. The result is shown in Fig. 4.2 for the case of combined Gaussian and Poisson noise.



**Figure 4.2:** Cramer-Rao bounds on the root-mean-square (RMS) estimation error for unbiased estimators of the degree of polarization from four-channel polarimeter intensity measurements with various noise model; for Gaussian noise case  $\bar{I} = 100$ ,  $\sigma = 30$ ; for Poisson noise case,  $\bar{I} = 15$ /photon counts,  $a = 1$ ,  $b = 0$ ; for Gaussian and Poisson noise case,  $\bar{I} = 50$ /photon counts,  $\sigma = 10$ ,  $a = 1$ ,  $b = 0$ .

#### 4.1.5 Plot of Bounds

The performance bounds for unbiased estimators for four-channel polarimeter intensity data are shown in Fig. 4.2. For comparison, the parameters are chosen as explained in section 4.4.1. The coordinate rotation angle  $\theta$  and the relative phase shift  $\phi_{xy}$  in Eq. (2.4) are set to be  $\pi/3$  and  $\pi/4$  respectively in all cases. For the Gaussian noise case,  $\bar{I} = 100$ ,  $\sigma = 30$ ; for Poisson noise case,  $\bar{I} = 15$ /photon counts,  $a = 1$ ,  $b = 0$ ; for Gaussian and Poisson noise case,  $\bar{I} = 50$ /photon counts,  $\sigma = 10$ ,  $a = 1$ ,  $b = 0$ .

## 4.2 Orthogonal Intensities Sensing

From Eq. (6.9), we see that the probability density function (PDF) for two orthogonal intensity data  $D = \begin{bmatrix} D_x & D_y \end{bmatrix}^T$  is parameterized by  $J_{xx}$ ,  $J_{yy}$  and  $|J_{xy}|$ ; accordingly, we let these three parameters be our choice of parameterization

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} = \begin{bmatrix} J_{xx} \\ J_{yy} \\ |J_{xy}| \end{bmatrix}. \quad (4.24)$$

The Fisher information matrix for these parameters is then defined as

$$F_{m,k} = \mathbb{E} \left[ \frac{\partial}{\partial \theta_m} \ln p_{D_x, D_y}(d_x, d_y) \frac{\partial}{\partial \theta_k} \ln p_{D_x, D_y}(d_x, d_y) \right]. \quad (4.25)$$

Recall Eq. (2.5), and we have that

$$\text{DOP} = \sqrt{1 - 4 \frac{\det(J)}{[\text{tr}(J)]^2}} = \psi(\theta) = \sqrt{1 - 4 \frac{\theta_1 \theta_2 - \theta_3^2}{(\theta_1 + \theta_2)^2}}. \quad (4.26)$$

Again, to bound the unbiased estimator variance for the estimation of degree of polarization, using the transformation property Eq. (2.24), we have

$$\text{var}(\widehat{\text{DOP}}) \geq \left( \frac{\partial \psi(\theta)}{\partial \theta} \right)^T F^{-1} \frac{\partial \psi(\theta)}{\partial \theta}, \quad (4.27)$$



where

$$\frac{\partial \psi(\theta)}{\partial \theta} = \begin{bmatrix} \frac{\partial \psi(\theta)}{\partial \theta_1} \\ \frac{\partial \psi(\theta)}{\partial \theta_2} \\ \frac{\partial \psi(\theta)}{\partial \theta_3} \end{bmatrix}. \quad (4.28)$$

### 4.2.1 Speckle Only

When no detection noise is present, the joint PDF of the orthogonal intensity measurements

$D = \begin{bmatrix} D_x & D_y \end{bmatrix}^T$  is given by Eq. (6.9)

$$p_{D_x, D_y}(d_x, d_y) = \frac{1}{\det(J)} \exp \left( -\frac{J_{yy}d_x + J_{xx}d_y}{\det(J)} \right) I_0 \left( 2 \frac{|J_{xy}|}{\det(J)} \sqrt{d_x d_y} \right), \quad (4.29)$$

where  $I_0$  is the zero-order modified Bessel function of the first kind. By substituting Eq. (4.29) into Eq. (4.25), we can obtain the Fisher information matrix. Although analytic expressions are available for the derivatives, analytic expressions are not available for the expectations. However, similar to the method for polarimeter intensities data, we can evaluate the expectations by simulating data and computing the expectations with statistical averages. After we obtain the Fisher information matrix, the Cramer-Rao bound can be computed by substituting the Monte Carlo result into Eq. (4.27). The result is shown in Fig. 4.3.

### 4.2.2 Gaussian Noise

When read-out noise exists in the detectors, the measurements are corrupted by Gaussian noise

$$D = \begin{bmatrix} D_x \\ D_y \end{bmatrix} = \begin{bmatrix} I_x + N_x \\ I_y + N_y \end{bmatrix}, \quad (4.30)$$

where  $N_x$  and  $N_y$  are independent Gaussian random variables with mean zero and variance  $\sigma^2$ . We assume each detector has the same noise level, and they are mutually independent.

The joint PDF of the orthogonal intensity measurements is given by

$$p_{D_x, D_y}(d_x, d_y) = \int_{-\infty}^{d_y} \int_{-\infty}^{d_x} p_{I_x, I_y}(d_x - n_x, d_y - n_y) p_{N_x}(n_x) p_{N_y}(n_y) dn_x dn_y, \quad (4.31)$$

where  $p_{I_x, I_y}(d_x - n_x, d_y - n_y)$  is given by Eq. (6.9), and  $p_{N_x}$  and  $p_{N_y}$  are the PDFs for the independent Gaussian random variables with zero means and variances of  $\sigma^2$ . By substituting Eq. (4.31) into Eq. (4.25), we can obtain the Fisher information matrix. Similar to the four-channel polarimeter, analytic expressions for neither derivatives nor expectations is available; however, we can approximate the derivatives by the local slope of the secant line and evaluate the expectations by simulating data and computing the statistical average. After we obtain the Fisher information matrix, the Cramer-Rao bound can be computed by substituting the Monte Carlo result into Eq. (4.27). The result is shown in Fig. 4.3.

### 4.2.3 Poisson Noise

Let us consider the Poisson noise case:

$$D = \begin{bmatrix} D_x \\ D_y \end{bmatrix} = \begin{bmatrix} \text{Poisson}(aI_x + b) \\ \text{Poisson}(aI_y + b) \end{bmatrix}, \quad (4.32)$$

where  $a$  is the gain of the amplifier and  $b$  is the dark current. The conditional probability mass function of the data (number of photons collected) is then given by

$$p_{D_x, D_y | I_x, I_y}(d_x, d_y | i_x, i_y) = \frac{\exp[-(ai_x + b)](ai_x + b)^{d_x}}{d_x!} \frac{\exp[-(ai_y + b)](ai_y + b)^{d_y}}{d_y!}. \quad (4.33)$$

Because the joint PDF of  $I_x$  and  $I_y$  is known, according to the definition of conditional probability density function, we have

$$p_{D_x, D_y, I_x, I_y}(d_x, d_y, i_x, i_y) = p_{D_x, D_y | I_x, I_y}(d_x, d_y | i_x, i_y) P_{I_x, I_y}(i_x, i_y). \quad (4.34)$$

Therefore, after computing the marginal PDF, the joint probability mass function of  $D_x$  and  $D_y$  is given by

$$p_{D_x, D_y}(d_x, d_y) = \int_0^\infty \int_0^\infty p_{D_x, D_y, I_x, I_y}(d_x, d_y, i_x, i_y) di_x di_y. \quad (4.35)$$

We know  $p_{D_x, D_y | I_x, I_y}(d_x, d_y | i_x, i_y)$  and  $P_{I_x, I_y}(i_x, i_y)$ . Then, using Eq.(4.34), we have:

$$p_{D_x, D_y}(d_x, d_y) = \int_0^\infty \int_0^\infty \frac{\exp[-(ai_x + b)](ai_x + b)^{d_x}}{d_x!} \frac{\exp[-(ai_y + b)](ai_y + b)^{d_y}}{d_y!} \\ \times \frac{1}{\det(J)} \exp\left(-\frac{J_{yy}i_x + J_{xx}i_y}{\det(J)}\right) I_0\left(2\frac{|J_{xy}|}{\det(J)}\sqrt{i_x i_y}\right) di_x di_y. \quad (4.36)$$

The integral can be computed by numerical integration, and, after we have the joint probability density function of the two orthogonal intensities data, we compute the Cramer-Rao bound by the Monte Carlo method used in Gaussian noise case. The result is shown in Fig. 4.3.

#### 4.2.4 Combined Gaussian and Poisson Noise

Follow the detector noise model we discussed in section 2.3, and we have

$$D = \begin{bmatrix} D_x \\ D_y \end{bmatrix} = \begin{bmatrix} \text{Poisson}(aI_x + b) + N_x \\ \text{Poisson}(aI_y + b) + N_y \end{bmatrix} = \begin{bmatrix} L_x + N_x \\ L_y + N_y \end{bmatrix}, \quad (4.37)$$

where  $N_x$  and  $N_y$  are independent Gaussian random variables with zero means and variances of  $\sigma^2$ ,  $a$  is the gain of the amplifier, and  $b$  is the dark current. Use the same approach in

section 4.3.4, we have

$$p_{L_x, L_y | I_x, I_y}(l_x, l_y | i_x, i_y) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{\exp[-(ai_x + b)](ai_x + b)^{l_x}}{l_x!} \times \frac{\exp[-(ai_y + b)](ai_y + b)^{l_y}}{l_y!} \delta(l_x - m, l_y - n) \quad (4.38)$$

and

$$p_{N_x, N_y | I_x, I_y}(n_x, n_y | i_x, i_y) = p_{N_x, N_y}(n_x, n_y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n_x^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n_y^2}{2\sigma^2}\right). \quad (4.39)$$

Poisson noise and Gaussian noise are independent; therefore

$$\begin{aligned} p_{D_x, D_y | I_x, I_y}(d_x, d_y | i_x, i_y) &= \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} P(l_x = m, l_y = k, n_x = d_x - m, n_y = d_y - k | i_x, i_y) \\ &= \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{(ai_x + b)^m \exp[-(ai_x + b)]}{m!} \frac{(ai_y + b)^k \exp[-(ai_y + b)]}{k!} \\ &\quad \times \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d_x - m)^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d_y - k)^2}{2\sigma^2}\right). \end{aligned} \quad (4.40)$$

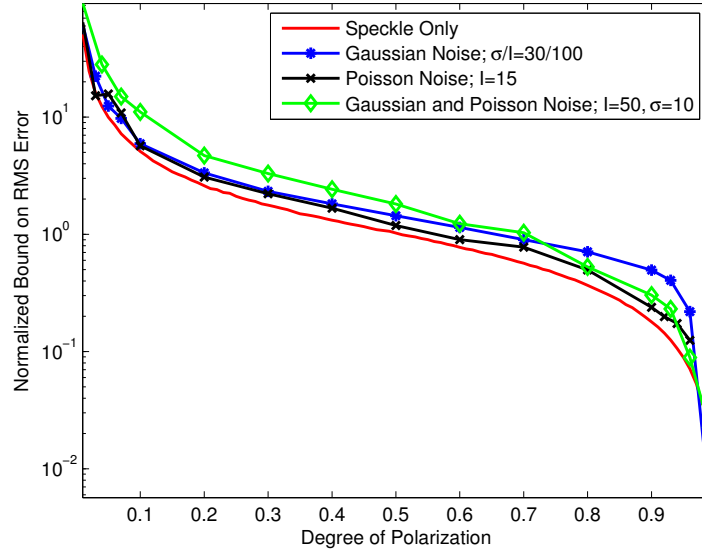
According to the definition of conditional probability density function, the joint PDF of  $D_x$  and  $D_y$  is given by

$$\begin{aligned}
p_{D_x, D_y}(d_x, d_y) &= \int_0^\infty \int_0^\infty p_{D_x, D_y | I_x, I_y}(d_x, d_y | i_x, i_y) p_{I_x, I_y}(i_x, i_y) di_x di_y \\
&= \int_0^\infty \int_0^\infty \sum_{m=0}^\infty \sum_{k=0}^\infty \frac{(ai_x + b)^m \exp[-(ai_x + b)]}{m!} \frac{(ai_y + b)^k \exp[-(ai_y + b)]}{k!} \\
&\quad \times \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d_x - m)^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d_y - k)^2}{2\sigma^2}\right) \\
&\quad \times \frac{1}{\det(J)} \exp\left(-\frac{J_{yy}i_x + J_{xx}i_y}{\det(J)}\right) I_0\left(2\frac{|J_{xy}|}{\det(J)}\sqrt{i_x i_y}\right) di_x di_y. \quad (4.41)
\end{aligned}$$

As the previous cases of Gaussian and Poisson noise, the expression of the joint PDF of the measurements is too complicated to find the derivative and expectation. We adopt the same approach to compute the Cramer-Rao bound. The result is shown in Fig. 4.3.

#### 4.2.5 Plot of Bounds

The performance bounds for unbiased estimators for two orthogonal intensity data are shown in Fig. 4.3. For comparison, the parameters are chosen as explained in section 4.4.1. The coordinate rotation angle  $\theta$  and the relative phase shift  $\phi_{xy}$  are  $\pi/3$  and  $\pi/4$  respectively. For Gaussian noise case  $\bar{I} = 100$ ,  $\sigma = 30$ ; for Poisson noise case,  $\bar{I} = 15/\text{photon counts}$ ,  $a = 1$ ,  $b = 0$ ; for Gaussian and Poisson noise case,  $\bar{I} = 50/\text{photon counts}$ ,  $\sigma = 10$ ,  $a = 1$ ,  $b = 0$ .



**Figure 4.3:** Cramer-Rao bounds on the root-mean-square (RMS) estimation error for unbiased estimators of the degree of polarization from two orthogonal intensities measurements with various noise model; for Gaussian noise case  $\bar{I} = 100$ ,  $\sigma = 30$ ; for Poisson noise case,  $\bar{I} = 15$ /photon counts,  $a = 1$ ,  $b = 0$ ; for Gaussian and Poisson noise case,  $\bar{I} = 50$ /photon counts,  $\sigma = 10$ ,  $a = 1$ ,  $b = 0$ .

### 4.3 Total Intensity Sensing

Recall the probability density function (PDF) for the total intensity data, Eq. (2.12). We note that the PDF is a function of the two real-valued eigenvalues of the coherency matrix; therefore, we choose the parametrization as

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}. \quad (4.42)$$

The Fisher information matrix for total intensity data  $D$  is then defined as

$$F_{m,k} = \text{E} \left[ \frac{\partial}{\partial \theta_m} \ln p_D(d) \frac{\partial}{\partial \theta_k} \ln p_D(d) \right], \quad (4.43)$$

and the degree of the polarization is defined as

$$\text{DOP} = \frac{|\lambda_1 - \lambda_2|}{\lambda_1 + \lambda_2} = \frac{|\theta_1 - \theta_2|}{\theta_1 + \theta_2}. \quad (4.44)$$

To bound the unbiased estimator variance for the estimation of degree of polarization, using the transformation property Eq. (2.24), we again have

$$\text{var}(\widehat{\text{DOP}}) \geq \left( \frac{\partial \psi(\theta)}{\partial \theta} \right)^T F^{-1} \frac{\partial \psi(\theta)}{\partial \theta}, \quad (4.45)$$

where

$$\frac{\partial \psi(\theta)}{\partial \theta} = \begin{bmatrix} \frac{\partial \psi(\theta)}{\partial \theta_1} \\ \frac{\partial \psi(\theta)}{\partial \theta_2} \end{bmatrix}. \quad (4.46)$$

### 4.3.1 Speckle Only

When no detection noise is considered, the PDF of the total intensity measurements  $D$  is given by Eq. (2.12)

$$p_D(d) = \frac{1}{\lambda_1 - \lambda_2} \left[ \exp\left(-\frac{d}{\lambda_1}\right) - \exp\left(-\frac{d}{\lambda_2}\right) \right]. \quad (4.47)$$



By substituting Eq. (4.47) into Eq. (4.43), we can obtain the Fisher information matrix. As the previous two sensing modalities, analytic expressions are available for the derivatives while analytic expressions are not easily available for the expectations. The analytic expressions for the expectations are given in [29], which has a generalized Riemann zeta function that can only be approximated. We evaluate the expectations by simulating data and computing the expectations with statistical averages. Once we Obtain the Fisher information matrix, the Cramer-Rao bound can be computed by substituting the simulation result into Eq. (4.45). The result is shown in Fig. 4.4.

### 4.3.2 Gaussian Noise

When the total intensity measurement is corrupted by Gaussian read-out noise, we have

$$D = I + N, \quad (4.48)$$

where  $N \sim N(0, \sigma^2)$ . PDF of the total intensity measurement is given by

$$\begin{aligned} p_D(d) &= \int_{-\infty}^d p_I(d-n) p_N(n) dn \\ &= \int_{-\infty}^d \frac{1}{\lambda_1 - \lambda_2} \left[ \exp\left(-\frac{d-n}{\lambda_1}\right) - \exp\left(-\frac{d-n}{\lambda_2}\right) \right] \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n^2}{\sigma^2}\right) dn \\ &= \frac{1}{\lambda_1 - \lambda_2} \left[ \exp\left(-\frac{2\lambda_1 d - \sigma^2}{2\lambda_1^2}\right) \Phi\left(\frac{d}{\sigma} - \frac{\sigma}{\lambda_1}\right) - \exp\left(-\frac{2\lambda_2 d - \sigma^2}{2\lambda_2^2}\right) \Phi\left(\frac{d}{\sigma} - \frac{\sigma}{\lambda_2}\right) \right]. \end{aligned} \quad (4.49)$$

Substituting Eq. (4.49) into Eq. (4.43), we obtain the Fisher information matrix. Analytic expressions for neither derivatives nor expectations is available. As discussed before, we again approximate the derivatives by the local slope of the secant line, and evaluate the expectations by simulating data and computing the statistical average. The result is shown in Fig. 4.4.

### 4.3.3 Poisson Noise

In the presence of Poisson noise, the total intensity measurement is

$$D = \text{Poisson}(aI + b) = \text{Poisson}[a(I_x + I_y) + b], \quad (4.50)$$

where  $a$  is the combined gain of the amplifier and  $b$  is the dark current. The conditional probability mass function of the data (number of photons collected) is then given by

$$p_{D|I}(d|i) = \frac{\exp[-(ai + b)](ai + b)^d}{d!}. \quad (4.51)$$

According to the definition of conditional probability density function, we have

$$p_{D,I}(d, i) = p_{D|I}(d|i)P_I(i). \quad (4.52)$$

Therefore, by computing the marginal PDF, the probability mass function of  $D$  is given by

$$\begin{aligned}
p_D(d) &= \int_0^\infty p_{D,I}(d, i) di \\
&= \int_0^\infty \frac{\exp[-(ai+b)](ai+b)^d}{d!} \frac{1}{\lambda_1 - \lambda_2} \left[ \exp\left(-\frac{i}{\lambda_1}\right) - \exp\left(-\frac{i}{\lambda_2}\right) \right] di \\
&= \frac{\exp\left(\frac{b}{a\lambda_1}\right) \text{Ei}(d, b, X)}{a(\lambda_1 - \lambda_2)d!} - \frac{\exp\left(\frac{b}{a\lambda_2}\right) \text{Ei}(d, b, Y)}{a(\lambda_1 - \lambda_2)d!},
\end{aligned} \tag{4.53}$$

where  $X = \left(1 + \frac{1}{a\lambda_1}\right)$  and  $Y = \left(1 + \frac{1}{a\lambda_2}\right)$ , and the exponential integral function, denoted as  $\text{Ei}(n, b, z)$ , is given by [30]

$$\text{Ei}(n, b, z) = \int_b^\infty \lambda^n \exp(-z\lambda) d\lambda = \begin{cases} z^{-n-1} \Gamma(n+1, bz) & b > 0 \\ z^{-n-1} n! & b = 0 \end{cases}, \tag{4.54}$$

where  $\Gamma(n, z)$  is the incomplete Gamma function.

