

BANDWIDTH AND NOISE IN SPATIO-TEMPORALLY MODULATED  
MUELLER MATRIX POLARIMETERS

by

Israel Jacob Vaughn



A Dissertation Submitted to the Faculty of the

COLLEGE OF OPTICAL SCIENCES

In Partial Fulfillment of the Requirements  
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2016

ProQuest Number: 3746004

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 3746004

Published by ProQuest LLC (2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 - 1346



THE UNIVERSITY OF ARIZONA  
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Israel Jacob Vaughn entitled "Bandwidth and Noise in Spatio-temporally Modulated Mueller Matrix Polarimeters" and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

\_\_\_\_\_  
Russell A. Chipman

Date: 7 December 2015

\_\_\_\_\_  
Matthew A. Kupinski

Date: 7 December 2015

\_\_\_\_\_  
J. Scott Tyo

Date: 7 December 2015

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.  
I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

\_\_\_\_\_  
Dissertation Director: J. Scott Tyo

Date: 7 December 2015

## STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. This work is licensed under the Creative Commons Attribution-No Derivative Works 3.0 United States License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nd/3.0/us/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

SIGNED: Israel Jacob Vaughn

## ACKNOWLEDGEMENTS

I would like to thank my wife, Erin, for always being there for support, and being part of team Vaughn. I would like to thank my daughters, Teagan and Farrah, for bringing me joy during arduous times. I also acknowledge my dad, my mom, and my siblings, for instilling a sense of freedom, creativity, and independence in me.

I would like to thank Prof. J. Scott Tyo, my advisor, for giving me the tools to accomplish research, while still allowing me the independence to choose my own research directions, all while giving me technical insight and sound advice. I would also like to thank Prof. Russell Chipman and Matthew Kupinski for assessing my dissertation material and for providing advice; a few of their questions have led to new research which I had not thought to pursue.

I would like to thank my current and former colleagues in the Advanced Sensing Laboratory, especially Andrey Alenin, Oscar Rodríguez-Herrera, Charles LaCasse, Mohan Xu, and Anael Guilmo for discussions about topics ranging from statistical decision theory to geometric optics. I would like to thank one of our undergraduate researchers in particular, Rafael Rojas, who helped with much of the instrument calibration.

Finally, I would like to thank all of my friends and colleagues.

## DEDICATION

*For Carrie Meharg, who taught me that science is inspiring;*

*for my wife, Erin;*

*and for my father, Stephen, my mother, Leslie.*

## TABLE OF CONTENTS

LIST OF FIGURES . . . . .	9
LIST OF TABLES . . . . .	18
ABSTRACT . . . . .	19
CHAPTER 1 INTRODUCTION . . . . .	21
1.1 Mathematical notation . . . . .	22
1.2 Motivation . . . . .	22
CHAPTER 2 REVIEW . . . . .	25
2.1 Polarized light . . . . .	25
2.2 Mueller-Stokes formalism . . . . .	31
2.3 Polarimetric instruments . . . . .	35
2.3.1 Passive polarimeters . . . . .	35
2.3.2 Active polarimeters . . . . .	37
2.4 Channeled systems . . . . .	39
2.4.1 Bandwidth in channeled systems . . . . .	41
2.4.2 Notation . . . . .	44
2.5 Applications . . . . .	47
CHAPTER 3 Bandwidth in spatio-temporal channeled systems . . . . .	49
3.1 Introduction . . . . .	49
3.2 Spatio-temporal channels . . . . .	49
3.3 Physical constraints . . . . .	52
3.4 Bandwidth . . . . .	56
3.4.1 Crosstalk . . . . .	58
3.5 Filtering . . . . .	61
3.6 Maximizing bandwidth . . . . .	62
3.7 Optimization . . . . .	65
3.8 Channel cancellation . . . . .	67
3.9 Bandwidth Optimization Discussion . . . . .	69
3.10 An Example . . . . .	70
3.10.1 Instrument layout . . . . .	71
3.10.2 Derivation of analyzer and generator equations . . . . .	72
3.10.3 Fourier transform of analyzer and generator equations . . . . .	74

TABLE OF CONTENTS – *Continued*

3.10.4 Discussion of parameters . . . . .	78
3.11 HyDMIP Results . . . . .	78
3.11.1 Other spatio-temporal systems . . . . .	83
3.12 Conclusions . . . . .	87
CHAPTER 4 Noise and Systematic Error in Channeled Polarimeters . . . . .	88
4.1 Introduction . . . . .	88
4.2 Noise and Error . . . . .	88
4.2.1 Systematic Error . . . . .	89
4.2.2 Random Noise . . . . .	90
4.3 Random noise in channeled systems . . . . .	91
4.3.1 Sensor noise . . . . .	93
4.3.2 Other noise . . . . .	95
4.4 Systematic error in channeled systems . . . . .	98
4.5 Spatio-temporal example . . . . .	103
4.5.1 Systematic error . . . . .	103
4.5.2 Noise . . . . .	108
4.6 Optimization . . . . .	112
4.7 Conclusion . . . . .	115
CHAPTER 5 Hybrid domain modulated imaging polarimeter (HyDMIP) . . . . .	116
5.1 Introduction . . . . .	116
5.2 Physical design . . . . .	117
5.2.1 Polarization state generator . . . . .	117
5.2.2 Polarization state analyzer . . . . .	125
5.2.3 Triggering . . . . .	130
5.2.4 Overview . . . . .	131
5.3 Control system . . . . .	131
5.3.1 Stage control . . . . .	133
5.3.2 Camera control . . . . .	135
5.3.3 Pre-processing . . . . .	137
5.3.4 Reconstruction . . . . .	137
5.4 Calibration . . . . .	138
5.4.1 Micropolarizer array . . . . .	138
5.4.2 PSA and PSG reference frame . . . . .	139
5.5 Semi-empirical model . . . . .	144
5.6 Filtering and sampling . . . . .	146
5.6.1 Spectral leakage . . . . .	147
5.6.2 Filter types . . . . .	151

TABLE OF CONTENTS – *Continued*

5.6.3	Implementation . . . . .	153
5.7	Validation . . . . .	153
5.8	Imaging results . . . . .	156
5.9	Discussion . . . . .	162
CHAPTER 6	Closing Remarks . . . . .	163
APPENDIX A	Derivations of Quad-retarder + micropolarizer array equations .	165
A.1	Two retarder Mueller matrix . . . . .	165
A.2	Fourier transform of two retarder Mueller matrix elements . . . . .	166
APPENDIX B	Instrument details . . . . .	169
B.0.1	PSG custom parts . . . . .	169
B.0.2	PSA custom parts . . . . .	171
B.0.3	Mounting and display hardware . . . . .	171
B.0.4	Code . . . . .	179
REFERENCES	. . . . .	181

## LIST OF FIGURES

2.1	Electromagnetic radiation can be conceptualized as a wave traveling in some direction, with components in the planes perpendicular to that propagation direction. The plane components define the polarization. Optical polarimeters indirectly measure a quantity proportional to the time averaged red ellipse and its direction shown in (b).	26
2.2	A Fourier photopolarimeter capable of measuring all 16 elements of the Mueller matrix of an optical system. This is a DRR configuration. Reproduced from Azzam [46].	40
2.3	Encoded spectrum and resultant channels from the Oka and Kato spectropolarimeter design. Reproduced from Oka and Kato [47].	40
2.4	A modulated measurement of the Stokes parameters using a rotating analyzer polarimeter, with each parameter band limited such that the signal can be ideally reconstructed. $W_B$ is bandwidth of each parameter, $f_0$ is the frequency of analyzer rotation, and $f_s$ is the detector sampling frequency. The dashed blue line indicates that $s_2$ is in the quadrature component of the side band. The configuration with maximum allowed bandwidth is shown. Reproduced from LaCasse <i>et al</i> [43].	42
2.5	The four components of the estimated Stokes parameters in the Fourier domain given by $\mathcal{F}\{w(t) * \mathbf{ZA}(t)I(t)\}_{t \rightarrow f}$ with $w(n) = \text{rect}\left(\frac{n}{16}\right)$ , a 16 sample rect window. (a) $\mathcal{F}\{\hat{s}_0\}$ , (b) $\mathcal{F}\{\hat{s}_1\}$ , (c) $\mathcal{F}\{\hat{s}_2\}$ , (d) $\mathcal{F}\{\hat{s}_3\}$ . Differences from a triangle function in the Fourier domain indicate errors in the data reduction method. Reproduced from LaCasse <i>et al</i> [43].	43
2.6	The four components of the estimated Stokes parameters given by $\mathcal{F}\{\mathbf{s}(t)\}_{t \rightarrow f} = \mathcal{F}\{\mathbf{ZA}(t)I(t)\}_{t \rightarrow f}$ prior to low pass filtering with $w(t)$ according to Equation (2.4.4). Also shown in the dotted line is the rectangular low pass filter that ideally reconstructs the correct individual Stokes parameters. The marking $\gamma$ is an example of self-error, while the marking $\epsilon$ is an example of cross-error. If the low pass filter $w(t)$ does not reject these frequencies outside of the shaded base band artifacts will arise due to these error terms. Reproduced from LaCasse <i>et al</i> [43].	43



LIST OF FIGURES – *Continued*

2.7	Locations of modulated side bands in the frequency domain for four imaging polarimeters: (a) a full micropolarizer array polarimeter, (b) a rotating retarder polarimeter, (c) a polarimeter consisting of a rotating retarder followed by a micropolarizer array polarimeter, (d) and a half wave retarder rotating $90^\circ$ per temporal sample in front of a micropolarizer array, which creates a polarimeter with maximum spatio-temporal bandwidth. Reproduced from LaCasse <i>et al</i> [33]. . . . .	45
3.1	Example of a spatio-temporal channel structure with $\delta$ -functions specific to $m_{23}$ . The maximum bandwidth corresponds to the minimum distance between two adjacent channels, taken over all possible adjacent channel pairs. . . . .	53
3.2	Optimal equal sphere packing arrangement for 3 dimensions : hexagonal close packed (hcp). Animated in the electronic version. . . . .	57
3.3	An example of convolution of data with a channel. The gray band represents the range of $\epsilon_c$ , resulting in data being outside of the $\epsilon_c$ range for $-0.5 < \nu < 1.5$ . This implies a bandwidth of 2 arb. units for the Mueller data. . . . .	59
3.4	An example of channel crosstalk. Mueller data is placed at two channels, with the distance (bandwidth) between them less than the bandwidth of the Mueller data. When added, the Mueller data from different channels adds together, leaving no remedy to differentiate data between channels in the region of bandwidth crossover (the neighborhood of 0) when given arbitrary Mueller data. . . . .	60
3.5	An example of how varying retardance parameters change channels in the channel space. Animated in the electronic version, use the controls displayed to view. . . . .	63
3.6	An example of how varying starting offset parameters change channels in the channel space. Animated in the electronic version, use the controls displayed to view. . . . .	64
3.7	An example of varying frequency parameters to move channels around in the channel space. Note the cancellation/addition at certain parameter values which opens up more bandwidth between the channels. Animated in the electronic version, use the controls displayed to view. . . . .	68

LIST OF FIGURES – *Continued*

3.9	Spatio-temporally channelled polarimeter schematic, <b>lp</b> =linear polarizer, <b>lr</b> =linear retarder, blue components denote the polarization state generator (PSG), green components denote the polarization state analyzer (PSA). The micropolarizer array is the conventional tiling, shown in (b). (a) was derived and modified from a figure created by Andrey Alenin. . . . .	71
3.10	Optimal channel structure for a specific quad retarder + micropolarizer system. The system is balanced (equal channel bandwidth). . . .	79
3.11	Close to optimal channel structure for a specific quad retarder + micropolarizer system. The system is balanced (equal channel bandwidth). The axes are identical to those in Figure 3.1. $\delta_2 = \pi$ here instead of the optimized $\delta_2 = \pi - \cos^{-1} \frac{1}{\sqrt{3}}$ , this retardance is what was available to test these types of systems in our physical instrument.	80
3.12	Optimized channel structure for a specific dual retarder + micropolarizer system. The system is balanced (equal channel bandwidth). The axes are identical to those in Figure 3.1. . . . .	81
3.13	Optimized channel structure for a dual retarder + micropolarizer array Stokes system. $\delta_1 = 3\pi/2, \delta_2 = \pi, \epsilon_1 = 0, \epsilon_2 = -0.03125, \nu_1 = \nu_0, \nu_2 = \nu_0/2$ . The axes are identical to those in Figure 3.1. . . . .	84
3.14	Optimized channel structure for a specific unbalanced quad retarder + micropolarizer system. The axes are identical to those in Figure 3.1.	85
3.15	Optimized channel structure for a specific unbalanced quad retarder + micropolarizer rank 10 pMMP system. The axes are identical to those in Figure 3.1. . . . .	86
4.1	$8 \times 8$ pixel array and 4 time point (256 total measurements) noise covariance matrix for thermal and read noise only (no shot noise). 100,000 realizations were Fourier transformed and the subsequent covariance matrix was computed between all Fourier coefficients. The imaginary part is not shown, since all imaginary parts of the matrix elements were less than 0.01. Both the thermal and read noise were assumed to be sampled from Gaussian distributions with means and standard deviations of $75e-$ . . . . .	95
4.2	$8 \times 8$ pixel array spatial Gaussian irradiance profile. The peak value is $25,000e-$ . . . . .	96
4.3	$8 \times 8$ pixel array Poisson noise realization from a spatial Gaussian profile in photoelectrons $e-$ . . . . .	96

LIST OF FIGURES – *Continued*

- 4.4  $8 \times 8$  pixel array and 4 time point (256 total measurements) noise covariance matrix for thermal, read noise, and shot noise with a Gaussian irradiance profile illuminating the detector. 100,000 realizations were Fourier transformed and the subsequent covariance matrix was computed between all Fourier coefficients. The imaginary part is not shown, since all imaginary parts of the matrix elements are small. The largest off off-diagonal magnitudes are about 0.16. . . . . 97
- 4.5 Channel effects as rotation wander noise increases for a carrier generation element like a retarder. On the left is the carrier function with noise, and on the right are the channels resulting from the carrier. The noise free carrier resulted in channels with an amplitude of 512. Animated in the electronic version. . . . . 99
- 4.6 An example of systematic error in a real instrument, the small triangles represent extraneous channels which are present due to retardance deviation from the specifications. . . . . 100
- 4.7 Channel effects as rotation drift occurs for a carrier generation element like a retarder. On the left is the carrier function with random biased drift applied, and on the right are the channels resulting from the carrier, blue, —, is the real part and red, —, is the imaginary part. Without drift there are two real (blue) channels with no imaginary components. Animated in the electronic version. . . . . 101
- 4.8 Channel effects of 3 missed trigger events in a physical instrument, likely due to a Microsoft Windows system interrupt. Note the spectral leakage present when the time shift occurs in the DFT window; when the DFT window is before any missed triggers there is little to no spectral leakage, and when the DFT window is after any missed triggers there is also little to no spectral leakage, but the channel structure between the two is inconsistent. Animated in the electronic version. . . . 102
- 4.9 Specified (left) and the mean of the physical (right) micropolarizer array tiling unit cells. Array adapted from a figure by Andrey Alenin. . 103

LIST OF FIGURES – *Continued*

- 4.10 Physical micropolarizer array channels vs ideal micropolarizer array channels. Each graph is a 2-dimensional DFT of the physical and ideal micropolarizer arrays. The ideal array parameters used the mean polarizer angle values taken from the physical microgrid. In the figures the axes ranges are  $-1/2 \leq \xi, \eta \leq 1/2$  and  $-6 \leq z \leq 6$ ; the  $\xi$  axis is on the lower right and the  $\eta$  axis is on the lower left. We can observe some extraneous spatial micropolarizer channels which occur at the  $(0, 0)$  and  $(-1/2, -1/2)$  locations due to the mean angles being slightly different from  $0^\circ, +45^\circ, -45^\circ, 90^\circ$ . The vertical  $z$ -axis is inverse hyperbolic sine scaled. The array size was  $980 \times 980$ . Note that the graphs have been downsampled for visualisation purposes, the actual  $\delta$ -functions are much thinner and difficult to view. . . . . 105
- 4.11 Physical micropolarizer array channels vs ideal micropolarizer array channels. Each graph is taken from the 2-dimensional DFT (Figure 4.10) along  $\eta = 0$  and  $\xi = 0$  respectively, since all the relevant channels lie on these lines. The vertical axis is inverse hyperbolic sine scaled. The array size was  $980 \times 980$ . Blue, —, denotes the ideal array and red, —, denotes the physical array. . . . . 106
- 4.12 An example of the effect of Gaussian illumination on the channel structure. Left pane shows the real part of the channel structure with Gaussian illumination in the  $\nu = 0$  plane for the system shown in Figure 3.10 and the channels for  $m_{00}, m_{11}, m_{12}$ , with the magnitude inversely hyperbolic sine scaled. The right pane is the spatial Gaussian irradiance profile at the object plane. Note that a narrow Gaussian profile was used to exaggerate the effect on the channels. Animated in the electronic version, with position of the Gaussian profile varying. 107
- 4.13 Gaussian detector noise effects on channel structure. Note that these are channels similar to the channels as shown in Fig. 3.10, but plotted along the  $\nu$  lines where the channels are located. They are not exactly the same channels because this model incorporated many of the systematic errors of the real instrument, resulting in different channel magnitudes. The left column is the real part, the right column is the imaginary part, there are 16 colors in each graph representing each Mueller matrix element. The first row is  $\nu$  at  $\xi = 0, \eta = -0.5$ , the second row is  $\nu$  at  $\xi = -0.5, \eta = 0$  and the third row is  $\nu$  at  $\xi = 0, \eta = 0$ . Animated in the electronic version. The animation is a spectrogram with window length 64 samples. . . . . 109

LIST OF FIGURES – *Continued*

4.14	Normalized error between a Poisson distribution with mean $N$ , denoted $pdf_{P(N)}$ , and the approximate Gaussian distribution with mean $\mu = N$ and variance $N$ , denoted $pdf_{G(N,N)}$ , as a function of the mean in photoelectrons, $e^-$ . . . . .	110
4.15	Uniform irradiance profile with Poisson noise vs Gaussian irradiance profile with Poisson noise for the $m_{00}$ channel. The y-axis is inversely hyperbolic sine scaled. Other channels have similar differences. All curves are taken in the $\nu = 0$ plane, along $\xi$ , i.e., the $\eta = 0$ line. (a) shows the case where the peak pixel irradiance corresponds to $6e^-$ detected per pixel, and (b) shows the case where the peak pixel irradiance corresponds to $40,006e^-$ detected per pixel. . . . .	111
4.16	An example of varying the starting position of a retarder, $\epsilon_2$ to optimize the impact of extraneous channels on the bandwidth optimized system shown in Figure 3.10 for $m_{22}$ . . . . .	113
4.17	Re-optimized channel structure for a specific quad retarder + micropolarizer system with extraneous channels. The axes are identical to those in Figure 3.1. . . . .	114
5.1	Aplanatic lens system. The specified lens has $f = 51mm, D = 31.7mm, L = 25.4mm, WD = 35mm$ . These specifications are for $633nm$ , so the focal length will change slightly for $671nm$ light. Courtesy of CVI Melles Griot. . . . .	121
5.2	CVI lens assembly mount. Threaded to Thorlabs SM-1 specifications so that off the shelf retainer rings can be used. . . . .	123
5.3	Source assembly, zoom in for detail. The laser beam is expanded to about a $8mm$ beam diameter prior to passing through the retarders. . . . .	124
5.4	Receiver lens design. The left plane is the intermediate image with a total height of $24mm$ , the second plane is $L_2$ as a thin lens, the third plane is $L_3$ as a thin lens, and the fourth plane is the sensor plane with a total height of $\sim 10.5mm$ . The actual sensor is slightly smaller with a height of $9.1mm$ , but this was the closest match available with the focal lengths available from Pentax. . . . .	127
5.5	Receiver afocal lens mount body design. The Samsung adapters are not shown, but attach to either side. Zoom in for detail. . . . .	128
5.6	Receiver afocal lens mount assembly. The modified Samsung adapters are shown exploded in the bottom right view. Zoom in for detail. . . . .	129
5.7	Receiver (PSA) assembly, zoom in for detail. . . . .	130
5.8	Hybrid Domain Modulated Imaging Polarimeter - Portable (HyDMIP-P)	131

LIST OF FIGURES – *Continued*

5.9	Control system design overview flowchart. RS denotes rotation stage. The motor control and camera control modules are fully implemented. The remaining modules are only partially or not implemented at this time. The reconstruction module is implemented in Matlab, but cannot operate in real time at the moment. . . . .	134
5.10	Class diagram for the camera module. The lines with arrows denote inheritance. . . . .	136
5.11	Histograms of angles of the micropolarizer elements. The $-45^\circ$ histogram is much narrower than the rest, with the center bin reaching around 70,000 counts, exceeding the range of the graph. Each histogram shows the angular spread from the specified values over the pixels. . . . .	140
5.12	Histograms of the entire sensor with $T_{\min} = a$ from Equation (5.4.1) in red and $T_{\max} = a + b$ from Equation (5.4.1) in blue. The odd shape of the $T_{\max}$ histogram is explained in Figure 5.13. . . . .	141
5.13	Histograms of $T_{\max} = a + b$ from Equation (5.4.1) for each micropolarizer orientation. Notice the much higher overall $T_{\max}$ for the $90^\circ$ elements. This is a result of the micropolarizer etching and manufacturing process [136]. . . . .	142
5.14	Sum of cosines curve sampled at the Nyquist frequency (blue dots) and slightly higher than the Nyquist frequency (red squares). . . . .	148
5.15	Fast Fourier transforms of the sampled sum of cosines curve, blue corresponds to the Nyquist sampling rate, and red corresponds to the slightly greater than Nyquist rate. . . . .	149
5.16	Space-time $(x, y, t)$ modulated data "cube" transformed into Fourier domain. . . . .	152
5.17	Planck-taper filters used for our system. In (b) yellow represents a magnitude of 1 and blue represents a magnitude close to, but greater than, 0. . . . .	154
5.18	Filtered data in the Fourier domain, filtered via Planck taper filters. The figure depicts an isosurface (surface of equal magnitude) to show the data. . . . .	155