After we have the probability density function of the total intensities measurement, we can compute the Cramer-Rao bound by using the Monte Carlo method. The result is shown in Fig. 4.4.

#### 4.3.4 Combined Gaussian and Poisson Noise

If we use the detector noise model Eq. (2.25) in section 2.3, the total intensity data is

$$D = \text{Poisson}(aI + b) + N, \quad (4.55)$$

where  $N \sim N(0, \sigma^2)$ ,  $a$  is the gain of the amplifier and  $b$  is the dark current. Let us denote the photon counts  $\text{Poisson}(aI + b)$  as  $L$ , then  $D = L + N$ . First of all, we know the conditional probability mass function:

$$p_{L|I}(l, i) = \sum_{m=0}^{\infty} \frac{(ai + b)^l \exp[-(ai + b)]}{l!} \delta(l - m), \quad (4.56)$$

and the PDF of the Gaussian noise

$$p_N(n) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n^2}{2\sigma^2}\right). \quad (4.57)$$

Because the Gaussian noise are independent of the total intensity and the Poisson noise, we have

$$\begin{aligned} p_{D,N|I}(d, n|i) &= p_{L|I}(l, i) p_N(n) \\ &= \sum_{m=0}^{\infty} \frac{(ai + b)^{d-n} \exp[-(ai + b)]}{(d-n)!} \delta(d - n - m) \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n^2}{2\sigma^2}\right), \end{aligned} \quad (4.58)$$

and the marginal PDF:

$$\begin{aligned}
p_{D|I}(d|i) &= \int_{-\infty}^{dx} p_{D,N|I}(d,n|i) dn \\
&= \int_{-\infty}^{dx} \sum_{m=0}^{\infty} \frac{(ai+b)^{d-n} \exp[-(ai+b)]}{(d-n)!} \delta(d-n-m) \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n^2}{2\sigma^2}\right) dn \\
&= \sum_{m=0}^{\infty} \frac{(ai+b)^m \exp[-(ai+b)]}{m!} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d-m)^2}{2\sigma^2}\right).
\end{aligned} \tag{4.59}$$

After we have the above conditional PDF, according to the definition of conditional possibility density function, the PDF of the total intensity data is

$$\begin{aligned}
p_D(d) &= \int_0^{\infty} p_{D|I}(d|i) p_I(i) di \\
&= \int_0^{\infty} \sum_{m=0}^{\infty} \frac{(ai+b)^m \exp[-(ai+b)]}{m!} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d-m)^2}{2\sigma^2}\right) \\
&\quad \times \frac{1}{\lambda_1 - \lambda_2} \left[ \exp\left(-\frac{i}{\lambda_1}\right) - \exp\left(-\frac{i}{\lambda_2}\right) \right] di \\
&= \sum_{m=0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d-m)^2}{2\sigma^2}\right) \int_0^{\infty} \frac{(ai+b)^m \exp[-(ai+b)]}{m!} \\
&\quad \times \frac{1}{\lambda_1 - \lambda_2} \left[ \exp\left(-\frac{i}{\lambda_1}\right) - \exp\left(-\frac{i}{\lambda_2}\right) \right] di \\
&= \sum_{m=0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d-m)^2}{2\sigma^2}\right) \left[ \frac{\exp\left(\frac{b}{a\lambda_1}\right) \text{Ei}(m, b, X)}{a(\lambda_1 - \lambda_2)m!} - \frac{\exp\left(\frac{b}{a\lambda_2}\right) \text{Ei}(m, b, Y)}{a(\lambda_1 - \lambda_2)m!} \right].
\end{aligned} \tag{4.60}$$

As in Eq. (4.53),  $X = \left(1 + \frac{1}{a\lambda_1}\right)$  and  $Y = \left(1 + \frac{1}{a\lambda_2}\right)$ , and we denote the exponential integral function as

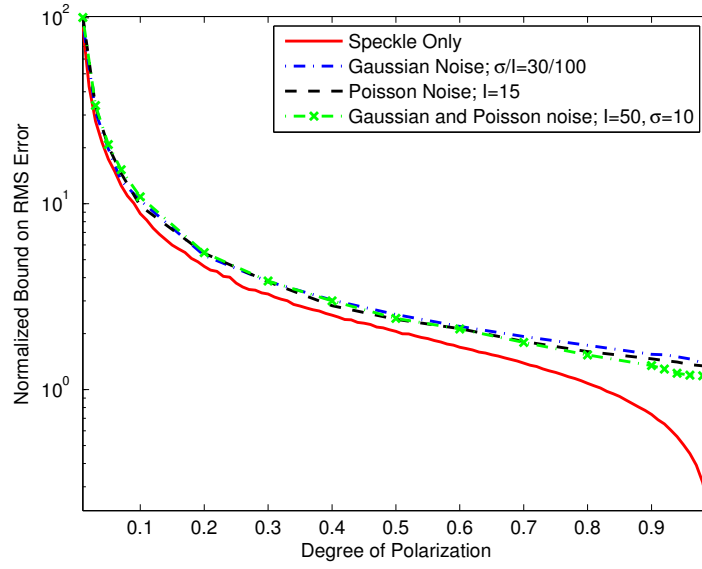
$$\text{Ei}(n, b, z) = \int_b^\infty \lambda^n \exp(-z\lambda) d\lambda = \begin{cases} z^{-n-1} \Gamma(n+1, bz) & b > 0 \\ z^{-n-1} n! & b = 0 \end{cases}, \quad (4.61)$$

where  $\Gamma(n, z)$  is the incomplete Gamma function.

To compute the Cramer-Rao bound for the estimation from total intensity data, we use the same numerical approach. The result is shown in Fig. 4.4.

### 4.3.5 Plot of Bounds

The lower performance bounds for unbiased estimators for total intensity data are shown in Fig. 4.4. For comparison, the parameters are chosen as explained in section 4.4.1. The coordinate rotation angle  $\theta$  and the relative phase shift  $\phi_{xy}$  are  $\pi/3$  and  $\pi/4$  respectively. For Gaussian noise case  $\bar{I} = 100$ ,  $\sigma = 30$ ; for Poisson noise case,  $\bar{I} = 15$ /photon counts,  $a = 1$ ,  $b = 0$ ; for Gaussian and Poisson noise case,  $\bar{I} = 50$ /photon counts,  $\sigma = 10$ ,  $a = 1$ ,  $b = 0$ .

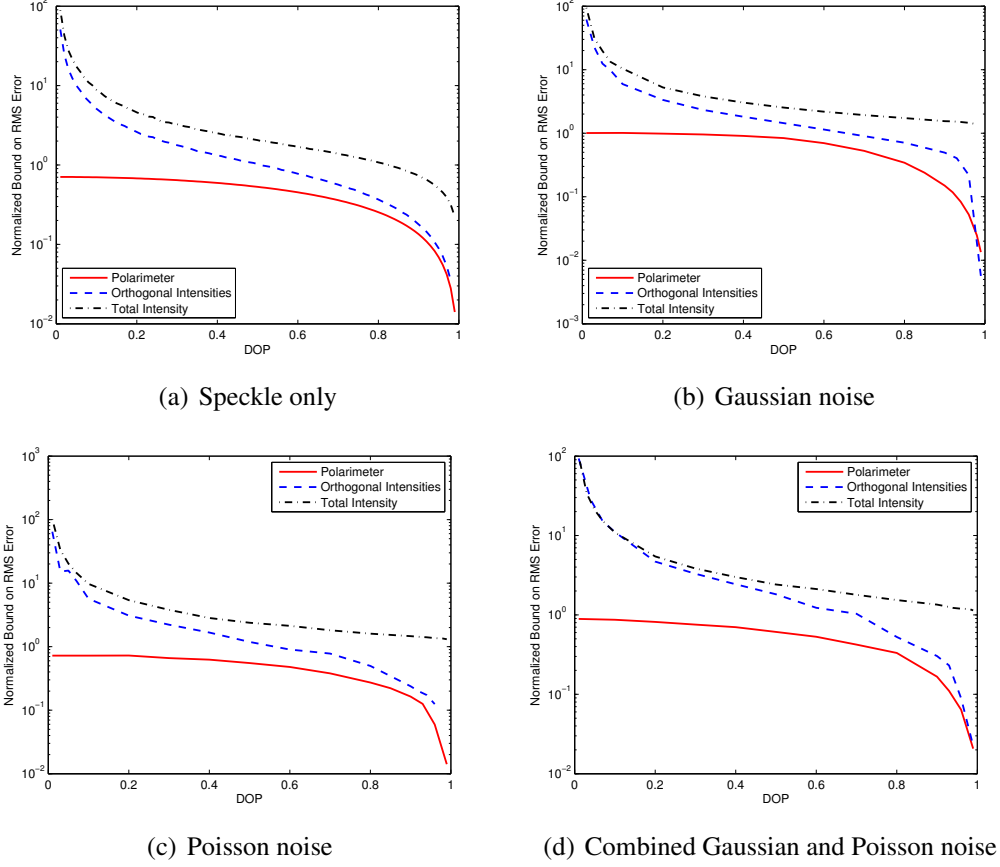


**Figure 4.4:** Cramer-Rao bounds on the root-mean-square (RMS) estimation error for unbiased estimators of the degree of polarization from total intensity measurements with various noise model; for Gaussian noise case  $\bar{I} = 100$ ,  $\sigma = 30$ ; for Poisson noise case,  $\bar{I} = 15$ /photon counts,  $a = 1$ ,  $b = 0$ ; for Gaussian and Poisson noise case,  $\bar{I} = 50$ /photon counts,  $\sigma = 10$ ,  $a = 1$ ,  $b = 0$ .

## 4.4 Discussion

### 4.4.1 Comparison Of Bounds

The Cramer-Rao lower bounds on the normalized root-mean-square (RMS) error for unbiased estimators of the degree of polarization (DOP) are shown in Fig. 4.4.1, and grouped by sensing noise models: (a) noise free; (b) Gaussian read out noise; (c) Poisson shot noise; (d) combined Gaussian and Poisson noise. The bounds are computed by the methods discussed in the previous sections. For the sake of simplicity here, we choose the unit as photons, ignore the dark current, and assume the *quantum efficiency* [16] as one. Read



**Figure 4.5:** Comparison of normalized Cramer-Rao bounds on RMS estimation error for unbiased estimators of the DOP. (a) detector noise free; (b) Gaussian read out noise.  $\bar{I} = 100, \sigma = 30$ ; (c) Poisson shot noise.  $\bar{I} = 15$ ; (d) combined Gaussian and Poisson noise.  $\bar{I} = 50, \sigma = 10$

noise in widely available visible wavelength cameras typically lies in the range of 5 to 100 RMS electrons per pixel per readout. Therefore, for a typical Gaussian noise case, the simulation parameters are set as: average total intensity is  $\bar{I} = 100$  photons per pixel per readout and the read out noise level is  $\sigma = 30$  RMS photons per pixel per readout. For Poisson noise case, we simulate a low incident light situation where  $\bar{I} = 15$  photons per pixel per readout. For combined Gaussian and Poisson noise case,  $\bar{I} = 50$  photons per pixel per readout and  $\sigma = 10$  RMS photons per pixel per readout. The coordinate rotation an-

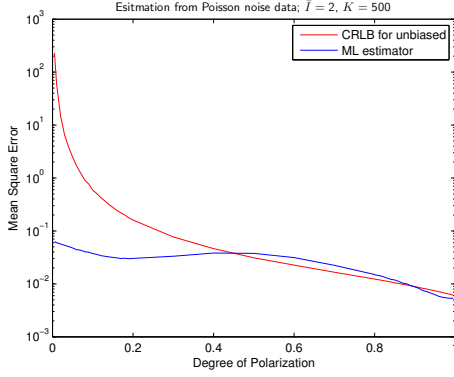


gle  $\theta$  and the relative phase shift  $\phi_{xy}$  in Eq. (2.4) are set to be  $\pi/3$  and  $\pi/4$  respectively in all simulations. However, our simulation shows that the coordinate angle  $\theta$  affects the performance bound. The two angles are chosen arbitrarily as an example.

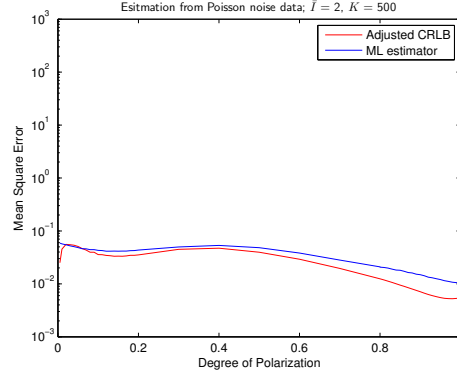
All bounds are normalized by dividing  $\sqrt{K}$ , where  $K$  is the number of independent speckle realizations to form the estimate. Note that, when scaling the bounds by  $\sqrt{K}$ , we assume that for each speckle realization the ratio of noise level and average intensity remains the same, rather than that the average intensity decreases but noise level is fixed when we have more realizations ( $K$  increases).

The performance bounds show the expected hierarchy that polarimeter estimators perform best while orthogonal intensities estimators do better than total intensity estimators, because total intensity data can be created from the sum of orthogonal intensities data and orthogonal intensity data can be created from the four intensities data.

For all three data-collecting modalities, when the degree of polarization is below 0.2, having four intensities data has the great advantage of smaller RMS error of  $10^{-1}$  or less over having orthogonal intensities data or total intensity data. When the degree of polarization is larger than 0.8, the bounds are close to each other and such advantage over having orthogonal intensities data becomes less and is eventually lost when DOP reaches 1. Meanwhile, through out the range of DOP from zero to one, having four intensities data gives an improvement on the bounds to the order of  $10^1$  or more compared to bounds of the total intensity data only.



(a) CRLB for unbiased estimator



(b) Adjusted CRLB for biased estimator

**Figure 4.6:** CRLBs for both unbiased and biased estimators for total intensity data with Poisson noise. Average Intensity  $\bar{I} = 2$ ,  $K = 500$  speckle realizations.

#### 4.4.2 Adjusting Cramer-Rao bound for Bias

We note that for total intensity and orthogonal intensities data, the bounds for unbiased estimators of the DOP approaches to infinity when the DOP approaches to zero. However, in practice, an estimator will have the implicit or explicit constraint that its estimate falls into the range between 0 and 1 [13]; therefore, the estimator is biased. As a result, both the variance and the mean squared error may be below the unbiased Cramer-Rao lower bound. However, we still want to assess the performance of a certain biased estimator. For a biased estimator, the CRLB is adjusted by

$$b^2(p) + \text{CRLB} \left( 1 + \frac{\partial b(p)}{\partial p} \right)^2, \quad (4.62)$$

where  $b(p) = E[\hat{p}] - p$  is the bias. Fig.4.6 shows the CRLB for both unbiased and biased estimators for total intensity data with Poisson noise. The average total intensity  $\bar{I} = 2$

photons per pixel per readout, and sample size is 500 speckle realizations. The mean-square error (MSE) of the maximum likelihood (ML) estimator (expectation-maximization algorithm) is compared to both bounds. The MSE of the ML estimator is very close to the adjusted CRLB, therefore, the ML estimator is close to efficient.

### 4.4.3 Signal-to-noise Ratio

The signal-to-noise ratio (SNR) is a key parameter for any signal processing system, which determines the overall performance of the system. For our degree of polarization estimation system, we define the signal-to-noise ratio as

$$\text{SNR} = \frac{\mu \mu^T}{\text{tr}(\Sigma)}, \quad (4.63)$$

where  $\mu = \text{E}[D]$  is the mean of measurements and

$$\Sigma = \text{E} \left[ (D - \mu)(D - \mu)^T \right] \quad (4.64)$$

is the covariance matrix of the measurements.

For simplicity, we discuss the case for total intensity data. Goodman has given the SNR for speckle only case in [1]:

$$\text{SNR} = \frac{2}{1 + \text{DOP}^2}. \quad (4.65)$$

The SNR for Gaussian noise case is

$$\text{SNR}_G = 1 + \frac{2J_{xx}J_{yy} - \sigma^2}{J_{xx}^2 + J_{yy}^2 + \sigma^2}, \quad (4.66)$$

and the SNR for Poisson noise case is

$$\text{SNR}_P = 1 + \frac{2J_{xx}J_{yy} - \bar{I}}{J_{xx}^2 + J_{yy}^2 + \bar{I}}. \quad (4.67)$$

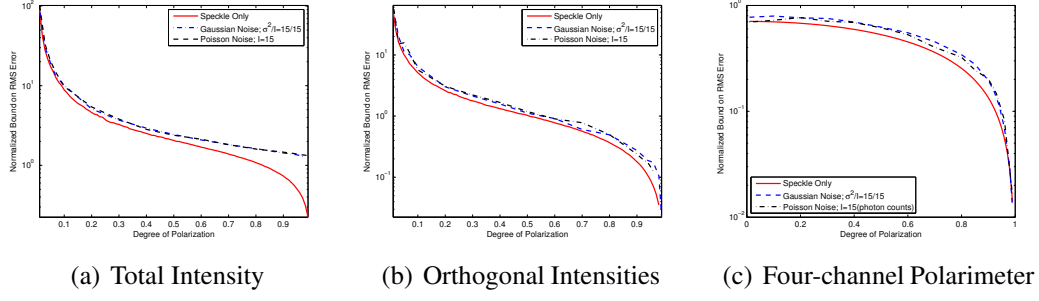
As shown in Fig. 4.7(a), we have the same performance bounds as expected when we set  $\bar{I} = \sigma^2 = 15$ , which makes  $\text{SNR}_G = \text{SNR}_P$ . Similarly, we find that the condition for equal SNR of orthogonal intensity measurements with Gaussian and Poisson noise is

$$\bar{I} = \sigma_x^2 + \sigma_y^2 = \sigma^2, \quad (4.68)$$

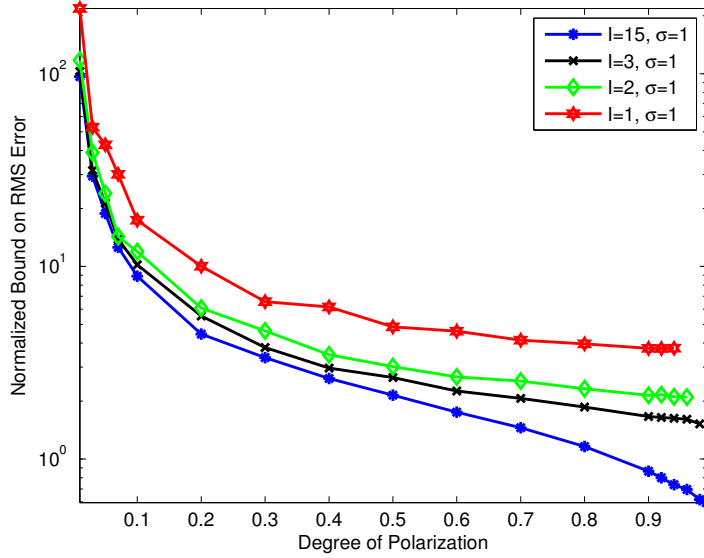
and the condition for four-channel polarimeter is

$$\bar{I} = \sigma_0^2 + \sigma_1^2 + \sigma_2^2 + \sigma_3^2 = \sigma^2. \quad (4.69)$$

Fig. 4.7(b) and Fig. 4.7(c) show the performance bounds for orthogonal intensity estimators and four-channel polarimeter estimators with equal SNRs. Fig. 4.8 shows the diagram of performance bounds vs. SNR. As predicted, the bound moves down while SNR becomes larger.