LIST OF FIGURES – *Continued*

5.19	Polarizer channel response from the semi-empirical model (red, dashed) versus the physical instrument (blue, solid). The left column are the real parts and the right column are the imaginary parts. The first row is taken along the center column line, $v$ , of the channel structure where $\xi = \eta = 0$ , the second row is taken along the line $\xi = 0, \eta = -0.5$ and the third row is taken along the line $\xi = -0.5, \eta = 0$ . The channels match well, but there are some differences. . . . .	157
5.20	Polarizer channel response differences between the semi-empirical (SE) model and the physical instrument, the physical instrument data was subtracted from the SE model data. The left column are the real parts and the right column are the imaginary parts. The first row is taken along the center column line, $v$ , of the channel structure where $\xi = \eta = 0$ , the second row is taken along the line $\xi = 0, \eta = -0.5$ and the third row is taken along the line $\xi = -0.5, \eta = 0$ . The channels match well, but there are some differences. . . . .	158
5.21	Moving hand Mueller matrix reconstruction. All elements except for $m_{00}$ are normalized by $m_{00}$ . Animated in the electronic version. . . . .	159
5.22	Moving hand reconstruction. On the left is the reconstruction of $m_{01}$ and on the right is a downsampled raw camera image. Animated in the electronic version. . . . .	160
B.1	Renderings of the source and receiver for HyDMIP. . . . .	169
B.2	PSG lightweight breadboard. . . . .	170
B.3	PSG beam expander mount. . . . .	170
B.4	PSG laser mount. . . . .	171
B.5	PSA lightweight breadboard. . . . .	172
B.6	Modified Samsung K-mount adapter, additional detail and model. . . . .	172
B.7	4D Technologies <sup>TM</sup> micropolarizer camera mount. . . . .	173
B.8	Dual rotation stage mounting plate. Mounts 2 IntelLiDrives stages back to back with high tolerances between the two mounting surfaces. Used in both the PSG and PSA. . . . .	174
B.9	Retarder mounting fixture. Bolts to IntelLiDrives stages. The interior is threaded to ThorLabs SM-2 for the use of COTS retaining rings. Used in both the PSG and PSA. . . . .	175
B.10	Angle bracket for monitor mount. . . . .	176

LIST OF FIGURES – *Continued*

- B.11 Multi-mount bracket. Mounts the PSG and PSA assembly together on the same plane. Allows for the angle between the PSA and PSG assembly to be adjusted. Can be mounted to a standard video (television) grade tri-pod. . . . . 177
- B.12 Multi-mount bracket assembly. Mounts the PSA and PSG and viewing monitor together, and mounts to a tripod. . . . . 178



## LIST OF TABLES

2.1	Notation for channels. . . . .	46
3.1	Table of COTS polarization components. . . . .	50
3.2	Specifications for optimized quad retarder + micropolarizer system. .	82
3.3	Specifications for optimized dual retarder + micropolarizer system. $e$ in $2\pi/e$ here denotes the natural number. . . . .	82
3.4	Performance gains and tradeoffs for various system types. Note that temporal bandwidth here is given as a fraction of some maximum fre- quency set to $1/s$ . When sampling and Nyquist is taken into account, the actual framerate of the instrument will be $1/2$ of temporal band- width listed in the table, multiplied by the base camera framerate. Sensor resolution is the reconstructable resolution of the Mueller ma- trix images, as a fraction of the base full sensor size. . . . .	83
5.1	KAI-1010 sensor specifications for $671nm$ . . . . .	119
5.2	System specifications . . . . .	132

## ABSTRACT

Polarimetric systems design has seen recent utilization of linear systems theory for system descriptions. Although noise optimal systems have been shown, bandwidth performance has not been addressed in depth generally and is particularly lacking for Mueller matrix (active) polarimetric systems. Bandwidth must be considered in a systematic way for remote sensing polarimetric systems design. The systematic approach facilitates both understanding of fundamental constraints and design of higher bandwidth polarimetric systems. Fundamental bandwidth constraints result in production of polarimetric “artifacts” due to channel crosstalk upon Mueller matrix reconstruction.

This dissertation analyzes bandwidth trade-offs in spatio-temporal channeled Mueller matrix polarimetric systems. Bandwidth is directly related to the geometric positioning of channels in the Fourier (channel) space, however channel positioning for polarimetric systems is constrained both physically and by design parameters like domain separability. We present the physical channel constraints and the constraints imposed when the carriers are separable between space and time.

Polarimetric systems are also constrained by noise performance, and there is a trade-off between noise performance and bandwidth. I develop cost functions which account for the trade-off between noise and bandwidth for spatio-temporal polarimetric systems. The cost functions allow a systems designer to jointly optimize systems with good bandwidth and noise performance. Optimization is implemented for a candidate spatio-temporal system design, and high temporal bandwidth systems resulting from the optimization are presented. Systematic errors which impact the bandwidth performance and mitigation strategies for these systematic errors are also presented.

Finally, a portable imaging Mueller matrix system is built and analyzed based on the theoretical bandwidth analysis and system bandwidth optimization. Temporal bandwidth performance is improved by 300% over a conventional dual rotating retarder Mueller matrix polarimeter. Reconstruction results from the physical instrument are presented, and issues with the implemented system design are discussed.

## CHAPTER 1

### INTRODUCTION

Research included in this dissertation comprises published work, in press work, or in preparation work for submission to journals. *Portions of this work will be reproduced here and may be copied verbatim, however quotes will not be used [1–8].* A list of the novel contributions contained in this dissertation are below:

- Cost functions characterizing noise and channel bandwidth,
- A channel model for a specific spatio-temporal system,
- Analysis of geometric constraints on spatio-temporal channel structure due to carrier function separability,
- Optimization for temporal bandwidth on a spatio-temporal model,
- A set of optimized systems showing temporal bandwidth improvements over conventional systems,
- Analysis of noise and systematic error analysis on channeled systems, including spatial illumination variation effects,
- The design and implementation of a candidate portable Mueller matrix polarimeter,
- A semi-empirical instrument model,
- Conditions on temporal carriers for spectral leakage free channels,
- Reconstructed Mueller matrix image data from the implemented instrument.

## 1.1 Mathematical notation

This dissertation will use typical or standard notation from mathematics, physics, and optics. Matrices will be represented as uppercase bold,  $\mathbf{M}$ , vectors or parameter/variable packs will be lowercase bold,  $\mathbf{x}$ ,  $i$  typically denotes the standard imaginary unit (although in context it may be used as an index), real and imaginary parts of complex quantities are represented by  $\Re\{z\}$  and  $\Im\{z\}$  respectively. For brevity the element of symbol,  $\in$ , will occasionally be used, e.g.,  $z \in \mathbb{C}$  for  $z$  is a complex number. Blackboard bold represents a number set type, e.g.,  $\mathbb{R}^4$  for a 4-dimensional real field. Additional notation will be introduced as necessary.

## 1.2 Motivation

Many attempts have been made to utilize polarimetric imaging and polarimetric measurements for remote sensing purposes [9–18]. Polarimetric imaging is intrinsically promising to fill voids in current imaging schemes; the underlying phenomenology and imaging physics can yield improvements in target classification tasks which cannot be easily realized by other modalities such as hyperspectral. Issues with polarimetric imaging include instrument complexity, computational limitations, lack of statistically robust polarimetric datasets, lack of generalization of contrast improvements from the lab to the field, and measurement bandwidth/speed limitations of polarimetric instruments.

Even with the drawbacks outlined above, imaging polarimeters have proven to be powerful tools for a wide variety of remote sensing tasks. Active and passive polarimeters for remote sensing have been developed for usage in many regions of the optical spectrum. Polarimeters have been used to detect targets with clutter [19], for target identification [20], to penetrate scattering media [20–23] and to aid in three-dimensional image reconstruction [24]. Polarimeters have been utilized in atmospheric sensing applications, including determination of aerosol properties [16], discrimination of ice/water phase particulates in clouds [25], and observation of plasmas in rocket engine exhaust [26]. Polarimeters have been used to estimate vegetation height, type, and quantity [27, 28]. Polarimeters have been also been used in industry to detect defects in and calibrate liquid crystal displays [29, 30].

Recent advances in polarimetric instrument design [31–35] have facilitated *systematic* instead of *ad-hoc* designs for specific tasks using a linear systems framework. Particularly Alenin and Tyo [34] have provided a general framework to design any periodically modulated polarimetric instrument. An instrument will be presented here which utilizes this framework to design an active polarimeter optimized for temporal bandwidth, however more general case(s) of spatio-temporal channels are also discussed. Motivating examples of the need for more temporal bandwidth in active polarimeters are polarimetric detection of optically active materials [36], dermatology [37], polarimetric retinal imaging [38–40] and skin cancer detection [41, 42].

Polarimetric imaging for remote sensing assumes classes of objects/targets of interest which the operator has at most only partial control over (a medical patient can be asked to sit still, but will still involuntarily move, and satellites measure moving objects which are under no control whatsoever). The lack of control of and ability to directly measure movement, surface structure, chemical composition, among other factors requires remote sensing polarimeters to maximize bandwidth in order to make useful measurements. The *canonical* approach to bandwidth in polarimetric instruments is linear systems and channeled system design [34]. Faster instruments minimize the need for complex image registration algorithms and processing. Higher resolution instruments may allow for faster acquisition, or for more detail being resolved polarimetrically.

Prior to the work by LaCasse *et al* [33, 43, 44], bandwidth in channeled polarimetric systems had not been addressed in depth, or only addressed as a consequence of instrumental “error.” Additionally, prior to Alenin and Tyo [34] channeled systems were designed in an *ad-hoc* manner. In this dissertation I address bandwidth using the *systematic* design tools introduced by Alenin and Tyo [34] for a hybrid spatio-temporally channeled Mueller matrix polarimetric system.

Although this dissertation primarily addresses bandwidth in channeled systems, and channels of specific systems, noise in channeled systems must be taken into account. Alenin and Tyo have addressed noise in a systematic way for channeled systems, and have presented channeled systems which are optimized for noise performance [34, 45]. Channeled systems which have been optimized for bandwidth must also address noise and system conditioning to noise, or be optimized under some

system conditioning constraints. If system conditioning is not considered, then bandwidth may be high for the system, but noise effects may also be high, resulting in poor Stokes parameter or Mueller matrix reconstruction.

## CHAPTER 2

### REVIEW

Channeled Mueller matrix polarimeters and the concept of using these channels was first introduced by Azzam [46]. Azzam published a very specific case, 1) a specific temporal framework was analyzed, 2) an implicit assumption about the object was made, *the object had no temporal bandwidth*, i.e., the object was constant in time. Oka, Sabatke, Derniak, Kudenov, and Hagen then demonstrated both spectrally channeled and spatially (over spectrum) channeled systems [31, 47–51], mostly Stokes polarimeters. Dubreuil *et al* [52] and Hagen *et al* [53] presented spectrally channeled Mueller matrix polarimeters, which are non-imaging since the focal plane arrays are used to resolve the spectrum. LaCasse, Chipman, Tyo, and LeMaster and Hirakawa [35, 43, 44] described micropolarizer array partial Stokes polarimeters as channeled systems, and LaCasse *et al* presented a spatio-temporally modulated hybrid channeled Stokes system [33], and subsequently both Myhre *et al* [54] and Zhao *et al* [55] presented spatially modulated full Stokes polarimeters. Finally Alenin and Tyo [34] formalized a general framework which describes channeled polarimeters almost completely, both Mueller and Stokes.