**Figure 4.7:** Normalized Cramer-Rao bounds on RMS error for estimators of the degree of polarization from measurements corrupted by Gaussian and Poisson noise with the same SNRs;  $\bar{I} = 15$  and  $\sigma^2 = 15$ .



**Figure 4.8:** Performance bounds with different SNRs for total intensity measurements with Gaussian noise.

#### 4.4.4 Summary

We have compared the performance bounds for the estimation of the degree of polarization from various sensing modalities. For each modality, we have also included noise models and utilize realistic data-collecting models. The simulated bounds show the expected

hierarchy of performances. Although estimators from four-channel polarimeter (intensity interferometer) data have a better performance than from orthogonal intensities data and total intensity data, processing the four intensities data from polarimeter requires complicated optical device, alignment, and four CCD detectors. It only requires one or two detectors and a computer to process orthogonal intensities data and total intensity data, and the bounds demonstrate that reasonable estimates may still be obtained from orthogonal intensities or total intensity data. Therefore, computational sensing is a promising way to estimate the degree of polarization. The proposed computational sensing estimators (maximum-likelihood estimators) are presented in the following chapters.



## Chapter 5

# Maximum-Likelihood Estimation from Total Intensity Data

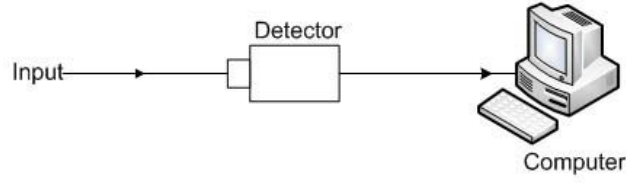
The schematic diagram for estimating the degree of polarization from total intensity data is shown in Fig. 5.1. A scene is illuminated with an active laser. A Charge-Coupled Device (CCD) camera collects the intensity data, and a computer processes the data.

We showed in section 3.3 that the degree of polarization can be estimated by the algebraic estimator:

$$\widehat{\text{DOP}} = \sqrt{\frac{2\text{VAR}[D]}{a^2\bar{I}^2} - \frac{2}{a\bar{I}} - \frac{2(b + \sigma^2)}{a^2\bar{I}^2} - 1}. \quad (5.1)$$

Because the degree of polarization is an unknown but deterministic quantity, the estimation can also be formed by using maximum-likelihood (ML) methods. The log-likelihood func-





**Figure 5.1:** Diagram of a computational sensor for total intensity data

tion of the total intensity data can be derived from the probability density function (PDF) of the data, which is given in Eq.(4.60). Then, a two-dimension active-set algorithm will find the ML estimate for the coherency matrix ( $J_{xx}$  and  $J_{yy}$ ).

$$\hat{J}_{ml} = \arg \max_J \sum_{k=1}^K \log P_D(d_k|J). \quad (5.2)$$

Then, the degree of polarization is

$$\widehat{\text{DOP}} = \frac{|\hat{J}_{xx} - \hat{J}_{yy}|}{\hat{J}_{xx} + \hat{J}_{yy}} \quad (5.3)$$

Alternatively, we can find the estimate by expectation-maximization (EM) algorithms. The algorithms for different noise models are presented in the following section.

## 5.1 EM algorithms for Total Intensity Sensing

Recalling Eq.(2.12), we note that the probability density function of the total intensity depends only on the eigenvalues of the coherence matrix and does not depend on the phase shift and coordinate rotation. Therefore, without loss of generality, we can assume the coherence matrix is diagonalized, i.e., the eigenvalues are  $J_{xx}$  and  $J_{yy}$ .

$$J = \begin{bmatrix} J_{xx} & 0 \\ 0 & J_{yy} \end{bmatrix}. \quad (5.4)$$

Then, the probability density function is

$$p_I(i) = \frac{1}{J_{xx} - J_{yy}} \left[ \exp\left(-\frac{i}{J_{xx}}\right) - \exp\left(-\frac{i}{J_{yy}}\right) \right]. \quad (5.5)$$

Consider the observed total intensity data,  $D$  as incomplete data and the intensities of the two orthogonal components,  $I_x$  and  $I_y$  as complete data. The joint probability density function of  $I_x$  and  $I_y$  is

$$\begin{aligned} p_{I_x, I_y}(i_x, i_y) &= \frac{1}{\det(J)} \exp\left(-\frac{J_{yy}i_x + J_{xx}i_y}{\det(J)}\right) I_0\left(2\frac{|J_{xy}|}{\det(J)}\sqrt{i_x i_y}\right) \\ &= \frac{1}{J_{xx}J_{yy}} \exp\left(-\frac{J_{yy}i_x + J_{xx}i_y}{J_{xx}J_{yy}}\right). \end{aligned} \quad (5.6)$$

The log-likelihood function of the complete data is

$$L_{CD} = \ln \left( \prod_{k=1}^K p_{I_x, I_y}(i_{x_k}, i_{y_k} | J_{xx}, J_{yy}) \right) = -K \ln(J_{xx}) - K \ln(J_{yy}) - \sum_{k=1}^K \frac{i_{x_k}}{J_{xx}} - \sum_{k=1}^K \frac{i_{y_k}}{J_{yy}}, \quad (5.7)$$

where  $i_{x_k}$  and  $i_{y_k}$  are the samples of the orthogonal intensities, and  $K$  is the total number of samples.

### Expectation Step:

$$\begin{aligned} Q(J_{xx}, J_{yy} | J_{xx}^{[n]}, J_{yy}^{[n]}) &= E \left[ L_{CD} | D, J_{xx}^{[n]}, J_{yy}^{[n]} \right] \\ &= -K \ln(J_{xx}) - K \ln(J_{yy}) - \sum_{k=1}^K \frac{E \left[ I_{xk} | D_k, J_{xx}^{[n]}, J_{yy}^{[n]} \right]}{J_{xx}} - \sum_{k=1}^K \frac{E \left[ I_{yk} | D_k, J_{xx}^{[n]}, J_{yy}^{[n]} \right]}{J_{yy}} \end{aligned} \quad (5.8)$$

where  $D = \begin{bmatrix} D_1 & D_2 & \dots & D_K \end{bmatrix}$  is the observed data, sampled total intensity.

### Maximization Step:

Let  $J_{xx}^{[n+1]}, J_{yy}^{[n+1]}$  be the values that maximize  $Q(J_{xx}, J_{yy} | J_{xx}^{[n]}, J_{yy}^{[n]})$

$$\begin{bmatrix} J_{xx}^{[n+1]} & J_{yy}^{[n+1]} \end{bmatrix} = \arg \max_{J_{xx}, J_{yy}} Q(J_{xx}, J_{yy} | J_{xx}^{[n]}, J_{yy}^{[n]}). \quad (5.9)$$

If we compute the derivative with respect to  $J_{xx}$  and  $J_{yy}$  and equate the results to zero, then

we have

$$J_{xx}^{[n+1]} = \frac{1}{K} \sum_{k=1}^K \mathbb{E} \left[ I_{xx} | D_k, J_{xx}^{[n]}, J_{yy}^{[n]} \right] \quad (5.10a)$$

$$J_{yy}^{[n+1]} = \frac{1}{K} \sum_{k=1}^K \mathbb{E} \left[ I_{yy} | D_k, J_{xx}^{[n]}, J_{yy}^{[n]} \right] \quad (5.10b)$$

The two steps are iterated until convergence, which gives the ML estimate,  $\hat{J}_{xx}$  and  $\hat{J}_{yy}$ , then, the estimate for the degree of polarization is

$$\widehat{\text{DOP}} = \frac{|\hat{J}_{xx} - \hat{J}_{yy}|}{\hat{J}_{xx} + \hat{J}_{yy}} \quad (5.11)$$

### 5.1.1 Laser Speckle Only

For an ideal detector (no detection noise exists), only speckle noise affects our estimates.

In such case, the observation is the total intensity of the back-scattered field,  $D = I_t$ . The conditional expectations are

$$\mathbb{E} \left[ I_{xx} | D_k, J_{xx}^{[n]}, J_{yy}^{[n]} \right] = D_k \frac{\exp \left( -\frac{D_k}{J_{xx}^{[n]}} \right)}{\exp \left( -\frac{D_k}{J_{xx}^{[n]}} \right) - \exp \left( -\frac{D_k}{J_{yy}^{[n]}} \right)} - \frac{J_{xx}^{[n]} J_{yy}^{[n]}}{J_{xx}^{[n]} - J_{yy}^{[n]}} \quad (5.12a)$$

$$\mathbb{E} \left[ I_{yy} | D_k, J_{xx}^{[n]}, J_{yy}^{[n]} \right] = D_k \frac{\exp \left( -\frac{D_k}{J_{yy}^{[n]}} \right)}{\exp \left( -\frac{D_k}{J_{yy}^{[n]}} \right) - \exp \left( -\frac{D_k}{J_{xx}^{[n]}} \right)} - \frac{J_{xx}^{[n]} J_{yy}^{[n]}}{J_{yy}^{[n]} - J_{xx}^{[n]}} \quad (5.12b)$$

The detailed derivation is given in Appendix B.2.1.

### 5.1.2 Laser Speckle with Gaussian Noise

If we consider the Gaussian read-out noise, the total intensity data is corrupted by zero-mean Gaussian noise

$$D = I_x + I_y + N = I_t + N \quad (5.13)$$

where  $N \sim (0, \sigma^2)$  is a Gaussian random variable with mean zero and variance  $\sigma^2$ .

The conditional expectations for updating the coherency matrix are:

$$\begin{aligned} & \mathbb{E} \left[ I_{x_k} | D_k, J_{xx}^{[n]}, J_{yy}^{[n]} \right] \quad (5.14a) \\ &= \frac{\exp \left( -\frac{D_k}{J_{xx}^{[n]}} + \frac{\sigma^2}{2J_{xx}^{[n]2}} \right) \left( D_k - \frac{\sigma^2}{J_{xx}^{[n]}} - \frac{J_{xx}^{[n]} J_{yy}^{[n]}}{J_{xx}^{[n]} - J_{yy}^{[n]}} \right) + \frac{J_{xx}^{[n]} J_{yy}^{[n]}}{J_{xx}^{[n]} - J_{yy}^{[n]}} \exp \left( -\frac{D_k}{J_{yy}^{[n]}} + \frac{\sigma^2}{2J_{yy}^{[n]2}} \right)}{\exp \left( -\frac{D_k}{J_{xx}^{[n]}} + \frac{\sigma^2}{2J_{xx}^{[n]2}} \right) \Phi \left( \frac{D_k}{\sigma} - \frac{\sigma}{J_{xx}^{[n]}} \right) - \exp \left( -\frac{D_k}{J_{yy}^{[n]}} + \frac{\sigma^2}{2J_{yy}^{[n]2}} \right) \Phi \left( \frac{D_k}{\sigma} - \frac{\sigma}{J_{yy}^{[n]}} \right)} \end{aligned}$$

$$\begin{aligned} & \mathbb{E} \left[ I_{y_k} | D_k, J_{xx}^{[n]}, J_{yy}^{[n]} \right] \quad (5.14b) \\ &= \frac{\exp \left( -\frac{D_k}{J_{yy}^{[n]}} + \frac{\sigma^2}{2J_{yy}^{[n]2}} \right) \left( D_k - \frac{\sigma^2}{J_{yy}^{[n]}} - \frac{J_{yy}^{[n]} J_{xx}^{[n]}}{J_{yy}^{[n]} - J_{xx}^{[n]}} \right) + \frac{J_{yy}^{[n]} J_{xx}^{[n]}}{J_{yy}^{[n]} - J_{xx}^{[n]}} \exp \left( -\frac{D_k}{J_{xx}^{[n]}} + \frac{\sigma^2}{2J_{xx}^{[n]2}} \right)}{\exp \left( -\frac{D_k}{J_{yy}^{[n]}} + \frac{\sigma^2}{2J_{yy}^{[n]2}} \right) \Phi \left( \frac{D_k}{\sigma} - \frac{\sigma}{J_{yy}^{[n]}} \right) - \exp \left( -\frac{D_k}{J_{xx}^{[n]}} + \frac{\sigma^2}{2J_{xx}^{[n]2}} \right) \Phi \left( \frac{D_k}{\sigma} - \frac{\sigma}{J_{xx}^{[n]}} \right)} \end{aligned}$$

where the function  $\Phi(x)$  is the standard normal cumulative distribution function. The detailed derivation is given in Appendix B.2.2.

### 5.1.3 Laser Speckle with Poisson Noise

If we consider Poisson noise exists and ignore Gaussian read-out noise, we have the following statistical model:

$$\begin{aligned}
 D &= \text{Poisson}(a(I_x + I_y) + b) \quad a > 0, b \geq 0 \\
 &= \text{Poisson}(aI_x) + \text{Poisson}(aI_y) + \text{Poisson}(b) \\
 &= N_x + N_y + N_b
 \end{aligned} \tag{5.15}$$

where  $a$  is the gain of the CCD detector and  $b$  is the bias (dark current) of the detector.

For Poisson noise case, we choose  $N_x$  and  $N_y$  as the complete data and assume  $a = 1$  and there is no bias. In this case,  $D = N_x + N_y$ . The log-likelihood function of the complete data is

$$L_{CD} = \ln \left( \prod_{k=1}^K p_{N_x, N_y}(n_{x_k}, n_{y_k} | J_{xx}, J_{yy}) \right) \tag{5.16}$$

where  $n_{x_k}$  and  $n_{y_k}$  are the samples of the orthogonal intensities, and  $K$  is the total number of samples. Because  $I_x$  and  $I_y$  are independent, from Eq.(5.6), we have

$$p_{I_x}(i) = \frac{1}{J_{xx}} \exp \left( \frac{-i}{J_{xx}} \right) \tag{5.17}$$

$$p_{I_y}(i) = \frac{1}{J_{yy}} \exp\left(\frac{-i}{J_{yy}}\right) \quad (5.18)$$

The probability mass function (PMF) for  $N_x$  is

$$\begin{aligned} p_{N_x}(n) &= \int_0^\infty p_{N_x|I_x}(n|i) p_{I_x}(i) di \\ &= \int_0^\infty \frac{\exp(-i)i^n}{n!} \times \frac{1}{J_{xx}} \exp\left(\frac{-i}{J_{xx}}\right) di \\ &= \frac{(1 + J_{xx}^{-1})^{-n-1}}{J_{xx}}. \end{aligned} \quad (5.19)$$

Similarly, we have the PMF for  $N_y$

$$p_{N_y}(n) = \frac{(1 + J_{yy}^{-1})^{-n-1}}{J_{yy}}. \quad (5.20)$$

Therefore, the log-likelihood function of the complete data is

$$L_{\text{cd}} = -K \ln(J_{xx}) - K \ln(J_{yy}) - \sum_{k=1}^K (n_{x_k} + 1) \ln(1 + J_{xx}^{-1}) - \sum_{k=1}^K (n_{y_k} + 1) \ln(1 + J_{yy}^{-1}). \quad (5.21)$$

**Expectation Step:**

$$\begin{aligned} Q(J_{xx}, J_{yy} | J_{xx}^{[n]}, J_{yy}^{[n]}) &= E \left[ L_{CD} | D, J_{xx}^{[n]}, J_{yy}^{[n]} \right] \\ &= -K \ln(J_{xx}) - K \ln(J_{yy}) - \sum_{k=1}^K \left( E[n_{x_k} | D_k, J_{xx}^{[n]}, J_{yy}^{[n]}] + 1 \right) \ln(1 + J_{xx}^{-1}) \\ &\quad - \sum_{k=1}^K \left( E[n_{y_k} | D_k, J_{xx}^{[n]}, J_{yy}^{[n]}] + 1 \right) \ln(1 + J_{yy}^{-1}). \end{aligned} \quad (5.22)$$

**Maximization Step:**

Let  $J_{xx}^{[n+1]}, J_{yy}^{[n+1]}$  be the values that maximize  $Q(J_{xx}, J_{yy} | J_{xx}^{[n]}, J_{yy}^{[n]})$

$$\left[ J_{xx}^{[n+1]}, J_{yy}^{[n+1]} \right] = \arg \max_{J_{xx}, J_{yy}} Q(J_{xx}, J_{yy} | J_{xx}^{[n]}, J_{yy}^{[n]}) \quad (5.23)$$

Compute the derivative with respect to  $J_{xx}$  and  $J_{yy}$  and equate the results to zero. We find

$$J_{xx}^{[n+1]} = \frac{1}{K} \sum_{k=1}^K E \left[ N_{xk} | D_k, J_{xx}^{[n]}, J_{yy}^{[n]} \right] \quad (5.24a)$$

$$J_{yy}^{[n+1]} = \frac{1}{K} \sum_{k=1}^K E \left[ N_{yk} | D_k, J_{xx}^{[n]}, J_{yy}^{[n]} \right] \quad (5.24b)$$

We show in Appendix B.2.3 that the above conditional expectations are

$$J_{xx}^{[n+1]} = \frac{r}{1-r} - \frac{1}{K} \sum_{k=1}^K (D_k + 1) \frac{r^{D_k+1}}{1-r^{D_k+1}} \quad (5.25a)$$

$$J_{yy}^{[n+1]} = \frac{r^{-1}}{1-r^{-1}} - \frac{1}{K} \sum_{k=1}^K (D_k + 1) \frac{r^{-D_k-1}}{1-r^{-D_k-1}} \quad (5.25b)$$

where  $r = \frac{1+1/J_{yy}^{[n]}}{1+1/J_{xx}^{[n]}}$ .



### 5.1.4 Laser Speckle with Combined Gaussian and Poisson Noise

Considering the noise model discussed in Sec 2.3, which takes account of both Gaussian and Poisson noise, the statistics of the collected noisy intensity data is

$$D = I + N, \quad (5.26)$$

where

$$I \sim \text{Poisson}(\lambda) = \text{Poisson}(a(I_x + I_y) + b), \quad (5.27)$$

$$N \sim N(0, \sigma^2), \quad (5.28)$$

$a$  is the gain of the detector and  $b$  is the bias of the detector.

Following the above statistical model, we find the conditional expectations are

$$\begin{aligned} & \mathbb{E} \left[ I_{xk} | D_k, J_{xx}^{[n]}, J_{yy}^{[n]} \right] \\ &= \frac{\sum_{m=0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(D_k - m)^2}{2\sigma^2} \right] \exp \left( \frac{b}{aJ_{xx}^{[n]}} \right) [\text{Ei}(m+1, b, X) - b\text{Ei}(m, b, X)]}{a \sum_{m=0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(D_k - m)^2}{2\sigma^2} \right] \left[ \exp \left( \frac{b}{aJ_{xx}^{[n]}} \right) \text{Ei}(m, b, X) - \exp \left( \frac{b}{aJ_{yy}^{[n]}} \right) \text{Ei}(m, b, Y) \right]} - \frac{J_{xx}^{[n]} J_{yy}^{[n]}}{J_{xx}^{[n]} - J_{yy}^{[n]}} \end{aligned} \quad (5.29)$$

$$\begin{aligned}
& \mathbb{E} \left[ I_{y_k} | D_k, J_{xx}^{[n]}, J_{yy}^{[n]} \right] \\
&= \frac{\sum_{m=0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(D_k-m)^2}{2\sigma^2} \right] \exp \left( \frac{b}{aJ_{yy}^{[n]}} \right) [\text{Ei}(m+1, b, Y) - b\text{Ei}(m, b, Y)]}{a \sum_{m=0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(D_k-m)^2}{2\sigma^2} \right] \left[ \exp \left( \frac{b}{aJ_{yy}^{[n]}} \right) \text{Ei}(m, b, Y) - \exp \left( \frac{b}{aJ_{xx}^{[n]}} \right) \text{Ei}(m, b, X) \right]} - \frac{J_{xx}^{[n]} J_{yy}^{[n]}}{J_{yy}^{[n]} - J_{xx}^{[n]}}
\end{aligned} \tag{5.30}$$

and

$$X = \left( 1 + \frac{1}{aJ_{xx}^{[n]}} \right), \quad Y = \left( 1 + \frac{1}{aJ_{yy}^{[n]}} \right), \tag{5.31a}$$

$$\text{Ei}(n, b, z) = \int_b^{\infty} \lambda^n \exp(-z\lambda) d\lambda = \begin{cases} z^{-n-1} \Gamma(n+1, bz) & b > 0 \\ z^{-n-1} n! & b = 0 \end{cases} \tag{5.31b}$$

The detailed derivation is given in Appendix B.2.4.

### 5.1.5 Summary

The EM algorithm can be summarized as:

1. Collect total intensity samples,  $D = \begin{bmatrix} D_1 & D_2 & \dots & D_K \end{bmatrix}$
2. Choose initial values for the parameters,  $J_{xx}^{[0]}$  and  $J_{yy}^{[0]}$ .
3. Calculate the conditional expectations by using the equations given above for the

corresponding noise model.

4. Update the parameters with the conditional expectations
5. Repeat 3 and 4 until convergence (the stopping criteria is met) and record the convergence results as  $\hat{J}_{xx}$  and  $\hat{J}_{yy}$ .
6. Compute the degree of polarization from the estimated coherency matrix

$$\widehat{\text{DOP}} = \frac{|\hat{J}_{xx} - \hat{J}_{yy}|}{\hat{J}_{xx} + \hat{J}_{yy}}.$$

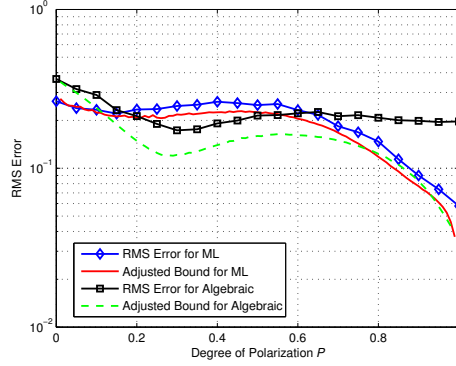
As shown, however, the EM algorithms for Gaussian noise case and combined Gaussian and Poisson noise case are rather complicated, which require large computing power. Meanwhile, because of the numerical error of the standard normal cumulative distribution function  $\Phi(x)$ , the EM algorithm is unstable to converge or gives a bad estimate. Therefore, for these two cases, we choose an active-set algorithm to maximize the log-likelihood function directly. The simulation results are given in the following section.

## 5.2 Simulation

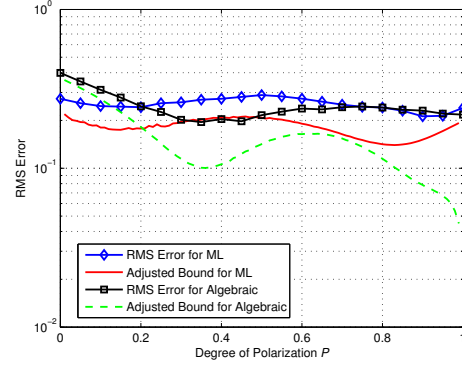
Fig.5.2 shows the performances of the estimators and comparison with corresponding modified Cramer-Rao bounds on RMS estimation error for biased estimators of the degree of polarization. Each sub-figure shows the performances and bounds for different noise mod-

els: noise free, Gaussian, Poisson and combined Gaussian and Poisson. For the sake of simplicity here, we choose the unit as photons, ignore the dark current, and assume the *quantum efficiency* [16] as one. Read noise in widely available visible wavelength cameras typically lies in the range of 5 to 100 RMS electrons per pixel per readout. Therefore, for a typical Gaussian noise case, the simulation parameters are set as: average total intensity is  $\bar{I} = 100$  photons per pixel per readout and the read out noise level is  $\sigma = 30$  RMS photons per pixel per readout. For Poisson noise case, we simulate a low incident light situation where  $\bar{I} = 15$  photons per pixel per readout. For combined Gaussian and Poisson noise case,  $\bar{I} = 50$  photons per pixel per readout and  $\sigma = 10$  RMS photons per pixel per readout. The coordinate rotation angle  $\theta$  and the relative phase shift  $\phi_{xy}$  in Eq. (2.4) are set to be  $\pi/3$  and  $\pi/4$  respectively in all simulations. Each performance bound is adjusted for the bias of each estimator.

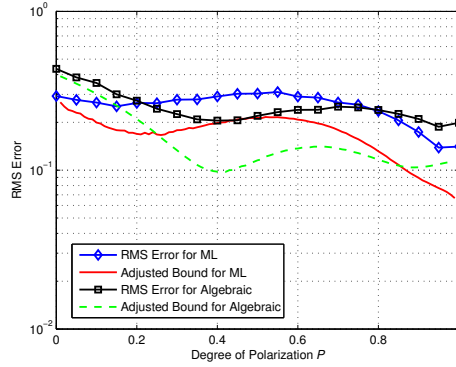
Algebraic estimators outperform maximum-likelihood estimators when the degree of polarization is somewhere between 0.2 and 0.7. Maximum-likelihood estimators are more efficient because the root mean square errors are closer to the bounds. We usually have more interest in highly polarized light (DOP is close to 1) which brings out the features of the scene, such as the first example in Chapter 1. In this region, maximum likelihood estimator is not only more efficient but also have smaller RMS error. If the degree of polarization is more likely to fall in the middle region between 0 and 1, one may choose the algebraic estimator.



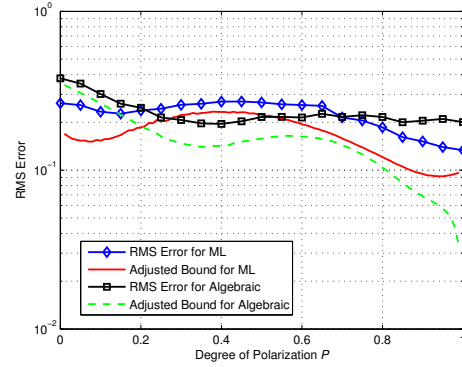
(a) Speckle only



(b) Gaussian noise



(c) Poisson noise



(d) Combined Gaussian and Poisson noise

**Figure 5.2:** Comparison of the performances of ML estimators and algebraic estimators with corresponding modified Cramer-Rao bounds on RMS estimation error for biased estimators of the DOP. (a) detector noise free. 100 speckle realizations; (b) Gaussian read out noise.  $\bar{I} = 100$ ,  $\sigma = 30$ , 100 speckle realizations; (c) Poisson shot noise.  $\bar{I} = 15$ , 100 speckle realizations; (d) combined Gaussian and Poisson noise.  $\bar{I} = 50$ ,  $\sigma = 10$ , 100 speckle realizations.

## Chapter 6

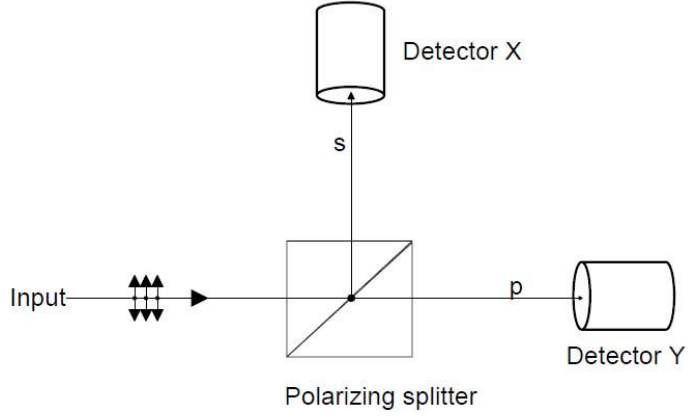
### Estimation from Orthogonal Intensity

#### Data

The degree of polarization of a light beam is in principle determined from measurements of the correlation in the outputs of two photoelectric detectors illuminated by the orthogonal components of the beam [19]. Fig.6.1 shows the schematic diagram of the system. The orthogonal intensities data are denoted as

$$D = \begin{bmatrix} D_x \\ D_y \end{bmatrix} = \begin{bmatrix} \text{Poisson}(aI_x + b) + N_x \\ \text{Poisson}(aI_y + b) + N_y \end{bmatrix} = \begin{bmatrix} L_x + N_x \\ L_y + N_y \end{bmatrix}, \quad (6.1)$$

where  $N_x$  and  $N_y$  are independent Gaussian random variables with zero means and variances of  $\sigma^2$ ,  $a$  is the gain of the amplifier, and  $b$  is the dark current. Here, we assume two detectors



**Figure 6.1:** Schematic diagram of orthogonal intensities sensing

are identical with the same gain, bias and read out noise.

## 6.1 Algebraic Estimator

Having both the speckle model and the detector noise model, we find that the means of the orthogonal intensity data are

$$E[D_x] = aI_x + b = aJ_{xx} + b, \quad (6.2)$$

$$E[D_y] = aI_y + b = aJ_{yy} + b. \quad (6.3)$$

and correlation between the orthogonal intensities data is

$$E[D_x D_y] = a^2 J_{xx} J_{yy} + a^2 |J_{xy}|^2 + ab(J_{xx} + J_{yy}) + b^2. \quad (6.4)$$

Solving the above three equations, we have

$$J_{xx} = \frac{E[D_x] - b}{a}, \quad (6.5a)$$

$$J_{yy} = \frac{E[D_y] - b}{a}, \quad (6.5b)$$

$$|J_{xy}|^2 = \frac{E[D_x D_y]}{a^2} - J_{xx} J_{yy} - \frac{b}{a}(J_{xx} + J_{yy}) + \frac{b^2}{a^2}. \quad (6.5c)$$

By using the sample means and correlation, we have an *algebraic estimate* of the degree of polarization [31]:

$$\begin{aligned} \text{DOP} &= \sqrt{1 - 4 \frac{\det(J)}{[\text{tr}(J)]^2}} \\ &= \sqrt{1 - 4 \frac{J_{xx} J_{yy} - |J_{xy}|^2}{(J_{xx} + J_{yy})^2}} \end{aligned} \quad (6.6)$$

and  $J_{xx}$ ,  $J_{yy}$  and  $|J_{xy}|^2$  are given in Eq.(6.5)

## 6.2 Maximum-Likelihood Estimator

The instantaneous intensities for the two orthogonal components are defined as

$$I_x = |U_x|^2 \quad (6.7)$$

$$I_y = |U_y|^2, \quad (6.8)$$



and the joint PDF for the orthogonal intensities is [12]

$$p_{I_x, I_y}(i_x, i_y) = \frac{1}{\det(J)} \exp\left(-\frac{J_{yy}i_x + J_{xx}i_y}{\det(J)}\right) I_0\left(2\frac{|J_{xy}|}{\det(J)}\sqrt{i_x i_y}\right), \quad (6.9)$$

where  $I_0$  is the zero-order modified Bessel function of the first kind. Note that the joint PDF for orthogonal intensities depends only on  $J_{xx}$ ,  $J_{yy}$  and the magnitude of  $J_{xy}$ . Considering the detector noise, we found the joint PDF of orthogonal intensities data in section 4.2.4:

$$\begin{aligned} p_{D_x, D_y}(d_x, d_y) &= \int_0^\infty \int_0^\infty \sum_{m=0}^\infty \sum_{k=0}^\infty \frac{(ai_x + b)^m \exp[-(ai_x + b)]}{m!} \frac{(ai_y + b)^k \exp[-(ai_y + b)]}{k!} \\ &\quad \times \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d_x - m)^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d_y - k)^2}{2\sigma^2}\right) \\ &\quad \times \frac{1}{\det(J)} \exp\left(-\frac{J_{yy}i_x + J_{xx}i_y}{\det(J)}\right) I_0\left(2\frac{|J_{xy}|}{\det(J)}\sqrt{i_x i_y}\right) di_x di_y. \end{aligned} \quad (6.10)$$

$J_{xx}$  and  $J_{yy}$  can be estimated by Eq.(6.5), and  $|J_{xy}|$  can be estimated by maximizing the log-likelihood function

$$L(|J_{xy}|) = \sum_{k=1}^K \log p_{D_x, D_y}(d_{x_k}, d_{y_k}). \quad (6.11)$$

Therefore, the maximum-likelihood estimator from orthogonal intensities data is

$$\hat{J}_{xx} = \frac{E[D_x] - b}{a}, \quad (6.12a)$$

$$\hat{J}_{yy} = \frac{E[D_y] - b}{a}, \quad (6.12b)$$

$$|\hat{J}_{xy}| = \arg \max_{|J_{xy}|} L(|J_{xy}|) \quad (6.12c)$$

### 6.2.1 Expectation-Maximization Algorithm

Similar to maximum-likelihood estimators for total intensity sensing, a maximum-likelihood estimate can be formed by an alternative expectation-maximization (EM) algorithm. The first order statistical model of the field's complex amplitudes is:

$$p_U(u|J) = \frac{1}{\pi^2 \det(J)} \exp(-u^\dagger J^{-1} u), \quad (6.13)$$

where  $\det(J) = J_{xx}J_{yy} - |J_{xy}|^2$  is the determinant of the coherency matrix, and  $J^{-1}$  is its inverse. The log-likelihood function of the complete data is:

$$\begin{aligned} L_{CD} &= \ln \left( \prod_{k=1}^K p_U(u_k|J) \right) \\ &= -2K \ln \pi - K \ln \det(J) - \sum_{k=1}^K u_k^\dagger J^{-1} u_k \\ &= -2K \ln \pi - K \ln \det(J) - \text{Trace} \left[ J^{-1} \sum_{k=1}^K u_k^\dagger u_k \right] \end{aligned} \quad (6.14)$$

E-step: Compute:

$$\begin{aligned} Q(J|J^{[n]}) &= E \left[ L_{CD} | I, J^{[n]} \right] \\ &= -2K \ln \pi - K \ln \det(J) - \text{Trace} \left[ J^{-1} \sum_{k=1}^K E \left[ u_k^\dagger u_k | I_k, J^{[n]} \right] \right] \end{aligned} \quad (6.15)$$

where

$$I_k = \begin{bmatrix} I_{xk} \\ I_{yk} \end{bmatrix} = \begin{bmatrix} |u_{xk}|^2 \\ |u_{yk}|^2 \end{bmatrix} \quad (6.16)$$

is the mean of the sampled orthogonal intensities, which are our observed data.

M-Step: Let  $J^{[n+1]}$  be the value of  $J$  that maximizes  $Q(J|J^{[n]})$ :

$$J^{[n+1]} = \arg \max_{\theta} Q \left( J | J^{[n]} \right) \quad (6.17)$$

We compute the derivative with respect to  $J$  and equate the result to zero:

$$\frac{\partial}{\partial J} \left( -2K \ln \pi - K \ln \det(J) - \text{Trace} \left[ J^{-1} \sum_{k=1}^K E \left[ u_k^\dagger u_k | I_k, J^{[n]} \right] \right] \right) = 0, \quad (6.18)$$

and we have

$$J^{[n+1]} = \frac{1}{K} \sum_{k=1}^K E \left[ u_k^\dagger u_k | I_k, J^{[n]} \right]. \quad (6.19)$$

Therefore,

$$\begin{aligned} J^{[n+1]} &= \frac{1}{K} \sum_{k=1}^K E \begin{bmatrix} |u_{xk}|^2 & u_{xk} u_{yk}^* \\ u_{xk}^* u_{yk} & |u_{yk}|^2 \end{bmatrix} \bigg|_{I_{xk}, I_{yk}, J^{[n]}} \\ &= \frac{1}{K} \sum_{k=1}^K \begin{bmatrix} I_{xk} & E \left[ u_{xk} u_{yk}^* | I_{xk}, I_{yk}, J^{[n]} \right] \\ \left( E \left[ u_{xk} u_{yk}^* | I_{xk}, I_{yk}, J^{[n]} \right] \right)^* & I_{yk} \end{bmatrix}. \end{aligned} \quad (6.20)$$

The conditional expectations for different noise models are derived in details in Appendix

B.1. After having the conditional expectations for different noise models, we have the the EM algorithms for each noise model.

### 6.2.1.1 Speckle Only

The EM algorithm for speckle only is

$$J^{[n+1]} = \frac{1}{K} \sum_{k=1}^K \begin{bmatrix} I_{xk} & E \left[ u_{xk} u_{yk}^* | I_{xk}, I_{yk}, J^{[n]} \right] \\ (E \left[ u_{xk} u_{yk}^* | I_{xk}, I_{yk}, J^{[n]} \right])^* & I_{yk} \end{bmatrix} \quad (6.21)$$

where

$$E \left[ u_{xk} u_{yk}^* | I_{xk}, I_{yk}, J^{[n]} \right] = \sqrt{I_{xk} I_{yk}} \frac{I_1 \left( \frac{2|J_{xy}^{[n]}| \sqrt{I_{xk} I_{yk}}}{\det(J^{[n]})} \right)}{I_0 \left( \frac{2|J_{xy}^{[n]}| \sqrt{I_{xk} I_{yk}}}{\det(J^{[n]})} \right)} \exp(j \angle J_{xy}^{[n]}). \quad (6.22)$$