#### 2.1 Polarized light

The “polarization” of light arises from the 3 spatial degrees of freedom allowed by Maxwell’s equations [56]. The propagation of light can be formalized via the *Poynting vector*,  $\mathbf{s}(\mathbf{r}, t)$ , which quantifies the directional energy flux density, with units  $W/m^2$ . The vector  $\mathbf{r} = [x \ y \ z]^T$  denotes the spatial coordinates and  $t$  denotes time. At any given instant in time and at some given point in space, the complex valued  $\mathbf{k}$ -vector will have a direction, and the polarization properties of the light are defined to be in the plane perpendicular to the real part of the  $\mathbf{k}$ -vector,  $\Re\{\mathbf{k}\}$ . More generally, as time advances, the polarization is the dynamic behavior of the vector  $\mathbf{E}$ -field in the plane perpendicular to  $\Re\{\mathbf{k}\}$  [56]. For the case of a propagating plane wave in isotropic,



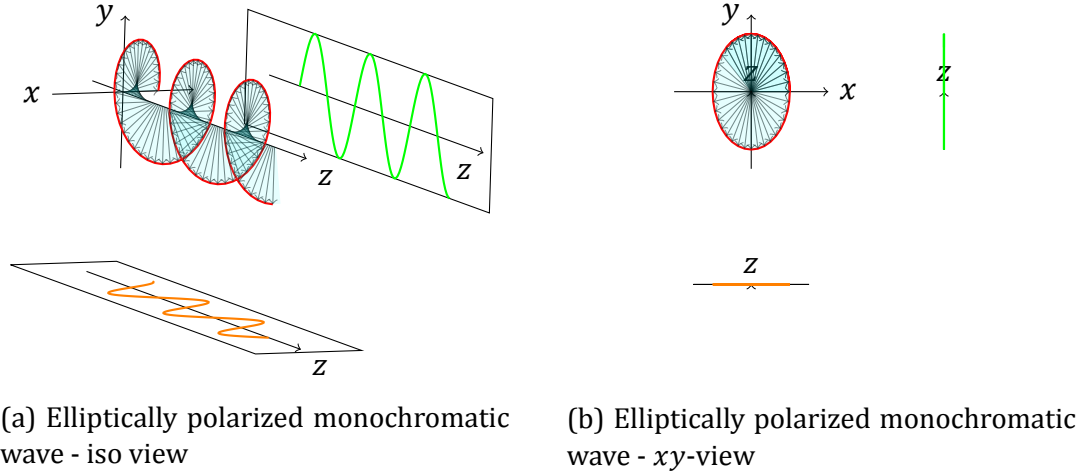


Figure 2.1: Electromagnetic radiation can be conceptualized as a wave traveling in some direction, with components in the planes perpendicular to that propagation direction. The plane components define the polarization. Optical polarimeters indirectly measure a quantity proportional to the time averaged red ellipse and its direction shown in (b).

homogeneous, linear media,  $\mathbf{s}(\mathbf{r}, t)$  and  $\mathbf{k}$ , the complex valued wave-vector [56], are parallel vectors when  $\mathbf{s}(\mathbf{r}, t) \neq \mathbf{0}$ . For the plane wave shown in Figure 2.1, the Poynting vector is parallel to the  $+z$  direction, and the polarization behavior is described in the set of  $xy$ -planes. Here, elliptical polarization behavior is shown, elliptical because of the shape that is drawn out when looking down the axis of propagation, as shown in Figure 2.1b.

Optical detectors cannot measure the phase of light at optical wavelengths directly. They can, however, measure a quantity proportional to the space integrated time-averaged Poynting vector, which is equivalent to space integrated irradiance. Additionally, light incident on a detector or detector element (pixel) is not a monochromatic plane wave. Any vector field whose components are elements of a separable Hilbert space (of functions) can be represented as a possibly infinite, but countable, superposition of plane waves [57], so a general field with specific spatial properties and spectral properties,  $\mathbf{f}(\mathbf{r}, t)$  can be represented as

$$\mathbf{f}(\mathbf{r}, t) = \sum_j \mathbf{e}_j \exp [i (\mathbf{k}_j \cdot \mathbf{r} - \omega_j t)] \quad (2.1.1)$$

where the  $\mathbf{e}_j = (E_{j_1}, E_{j_2}, E_{j_3})$  are possibly complex field amplitudes. The sum can

be reordered since it is countable and uniformly convergent (if it wasn't then an arbitrary field couldn't be represented as a superposition of plane waves; this restricts the forms the sum can take mathematically). In this specific case, it can be assumed that  $\mathbf{f}(\mathbf{r}, t)$  is statistically stationary in time, given approximately constant illumination of an object over the integration time of the sensor, and for coherent or nearly coherent illumination, given that the number of the speckles on each detector or detector element is large. For example, in a sun-lit scene, the sun position with respect to the observed object is moving very slowly compared with acquisition time and the light illuminating the object being imaged has undergone multiple scattering events. This type of illumination which can be modeled using stochastic processes. The Poynting vector is then

$$\mathbf{s}(\mathbf{r}, t) = \Re \{ \mathbf{e}(\mathbf{r}, t) \} \times \Re \{ \mathbf{h}(\mathbf{r}, t) \} \quad (2.1.2)$$

where  $\times$  denotes the cross product,  $\mathbf{e}$  is the complex electric field, and  $\mathbf{h}$  is the complex magnetic (induction) field. For an electromagnetic plane wave in a homogeneous, isotropic, linear medium, i.e.

$$\begin{aligned} \text{Electric vector field} &= \Re \{ \mathbf{e}_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \} \\ \text{Magnetic vector field} &= \Re \{ \mathbf{h}_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \}, \end{aligned} \quad (2.1.3)$$

where  $\mathbf{e}_0 = \mathbf{e}'_0 + i\mathbf{e}''_0$  is the possibly complex electric amplitude,  $\mathbf{h}_0 = \mathbf{h}'_0 + i\mathbf{h}''_0$  is the possibly complex magnetic amplitude, and  $\mathbf{k} = \mathbf{k}' + i\mathbf{k}''$  is the possibly complex wave vector; it can be shown that the time average of a Poynting vector can be represented in the following form:

$$\langle \mathbf{s}(\mathbf{r}) \rangle_t = \frac{e^{-2\mathbf{k}'' \cdot \mathbf{r}} (\|\mathbf{e}'_0\|^2 + \|\mathbf{e}''_0\|^2) (\mu' \mathbf{k}' + \mu'' \mathbf{k}'') - 2(\mu' \mathbf{k}'' - \mu'' \mathbf{k}') \times (\mathbf{e}'_0 \times \mathbf{e}''_0)}{2 z_0 (\omega/c) (\mu'^2 + \mu''^2)} \quad (2.1.4)$$

where  $\mu(\omega) = \mu'(\omega) + i\mu''(\omega) = 1 + \chi_m(\omega)$  is the complex valued relative magnetic permeability and  $z_0$  is the impedance of free space [56]. This assumes that  $\mathbf{s}(\mathbf{r}, t)$  is statistically stationary. Note that the  $\times$  symbols in the above equation denote cross products. If the plane wave is in a typical optical medium then  $\epsilon''(\omega) \approx 0$  and  $\mu''(\omega) \approx 0$ , where  $\epsilon(\omega) = \epsilon'(\omega) + i\epsilon''(\omega) = 1 + \chi_e(\omega)$  is the complex valued

relative permittivity. Additionally if  $\mathbf{k}'' = 0$  then Equation (2.1.4) can be simplified to

$$\langle \mathbf{s}(\mathbf{r}) \rangle_t = \frac{1}{2} \frac{(\|\mathbf{e}_0'\|^2 + \|\mathbf{e}_0''\|^2)}{z_0 \sqrt{\mu(\omega)/\epsilon(\omega)}} \hat{\mathbf{k}} \quad (2.1.5)$$

where  $\hat{\mathbf{k}} = \mathbf{k}/\|\mathbf{k}\|$  [56]. The irradiance at a single point,  $\mathbf{r}$ , is usually defined as the time average of the Poynting vector.

Equation (2.1.5) is for a single plane wave, so we must generalize the equation for an arbitrary superposition of plane waves. For an arbitrary superposition of plane waves

$$\mathbf{s}(\mathbf{r}, t) = \Re \left\{ \sum_j \mathbf{e}_j \exp [i(\mathbf{k}_j \cdot \mathbf{r} - \omega_j t)] \right\} \times \Re \left\{ \sum_j \mathbf{h}_j \exp [i(\mathbf{k}_j \cdot \mathbf{r} - \omega_j t)] \right\}. \quad (2.1.6)$$

Since

$$\Re \{z\} = \frac{1}{2} (z + z^*) \quad (2.1.7)$$

where  $*$  denotes the complex conjugate Equation (2.1.6) can be rewritten

$$\mathbf{s}(\mathbf{r}, t) = \frac{1}{4} \sum_j \sum_{j'} \left( \begin{array}{l} \mathbf{e}_j \times \mathbf{h}_{j'} \exp [i(\mathbf{k}_j + \mathbf{k}_{j'}) \cdot \mathbf{r} - i(\omega_j + \omega_{j'})t] \\ + \mathbf{e}_j \times \mathbf{h}_{j'}^* \exp [i(\mathbf{k}_j - \mathbf{k}_{j'}) \cdot \mathbf{r} - i(\omega_j - \omega_{j'})t] \\ + \mathbf{e}_j^* \times \mathbf{h}_{j'} \exp [-i(\mathbf{k}_j - \mathbf{k}_{j'}) \cdot \mathbf{r} + i(\omega_j - \omega_{j'})t] \\ + \mathbf{e}_j^* \times \mathbf{h}_{j'}^* \exp [-i(\mathbf{k}_j + \mathbf{k}_{j'}) \cdot \mathbf{r} + i(\omega_j + \omega_{j'})t] \end{array} \right) \quad (2.1.8)$$

In the above equation the terms inside the parentheses are added because the equation will not fit on a single line, the parentheses do not represent a vector. When  $j = j'$  the time average of  $\mathbf{s}$  for that portion of the sum is the same as in Equation (2.1.4). What about when  $j \neq j'$ ? Then the electric field is not associated with the magnetic field, and hence the cross product is not perpendicular to either plane wave. The time average can be taken inside the sum and then the following four integrals must be evaluated:

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \exp [-i(\omega_j + \omega_{j'})t] dt \quad (2.1.9)$$

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \exp[-i(\omega_j - \omega_{j'})t] dt \quad (2.1.10)$$

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \exp[i(\omega_j - \omega_{j'})t] dt \quad (2.1.11)$$

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \exp[i(\omega_j + \omega_{j'})t] dt \quad (2.1.12)$$

Since they are sinusoids, they average to zero in all except the following special cases:

$$\omega_j = \omega_{j'}, \quad (2.1.13)$$

$$\omega_j = -\omega_{j'}. \quad (2.1.14)$$

From Maxwell's equations, each plane wave component has the following property

$$\mathbf{k}_j \cdot \mathbf{k}_j = \mathbf{k}'_j \cdot \mathbf{k}'_j - \mathbf{k}''_j \cdot \mathbf{k}''_j + 2i\mathbf{k}'_j \cdot \mathbf{k}''_j = \omega_j^2 \epsilon_0 \mu_0 \epsilon(\omega_j) \mu(\omega_j) = \left(\frac{\omega_j}{c}\right)^2 \epsilon(\omega_j) \mu(\omega_j) \quad (2.1.15)$$

but for imaging systems the field is usually propagating in air, which is close enough to free space such that  $\epsilon(\omega_j) = \mu(\omega_j) \approx 1$ . This implies there is no imaginary component of the equation so

$$\mathbf{k}_j \cdot \mathbf{k}_j = \mathbf{k}'_j \cdot \mathbf{k}'_j = \left(\frac{\omega_j}{c}\right)^2 \quad (2.1.16)$$

If  $\omega_j = \omega_{j'}$  then

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \exp[-i(\omega_j + \omega_{j'})t] dt = 0 \quad (2.1.17)$$