### 6.2.1.2 Speckle with Gaussian Noise

The EM algorithm for estimation of the DOP from orthogonal intensities data with Gaussian noise is:

$$J^{[n+1]} = \frac{1}{M} \sum_{m=1}^M \begin{bmatrix} D_{x_m} & E \left[ u_{x_m} u_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]} \right] \\ \left( E \left[ u_{x_m} u_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]} \right] \right)^* & D_{y_m} \end{bmatrix} \quad (6.23)$$

where

$$E \left[ U_{x_m} U_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]} \right] = \exp(j\phi_{xy}) \frac{\sum_{k=0}^{\infty} \frac{\left( \frac{|J_{xy}^{[n]}|}{|J^{[n]}|} \right)^{2k+1}}{k!(k+1)!} \text{Inx}[k+1] \text{Iny}[k+1]}{\sum_{k=0}^{\infty} \frac{\left( \frac{|J_{xy}^{[n]}|}{|J^{[n]}|} \right)^{2k}}{k!k!} \text{Inx}[k] \text{Iny}[k]} \quad (6.24)$$

and

$$\text{Inx}[k+2] = (k+1)\sigma_x^2 \text{Inx}[k] + \mu_x \text{Inx}[k+1] \quad (6.25a)$$

$$\text{Iny}[k+2] = (k+1)\sigma_y^2 \text{Inx}[k] + \mu_y \text{Iny}[k+1] \quad (6.25b)$$

$$\text{Inx}[0] = 1 - \Phi\left(\frac{\mu_x}{\sigma_x}\right) \quad (6.25c)$$

$$\text{Inx}[1] = \sigma_x^2 \exp\left(-\frac{\mu_x^2}{2\sigma_x^2}\right) + \mu_x \text{Inx}[0] \quad (6.25d)$$

$$\text{Iny}[0] = 1 - \Phi\left(\frac{\mu_y}{\sigma_y}\right) \quad (6.25e)$$

$$\text{Iny}[1] = \sigma_y^2 \exp\left(-\frac{\mu_y^2}{2\sigma_y^2}\right) + \mu_y \text{Iny}[0] \quad (6.25f)$$

$$\mu_{x_m} = D_{x_m} - \frac{J_{yy}\sigma_x^2}{|J|} \quad (6.25g)$$

$$\mu_{y_m} = D_{y_m} - \frac{J_{xx}\sigma_y^2}{|J|} \quad (6.25h)$$

$$\phi_{xy} = \angle J_{xy} \quad (6.25i)$$

### 6.2.1.3 Speckle with Poisson Noise

The EM algorithm for estimation of the DOP from orthogonal intensities data with Poisson detector noise is:

$$J^{[n+1]} = \frac{1}{M} \sum_{m=1}^M \begin{bmatrix} D_{x_m} & E[u_{x_m} u_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]}] \\ \left( E[u_{x_m} u_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]}] \right)^* & D_{y_m} \end{bmatrix} \quad (6.26)$$

where

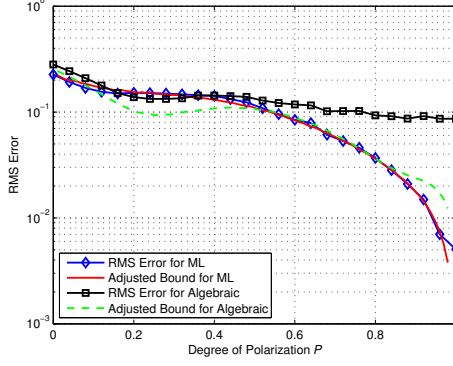
$$E[u_{x_m} u_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]}] = \exp(j\phi_{xy}) \frac{\sum_{k=0}^{\infty} \left( \frac{|J_{xy}^{[n]}|}{|J^{[n]}|} \right)^{2k+1} \frac{(D_{x_m}+k+1)!(D_{y_m}+k+1)!}{(k+1)!k!} A^{-k-2} B^{-k-2}}{\sum_{k=0}^{\infty} \left( \frac{|J_{xy}^{[n]}|}{|J^{[n]}|} \right)^{2k} \frac{(D_{x_m}+k)!(D_{y_m}+k)!}{k!k!} A^{-k-1} B^{-k-1}} \quad (6.27)$$

and

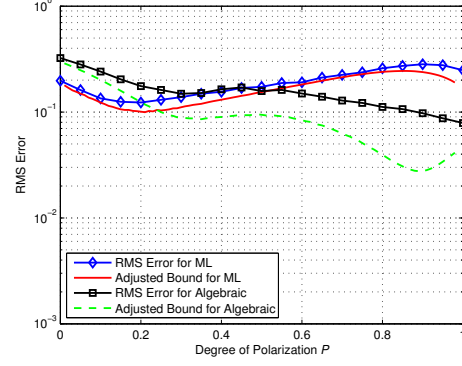
$$A = a + J_{yy}^{[n]} / |J|, \quad B = a + J_{xx}^{[n]} / |J|. \quad (6.28)$$

## 6.3 Simulation

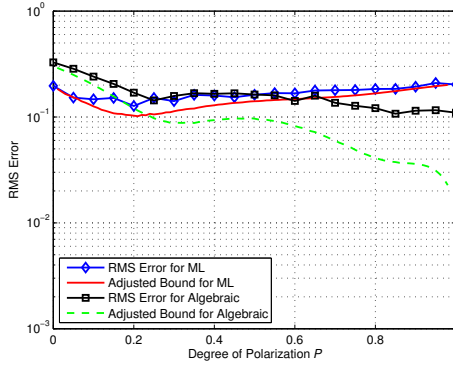
Although we have the analytical expressions for the maximum-likelihood estimators, it is difficult, therefore, impractical to implement. The log likelihood function is a double inte-



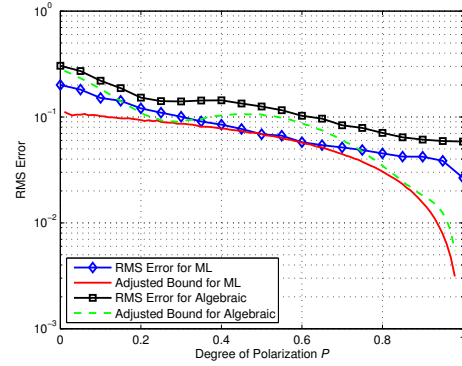
(a) Speckle only



(b) Gaussian noise



(c) Poisson noise



(d) Combined Gaussian and Poisson noise

**Figure 6.2:** Comparison of the performances of ML estimator and algebraic estimators with corresponding modified Cramer-Rao bounds on RMS estimation error for biased estimators of the DOP. (a) detector noise free. 100 speckle realizations; (b) Gaussian read out noise.  $\bar{I} = 100$ ,  $\sigma = 30$ , 100 speckle realizations; (c) Poisson shot noise.  $\bar{I} = 15$ , 100 speckle realizations; (d) combined Gaussian and Poisson noise.  $\bar{I} = 50$ ,  $\sigma = 10$ , 100 speckle realizations.

gral of the summation of infinite terms, and EM algorithms with detector noise (Gaussian or Poisson) are not closed-form expressions. The EM algorithm for the speckle only case, however, is very simple. We can use this algorithm to process the speckle data with detector noise. Simulation results are showed in Fig.6.2, and the simulation parameters are the same as total intensity sensing.

Fig.6.2 shows the performances of the estimators and comparison with corresponding modified Cramer-Rao bounds on RMS estimation error for biased estimators of the degree of polarization. Each sub-figure shows the estimation performances and the adjusted bounds for different noise models: noise free, Gaussian, Poisson and combined Gaussian and Poisson. As total intensity sensing, we choose the unit as photons, ignore the dark current, and assume the *quantum efficiency* as one. For a typical Gaussian noise case, the simulation parameters are set as: average total intensity is  $\bar{I} = 100$  photons per pixel per readout and the read out noise level is  $\sigma = 30$  RMS photons per pixel per readout. For Poisson noise case, we simulate a low incident light situation where  $\bar{I} = 15$  photons per pixel per readout. For combined Gaussian and Poisson noise case,  $\bar{I} = 50$  photons per pixel per readout and  $\sigma = 10$  RMS photons per pixel per readout. The coordinate rotation angle  $\theta$  and the relative phase shift  $\phi_{xy}$  in Eq. (2.4) are set to be  $\pi/3$  and  $\pi/4$  respectively in all simulations. Each performance bound is adjusted for the bias of each estimator.

The performances of the simple EM algorithms are close to efficient. The root-mean-square (RMS) errors of algebraic estimators, however, can be smaller than EM estimators if the degree of polarization is between 0.5 and 1 and there are only Gaussian or only Poisson noise. If we consider both Gaussian and Poisson noise, which is the general case, the EM estimator outperforms the algebraic estimator.





# Chapter 7

## Conclusion

We have computed the performance bounds for the estimation of the degree of polarization from various sensing modalities: fully polarimetric sensing, orthogonal intensities sensing, and total intensity sensing. For each modality, we have also included noise models and utilized realistic data-collecting models.

The performance bounds show the expected hierarchy of performances. Estimators for four-channel polarimeter (intensity interferometer) sensing have a better performance than from orthogonal intensities data and total intensity data. Processing the four intensities data from polarimeter, however, requires complicated optical devices, alignment, and four CCD detectors. It only requires one or two detectors and a computer to process orthogonal intensities data and total intensity data, and the bounds demonstrate that reasonable esti-

mates may still be obtained from orthogonal intensities or total intensity data. Therefore, computational sensing is a promising way to estimate the degree of polarization. We proposed computational sensing estimators (algebraic and maximum-likelihood estimators) to replace the optical devices.

The performances of the estimators are compared with each other and with their corresponding Cramer-Rao bounds on RMS estimation error for biased estimators of the degree of polarization. The simulation parameters are set for typical situations: for Gaussian noise case, the simulation parameters are set as: average total intensity is  $\bar{I} = 100$  photons per pixel per readout and the read out noise level is  $\sigma = 30$  RMS photons per pixel per readout; for Poisson noise case, we simulate a low incident light situation where  $\bar{I} = 15$  photons per pixel per readout; and for combined Gaussian and Poisson noise case,  $\bar{I} = 50$  photons per pixel per readout and  $\sigma = 10$  RMS photons per pixel per readout.

Total intensity estimators are discussed in Chapter 5. The algebraic estimators outperform the maximum-likelihood estimators when the degree of polarization is somewhere between 0.2 and 0.7, however, maximum-likelihood estimators are more efficient because the root mean square errors are closer to the bounds. We usually have more interest in highly polarized light (DOP is close to 1) which brings out the features of the scene. In this region, the maximum likelihood estimator is not only more efficient but also has a smaller RMS error. If the degree of polarization is more likely to fall in the middle region between 0 and 1, one may choose the algebraic estimator.

Orthogonal intensities estimators are discussed in Chapter 6. Although we have the analytical expressions for the maximum-likelihood estimators, it is difficult, therefore, impractical to implement. We can use the simple EM algorithm for speckle only data to process the speckle data with detector noise. The performances of the simple EM algorithm are close to efficient, but the root-mean-square (RMS) errors of algebraic estimators can be smaller than EM estimators if the degree of polarization is between 0.5 and 1 and there are only Gaussian or only Poisson noise. In the general case (combined Gaussian and Poisson noise), the EM estimator outperforms the algebraic estimator. The EM algorithm is simple and converges fast, and the performance gain from the EM algorithm is worth the sacrifice of more computing time compared to the algebraic estimator.



# Appendix A

## Joint PDF of the instantaneous Stokes parameters

In Eq. (4.2), it is obvious that the four instantaneous Stokes parameters have the relation:

$$S_0^2 = S_1^2 + S_2^2 + S_3^2. \quad (\text{A1})$$

We also note that in Eq. (4.2), the instantaneous Stokes parameters depend only on amplitudes  $a_x$  and  $a_y$ , and relative phase  $\Delta$ , which means a many-to-one mapping between the two sets of parameters. In other words, given certain instantaneous Stokes parameters, there are infinite possible fields with the same amplitudes and relative phase  $\Delta$  but different  $\theta_x$  or  $\theta_y$ . Therefore, the Jacobian  $J\left(\frac{S_0, S_1, S_2, S_3}{a_x, a_y, \theta_x, \theta_y}\right)$  is zero.

To find the joint probability density function (PDF) of the instantaneous Stokes parameters, we start from the definition of conditional probability density function

$$p_{S_0, S_1, S_2, S_3}(s_0, s_1, s_2, s_3) = p_{S_0|S_1, S_2, S_3}(s_0|s_1, s_2, s_3)p_{S_1, S_2, S_3}(s_1, s_2, s_3). \quad (\text{A2})$$

Because of the relation we have, Eq.(A1), we have

$$p_{S_0|S_1, S_2, S_3}(s_0|s_1, s_2, s_3) = \delta\left(s_0 - \sqrt{s_1^2 + s_2^2 + s_3^2}\right). \quad (\text{A3})$$

The PDF for the field's complex amplitude, Eq. (2.6), can also be written as

$$p_U(u_x, u_y) = \frac{1}{\pi^2 \det(J)} \exp\left(-\frac{1}{\det(J)} [J_{yy}|u_x|^2 + J_{xx}|u_y|^2 - 2\text{Re}(J_{xy}u_x^*u_y)]\right). \quad (\text{A4})$$

The Jacobian of the transformation,

$$\begin{cases} u_x^r &= u_x^r, \\ s_1 &= |u_x|^2 - |u_y|^2, \\ s_2 &= u_x u_y^* + u_x^* u_y, \\ s_3 &= j(u_x u_y^* - u_x^* u_y), \end{cases} \quad (\text{A5})$$

is

$$J\left(\frac{u_x^r, s_1, s_2, s_3}{u_x^r, u_x^i, u_y^r, u_y^i}\right) = 8u_x^i(|u_x|^2 + |u_y|^2), \quad (\text{A6})$$

where  $u_x^r$ ,  $u_x^i$ ,  $u_y^r$  and  $u_y^i$  are the real and imaginary parts of the complex amplitudes  $u_x$  and  $u_y$ . Finding the inverse transformation and using the Jacobian, we have

$$p_{U_x^r, S_1, S_2, S_3}(u_x^r, s_1, s_2, s_3) = \left( 8\sqrt{\frac{1}{2}(s_0 + s_1) - (u_x^r)^2} \sqrt{s_1^2 + s_2^2 + s_3^2} \right)^{-1} \frac{1}{\pi^2 \det(J)} \\ \times \exp \left( -\frac{1}{\det(J)} \left[ J_{yy} \frac{s_0 + s_1}{2} + J_{xx} \frac{s_0 - s_1}{2} - 2\text{Re}(J_{xy} \frac{s_2 + js_3}{2}) \right] \right). \quad (\text{A7})$$

Integrating  $u_x^r$  on the range from  $-\sqrt{\frac{1}{2}(s_0 + s_1)}$  to  $\sqrt{\frac{1}{2}(s_0 + s_1)}$ , we get

$$p_{S_1, S_2, S_3}(s_1, s_2, s_3) = \int_{-\sqrt{(s_0+s_1)/2}}^{\sqrt{(s_0+s_1)/2}} p_{U_x^r, S_1, S_2, S_3}(u_x^r, s_1, s_2, s_3) du_x^r \\ = \frac{1}{8\pi d \sqrt{s_1^2 + s_2^2 + s_3^2}} \exp \left[ -\frac{1}{2d} \left( \langle S_0 \rangle \sqrt{s_1^2 + s_2^2 + s_3^2} - \langle S_1 \rangle s_1 - \langle S_2 \rangle s_2 - \langle S_3 \rangle s_3 \right) \right]. \quad (\text{A8})$$

Here we choose the elements of Stokes vector  $\langle S \rangle$  to parameterize the probability density function; therefore, the determinant of the coherency matrix is

$$d = \det(J) = \frac{1}{4} \left( \langle S_0 \rangle^2 - \langle S_1 \rangle^2 - \langle S_2 \rangle^2 - \langle S_3 \rangle^2 \right). \quad (\text{A9})$$

By substituting Eq.(A3) and Eq. (A8) into Eq. (A2), we have the joint probability density



function of the instantaneous Stokes parameters

$$p_{S_0, S_1, S_2, S_3}(s_0, s_1, s_2, s_3) = \frac{\delta\left(s_0 - \sqrt{s_1^2 + s_2^2 + s_3^2}\right)}{8\pi ds_0} \times \exp\left[-\frac{1}{2d}(\langle S_0 \rangle s_0 - \langle S_1 \rangle s_1 - \langle S_2 \rangle s_2 - \langle S_3 \rangle s_3)\right]. \quad (\text{A10})$$

As mentioned before, the four-channel polarimeter intensity data  $I$  is related to instantaneous Stokes vector  $S$  by Eq.(4.3). Using the Jacobian for the transformation

$$J\left(\frac{i_0, i_1, i_2, i_3}{s_0, s_1, s_2, s_3}\right) = 0.012, \quad (\text{A11})$$

we obtain the joint PDF of the four intensity data

$$p_{I_0, I_1, I_2, I_3}(i_0, i_1, i_2, i_3) = \frac{125\delta\left(s_0 - \sqrt{s_1^2 + s_2^2 + s_3^2}\right)}{12\pi ds_0} \times \exp\left[-\frac{1}{2d}(\langle S_0 \rangle s_0 - \langle S_1 \rangle s_1 - \langle S_2 \rangle s_2 - \langle S_3 \rangle s_3)\right], \quad (\text{A12})$$

where the relation of  $s$  and  $i$  is given in Eq. (4.3), i.e.

$$s_0 = i_0 + i_1 + i_2 + i_3, \quad (\text{A13a})$$

$$s_1 = \frac{5}{3}(i_0 + i_1 - i_2 - i_3), \quad (\text{A13b})$$

$$s_2 = -2.5(i_0 - i_1), \quad (\text{A13c})$$

$$s_3 = -2.5(i_2 - i_3). \quad (\text{A13d})$$

# Appendix B

## Derivation of EM Algorithms

### B.1 Estimation from Orthogonal Intensities

#### B.1.1 Laser Speckle Only

The first order statistical model of the field's complex amplitudes is:

$$p_U(u|J) = \frac{1}{\pi^2 \det(J)} \exp(-u^\dagger J^{-1} u), \quad (\text{A1})$$

where  $\det(J) = J_{xx}J_{yy} - |J_{xy}|^2$  is the determinant of the coherency matrix, and  $J^{-1}$  is its inverse. The log-likelihood function of the complete data is:

$$\begin{aligned}
L_{CD} &= \ln \left( \prod_{k=1}^K p_U(u_k|J) \right) \\
&= -2K \ln \pi - K \ln \det(J) - \sum_{k=1}^K u_k^\dagger J^{-1} u_k \\
&= -2K \ln \pi - K \ln \det(J) - \text{Trace} \left[ J^{-1} \sum_{k=1}^K u_k^\dagger u_k \right]
\end{aligned} \tag{A2}$$

E-step: Compute:

$$\begin{aligned}
Q(J|J^{[n]}) &= E \left[ L_{CD} | I, J^{[n]} \right] \\
&= -2K \ln \pi - K \ln \det(J) - \text{Trace} \left[ J^{-1} \sum_{k=1}^K E \left[ u_k^\dagger u_k | I_k, J^{[n]} \right] \right]
\end{aligned} \tag{A3}$$

where

$$I_k = \begin{bmatrix} I_{xk} \\ I_{yk} \end{bmatrix} = \begin{bmatrix} |u_{xk}|^2 \\ |u_{yk}|^2 \end{bmatrix} \tag{A4}$$

is the mean of the sampled orthogonal intensities, which are our observed data.

M-Step: Let  $J^{[n+1]}$  be that value of  $J$  that maximizes  $Q(J|J^{[n]})$ :

$$J^{[n+1]} = \arg \max_{\theta} Q(J|J^{[n]}) \tag{A5}$$

We Compute the derivative with respect to  $J$  and equate the result to zero:

$$\begin{aligned} \frac{\partial}{\partial J} \left( -2K \ln \pi - K \ln \det(J) - \text{Trace} \left[ J^{-1} \sum_{k=1}^K E \left[ u_k^\dagger u_k | I_k, J^{[n]} \right] \right] \right) &= 0 \\ \Rightarrow J^{[n+1]} &= \frac{1}{K} \sum_{k=1}^K E \left[ u_k^\dagger u_k | I_k J^{[n]} \right] \end{aligned} \quad (\text{A6})$$

This gives

$$\begin{aligned} J^{[n+1]} &= \begin{bmatrix} J_{xx} & J_{xy} \\ J_{xy}^* & J_{yy} \end{bmatrix} \\ &= \frac{1}{K} \sum_{k=1}^K E \begin{bmatrix} |u_{xk}|^2 & u_{xk} u_{yk}^* \\ u_{xk}^* u_{yk} & |u_{yk}|^2 \end{bmatrix} \Big|_{I_{xk}, I_{yk}, J^{[n]}} \\ &= \frac{1}{K} \sum_{k=1}^K \begin{bmatrix} I_{xk} & E \left[ u_{xk} u_{yk}^* | I_{xk}, I_{yk}, J^{[n]} \right] \\ \left( E \left[ u_{xk} u_{yk}^* | I_{xk}, I_{yk}, J^{[n]} \right] \right)^* & I_{yk} \end{bmatrix} \end{aligned} \quad (\text{A7})$$

Therefore, the updates to the coherency matrix  $J$  can be found by computing the conditional expectation

$$\begin{aligned} E \left[ u_{xk} u_{yk}^* | I_{xk}, I_{yk}, J^{[n]} \right] &= E \left[ \sqrt{I_{xk}} \exp(j\theta_{xk}) \sqrt{I_{yk}} \exp(-j\theta_{yk}) | I_{xk}, I_{yk}, J^{[n]} \right] \\ &= \sqrt{I_{xk} I_{yk}} E \left[ \exp(j\theta_{xk}) \exp(-j\theta_{yk}) | I_{xk}, I_{yk}, J^{[n]} \right] \end{aligned} \quad (\text{A8})$$

where

$$u_k = \begin{bmatrix} u_{xk} \\ u_{yk} \end{bmatrix} = \begin{bmatrix} \sqrt{I_{xk}} \exp(j\theta_{xk}) \\ \sqrt{I_{yk}} \exp(j\theta_{yk}) \end{bmatrix} \quad (\text{A9})$$

from eq (A1), we find the joint probability distribution function of the intensities and phase is [32]

$$\begin{aligned} p_{I_x, I_y, \Theta_x, \Theta_y}(i_x, i_y, \theta_x, \theta_y) &= \frac{\det(S)}{4\pi^2} \exp(-u^\dagger J^{-1} u) \\ &= \frac{\det(S)}{4\pi^2} \exp\left(-\left(S_{11}i_x + S_{22}i_y + 2|S_{12}|\sqrt{i_x i_y} \cos(\theta_x - \theta_y - \psi_{12})\right)\right) \end{aligned} \quad (\text{A10})$$

where  $J^{-1} = S = \begin{bmatrix} S_{11} & S_{12} \\ S_{12}^* & S_{22} \end{bmatrix}$  and  $\psi_{12} = \angle S_{12}$ . The joint probability distribution function of the two orthogonal intensities is [32]:

$$p_{I_x, I_y}(i_x, i_y) = \det(S) \exp(-S_{11}i_x - S_{22}i_y) I_0(2|S_{12}|\sqrt{i_x i_y}) \quad (\text{A11})$$

where  $I_0$  is the modified Bessel function of the first kind, order zero. Therefore, the conditional probability distribution is given by

$$\begin{aligned} p_{\Theta_x, \Theta_y | I_x, I_y}(\theta_x, \theta_y | i_x, i_y) &= \frac{p_{I_x, I_y, \Theta_x, \Theta_y}(i_x, i_y, \theta_x, \theta_y)}{p_{I_x, I_y}(i_x, i_y)} \\ &= \frac{1}{4\pi^2 I_0(2|S_{12}|\sqrt{i_x i_y})} \exp(-2|S_{12}|\sqrt{i_x i_y} \cos(\theta_x - \theta_y - \psi_{12})) \end{aligned} \quad (\text{A12})$$

We obtain the conditional expectation (eq (A8)) by computing

$$\begin{aligned}
E [\exp(j\theta_{xk}) \exp(-j\theta_{yk}) | I_{xk}, I_{yk}] &= \int_0^{2\pi} \int_0^{2\pi} \exp(j(\theta_x - \theta_y)) p_{\Theta_x, \Theta_y | I_x, I_y}(\theta_x, \theta_y | I_x, I_y) d\theta_x d\theta_y \\
&= \int_0^{2\pi} \int_0^{2\pi} \frac{1}{4\pi^2 I_0(2|S_{12}|\sqrt{I_x I_y})} \exp(-2|S_{12}|\sqrt{I_x I_y} \cos(\theta_x - \theta_y - \psi_{12}) + j(\theta_x - \theta_y)) d\theta_x d\theta_y \\
&= \int_0^{2\pi} \int_0^{2\pi} \frac{\exp(j(\psi_{12} - \pi))}{4\pi^2 I_0(2|S_{12}|\sqrt{I_x I_y})} \\
&\quad \times \exp(2|S_{12}|\sqrt{I_x I_y} \cos(\theta_x - \theta_y - \psi_{12} + \pi)) \exp(j(\theta_x - \theta_y - \psi_{12} + \pi)) d\theta_x d\theta_y \\
&= \int_0^{2\pi} \int_0^{2\pi} \frac{\exp(j(\psi_{12} - \pi))}{4\pi^2 I_0(2|S_{12}|\sqrt{I_x I_y})} \\
&\quad \times \exp(2|S_{12}|\sqrt{I_x I_y} \cos(\theta_x - \theta_y - \psi_{12} + \pi)) \cos(\theta_x - \theta_y - \psi_{12} + \pi) d\theta_x d\theta_y \\
&\quad + j \int_0^{2\pi} \int_0^{2\pi} \frac{\exp(j(\psi_{12} - \pi))}{4\pi^2 I_0(2|S_{12}|\sqrt{I_x I_y})} \\
&\quad \times \exp(2|S_{12}|\sqrt{I_x I_y} \cos(\theta_x - \theta_y - \psi_{12} + \pi)) \sin(\theta_x - \theta_y - \psi_{12} + \pi) d\theta_x d\theta_y
\end{aligned} \tag{A13}$$

We notice that the second integrated function is an odd function, therefore, the second

integration is zero.

$$\begin{aligned}
& E \left[ \exp(j\theta_{xk}) \exp(-j\theta_{yk}) | I_{xk}, I_{yk} \right] \\
&= \int_0^{2\pi} \int_0^{2\pi} \frac{\exp(j(\psi_{12} - \pi))}{4\pi^2 I_0 (2|S_{12}| \sqrt{I_x I_y})} \\
&\quad \times \exp(2|S_{12}| \sqrt{I_x I_y} \cos(\theta_x - \theta_y - \psi_{12} + \pi)) \cos(\theta_x - \theta_y - \psi_{12} + \pi) d\theta_x d\theta_y \\
&= \int_0^{2\pi} \left( \int_{-\theta_y - \psi_{12} + \pi}^{-\theta_y - \psi_{12} + 3\pi} \frac{\exp(j(\psi_{12} - \pi))}{4\pi^2 I_0 (2|S_{12}| \sqrt{I_x I_y})} \exp(2|S_{12}| \sqrt{I_x I_y} \cos(\theta)) \cos(\theta) d\theta \right) d\theta_y \\
&\quad , \theta = \theta_x - \theta_y - \psi_{12} + \pi \\
&= \frac{\exp(j(\psi_{12} - \pi))}{4\pi^2 I_0 (2|S_{12}| \sqrt{I_x I_y})} \int_0^{2\pi} \left( \int_{-\pi}^{\pi} \exp(2|S_{12}| \sqrt{I_x I_y} \cos(\theta)) \cos(\theta) d\theta \right) d\theta_y \\
&= \frac{I_1(2|S_{12}| \sqrt{I_x I_y})}{I_0(2|S_{12}| \sqrt{I_x I_y})} \exp(j(\psi_{12} - \pi))
\end{aligned} \tag{A14}$$

By substituting eq(A14) into eq(A8) and eq(A7), we obtain the EM algorithm:

$$J^{[n+1]} = \frac{1}{K} \sum_{k=1}^K \begin{bmatrix} I_{xk} & E \left[ u_{xk} u_{yk}^* | I_{xk}, I_{yk}, J^{[n]} \right] \\ (E \left[ u_{xk} u_{yk}^* | I_{xk}, I_{yk}, J^{[n]} \right])^* & I_{yk} \end{bmatrix} \tag{A15}$$

where

$$E \left[ u_{xk} u_{yk}^* | I_{xk}, I_{yk}, J^{[n]} \right] = \sqrt{I_{xk} I_{yk}} \frac{I_1 \left( \frac{2|J_{xy}^{[n]}| \sqrt{I_{xk} I_{yk}}}{\det(J^{[n]})} \right)}{I_0 \left( \frac{2|J_{xy}^{[n]}| \sqrt{I_{xk} I_{yk}}}{\det(J^{[n]})} \right)} \exp(j\angle J_{xy}^{[n]}) \tag{A16}$$

### B.1.2 Gaussian Noise

In the presence of Gaussian noise in the detectors, we have the following statistical model for the measured intensity data

$$D = \begin{bmatrix} D_x \\ D_y \end{bmatrix} = \begin{bmatrix} I_x + N_x \\ I_y + N_y \end{bmatrix} \quad (\text{A17})$$

where  $N_x$  and  $N_y$  are Gaussian read out noise with zero mean and variance of  $\sigma_x^2$  and  $\sigma_y^2$ .

Similar to speckle only case, we need to find the conditional expectation  $E[u_x u_y^* | D_x, D_y, J]$  to form the EM estimate of the coherency matrix

$$J^{[n+1]} = \frac{1}{M} \sum_{m=1}^M \begin{bmatrix} D_{x_m} & E[u_{x_m} u_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]}] \\ \left(E[u_{x_m} u_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]}]\right)^* & D_{y_m} \end{bmatrix}. \quad (\text{A18})$$

In order to find the conditional expectation, let us find the joint probability density function (PDF) of the measured intensity data  $D_x$  and  $D_y$ . Since the read out Gaussian noise is



independent, we have

$$\begin{aligned}
& p_{D_x, D_y, N_x, N_y}(d_x, d_y, n_x, n_y) \\
&= \frac{1}{|J|} \exp \left[ -\frac{J_{xx}}{|J|}(d_y - n_y) - \frac{J_{yy}}{|J|}(d_x - n_x) \right] I_0 \left( \frac{2|J_{xy}|}{|J|} \sqrt{(d_x - n_x)(d_y - n_y)} \right) \\
& \quad \frac{1}{2\pi\sigma_x\sigma_y} \exp \left( -\frac{n_x^2}{2\sigma_x^2} - \frac{n_y^2}{2\sigma_y^2} \right)
\end{aligned} \tag{A19}$$

where  $I_0$  is the zero-order modified Bessel function of the first kind. Then, we can obtain the joint PDF of  $D_x$  and  $D_y$  by computing the marginal PDF

$$p_{D_x, D_y}(d_x, d_y) = \int_{-\infty}^{d_x} \int_{-\infty}^{d_y} p_{D_x, D_y, N_x, N_y}(d_x, d_y, n_x, n_y) dn_x dn_y \tag{A20}$$

Let us write the modified Bessel function into its ascending series form [33]

$$I_\nu(z) = \left( \frac{1}{2}z \right)^\nu \sum_{k=0}^{\infty} \frac{\left( \frac{1}{4}z^2 \right)^k}{k! \Gamma(\nu + k + 1)}. \tag{A21}$$

then, we have

$$\begin{aligned}
& p_{D_x, D_y}(d_x, d_y) \tag{A22} \\
&= \int_{-\infty}^{d_x} \int_{-\infty}^{d_y} \frac{1}{|J|} \exp \left[ -\frac{J_{xx}}{|J|}(d_y - n_y) - \frac{J_{yy}}{|J|}(d_x - n_x) \right] \sum_{k=0}^{\infty} \frac{\left(\frac{1}{4}z^2\right)^k}{k!k!} \\
&\quad \frac{1}{2\pi\sigma_x\sigma_y} \exp \left( -\frac{n_x^2}{2\sigma_x^2} - \frac{n_y^2}{2\sigma_y^2} \right) dn_x dn_y \\
&= \frac{1}{|J|} \exp \left( \frac{J_{yy}^2\sigma_x^2}{2|J|^2} - \frac{d_x J_{yy}}{|J|} \right) \exp \left( \frac{J_{xx}^2\sigma_y^2}{2|J|^2} - \frac{d_y J_{xx}}{|J|} \right) \\
&\quad \sum_{k=0}^{\infty} \frac{\left(\frac{|J_{xy}|}{|J|}\right)^{2k}}{k!k!} \int_0^{\infty} i_x^k \exp \left[ -\frac{(i_x - \mu_x)^2}{2\sigma_x^2} \right] di_x \int_0^{\infty} i_y^k \exp \left[ -\frac{(i_y - \mu_y)^2}{2\sigma_y^2} \right] di_y
\end{aligned}$$

where  $\mu_x = d_x - \frac{J_{yy}\sigma_x^2}{|J|}$  and  $\mu_y = d_y - \frac{J_{xx}\sigma_y^2}{|J|}$ . Note that

$$\begin{aligned}
& \int_0^{\infty} i_x^k \exp \left[ -\frac{(i_x - \mu_x)^2}{2\sigma_x^2} \right] di_x \tag{A23} \\
&= \frac{1}{k+1} \int_0^{\infty} i_x^{k+2} \exp \left[ -\frac{(i_x - \mu_x)^2}{2\sigma_x^2} \right] di_x - \frac{\mu_x}{(k+1)\sigma_x^2} \int_0^{\infty} i_x^{k+1} \exp \left[ -\frac{(i_x - \mu_x)^2}{2\sigma_x^2} \right] di_x.
\end{aligned}$$

Let us denote the integration

$$\int_0^{\infty} i_x^k \exp \left[ -\frac{(i_x - \mu_x)^2}{2\sigma_x^2} \right] di_x$$

by  $\text{Inx}[k]$ , then we have the following relation

$$\text{Inx}[k+2] = (k+1)\sigma_x^2 \text{Inx}[k] + \mu_x \text{Inx}[k+1], \tag{A24}$$

and

$$\text{Inx}[0] = 1 - \Phi\left(\frac{\mu_x}{\sigma_x}\right), \quad (\text{A25})$$

$$\text{Inx}[1] = \sigma_x^2 \exp\left(-\frac{\mu_x^2}{2\sigma_x^2}\right) + \mu_x \text{Inx}[0], \quad (\text{A26})$$

where  $\Phi(x)$  is the standard normal cumulative distribution function (CDF). Similarly, we have

$$\text{Iny}[k] = \int_0^\infty i_y^k \exp\left[-\frac{(i_y - \mu_y)^2}{2\sigma_y^2}\right] \text{d}i_y, \quad (\text{A27})$$

$$\text{Iny}[k+2] = (k+1)\sigma_y^2 \text{Inx}[k] + \mu_y \text{Iny}[k+1], \quad (\text{A28})$$

and

$$\text{Iny}[0] = 1 - \Phi\left(\frac{\mu_y}{\sigma_y}\right), \quad (\text{A29})$$

$$\text{Iny}[1] = \sigma_y^2 \exp\left(-\frac{\mu_y^2}{2\sigma_y^2}\right) + \mu_y \text{Iny}[0]. \quad (\text{A30})$$

Therefore, the joint PDF of  $D_x$  and  $D_y$  is

$$\begin{aligned} p_{D_x, D_y}(d_x, d_y) & \quad (\text{A31}) \\ &= \frac{1}{|J|} \exp\left(\frac{J_{yy}^2 \sigma_x^2}{2|J|^2} - \frac{d_x J_{yy}}{|J|}\right) \exp\left(\frac{J_{xx}^2 \sigma_y^2}{2|J|^2} - \frac{d_y J_{xx}}{|J|}\right) \sum_{k=0}^{\infty} \frac{\left(\frac{|J_{xy}|}{|J|}\right)^{2k}}{k!k!} \text{Inx}[k] \text{Iny}[k] \end{aligned}$$

Now let us derive the conditional expectation

$$\begin{aligned}
& \mathbb{E}[U_x U_y^* | D_x = d_x, D_y = d_y, J] \\
&= \frac{\int_0^\infty \int_0^\infty \int_0^{2\pi} \int_0^{2\pi} r_x r_y \exp[j(\theta_x - \theta_y)] p_{U_x, U_y, D_x, D_y}(r_x, r_y, \theta_x, \theta_y, d_x, d_y) dr_x dr_y d\theta_x d\theta_y}{p_{D_x, D_y}(d_x, d_y)} \\
&= \frac{\exp(j\phi_{xy})}{p_{D_x, D_y}(d_x, d_y)} \int_0^\infty \int_0^\infty \frac{1}{|J|} \exp\left(-\frac{J_{xx}}{|J|} i_x - \frac{J_{yy}}{|J|} i_y\right) I_1\left(\frac{2|J_{xy}|}{|J|} \sqrt{i_x i_y}\right) \\
&\quad \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{(d_x - i_x)^2}{2\sigma_x^2} - \frac{(d_y - i_y)^2}{2\sigma_y^2}\right) di_x di_y
\end{aligned} \tag{A32}$$

where  $\phi_{xy} = \angle J_{xy}$ . We notice that the above double integral is similar to Eq. (A19). The difference is that the zero order modified Bessel function becomes order one. By using the same approach, substituting the Bessel function with its ascending series form, we have

$$\mathbb{E}[U_x U_y^* | D_x, D_y, J] = \frac{\exp(j\phi_{xy})}{p_{D_x, D_y}(d_x, d_y)} \frac{1}{|J|} \exp\left(\frac{J_{yy}^2 \sigma_x^2}{2|J|^2} - \frac{d_x J_{yy}}{|J|}\right) \exp\left(\frac{J_{xx}^2 \sigma_y^2}{2|J|^2} - \frac{d_y J_{xx}}{|J|}\right) \tag{A33}$$

$$\begin{aligned}
& \sum_{k=0}^{\infty} \frac{\left(\frac{|J_{xy}|}{|J|}\right)^{2k+1}}{k!(k+1)!} \text{Inx}[k+1] \text{Iny}[k+1] \\
&= \exp(j\phi_{xy}) \frac{\sum_{k=0}^{\infty} \frac{\left(\frac{|J_{xy}|}{|J|}\right)^{2k+1}}{k!(k+1)!} \text{Inx}[k+1] \text{Iny}[k+1]}{\sum_{k=0}^{\infty} \frac{\left(\frac{|J_{xy}|}{|J|}\right)^{2k}}{k!k!} \text{Inx}[k] \text{Iny}[k]}
\end{aligned}$$

Now we have obtained the EM algorithm for estimation of DOP from orthogonal intensity

data with Gaussian noise:

$$J^{[n+1]} = \frac{1}{M} \sum_{m=1}^M \begin{bmatrix} D_{x_m} & \mathbb{E} \left[ u_{x_m} u_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]} \right] \\ \left( \mathbb{E} \left[ u_{x_m} u_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]} \right] \right)^* & D_{y_m} \end{bmatrix} \quad (\text{A34})$$

where

$$E \left[ U_{x_m} U_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]} \right] = \exp(j\phi_{xy}) \frac{\sum_{k=0}^{\infty} \frac{\left( \frac{|J_{xy}^{[n]}|}{|J^{[n]}|} \right)^{2k+1} \text{Inx}[k+1] \text{Iny}[k+1]}{k!(k+1)!}}{\sum_{k=0}^{\infty} \frac{\left( \frac{|J_{xy}^{[n]}|}{|J^{[n]}|} \right)^{2k} \text{Inx}[k] \text{Iny}[k]}{k!k!}} \quad (\text{A35})$$

and

$$\text{Inx}[k+2] = (k+1)\sigma_x^2 \text{Inx}[k] + \mu_x \text{Inx}[k+1] \quad (\text{A36a})$$

$$\text{Iny}[k+2] = (k+1)\sigma_y^2 \text{Inx}[k] + \mu_y \text{Iny}[k+1] \quad (\text{A36b})$$

$$\text{Inx}[0] = 1 - \Phi\left(\frac{\mu_x}{\sigma_x}\right) \quad (\text{A36c})$$

$$\text{Inx}[1] = \sigma_x^2 \exp\left(-\frac{\mu_x^2}{2\sigma_x^2}\right) + \mu_x \text{Inx}[0] \quad (\text{A36d})$$

$$\text{Iny}[0] = 1 - \Phi\left(\frac{\mu_y}{\sigma_y}\right) \quad (\text{A36e})$$

$$\text{Iny}[1] = \sigma_y^2 \exp\left(-\frac{\mu_y^2}{2\sigma_y^2}\right) + \mu_y \text{Iny}[0] \quad (\text{A36f})$$

$$\mu_{x_m} = D_{x_m} - \frac{J_{yy}\sigma_x^2}{|J|} \quad (\text{A36g})$$

$$\mu_{y_m} = D_{y_m} - \frac{J_{xx}\sigma_y^2}{|J|} \quad (\text{A36h})$$

$$\phi_{xy} = \angle J_{xy} \quad (\text{A36i})$$

### B.1.3 Poisson Noise

Let us consider the case of poisson noise or shot noise.

$$\begin{bmatrix} D_x \\ D_y \end{bmatrix} = \begin{bmatrix} \text{Poisson}(aI_x + b) \\ \text{Poisson}(aI_y + b) \end{bmatrix}; \quad a > 0, b \geq 0 \quad (\text{A37})$$

where  $a$  is the gain of the detector and  $b$  is the dark current (bias) of the detector.