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \exp[-i(\omega_j - \omega_{j'})t] dt = 1 \quad (2.1.18)$$

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \exp[i(\omega_j - \omega_{j'})t] dt = 1 \quad (2.1.19)$$

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \exp[i(\omega_j + \omega_{j'})t] dt = 0 \quad (2.1.20)$$

and if  $\omega_j = -\omega_{j'}$  then

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \exp[-i(\omega_j + \omega_{j'})t] dt = 1 \quad (2.1.21)$$

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \exp[-i(\omega_j - \omega_{j'})t] dt = 0 \quad (2.1.22)$$

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \exp[i(\omega_j - \omega_{j'})t] dt = 0 \quad (2.1.23)$$

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \exp[i(\omega_j + \omega_{j'})t] dt = 1. \quad (2.1.24)$$

This implies that

$$\begin{aligned} \langle \mathbf{s}(\mathbf{r}, t) \rangle_t &= \frac{1}{4} \sum_j \sum_{j'} \left( \begin{array}{l} \mathbf{e}_j \times \mathbf{h}_{j'} \exp(i(\mathbf{k}_j + \mathbf{k}_{j'}) \cdot \mathbf{r}) \\ + \mathbf{e}_j \times \mathbf{h}_{j'}^* \exp(i(\mathbf{k}_j - \mathbf{k}_{j'}) \cdot \mathbf{r}) \\ + \mathbf{e}_j^* \times \mathbf{h}_{j'} \exp(-i(\mathbf{k}_j - \mathbf{k}_{j'}) \cdot \mathbf{r}) \\ + \mathbf{e}_j^* \times \mathbf{h}_{j'}^* \exp(-i(\mathbf{k}_j + \mathbf{k}_{j'}) \cdot \mathbf{r}) \end{array} \right) \\ &= \frac{1}{4} \sum_j \sum_{j'} \left( \begin{array}{l} 2\Re\{\mathbf{e}_j \times \mathbf{h}_{j'} \exp(i(\mathbf{k}_j + \mathbf{k}_{j'}) \cdot \mathbf{r})\} \\ + 2\Re\{\mathbf{e}_j \times \mathbf{h}_{j'}^* \exp(i(\mathbf{k}_j - \mathbf{k}_{j'}) \cdot \mathbf{r})\} \end{array} \right) \end{aligned} \quad (2.1.25)$$

where the  $j'$  indices are restricted to only those where  $\omega_j = \omega_{j'}$  or  $\omega_j = -\omega_{j'}$ <sup>1</sup> for each fixed  $j$ . Equation (2.1.16) also forces the following condition:

$$\|\mathbf{k}_j\|^2 = \|\mathbf{k}_{j'}\|^2 \quad (2.1.26)$$

This means that any corresponding  $\mathbf{k}_{j'}$  vector must lie on the sphere of radius  $\|\mathbf{k}_j\|$ .

Finally, given a 2-dimensional detector (or pixel) with a spatial area defined by  $D(\mathbf{r})$ , typically square or rectangular for pixels, and sometimes round for more sensitive single pixel detectors, we must integrate spatially to obtain the integrated irradiance which is proportional to the value read out by the detector,  $V_d$ :

$$V_d = \int_{D(\mathbf{r})} \langle \mathbf{s}(\mathbf{r}, t) \rangle_t d^2A, \quad (2.1.27)$$

where  $d^2A$  is the appropriate differential area element for the 2-dimensional surface  $D(\mathbf{r}) \in \mathbb{R}^3$ .

Equation (2.1.27) describes the most general case, formally, for the irradiance of an electromagnetic field on an optical detector, given a measurement plane located

---

<sup>1</sup>Note that negative frequencies are allowed here and are required from the use of the Fourier transform to describe a real field, they are a mathematical construct denoting quadrature, a purely real trigonometric form of the Fourier transform could also be used, without any negative frequencies [58]

in isotropic, homogeneous, linear media, however we cannot fully reconstruct the incident field from  $\langle \mathbf{s}(\mathbf{r}, t) \rangle_t$ , since we have a many to one mapping (many field variables to a single  $V_d$  value). Note that this analysis assumes measuring the field at the detector plane, and propagation can be handled via the Fresnel propagation equations if needed. Given this fact, a subset of the complete polarization picture can be measured via irradiance (or radiance) modulation and correlations which extract information about the time averaged version of the ellipse shown in Figure 2.1, *over the weighted sum of all wavelengths which the detector is sensitive to*. This measurement defined polarization information has two commonly used formalisms, the coherence formalism [59] and the Mueller-Stokes formalism [60,61]. In the remote sensing and polarimetric systems community, the Mueller-Stokes formalism is typically used.

## 2.2 Mueller-Stokes formalism

A compressed overview of the Mueller-Stokes mathematical formalism is presented here, a complete description can be found elsewhere [4, 60, 61]. Stokes parameters and linear transformations of those parameters are described by [60]

$$\mathbf{s} = \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} \propto \begin{bmatrix} \langle |E_x|^2 \rangle + \langle |E_y|^2 \rangle \\ \langle |E_x|^2 \rangle - \langle |E_y|^2 \rangle \\ 2\Re\langle E_x E_y^* \rangle \\ 2\Im\langle E_x E_y^* \rangle \end{bmatrix}, \text{ where } s_0 > 0, \quad s_0^2 \geq s_1^2 + s_2^2 + s_3^2 \quad (2.2.1)$$

and

$$\mathbf{s}_{\text{out}} = \mathbf{M} \cdot \mathbf{s}_{\text{in}} \quad (2.2.2)$$

$s_0$  represents (is proportional to) the total irradiance,  $s_1$  represents the prevalence of horizontal linear polarization over vertical linear polarization,  $s_2$  represents the prevalence of linear polarization oriented at  $45^\circ$  over linear polarization oriented at  $-45^\circ$ , and  $s_3$  represents the prevalence of right circular polarization over left circular polarization. A Mueller matrix,  $\mathbf{M} \in \mathbb{R}^{4 \times 4}$ , represents a linear transformation of these four Stokes parameters and is therefore a four-by-four matrix.  $\mathbf{M} \in \mathbb{R}^{4 \times 4}$  has constraints on it in order to be a physically realizable Mueller matrix [62]. Active (aka Mueller) polarimetric instruments must encode polarization information

in irradiance, radiance, or radiant flux, etc. to infer the Mueller matrix of some object. A Mueller matrix object,  $\mathbf{M}_{\text{obj}}(\mathbf{x})$ , can then, in a generic way, be measured over the domain  $\mathbf{x} = [x \ y \ z \ t \ \sigma]^T$ , where  $x, y, z$  are spatial domains,  $t$  is time, and  $\sigma$  is wavenumber (also denoted  $k$ ) or some related domain like optical path difference (OPD) or wavelength,  $\lambda$ . Equation (2.2.2) then becomes [4]

$$\begin{bmatrix} s_{0,\text{out}}(\mathbf{x}) \\ s_{1,\text{out}}(\mathbf{x}) \\ s_{2,\text{out}}(\mathbf{x}) \\ s_{3,\text{out}}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} m_{00}(\mathbf{x}) & m_{01}(\mathbf{x}) & m_{02}(\mathbf{x}) & m_{03}(\mathbf{x}) \\ m_{10}(\mathbf{x}) & m_{11}(\mathbf{x}) & m_{12}(\mathbf{x}) & m_{13}(\mathbf{x}) \\ m_{20}(\mathbf{x}) & m_{21}(\mathbf{x}) & m_{22}(\mathbf{x}) & m_{23}(\mathbf{x}) \\ m_{30}(\mathbf{x}) & m_{31}(\mathbf{x}) & m_{32}(\mathbf{x}) & m_{33}(\mathbf{x}) \end{bmatrix} \cdot \begin{bmatrix} s_{0,\text{in}} \\ s_{1,\text{in}} \\ s_{2,\text{in}} \\ s_{3,\text{in}} \end{bmatrix} \quad (2.2.3)$$

where  $\mathbf{s}_{\text{in}}$  is fixed due to typical polarimetric instrument design. Usually polarization information encoding is implemented by applying a known Mueller matrix to some fixed  $\mathbf{s}_{\text{in}}$  via a polarization changing optical element. Even with multiple polarization changing elements, at some point in the instrument design a known and fixed  $\mathbf{s}_{\text{in}}$  will be input into the system. Optical detectors measure some quantity proportional to  $s_{0,\text{out}}(\mathbf{x})$  over the domain  $\mathbf{x}$ .

An active instrument system equation can then be mathematically defined by

$$\mathbf{s}_{\text{out}}(\mathbf{x}) = \mathbf{A}(\mathbf{x}) \cdot \mathbf{M}_{\text{obj}}(\mathbf{x}) \cdot \mathbf{G}(\mathbf{x}) \cdot \mathbf{s}_{\text{in}} \quad (2.2.4)$$

where  $\mathbf{G}(\mathbf{x})$ ,  $\mathbf{A}(\mathbf{x})$  are denoted the generator and analyzer Mueller matrices respectively, and  $\mathbf{M}_{\text{obj}}(\mathbf{x})$  is the unknown Mueller matrix being measured [60, 61]. Usually, it is assumed that  $\mathbf{M}_{\text{obj}}(\mathbf{x}) = \mathbf{M}_{\text{obj}}$ , a constant matrix. A typical measurement scheme entails measuring or sampling some set of points  $\{\mathbf{x}_n\}$  or since  $\mathbf{G}(\mathbf{x})$ ,  $\mathbf{A}(\mathbf{x})$  can be independently changed via two sets of discrete samples over the domain  $\{\mathbf{x}_n\}$ ,  $\{\mathbf{y}_\ell\}$ , and then inverting the equation shown in Equation (2.2.4) to obtain an estimate of  $\mathbf{M}_{\text{obj}}$ , i.e., measurements

$$\begin{aligned} \mathbf{s}_{\text{out}}(\mathbf{x}_0; \mathbf{y}_0) &= \mathbf{A}(\mathbf{x}_0) \cdot \mathbf{M}_{\text{obj}} \cdot \mathbf{G}(\mathbf{y}_0) \cdot \mathbf{s}_{\text{in}} \\ \mathbf{s}_{\text{out}}(\mathbf{x}_1; \mathbf{y}_0) &= \mathbf{A}(\mathbf{x}_1) \cdot \mathbf{M}_{\text{obj}} \cdot \mathbf{G}(\mathbf{y}_0) \cdot \mathbf{s}_{\text{in}} \\ &\vdots \\ \mathbf{s}_{\text{out}}(\mathbf{x}_n; \mathbf{y}_0) &= \mathbf{A}(\mathbf{x}_n) \cdot \mathbf{M}_{\text{obj}} \cdot \mathbf{G}(\mathbf{y}_0) \cdot \mathbf{s}_{\text{in}} \\ &\vdots \end{aligned}$$

$$\begin{aligned} \mathbf{s}_{\text{out}}(\mathbf{x}_n; \mathbf{y}_\ell) &= \mathbf{A}(\mathbf{x}_n) \cdot \mathbf{M}_{\text{obj}} \cdot \mathbf{G}(\mathbf{y}_\ell) \cdot \mathbf{s}_{\text{in}} \\ &\vdots \end{aligned}$$

are taken. However, for optical instruments, only the irradiance can be measured, i.e.,

$$\begin{aligned} s_{0,\text{out}}(\mathbf{x}_0; \mathbf{y}_0) &= \left\{ \mathbf{A}(\mathbf{x}_0) \cdot \mathbf{M}_{\text{obj}} \cdot \mathbf{G}(\mathbf{y}_0) \cdot \mathbf{s}_{\text{in}} \right\} \bigg|_{s_0 \text{ element}} \\ s_{0,\text{out}}(\mathbf{x}_1; \mathbf{y}_0) &= \left\{ \mathbf{A}(\mathbf{x}_1) \cdot \mathbf{M}_{\text{obj}} \cdot \mathbf{G}(\mathbf{y}_0) \cdot \mathbf{s}_{\text{in}} \right\} \bigg|_{s_0 \text{ element}} \\ &\vdots \\ s_{0,\text{out}}(\mathbf{x}_n; \mathbf{y}_\ell) &= \left\{ \mathbf{A}(\mathbf{x}_n) \cdot \mathbf{M}_{\text{obj}} \cdot \mathbf{G}(\mathbf{y}_\ell) \cdot \mathbf{s}_{\text{in}} \right\} \bigg|_{s_0 \text{ element}} \cdot \\ &\vdots \end{aligned}$$

Each  $\mathbf{A}(\mathbf{x}_n)$ ,  $\mathbf{G}(\mathbf{y}_\ell)$  is assumed to be known, which produces a set of linear equations. The number of unknowns in  $\mathbf{M}_{\text{obj}}$  is 16 when it is assumed to be constant, so at least 16 measurements must be acquired. The inversion requires computing  $\left\{ \mathbf{A}(\mathbf{x}_n) \cdot \mathbf{M}_{\text{obj}} \cdot \mathbf{G}(\mathbf{y}_\ell) \cdot \mathbf{s}_{\text{in}} \right\} \bigg|_{s_0 \text{ element}}$ . The generator modulation can be thought of as only a Stokes parameter modulation,  $\mathbf{s}_{\mathbf{G}}(\mathbf{y}_\ell) = \mathbf{G}(\mathbf{y}_\ell) \cdot \mathbf{s}_{\text{in}}$  [4], which simplifies the calculation [60]:

$$\begin{aligned} s_{0,\text{out}}(\mathbf{x}_n; \mathbf{y}_\ell) &= \left[ m_{00}a_{00}(\mathbf{x}_n) + m_{10}a_{01}(\mathbf{x}_n) + m_{20}a_{02}(\mathbf{x}_n) + m_{30}a_{03}(\mathbf{x}_n) \right] s_0(\mathbf{y}_\ell) \\ &+ \left[ m_{01}a_{00}(\mathbf{x}_n) + m_{11}a_{01}(\mathbf{x}_n) + m_{21}a_{02}(\mathbf{x}_n) + m_{31}a_{03}(\mathbf{x}_n) \right] s_1(\mathbf{y}_\ell) \\ &+ \left[ m_{02}a_{00}(\mathbf{x}_n) + m_{12}a_{01}(\mathbf{x}_n) + m_{22}a_{02}(\mathbf{x}_n) + m_{32}a_{03}(\mathbf{x}_n) \right] s_2(\mathbf{y}_\ell) \\ &+ \left[ m_{03}a_{00}(\mathbf{x}_n) + m_{13}a_{01}(\mathbf{x}_n) + m_{23}a_{02}(\mathbf{x}_n) + m_{33}a_{03}(\mathbf{x}_n) \right] s_3(\mathbf{y}_\ell), \end{aligned} \tag{2.2.5}$$

which can be rewritten as:

$$\boxed{s_{0,\text{out}}(\mathbf{x}_n; \mathbf{y}_\ell) = \sum_{i=0}^3 \sum_{j=0}^3 a_{0i}(\mathbf{x}_n) s_j(\mathbf{y}_\ell) m_{ij}(\mathbf{x}_n; \mathbf{y}_\ell)} \tag{2.2.6}$$



where the  $s_j(\mathbf{y}_\ell)$  elements are components of  $\mathbf{s}_G(\mathbf{y}_\ell)$ . Equation (2.2.5) can also be rewritten as a matrix equation for each  $n, \ell$  pair

$$s_{0,\text{out}}(\mathbf{x}_n; \mathbf{y}_\ell) = \begin{bmatrix} a_{00}(\mathbf{x}_n) & a_{01}(\mathbf{x}_n) & a_{02}(\mathbf{x}_n) & a_{03}(\mathbf{x}_n) \end{bmatrix} \cdot \mathbf{M}_{\text{obj}} \cdot \begin{bmatrix} s_0(\mathbf{y}_\ell) \\ s_1(\mathbf{y}_\ell) \\ s_2(\mathbf{y}_\ell) \\ s_3(\mathbf{y}_\ell) \end{bmatrix}. \quad (2.2.7)$$

See the next section for additional details about traditional Mueller matrix polarimeter equations.

For ideal channeled systems, it is assumed that the modulations,  $a_{01}(\mathbf{x}), a_{02}(\mathbf{x}), a_{03}(\mathbf{x})$  and  $s_1(\mathbf{x}), s_2(\mathbf{x}), s_3(\mathbf{x})$  are *periodic functions* of  $\mathbf{x}$ . For the remaining analysis we also assume that the first element of  $\mathbf{s}_G(\mathbf{x})$ , corresponding to the total irradiance/radiance of the input, is equal to 1<sup>2</sup>, i.e.,  $\mathbf{s}_G(\mathbf{x}) = \begin{bmatrix} 1 & s_1(\mathbf{x}) & s_2(\mathbf{x}) & s_3(\mathbf{x}) \end{bmatrix}^T$ .

A more convenient version of Equation (2.2.6) can be derived, suppose that Equation (2.2.4) is rewritten as :

$$\mathbf{s}_{\text{out}}(\mathbf{x}) = \mathbf{A}(\mathbf{x}) \cdot \mathbf{s}'_G(\mathbf{x}) \quad (2.2.8)$$

where

$$\mathbf{s}'_G(\mathbf{x}) = \mathbf{M}_{\text{obj}}(\mathbf{x}) \cdot \mathbf{s}_G(\mathbf{x}). \quad (2.2.9)$$

If Equation (2.2.8) is then expanded, and  $s_{0,\text{out}}(\mathbf{x})$ , the first element of  $\mathbf{s}_{\text{out}}(\mathbf{x})$ , is inspected, then:

$$s_{0,\text{out}}(\mathbf{x}) = \left( \mathbf{a}(\mathbf{x}), \mathbf{s}'_G(\mathbf{x}) \right) \quad (2.2.10)$$

where  $\mathbf{a}(\mathbf{x}) = \begin{bmatrix} a_{00}(\mathbf{x}) & a_{01}(\mathbf{x}) & a_{02}(\mathbf{x}) & a_{03}(\mathbf{x}) \end{bmatrix}$ , i.e., the inner product,  $(\cdot, \cdot)$  of the first row of  $\mathbf{A}$  and the Stokes vector exiting the sample  $\mathbf{s}'_G(\mathbf{x})$ . Then

$$s_{0,\text{out}}(\mathbf{x}) = \left( \mathbf{a}(\mathbf{x}), \mathbf{s}'_G(\mathbf{x}) \right) \quad (2.2.11)$$

---

<sup>2</sup>Irradiance/radiance could be modulated, but this may negate the utility of modulation in polarization. Further work will be needed to fully characterize the tradeoffs for modulation in both polarization and irradiance/radiance, however there is likely no utility in modulating irradiance/radiance.

$$= \left( \mathbf{a}(\mathbf{x}), \mathbf{M}_{\text{obj}}(\mathbf{x}) \cdot \mathbf{s}_G(\mathbf{x}) \right) \quad (2.2.12)$$

$$= \left( \mathbf{M}_{\text{obj}}^\dagger(\mathbf{x}) \cdot \mathbf{a}(\mathbf{x}), \mathbf{s}_G(\mathbf{x}) \right), \text{ from functional analysis} \quad (2.2.13)$$

$$= \left( \mathbf{M}_{\text{obj}}^T(\mathbf{x}) \cdot \mathbf{a}(\mathbf{x}), \mathbf{s}_G(\mathbf{x}) \right), \text{ since all Mueller matrices are purely real valued} \quad (2.2.14)$$

Equation (2.2.14) is useful for intuition and for determining specific Mueller matrices needed for instrument calibration. Equation (2.2.14) can be explicitly expanded to obtain Equation (2.2.5). Again, modulating in intensity is outside the scope of this work, so the assumption  $a_{00}(\mathbf{x}) \equiv 1$  is asserted.

### 2.3 Polarimetric instruments

Generally, there are two major types of optical polarimetric imaging; 1) passive, where the polarization of light being reflected from objects illuminated with a source which we cannot control such as the sun is measured, yielding only Stokes parameters, and 2) active, where the Stokes parameters of a source illuminating the scene are actively controlled/modulated, yielding the Mueller matrix of the object.

#### 2.3.1 Passive polarimeters

Passive polarimetric instruments have been in use for quite some time, including for remote sensing tasks. Passive instruments are also called Stokes polarimeters, and both “full Stokes” and partial Stokes instruments have been built and deployed. A full Stokes instrument measures the complete set of Stokes parameters of a scene or perhaps the incoming radiance. In optical wavelength measuring instruments, this is accomplished via some *analyzer* modulation scheme. The equation for a Stokes polarimeter is

$$\mathbf{s}_{\text{out}}(\mathbf{x}) = \mathbf{A}(\mathbf{x}) \cdot \mathbf{s}_{\text{obj}}(\mathbf{x}) \quad (2.3.1)$$

and in the optical wavelength regime the measured quantity which is proportional to irradiance is

$$s_{0,\text{out}}(\mathbf{x}) = a_{00}(\mathbf{x}) \cdot s_{0,\text{obj}}(\mathbf{x}) + a_{01}(\mathbf{x}) \cdot s_{1,\text{obj}}(\mathbf{x})$$

$$+ a_{02}(\mathbf{x}) \cdot s_{2,\text{obj}}(\mathbf{x}) + a_{03}(\mathbf{x}) \cdot s_{3,\text{obj}}(\mathbf{x}). \quad (2.3.2)$$