Similar to Gaussian noise case, the EM algorithm is

$$J^{[n+1]} = \frac{1}{M} \sum_{m=1}^M \begin{bmatrix} D_{x_m} & \text{E} \left[ u_{x_m} u_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]} \right] \\ \left( \text{E} \left[ u_{x_m} u_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]} \right] \right)^* & D_{y_m} \end{bmatrix} \quad (\text{A38})$$

To find the conditional expectation  $\text{E} \left[ u_{x_m} u_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]} \right]$ , we need to find the join PDF of  $D_x$  and  $D_y$ , which is given by

$$\begin{aligned} p_{D_x, D_y}(d_x, d_y) & \quad (\text{A39}) \\ &= \int_0^\infty \int_0^\infty p_{D_x, D_y | I_x, I_y}(d_x, d_y | i_x, i_y) p_{I_x, I_y}(i_x, i_y) di_x di_y \\ &= \int_0^\infty \int_0^\infty \frac{(ai_x + b)^{d_x} \exp(-ai_x - b)}{d_x!} \frac{(ai_y + b)^{d_y} \exp(-ai_y - b)}{d_y!} \frac{1}{|J|} \exp\left(-\frac{J_{yy}}{|J|}i_x - \frac{J_{xx}}{|J|}i_y\right) \\ & \quad I_0\left(\frac{2|J_{xy}|}{|J|}\sqrt{i_x i_y}\right) di_x di_y \end{aligned}$$

Expand the modified Bessel function to ascending series by using the Eq. (A21). Then, we

have

$$\begin{aligned}
& p_{D_x, D_y}(d_x, d_y) \tag{A40} \\
&= \int_0^\infty \int_0^\infty \frac{(ai_x + b)^{d_x} \exp(-ai_x - b)}{d_x!} \frac{(ai_y + b)^{d_y} \exp(-ai_y - b)}{d_y!} \frac{1}{|J|} \exp\left(-\frac{J_{yy}}{|J|}i_x - \frac{J_{xx}}{|J|}i_y\right) \\
&\quad \sum_{k=0}^\infty \frac{\left(\frac{J_{xy}}{|J|}\right)^{2k}}{k!k!} i_x^k i_y^k \mathbf{d}i_x \mathbf{d}i_y \\
&= \frac{1}{d_x! d_y! |J|} \sum_{k=0}^\infty \frac{\left(\frac{|J_{xy}|}{|J|}\right)^{2k}}{k!k!} \int_0^\infty (ai_x + b)^{d_x} \exp(-ai_x - b - \frac{J_{yy}}{|J|}i_x) i_x^k \mathbf{d}i_x \\
&\quad \int_0^\infty (ai_y + b)^{d_y} \exp(-ai_y - b - \frac{J_{xx}}{|J|}i_y) i_y^k \mathbf{d}i_y
\end{aligned}$$

To simplify the calculation, we ignore the dark current (bias) of the detector, ie.  $b = 0$ .

Then,

$$\begin{aligned}
& p_{D_x, D_y}(d_x, d_y) \tag{A41} \\
&= \frac{a^{d_x} a^{d_y}}{d_x! d_y! |J|} \sum_{k=0}^\infty \frac{\left(\frac{|J_{xy}|}{|J|}\right)^{2k}}{k!k!} \int_0^\infty i_x^{d_x+k} \exp(-ai_x - \frac{J_{yy}}{|J|}i_x) \mathbf{d}i_x \int_0^\infty i_y^{d_y+k} \exp(-ai_y - \frac{J_{xx}}{|J|}i_y) \mathbf{d}i_y \\
&= \frac{a^{d_x} a^{d_y}}{d_x! d_y! |J|} \sum_{k=0}^\infty \left(\frac{|J_{xy}|}{|J|}\right)^{2k} \frac{(d_x+k)!(d_y+k)!}{k!k!} A^{-d_x-k-1} B^{-d_y-k-1}
\end{aligned}$$

where  $A = a + J_{yy}/|J|$  and  $B = a + J_{xx}/|J|$ .

Once we have the joint PDF of  $D_x$  and  $D_y$ , the conditional expectation is given by

$$\begin{aligned}
& \mathbb{E}[U_x U_y^* | D_x = d_x, D_y = d_y, J] \tag{A42} \\
&= \frac{\int_0^\infty \int_0^\infty \int_0^{2\pi} \int_0^{2\pi} r_x r_y \exp[j(\theta_x - \theta_y)] p_{U_x, U_y, D_x, D_y}(r_x, r_y, \theta_x, \theta_y, d_x, d_y) dr_x dr_y d\theta_x d\theta_y}{p_{D_x, D_y}(d_x, d_y)} \\
&= \frac{\exp(j\phi_{xy})}{p_{D_x, D_y}(d_x, d_y)} \int_0^\infty \int_0^\infty \frac{1}{|J|} \exp\left(-\frac{J_{xx}}{|J|}i_x - \frac{J_{yy}}{|J|}i_y\right) I_1\left(\frac{2|J_{xy}|}{|J|}\sqrt{i_x i_y}\right) \\
&\quad \frac{(ai_x + b)^{d_x} \exp(-ai_x - b)}{d_x!} \frac{(ai_y + b)^{d_y} \exp(-ai_y - b)}{d_y!} \sqrt{i_x i_y} di_x di_y
\end{aligned}$$

where  $\phi_{xy} = \angle J_{xy}$ . Similar to calculation of PDF of  $D_x$  and  $D_y$ , we expanding the first order modified Bessel function, and the conditional expectation is given by

$$\mathbb{E}[U_x U_y^* | D_x = d_x, D_y = d_y, J] = \exp(j\phi_{xy}) \frac{\sum_{k=0}^\infty \left(\frac{|J_{xy}|}{|J|}\right)^{2k+1} \frac{(d_x+k+1)!(d_y+k+1)!}{(k+1)!k!} A^{-k-2} B^{-k-2}}{\sum_{k=0}^\infty \left(\frac{|J_{xy}|}{|J|}\right)^{2k} \frac{(d_x+k)!(d_y+k)!}{k!k!} A^{-k-1} B^{-k-1}} \tag{A43}$$

The EM algorithm for estimation of DOP from orthogonal intensity data with Poisson detector noise is

$$J^{[n+1]} = \frac{1}{M} \sum_{m=1}^M \begin{bmatrix} D_{x_m} & \mathbb{E}[u_{x_m} u_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]}] \\ \left(\mathbb{E}[u_{x_m} u_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]}]\right)^* & D_{y_m} \end{bmatrix} \tag{A44}$$

where



$$\mathbb{E}[U_{x_m} U_{y_m}^* | D_{x_m}, D_{y_m}, J^{[n]}] = \exp(j\phi_{xy}) \frac{\sum_{k=0}^{\infty} \left( \frac{|J_{xy}^{[n]}|}{|J^{[n]}|} \right)^{2k+1} \frac{(D_{x_m}+k+1)!(D_{y_m}+k+1)!}{(k+1)!k!} A^{-k-2} B^{-k-2}}{\sum_{k=0}^{\infty} \left( \frac{|J_{xy}^{[n]}|}{|J^{[n]}|} \right)^{2k} \frac{(D_{x_m}+k)!(D_{y_m}+k)!}{k!k!} A^{-k-1} B^{-k-1}} \quad (\text{A45})$$

and

$$A = a + J_{yy}^{[n]} / |J|, \quad B = a + J_{xx}^{[n]} / |J| \quad (\text{A46})$$

## B.2 Estimation from Total Intensity

### B.2.1 Laser Speckle Only

To find the conditional expectation, we first find the conditional probability distribution by using eq(5.5) and eq(5.6)

$$\begin{aligned} p_{I_x, I_y | I_t}(i_x, i_y | i_t) &= \frac{p_{I_x, I_y, I_t}(i_x, i_y, i_t)}{p_{I_t}(i_t)} \\ &= \frac{p_{I_x, I_y}(i_x, i_y) \delta(i_t - i_x - i_y)}{p_{I_t}(i_t)} \\ &= \frac{\frac{1}{J_{xx}J_{yy}} \exp\left(-\frac{J_{xx}i_y + J_{yy}i_x}{J_{xx}J_{yy}}\right) \delta(i_t - i_x - i_y)}{\frac{1}{J_{xx} - J_{yy}} \left[ \exp\left(-\frac{i_t}{J_{xx}}\right) - \exp\left(-\frac{i_t}{J_{yy}}\right) \right]} \end{aligned} \quad (\text{A47})$$

the conditional probability is therefore given by

$$\begin{aligned}
p_{I_x|I_t}(i_x|i_t) &= \int_0^{i_t} p_{I_x, I_y|I_t}(i_x, i_y|i_t) di_y \\
&= \int_0^{i_t} \frac{\frac{1}{J_{xx}J_{yy}} \exp\left(-\frac{J_{xx}i_y + J_{yy}i_x}{J_{xx}J_{yy}}\right) \delta(i_t - i_x - i_y)}{\frac{1}{J_{xx} - J_{yy}} \left[ \exp\left(-\frac{i_t}{J_{xx}}\right) - \exp\left(-\frac{i_t}{J_{yy}}\right) \right]} di_y \\
&= \frac{\frac{1}{J_{xx}J_{yy}} \exp\left(-\frac{J_{xx}(i_t - i_y) + J_{yy}i_x}{J_{xx}J_{yy}}\right)}{\frac{1}{J_{xx} - J_{yy}} \left[ \exp\left(-\frac{i_t}{J_{xx}}\right) - \exp\left(-\frac{i_t}{J_{yy}}\right) \right]} \\
&= \frac{J_{xx} - J_{yy}}{J_{xx}J_{yy}} \frac{\exp\left(-\frac{i_t}{J_{yy}}\right)}{\exp\left(-\frac{i_t}{J_{xx}}\right) - \exp\left(-\frac{i_t}{J_{yy}}\right)} \exp\left(\frac{J_{xx} - J_{yy}}{J_{xx}J_{yy}} i_x\right)
\end{aligned} \tag{A48}$$

the conditional expectation is obtained by its definition

$$\begin{aligned}
E[I_x|I_t] &= \int_0^{I_t} i_x p_{I_x|I_t}(i_x|i_t) di_x \\
&= \frac{J_{xx} - J_{yy}}{J_{xx}J_{yy}} \frac{\exp\left(-\frac{I_t}{J_{yy}}\right)}{\exp\left(-\frac{I_t}{J_{xx}}\right) - \exp\left(-\frac{I_t}{J_{yy}}\right)} \int_0^{I_t} i_x \exp\left(\frac{J_{xx} - J_{yy}}{J_{xx}J_{yy}} i_x\right) di_x \\
&= \frac{J_{xx} - J_{yy}}{J_{xx}J_{yy}} \frac{\exp\left(-\frac{I_t}{J_{yy}}\right)}{\exp\left(-\frac{I_t}{J_{xx}}\right) - \exp\left(-\frac{I_t}{J_{yy}}\right)} \left(\frac{J_{xx} - J_{yy}}{J_{xx}J_{yy}}\right)^{-2} \exp\left(\frac{J_{xx} - J_{yy}}{J_{xx}J_{yy}} i_x\right) \left(\frac{J_{xx} - J_{yy}}{J_{xx}J_{yy}} i_x - 1\right) \Bigg|_0^{I_t} \\
&= I_t \frac{\exp\left(-\frac{I_t}{J_{xx}}\right)}{\exp\left(-\frac{I_t}{J_{xx}}\right) - \exp\left(-\frac{I_t}{J_{yy}}\right)} - \frac{J_{xx}J_{yy}}{J_{xx} - J_{yy}}
\end{aligned} \tag{A49}$$

as the same derivation above, we find the other conditional expectation is

$$E[I_y|I_t] = I_t \frac{\exp\left(-\frac{I_t}{J_{yy}}\right)}{\exp\left(-\frac{I_t}{J_{yy}}\right) - \exp\left(-\frac{I_t}{J_{xx}}\right)} - \frac{J_{xx}J_{yy}}{J_{yy} - J_{xx}} \tag{A50}$$

### B.2.2 Laser Speckle and Gaussian Noise

In this case, we have additive Gaussian noise when we measure the total intensity.

$$D = I_x + I_y + N = I_t + N \quad (\text{A51})$$

where  $N(0, \sigma^2)$  and we have the following relations:

$$0 \leq I_x \leq D - N \quad (\text{A52a})$$

$$0 \leq I_y \leq D - N \quad (\text{A52b})$$

$$-\infty \leq N \leq D \quad (\text{A52c})$$

We first find the conditional probability distribution function. We have

$$p_{D|I_x, I_y}(d|i_x, i_y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d - i_x - i_y)^2}{2\sigma^2}\right) \quad (\text{A53})$$

since  $N$  is normal distributed with zero mean and variance of  $\sigma^2$ . We also have the joint probability distribution function of  $I_x$  and  $I_y$ , eq(5.6). Therefore,

$$p_{D, I_x, I_y}(d, i_x, i_y) = p_{D|I_x, I_y}(d|i_x, i_y) p_{I_x, I_y}(i_x, i_y) \quad (\text{A54})$$

and

$$p_{D,N,I_x,I_y}(d,n,i_x,i_y) = p_{D|I_x,I_y}(d|i_x,i_y)p_{I_x,I_y}(i_x,i_y)\delta[d - (i_x + i_y + n)] \quad (\text{A55})$$

The probability distribution function of  $D$  is then given by

$$\begin{aligned} p_D(d) &= \int_{-\infty}^d \int_0^{d-n} \int_0^{d-n} p_{D,N,I_x,I_y}(d,n,i_x,i_y) di_y di_x dn \\ &= \int_{-\infty}^d \int_0^{d-n} \int_0^{d-n} p_{D,I_x,I_y}(d,i_x,i_y) \delta[d - (i_x + i_y + n)] di_y di_x dn \\ &= \int_{-\infty}^d \int_0^{d-n} p_{D|I_x,I_y}(d|i_x,i_y = d - i_x - n) p_{I_x,I_y}(i_x,i_y = d - i_x - n) di_x dn \\ &= \int_{-\infty}^d \int_0^{d-n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d - i_x - (d - i_x - n))^2}{2\sigma^2}\right) \\ &\quad \times \frac{1}{J_{xx}J_{yy}} \exp\left(-\frac{J_{xx}(d - i_x - n) + J_{yy}i_x}{J_{xx}J_{yy}}\right) di_x dn \\ &= \int_{-\infty}^d \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n^2}{2\sigma^2}\right) \frac{1}{J_{xx} - J_{yy}} \left(\exp\left(-\frac{d - n}{J_{xx}}\right) - \exp\left(-\frac{d - n}{J_{yy}}\right)\right) dn \\ &= \frac{\exp\left(-\frac{d}{J_{xx}}\right)}{J_{xx} - J_{yy}} \int_{-\infty}^d \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n^2}{2\sigma^2} + \frac{n}{J_{xx}}\right) dn \\ &\quad - \frac{\exp\left(-\frac{d}{J_{yy}}\right)}{J_{xx} - J_{yy}} \int_{-\infty}^d \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n^2}{2\sigma^2} + \frac{n}{J_{yy}}\right) dn \\ &= \frac{1}{J_{xx} - J_{yy}} \left[ \exp\left(-\frac{2J_{xx}d - \sigma^2}{2J_{xx}^2}\right) \Phi\left(\frac{d}{\sigma} - \frac{\sigma}{J_{xx}}\right) - \exp\left(-\frac{2J_{yy}d - \sigma^2}{2J_{yy}^2}\right) \Phi\left(\frac{d}{\sigma} - \frac{\sigma}{J_{yy}}\right) \right] \end{aligned} \quad (\text{A56})$$

where the function  $\Phi(x)$  is the standard normal cumulative distribution function, with  $\mu = 0$

and  $\sigma = 1$ . The conditional probability distribution function is defined as

$$p_{I_x|D}(i_x|d) = \frac{p_{I_x,D}(i_x,d)}{p_D(d)} = \frac{\int_{-\infty}^d \int_0^{d-n} p_{D,N,I_x,I_y}(d,n,i_x,i_y) di_y dn}{p_D(d)} \quad (\text{A57})$$

therefore, the conditional expectation is

$$\begin{aligned}
E[I_x|D=d] &= \int_0^\infty i_x p_{I_x|D}(i_x|d) di_x \\
&= \frac{1}{p_D(d)} \int_0^\infty i_x \left( \int_{-\infty}^d \int_0^{d-n} p_{D,N,I_x,I_y}(d,n,i_x,i_y) di_y dn \right) di_x
\end{aligned} \tag{A58}$$

by switching the order of integrations of  $di_x$  and  $dn$ , we can simplify the calculation. But, we should note that the range of the integrations also change. Since  $i_x$  is from 0 to  $\infty$ ,  $D$  ranges from 0 to  $\infty$  and  $n$  ranges from  $-\infty$  to  $\infty$ .

$$\begin{aligned}
E[I_x|D=d] &= \frac{1}{p_D(d)} \int_{-\infty}^\infty \int_0^{d-n} \int_0^{d-n} i_x p_{D,N,I_x,I_y}(d,n,i_x,i_y) di_y di_x dn \\
&= \frac{1}{p_D(d)} \int_{-\infty}^\infty \int_0^{d-n} \int_0^{d-n} i_x p_{D,I_x,I_y}(d,i_x,i_y) \delta[d - (i_x + i_y + n)] di_y di_x dn \\
&= \frac{1}{p_D(d)} \int_{-\infty}^\infty \int_0^{d-n} i_x \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n^2}{2\sigma^2}\right) \frac{1}{J_{xx}J_{yy}} \exp\left(-\frac{J_{xx}(d-i_x-n) + J_{yy}i_x}{J_{xx}J_{yy}}\right) di_x dn \\
&= \frac{\exp\left(-\frac{d}{J_{yy}}\right)}{p_D(d)J_{xx}J_{yy}} \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n^2}{2\sigma^2}\right) \exp\left(\frac{n}{J_{yy}}\right) \int_0^{d-n} i_x \exp\left(\frac{J_{xx}-J_{yy}}{J_{xx}J_{yy}}i_x\right) di_x dn
\end{aligned} \tag{A59}$$