Real instruments take discrete measurements, so the left hand side (LHS) of Equation (2.3.2) becomes a discrete point measurement, then these measurements can be concatenated into a  $n \times 1$  vector,  $\mathbf{g}_{s_{0,\text{out}}}$ , for  $n$  measurements of  $s_{0,\text{out}}(\mathbf{x}_j)$ . If  $\mathbf{A}(\mathbf{x})$  is known, then the right hand side (RHS) of Equation (2.3.2) can be rewritten as

$$\begin{bmatrix} a_{00}(\mathbf{x}_j) & a_{01}(\mathbf{x}_j) & a_{02}(\mathbf{x}_j) & a_{03}(\mathbf{x}_j) \end{bmatrix} \cdot \begin{bmatrix} s_{0,\text{obj}}(\mathbf{x}_j) \\ s_{1,\text{obj}}(\mathbf{x}_j) \\ s_{2,\text{obj}}(\mathbf{x}_j) \\ s_{3,\text{obj}}(\mathbf{x}_j) \end{bmatrix} \quad (2.3.3)$$

for each  $\mathbf{x}_j$ . Note  $\begin{bmatrix} a_{00}(\mathbf{x}_j) & a_{01}(\mathbf{x}_j) & a_{02}(\mathbf{x}_j) & a_{03}(\mathbf{x}_j) \end{bmatrix}$  is often denoted the “analyzer” or “analyzer vector” due to the formal projection of the measured Stokes parameters onto it, and “analyze” refers to the projection itself. This equation can then be built up as:

$$\mathbf{g}_{s_{0,\text{out}}} = \begin{bmatrix} a_{00}(\mathbf{x}_0) & a_{01}(\mathbf{x}_0) & a_{02}(\mathbf{x}_0) & a_{03}(\mathbf{x}_0) \\ a_{00}(\mathbf{x}_1) & a_{01}(\mathbf{x}_1) & a_{02}(\mathbf{x}_1) & a_{03}(\mathbf{x}_1) \\ \vdots & \vdots & \vdots & \vdots \\ a_{00}(\mathbf{x}_{n-1}) & a_{01}(\mathbf{x}_{n-1}) & a_{02}(\mathbf{x}_{n-1}) & a_{03}(\mathbf{x}_{n-1}) \end{bmatrix} \cdot \begin{bmatrix} s_{0,\text{obj}}(\mathbf{x}_{[(n-1)/2]}) \\ s_{1,\text{obj}}(\mathbf{x}_{[(n-1)/2]}) \\ s_{2,\text{obj}}(\mathbf{x}_{[(n-1)/2]}) \\ s_{3,\text{obj}}(\mathbf{x}_{[(n-1)/2]}) \end{bmatrix} \quad (2.3.4)$$

$$= \mathbf{A} \cdot \mathbf{s} \quad (2.3.5)$$

provided that a fairly strong assumption holds;  $\mathbf{s}_{\text{obj}}(\mathbf{x})$  is approximately constant over the range  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{n-1}$ .  $[\cdot]$  denotes the nearest integer function and  $\mathbf{A}$  is an  $n \times 4$  matrix.  $\mathbf{s}_{\text{obj}}(\mathbf{x})$  constant is equivalent to  $\mathbf{s}_{\text{obj}}(\mathbf{x})$  having *zero bandwidth* extent in the modulation domain.

In a real system noise must be accounted for,

$$\mathbf{g}_{s_{0,\text{out}}} = \mathbf{A} \cdot \mathbf{s} + \mathbf{n} \quad (2.3.6)$$

where  $\mathbf{n}$  is a  $n \times 1$  vector which contains both random noise and systematic errors. Typically,  $\mathbf{s}$  can then be estimated as  $\hat{\mathbf{s}}$  by

$$\hat{\mathbf{s}} = \mathbf{A}^+ \cdot \mathbf{g}_{s_{0,\text{out}}} \quad (2.3.7)$$

where  $\mathbf{A}^+$  is the pseudoinverse of  $\mathbf{A}$ . If the noise distribution is known, then additional conditioning can be done to improve the estimate.

Partial Stokes polarimeters can also be described. A common scheme is the measurement of the first three Stokes parameters,  $s_0, s_1, s_2$  using either temporal (rotating linear polarizers) or spatial (micropolarizer array) polarization modulation schemes. Adapting Equation (2.3.4) to this case requires setting  $a_{03}(\mathbf{x}_j) = 0, \forall j$  where  $\forall$  denotes “for all.” This leads to the conclusion that this type of polarimeter cannot detect  $s_3$ , or equivalently that  $s_3$  is in the *null space* of the instrument operator. Equation (2.3.4) can then be abbreviated for this case to

$$\mathbf{g}_{s_{0,\text{out}}} = \begin{bmatrix} a_{00}(\mathbf{x}_0) & a_{01}(\mathbf{x}_0) & a_{02}(\mathbf{x}_0) \\ \vdots & \vdots & \vdots \\ a_{00}(\mathbf{x}_{n-1}) & a_{01}(\mathbf{x}_{n-1}) & a_{02}(\mathbf{x}_{n-1}) \end{bmatrix} \cdot \begin{bmatrix} s_{0,\text{obj}}(\mathbf{x}_{\lfloor (n-1)/2 \rfloor}) \\ s_{1,\text{obj}}(\mathbf{x}_{\lfloor (n-1)/2 \rfloor}) \\ s_{2,\text{obj}}(\mathbf{x}_{\lfloor (n-1)/2 \rfloor}) \end{bmatrix} \quad (2.3.8)$$

$$= \mathbf{A}' \cdot \mathbf{s}' \quad (2.3.9)$$

i.e. the last column of  $\mathbf{A}$  and  $s_{3,\text{obj}}(\mathbf{x}_{\lfloor (n-1)/2 \rfloor})$  can be ignored. There are a myriad of other types of Stokes and partial Stokes polarimeters, however this dissertation will focus on active, or Mueller matrix instruments and Stokes instruments will not be discussed in depth.

### 2.3.2 Active polarimeters

Active, or Mueller matrix polarimeters measure the Mueller matrix of an object,  $\mathbf{M}_{\text{obj}}$ . An analysis similar to the above for Stokes polarimeters is typically carried out for active instruments. Equation (2.3.11) shows the measurement scheme for a Mueller matrix polarimeter and is the multiple measurement version of Equation (2.2.7). Note that  $\mathbf{M}(\mathbf{x})$  is assumed to be approximately constant jointly over both the  $\mathbf{x}_0, \dots, \mathbf{x}_{n-1}$  and the  $\mathbf{y}_0, \dots, \mathbf{y}_{m-1}$  value ranges, with the average at  $\mathbf{x}_a$ . Equation (2.3.11) can then be rewritten as shown in Equation (2.3.12) where  $\mathbf{I}_{s_{0,\text{out}}}$  is size  $n \times m$ ,  $\mathbf{A}$  is size  $n \times 4$ ,  $\mathbf{G}$  is size  $4 \times m$ , and  $\mathbf{N}$ , the noise matrix, is also of size  $n \times m$ . The estimate of  $\mathbf{M}$  is typically accomplished via

$$\hat{\mathbf{M}} = \mathbf{A}^+ \cdot \mathbf{I}_{s_{0,\text{out}}} \cdot \mathbf{G}^+ \quad (2.3.10)$$

$$\mathbf{I}_{S_{0,\text{out}}} = \begin{bmatrix} a_{00}(\mathbf{x}_0) & a_{01}(\mathbf{x}_0) & a_{02}(\mathbf{x}_0) & a_{03}(\mathbf{x}_0) \\ a_{00}(\mathbf{x}_1) & a_{01}(\mathbf{x}_1) & a_{02}(\mathbf{x}_1) & a_{03}(\mathbf{x}_1) \\ \vdots & \vdots & \vdots & \vdots \\ a_{00}(\mathbf{x}_{n-1}) & a_{01}(\mathbf{x}_{n-1}) & a_{02}(\mathbf{x}_{n-1}) & a_{03}(\mathbf{x}_{n-1}) \end{bmatrix} \cdot \begin{bmatrix} m_{00}(\mathbf{x}_a) & m_{01}(\mathbf{x}_a) & m_{02}(\mathbf{x}_a) & m_{03}(\mathbf{x}_a) \\ m_{10}(\mathbf{x}_a) & m_{11}(\mathbf{x}_a) & m_{12}(\mathbf{x}_a) & m_{13}(\mathbf{x}_a) \\ m_{20}(\mathbf{x}_a) & m_{21}(\mathbf{x}_a) & m_{22}(\mathbf{x}_a) & m_{23}(\mathbf{x}_a) \\ m_{30}(\mathbf{x}_a) & m_{31}(\mathbf{x}_a) & m_{32}(\mathbf{x}_a) & m_{33}(\mathbf{x}_a) \end{bmatrix} \cdot \begin{bmatrix} s_{0,G}(\mathbf{y}_0) & \dots & s_{0,G}(\mathbf{y}_{m-1}) \\ s_{1,G}(\mathbf{y}_0) & \dots & s_{1,G}(\mathbf{y}_{m-1}) \\ s_{2,G}(\mathbf{y}_0) & \dots & s_{2,G}(\mathbf{y}_{m-1}) \\ s_{3,G}(\mathbf{y}_0) & \dots & s_{3,G}(\mathbf{y}_{m-1}) \end{bmatrix} \quad (2.3.11)$$

$$(2.3.12)$$

$$\Rightarrow \mathbf{I}_{S_{0,\text{out}}} = \mathbf{A} \cdot \mathbf{M} \cdot \mathbf{G} + \mathbf{N}$$

where  $\mathbf{G}^+$  and  $\mathbf{A}^+$  are the pseudoinverse of  $\mathbf{G}$  and  $\mathbf{A}$  respectively. A more compact form of Eqns. 2.3.12 and 2.3.10 can be used by “vectorizing” the matrix equations. Equation (2.3.12) can be rewritten as [63]

$$\begin{aligned} \mathbf{i}_{S_{0,\text{out}}} &= (\mathbf{G}^T \otimes \mathbf{A})\mathbf{m} + \mathbf{n} \\ &= \mathbf{W} \cdot \mathbf{m} + \mathbf{n} \end{aligned} \quad (2.3.13)$$

where  $\mathbf{i}_{S_{0,\text{out}}}$  is vectorized as a row by row concatenation of  $\mathbf{I}_{S_{0,\text{out}}}$ ,  $\mathbf{m}$  is vectorized as a row by row concatenation of  $\mathbf{M}$ , and  $\otimes$  denotes the Kronecker product.  $\mathbf{i}_{S_{0,\text{out}}}$  has size  $mn \times 1$ ,  $\mathbf{m}$  has size  $16 \times 1$  and  $\mathbf{W} = (\mathbf{G}^T \otimes \mathbf{A})$  has size  $mn \times 16$ . Equation (2.3.13) can then be inverted via

$$\mathbf{m} = \mathbf{W}^+ \mathbf{i}_{S_{0,\text{out}}} \quad (2.3.14)$$

where  $\mathbf{W}^+$  is the pseudoinverse of  $\mathbf{W}$ .

Active instruments have been utilized in various laboratory settings, for polarized bidirectional reflectance distribution (pBRDF) measurements, ellipsometry and index of refraction measurements, measurements of LCD panels and defects, etc. However, full Mueller matrix instruments have seen little use as deployable remote sensing instruments, although the idea has been discussed before [2, 64–69]. Aside from the instrument presented here, I know of only one other possible portable active Mueller matrix instrument which

has actually been built, by Advanced Optical Technologies, Inc. [70]. The polarimeter from Advanced Optical Technologies is shown in a video on their website, however to my knowledge, no public literature about the instrument exists other than the video.

One of the primary reasons for the dearth of active polarimetric instruments for remote sensing is the complexity of instrument design and theoretical limits on combinations of instrument speed (temporal bandwidth), image resolution (spatial bandwidth), and spectral resolution. Moreover, many active instrument designs require extremely precise alignment of multiple optical paths, precise timing of rotating or moving components, and multiple sensors or apertures. Opto-mechanical design complexity is a hurdle to inexpensive, portable, and well performing polarimetric instruments. Additionally, data coming from an instrument contain more information than data from a typical color or monochrome camera, specifically four times as much information for a passive system and sixteen times as much information for an active system. The data processing requirements are also demanding, especially for active (Mueller matrix) imaging polarimeters.