We find the inner integral first:

$$\begin{aligned}
\int_0^{d-n} i_x \exp\left(\frac{J_{xx}-J_{yy}}{J_{xx}J_{yy}}i_x\right) di_x &= \frac{J_{xx}J_{yy}}{J_{xx}-J_{yy}}(d-n) \exp\left(\frac{J_{xx}-J_{yy}}{J_{xx}J_{yy}}(d-n)\right) - \\
&\quad \left(\frac{J_{xx}J_{yy}}{J_{xx}-J_{yy}}\right)^2 \exp\left(\frac{J_{xx}-J_{yy}}{J_{xx}J_{yy}}(d-n)\right) + \left(\frac{J_{xx}J_{yy}}{J_{xx}-J_{yy}}\right)^2
\end{aligned} \tag{A60}$$

substitute eq(A60) and eq(A56) into eq(A59)

$$\begin{aligned}
E[I_x|D=d] &= \frac{1}{p_D(d)J_{xx}J_{yy}} \frac{J_{xx}J_{yy}}{J_{xx}-J_{yy}} \exp\left(-\frac{d}{J_{xx}} + \frac{\sigma^2}{2J_{xx}^2}\right) \left(d - \frac{\sigma^2}{J_{xx}} - \frac{J_{xx}J_{yy}}{J_{xx}-J_{yy}}\right) \\
&\quad + \frac{1}{p_D(d)J_{xx}J_{yy}} \left(\frac{J_{xx}J_{yy}}{J_{xx}-J_{yy}}\right)^2 \exp\left(-\frac{d}{J_{yy}} + \frac{\sigma^2}{2J_{yy}^2}\right) \\
&= \frac{\exp\left(-\frac{d}{J_{xx}} + \frac{\sigma^2}{2J_{xx}^2}\right) \left(d - \frac{\sigma^2}{J_{xx}} - \frac{J_{xx}J_{yy}}{J_{xx}-J_{yy}}\right) + \frac{J_{xx}J_{yy}}{J_{xx}-J_{yy}} \exp\left(-\frac{d}{J_{yy}} + \frac{\sigma^2}{2J_{yy}^2}\right)}{\exp\left(-\frac{d}{J_{xx}} + \frac{\sigma^2}{2J_{xx}^2}\right) \Phi\left(\frac{d}{\sigma} - \frac{\sigma}{J_{xx}}\right) - \exp\left(-\frac{d}{J_{yy}} + \frac{\sigma^2}{2J_{yy}^2}\right) \Phi\left(\frac{d}{\sigma} - \frac{\sigma}{J_{yy}}\right)} \quad (A61)
\end{aligned}$$

similarly, we can obtain the conditional expectation of  $I_y$  given  $D$

$$\begin{aligned}
E[I_y|D=d] &= \frac{\exp\left(-\frac{d}{J_{yy}} + \frac{\sigma^2}{2J_{yy}^2}\right) \left(d - \frac{\sigma^2}{J_{yy}} - \frac{J_{xx}J_{yy}}{J_{yy}-J_{xx}}\right) + \frac{J_{xx}J_{yy}}{J_{yy}-J_{xx}} \exp\left(-\frac{d}{J_{xx}} + \frac{\sigma^2}{2J_{xx}^2}\right)}{\exp\left(-\frac{d}{J_{yy}} + \frac{\sigma^2}{2J_{yy}^2}\right) \Phi\left(\frac{d}{\sigma} - \frac{\sigma}{J_{yy}}\right) - \exp\left(-\frac{d}{J_{xx}} + \frac{\sigma^2}{2J_{xx}^2}\right) \Phi\left(\frac{d}{\sigma} - \frac{\sigma}{J_{xx}}\right)} \quad (A62)
\end{aligned}$$

### B.2.3 Laser Speckle and Poisson Noise

In the case of laser speckle and Poisson noise, we have the following statistical model

$$D \sim \text{Poisson}(\lambda) = \text{Poisson}(a(I_x + I_y) + b) \quad a > 0, b \geq 0 \quad (A63)$$

where  $\lambda = a(I_x + I_y) + b$  with  $a$  is the gain of the CCD detector and  $b$  is the bias of the detector. Similar to the Gaussian noise case, we also assume the coherence matrix is diagonalized.

To find the conditional expectations, we first find the probability mass function (PMF) of  $D$

$$\begin{aligned}
p_D(d) &= \int_0^\infty \int_0^\infty p_{D|I_x, I_y}(d|i_x, i_y) p_{I_x, I_y}(i_x, i_y) di_x di_y \\
&= \int_0^\infty \int_0^\infty \frac{\exp(-\lambda) \lambda^d}{d!} \frac{1}{J_{xx} J_{yy}} \exp\left(-\frac{i_x}{J_{xx}} - \frac{i_y}{J_{yy}}\right) di_x di_y, \lambda = a(I_x + I_y) + b \\
&= \frac{1}{J_{xx} J_{yy} d!} \int_0^\infty \int_0^\infty \exp(-\lambda) \lambda^d \exp\left(-\frac{i_x}{J_{xx}} - \frac{i_y}{J_{yy}}\right) di_x di_y \\
&= \frac{\exp\left(\frac{b}{aJ_{xx}}\right)}{J_{xx} J_{yy} d!} \int_0^\infty \int_0^\infty \exp(-\lambda) \lambda^d \exp\left(-\frac{\lambda}{aJ_{xx}}\right) \exp\left(\frac{i_y}{J_{xx}} - \frac{i_y}{J_{yy}}\right) di_x di_y
\end{aligned} \tag{A64}$$

By making a change of variable,  $\lambda = a(I_x + I_y) + b$ , the above equation can be written as

$$\begin{aligned}
p_D(d) &= \frac{\exp\left(\frac{b}{aJ_{xx}}\right)}{aJ_{xx} J_{yy} d!} \int_b^\infty \int_0^{\frac{\lambda-b}{a}} \exp(-\lambda) \lambda^d \exp\left(-\frac{\lambda}{aJ_{xx}}\right) \exp\left(\frac{i_y}{J_{xx}} - \frac{i_y}{J_{yy}}\right) di_y d\lambda \\
&= \frac{\exp\left(\frac{b}{aJ_{xx}}\right)}{aJ_{xx} J_{yy} d!} \int_b^\infty \exp(-\lambda) \lambda^d \exp\left(-\frac{\lambda}{aJ_{xx}}\right) \frac{J_{xx} J_{yy}}{J_{yy} - J_{xx}} \left( \exp\left(\frac{J_{xx} J_{yy}}{J_{yy} - J_{xx}} \left(\frac{\lambda - b}{a}\right)\right) - 1 \right) d\lambda \\
&= \frac{\exp\left(\frac{b}{aJ_{xx}}\right)}{a(J_{xx} - J_{yy}) d!} \int_b^\infty \lambda^d \exp\left(-\left(1 + \frac{1}{aJ_{xx}}\right) \lambda\right) d\lambda \\
&\quad - \frac{\exp\left(\frac{b}{aJ_{yy}}\right)}{a(J_{xx} - J_{yy}) d!} \int_b^\infty \lambda^d \exp\left(-\left(1 + \frac{1}{aJ_{yy}}\right) \lambda\right) d\lambda \tag{A65}
\end{aligned}$$

let  $X = \left(1 + \frac{1}{aJ_{xx}}\right)$  and  $Y = \left(1 + \frac{1}{aJ_{yy}}\right)$ .

$$p_D(d) = \frac{\exp\left(\frac{b}{aJ_{xx}}\right) \text{Ei}(d, b, X)}{a(J_{xx} - J_{yy}) d!} - \frac{\exp\left(\frac{b}{aJ_{yy}}\right) \text{Ei}(d, b, Y)}{a(J_{xx} - J_{yy}) d!} \tag{A66}$$

where the exponential integral function, we denote it as  $\text{Ei}(n, b, z)$ , is given by [30]

$$\text{Ei}(n, b, z) = \int_b^\infty \lambda^n \exp(-z\lambda) d\lambda = \begin{cases} z^{-n-1} \Gamma(n+1, bz) & b > 0 \\ z^{-n-1} n! & b = 0 \end{cases}$$

Then, the conditional PDF of  $N_x$  given  $D$  is

$$\begin{aligned} p_{N_x|D}(n_x|D=d) &= \frac{p_{D|N_x}(d|n_x)p_{N_x}(n_x)}{p_D(d)} \\ &= \frac{(1+J_{xx}^{-1})^{-n_x-1}}{J_{xx}} \frac{(1+J_{yy}^{-1})^{-d+n_x-1}}{J_{yy}} \frac{J_{xx}-J_{yy}}{(1+J_{xx}^{-1})^{-n-1} - (1+J_{yy}^{-1})^{-n_x-1}} \\ &= \frac{1-r}{1-r^{d+1}} r^{n_x} \end{aligned} \quad (\text{A67})$$

where

$$r = \frac{1+J_{yy}^{-1}}{1+J_{xx}^{-1}} \quad (\text{A68})$$

Therefore, the conditional expectation is

$$\begin{aligned} \mathbb{E}[N_x|D=d] &= \sum_{n_x=0}^d n_x \frac{1-r}{1-r^{d+1}} r^{n_x} \\ &= \frac{r}{1-r} - (d+1) \frac{r^{d+1}}{1-r^{d+1}}. \end{aligned} \quad (\text{A69})$$

Similarly, we have

$$\mathbb{E}[N_y|D=d] = \frac{r^{-1}}{1-r^{-1}} - (d+1) \frac{r^{-d-1}}{1-r^{-d-1}} \quad (\text{A70})$$



### B.2.4 Laser Speckle Combined with Gaussian Noise and Poisson Noise

When Gaussian and Poisson noise exist together, we have the following model:

$$D = I + N, \quad (\text{A71})$$

where  $I \sim \text{Poisson}(\lambda) = \text{Poisson}(a(I_x + I_y) + b)$ ,  $N \sim \mathcal{N}(0, \sigma^2)$ ,  $a$  is the gain of the detector and  $b$  is the bias of the detector. Similarly, we also assume coherence matrix is diagonalized.

Because  $I = \text{Poisson}(a(I_x + I_y) + b)$  and  $N$  are independent, we have

$$p_{D,N|I_x,I_y}(d, n|i_x, i_y) = p_{I|I_x,I_y}(d - n|i_x, i_y)p_N(n) \quad (\text{A72})$$

then,

$$p_{D|I_x,I_y}(d, n|i_x, i_y) = \sum_{m=0}^{\infty} p_N(d - m)p_{I|I_x,I_y}(m|i_x, i_y) \quad (\text{A73})$$

we already have the joint probability distribution  $p_{I_x,I_y}(i_x, i_y)$  eq(5.6). Therefore, using

Bayes' theorem, we have

$$p_{I_x,I_y|D}(i_x, i_y|d) = \frac{p_{D|I_x,I_y}(d|i_x, i_y)p_{I_x,I_y}(i_x, i_y)}{\int_0^\infty \int_0^\infty p_{D|I_x,I_y}(d|i_x, i_y)p_{I_x,I_y}(i_x, i_y)di_x di_y} \quad (\text{A74})$$

$$= \frac{\sum_{m=0}^{\infty} p_N(d - m)p_{I|I_x,I_y}(m|i_x, i_y)p_{I_x,I_y}(i_x, i_y)}{\sum_{m=0}^{\infty} p_N(d - m) \int_0^\infty \int_0^\infty p_{I|I_x,I_y}(m|i_x, i_y)p_{I_x,I_y}(i_x, i_y)di_x di_y} \quad (\text{A75})$$

The denominator of the above formula is the probability distribution function  $p_D(d)$ . It is found to be

$$p_D(d) = \sum_{m=0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(d-m)^2}{2\sigma^2}\right] \left( \frac{\exp\left(\frac{b}{aJ_{xx}}\right) \text{Ei}(m, b, X)}{a(J_{xx} - J_{yy})m!} - \frac{\exp\left(\frac{b}{aJ_{yy}}\right) \text{Ei}(m, b, Y)}{a(J_{xx} - J_{yy})m!} \right) \quad (\text{A76})$$

The conditional expectation is given by

$$\begin{aligned} E[I_x|D=d] &= \int_0^{\infty} i_x p_{I_x|D}(i_x|d) di_x \\ &= \int_0^{\infty} i_x \int_0^{\infty} p_{I_x, I_y|D}(i_x, i_y|d) di_y di_x \end{aligned} \quad (\text{A77})$$

substitute Eq(A75) into above, we have

$$\begin{aligned} E[I_x|D=d] &= \frac{1}{p_D(d)} \sum_{m=0}^M p_N(d-m) \int_0^{\infty} \int_0^{\infty} i_x p_{I|I_x, I_y}(m|i_x, i_y) p_{I_x, I_y}(i_x, i_y) di_y di_x \\ &= \frac{1}{p_D(d)} \sum_{m=0}^M p_N(d-m) \frac{1}{m! J_{xx} J_{yy}} \\ &\quad \times \left( \frac{J_{xx} J_{yy} \exp\left(\frac{b}{aJ_{xx}}\right) \text{Ei}(m+1, b, X)}{a^2(J_{xx} - J_{yy})} - \frac{b J_{xx} J_{yy} \exp\left(\frac{b}{aJ_{xx}}\right) \text{Ei}(m, b, X)}{a^2(J_{xx} - J_{yy})} \right) \\ &\quad - \frac{1}{p_D(d)} \frac{1}{m! J_{xx} J_{yy}} \frac{J_{xx}^2 J_{yy}^2}{a(J_{xx} - J_{yy})^2} \left( \exp\left(\frac{b}{aJ_{xx}}\right) \text{Ei}(m, b, X) - \exp\left(\frac{b}{aJ_{yy}}\right) \text{Ei}(m, b, Y) \right) \end{aligned} \quad (\text{A78})$$

substitute  $p_D(d)$ , Eq.(A76), into above, and we have

$$\begin{aligned}
& E[I_x|D = d] \tag{A79} \\
&= \frac{\sum_{m=0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(d-m)^2}{2\sigma^2}\right] \exp\left(\frac{b}{aJ_{xx}}\right) [\text{Ei}(m+1, b, X) - b\text{Ei}(m, b, X)]}{a \sum_{m=0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(d-m)^2}{2\sigma^2}\right] \left[\exp\left(\frac{b}{aJ_{xx}}\right) \text{Ei}(m, b, X) - \exp\left(\frac{b}{aJ_{yy}}\right) \text{Ei}(m, b, Y)\right]} - \frac{J_{xx}J_{yy}}{J_{xx} - J_{yy}}
\end{aligned}$$

By switching the two orthogonal channel, we obtain the other conditional expectation by switching the subscript  $x$  and  $y$ . Therefore, we have

$$\begin{aligned}
& E[I_y|D = d] \tag{A80} \\
&= \frac{\sum_{m=0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(d-m)^2}{2\sigma^2}\right] \exp\left(\frac{b}{aJ_{yy}}\right) [\text{Ei}(m+1, b, Y) - b\text{Ei}(m, b, Y)]}{a \sum_{m=0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(d-m)^2}{2\sigma^2}\right] \left[\exp\left(\frac{b}{aJ_{yy}}\right) \text{Ei}(m, b, Y) - \exp\left(\frac{b}{aJ_{xx}}\right) \text{Ei}(m, b, X)\right]} + \frac{J_{xx}J_{yy}}{J_{xx} - J_{yy}}
\end{aligned}$$

# References

- [1] J. W. Goodman, *Statistical Optics* (Wiley-Interscience, 2000).
- [2] S. G. Demos and R. R. Alfano, “Optical polarization imaging,” *Appl. Opt.* **36**, 150–155 (1997).
- [3] J. E. Solomon, “Polarization imaging,” *Appl. Opt.* **20**, 1537–1544 (1981).
- [4] J. S. Tyo, M. P. Rowe, J. E. N. Pugh, and N. Engheta, “Target detection in optically scattering media by polarization-difference imaging,” *Appl. Opt.* **35**, 1855–1870 (1996).
- [5] L. Wolff, “Polarization-based material classification from specular reflection,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **12**, 1059–1071 (1990).
- [6] J. Zallat, P. Grabbling, and Y. Takakura, “Using polarimetric imaging for material classification,” *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on* **2**, II–827–30 vol.3 (2003).
- [7] V. Thilak, D. G. Voelz, and C. D. Creusere, “Polarization-based index of refraction

- and reflection angle estimation for remote sensing applications,” *Appl. Opt.* **46**, 7527–7536 (2007).
- [8] K. Ikeuchi, “Determining surface orientations of specular surfaces by using the photometric stereo method,” *IEEE Trans. Pattern Anal. Mach. Intell.* **3**, 661–669 (1981).
- [9] D. Miyazaki, M. Saito, Y. Sato, and K. Ikeuchi, “Determining surface orientations of transparent objects based on polarization degrees in visible and infrared wavelengths,” *J. Opt. Soc. Am. A* **19**, 687–694 (2002).
- [10] F. Goudail and P. Réfrégier, “Statistical algorithms for target detection in coherent active polarimetric images,” *J. Opt. Soc. Am. A* **18**, 3049–3060 (2001).
- [11] jurvetson, “<http://www.flickr.com/photos/jurvetson/42066452/sizes/l/>,” .
- [12] K. S. Miller, *Complex Stochastic Processes: An Introduction to Theory and Application* (Addison-Wesley, 1974).
- [13] T. J. Schulz, “Performance bounds for the estimation of the degree of polarization from active laser illumination,” (*SPIE*, 2005), vol. 5888, p. 58880N.
- [14] T. K. Moon and W. C. Stirling, *Mathematical methods and algorithms for signal processing* (Upper Saddle River, NJ : Prentice Hall, 2000).
- [15] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, vol. 1 (Prentice Hall, 1993).

- [16] B. Jähne, *Practical handbook on image processing for scientific applications* (Boca Raton, Fla. : CRC, 1997).
- [17] G. Healey and R. Kondepudy, “Radiometric CCD camera calibration and noise estimation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **16**, 267–276 (1994).
- [18] C. Brosseau, *Fundamentals of polarized light : a statistical optics approach* (New York : Wiley, 1998).
- [19] E. Wolf, “Correlation between photons in partially polarized light beams,” *Proceedings of the Physical Society* **76**, 424–426 (1960).
- [20] D. H. Goldstein, “Mueller matrix dual-rotating retarder polarimeter,” *Appl. Opt.* (1992).
- [21] R. M. A. Azzam, “Division-of-amplitude photopolarimeter (doap) for the simultaneous measurement of all four stokes parameters of light,” *Opt. Acta* (1982).
- [22] S. Krishnan, “Calibration, properties, and applications of the division-of-amplitude photopolarimeter at 632.8 and 1523 nm,” *J. Opt. Soc. Am. A* (1992).
- [23] R. M. A. Azzam, I. M. Elminyaw, and A. M. El-Saba, “General analysis and optimization of the four-detector photopolarimeter,” *J. Opt. Soc. Am. A* (1988).
- [24] P. Réfrégier, F. Goudail, and N. Roux, “Estimation of the degree of polarization in active coherent imagery by using the natural representation,” *J. Opt. Soc. Am. A* (2004).

- [25] F. Goudail, P. Réfrégier, and N. Roux, “Estimation of the degree of polarization of coherent light in the presence of uniform and nonuniform illumination,” *J. Opt. Soc. Am. A* (2006).
- [26] A. Al-Qasimi, O. Korotkova, D. James, and E. Wolf, “Definitions of the degree of polarization of a light beam,” *Optics Letters* (2007).
- [27] J. Fade, M. Roche, and P. Réfrégier, “Precision of moment-based estimation of the degree of polarization in coherent imagery without polarization device,” *J. Opt. Soc. Am. A* **25**, 483–492 (2008).
- [28] R. M. A. Azzam, I. M. Elminyaw, and A. M. El-Saba, “General analysis and optimization of the four-detector photopolarimeter,” *J. Opt. Soc. Am. A* **5**, 681–689 (1988).
- [29] P. Réfrégier, J. Fade, and M. Roche, “Estimation precision of the degree of polarization from a single speckle intensity image,” *Opt. Lett.* **32**, 739–741 (2007).
- [30] A. Jeffrey and D. Zwillinger, eds., *Table of Integrals, Series, and Products* (Academic Press, 2000), sixth ed.
- [31] F. Chatelain, J.-Y. Tourneret, M. Roche, and M. Alouini, “Estimating the polarization degree of polarimetric images in coherent illumination using maximum likelihood methods,” *J. Opt. Soc. Am. A* **26**, 1348–1359 (2009).

- [32] K. S. Miller, *Complex Stochastic Processes: An Introduction to Theory and Application* (Addison-Wesley, 1974).
- [33] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover Publications, 1964).