## 2.4 Channeled systems

Azzam [46] introduced the concept of channeled polarimetric systems in 1978 for temporal carriers, however his analysis made the assumption that the object has no temporal bandwidth. Error analysis was undertaken by Goldstein and Chipman on this type of temporal channeled system for a dual rotating retarder polarimeter (DRR) [71]. Other than the Azzam channeled DRR polarimeter, channeled polarimeters discussions in the literature were rare until Oka and Kato's letter on channeled spectropolarimetry [47]. This work described a system where the spectrally resolved Stokes parameters could be measured in a single spectrometer measurement. The Stokes parameter information is encoded into the spectrum as quasi-cosinusoidal functions of wavenumber,  $\sigma$ . The channels then emerge when the Fourier transform of the spectrometer measurement is taken, with the channel space domain being optical path difference (OPD). The measured spectrum and channels are shown in Figure 2.3. After the work of Oka and Kato, many spectrally channeled and spatial-spectrally channeled systems were designed and described in the literature

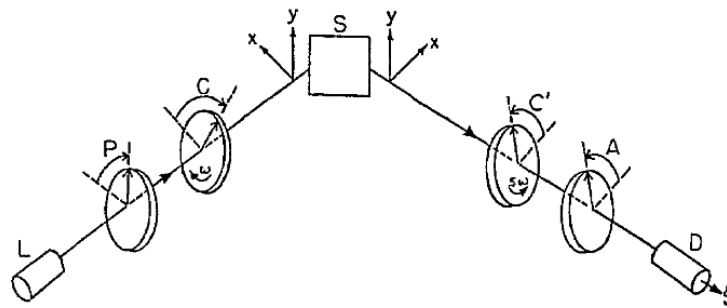


Figure 2.2: A Fourier photopolarimeter capable of measuring all 16 elements of the Mueller matrix of an optical system. This is a DRR configuration. Reproduced from Azzam [46].

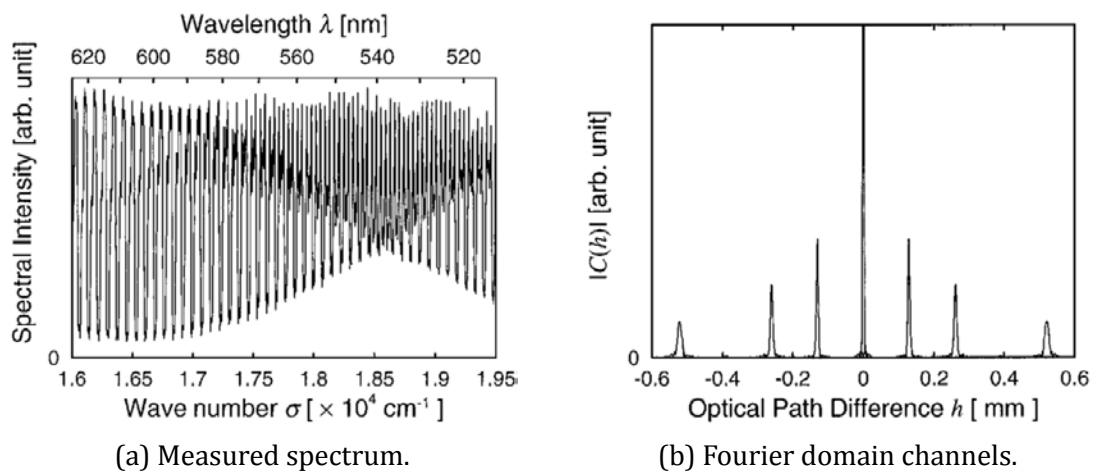


Figure 2.3: Encoded spectrum and resultant channels from the Oka and Kato spectropolarimeter design. Reproduced from Oka and Kato [47].

[52] [31, 32, 48–50, 52, 53], however none address bandwidth in a systematic way.

Recent work by Alenin and Tyo [34] has illuminated modulated (channeled) polarimeters and the underlying channel structure of a polarimeter in any modulation domain, generalizing the work of Oka and Kato, Hagen, Kudenov and Dereniak, Dubreil *et al*, etc. in the spectral and spectral-spatial modulation domains. This section will present a similar analysis and the underlying mathematics are derived. As in Alenin [34], the assumption is made that carriers are periodic in the underlying domain.

The Fourier transform of the general instrument equation Equation (2.2.6) is:

$$S_{0,\text{out}}(\boldsymbol{\rho}) = \sum_{i=0}^3 \sum_{j=0}^3 A_i(\boldsymbol{\rho}) * S_j(\boldsymbol{\rho}) * M_{ij}(\boldsymbol{\rho}) \quad (2.4.1)$$

where

$$\begin{aligned} S_{0,\text{out}}(\boldsymbol{\rho}) &= \mathcal{F} \{s_{0,\text{out}}(\mathbf{x})\}_{\mathbf{x} \rightarrow \boldsymbol{\rho}}, \quad A_i(\boldsymbol{\rho}) = \mathcal{F} \{a_{0i}(\mathbf{x})\}_{\mathbf{x} \rightarrow \boldsymbol{\rho}}, \\ M_{ij}(\boldsymbol{\rho}) &= \mathcal{F} \{m_{ij}(\mathbf{x})\}_{\mathbf{x} \rightarrow \boldsymbol{\rho}}, \quad S_j(\boldsymbol{\rho}) = \mathcal{F} \{s_j(\mathbf{x})\}_{\mathbf{x} \rightarrow \boldsymbol{\rho}}, \end{aligned} \quad (2.4.2)$$

\* indicates convolution, and the Fourier domain has  $\boldsymbol{\rho} = [\xi \ \eta \ \chi \ \nu \ \tau]^T$  corresponding to  $\mathbf{x}$ . Note that for periodic carriers each  $A_i$  will be a set of delta functions, and each  $S_j$  will also be a set of delta functions. When these sets are convolved, they will result in “channel splitting” [34]. The final set (a “cloud of delta functions” [34]) is then convolved with each respective Mueller matrix element  $M_{ij}$  in the Fourier domain. This results in a copy of the data for each Mueller matrix element centered on each delta function in the “cloud.” Alenin and Tyo’s results are extended by considering the available *bandwidth* between these channels. The complete set of  $\delta$ -functions for the system

$$\sum_{i=0}^3 \sum_{j=0}^3 A_i(\boldsymbol{\rho}) * S_j(\boldsymbol{\rho}) \quad (2.4.3)$$

are defined as the *channels* of the system, or the system’s *channel structure* [4].

#### 2.4.1 Bandwidth in channeled systems

LaCasse *et al* began to investigate bandwidth in Stokes polarimeters as channeled systems in 2011 [43] motivated by “polarization artifacts,” i.e., reconstructed im-



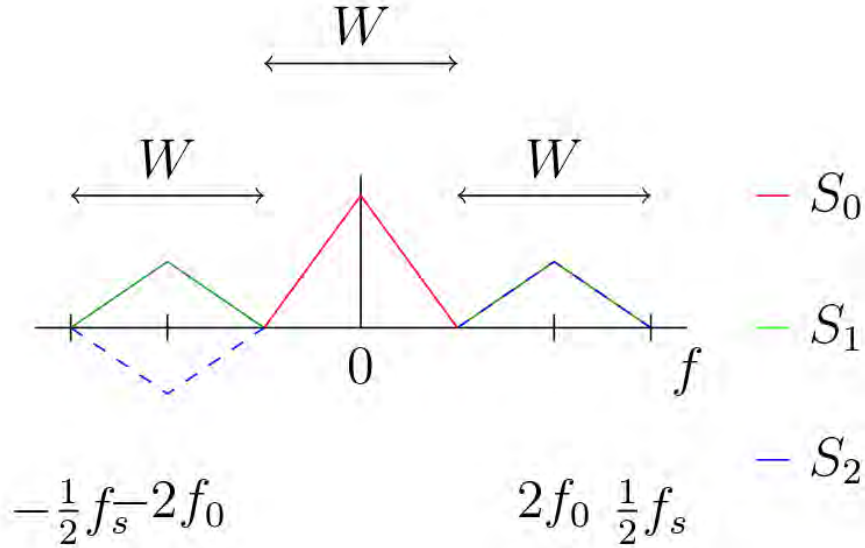


Figure 2.4: A modulated measurement of the Stokes parameters using a rotating analyzer polarimeter, with each parameter band limited such that the signal can be ideally reconstructed.  $W_b$  is bandwidth of each parameter,  $f_0$  is the frequency of analyzer rotation, and  $f_s$  is the detector sampling frequency. The dashed blue line indicates that  $s_2$  is in the quadrature component of the side band. The configuration with maximum allowed bandwidth is shown. Reproduced from LaCasse *et al* [43].

ages which contained non-physically realizable, erroneous, polarization information. These artifacts are particularly prevalent at edges or transition boundaries in the scene for partial Stokes micropolarizer array polarimeters [72]. A key result was that certain instrument designs and data processing used in conventional polarimetric data reduction matrix (DRM) schemes, as described in Equation (2.3.4), *inherently corrupt* the reconstructed data. The equation

$$\hat{\mathbf{s}}(\mathbf{x}) = \mathbf{W}^{-1} \{I(\mathbf{x})\} = \mathbf{w}(\mathbf{x}) * \mathbf{Z}^{-1} \mathbf{A}(\mathbf{x}) I(\mathbf{x}) \quad (2.4.4)$$

is derived where  $\mathbf{x} = (x, y, t, \lambda)$ ,  $\mathbf{Z}^{-1}$  is the constant modulator inner product inversion matrix,  $\mathbf{A}(\mathbf{x})$  is the polarimetric carrier generation matrix, and  $I(\mathbf{x})$  is the modulated irradiance or radiance. The elements,  $Z_{ij}$  of  $\mathbf{Z}$  are defined as

$$Z_{ij} = \int A_i(t) A_j(t) dt \quad (2.4.5)$$

where  $A_i(t), A_j(t) \in \mathbf{A}(t)$ . Here the elements of  $\mathbf{Z}^{-1}$  are assumed to be constant over the modulation [43]. LaCasse *et al* showed that the windowing function in the

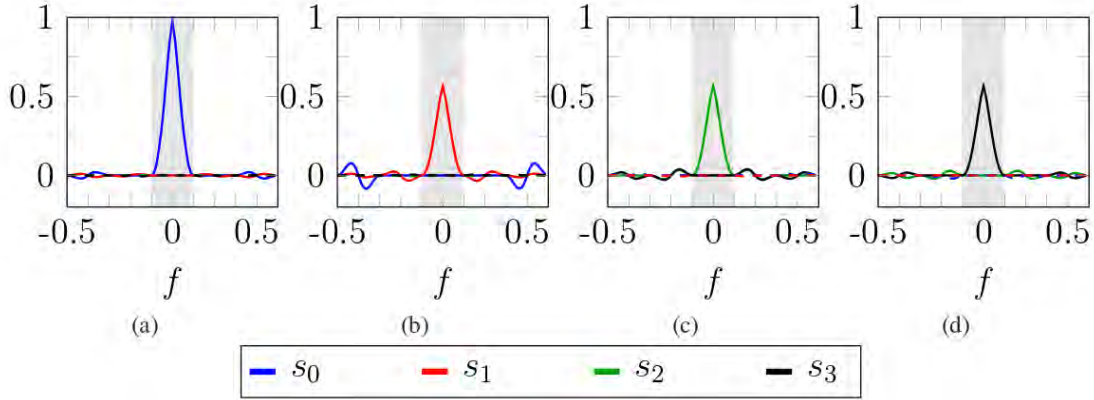


Figure 2.5: The four components of the estimated Stokes parameters in the Fourier domain given by  $\mathcal{F}\{w(t) * \mathbf{Z}\mathbf{A}(t)I(t)\}_{t \rightarrow f}$  with  $w(n) = \text{rect}\left(\frac{n}{16}\right)$ , a 16 sample rect window. (a)  $\mathcal{F}\{\hat{s}_0\}$ , (b)  $\mathcal{F}\{\hat{s}_1\}$ , (c)  $\mathcal{F}\{\hat{s}_2\}$ , (d)  $\mathcal{F}\{\hat{s}_3\}$ . Differences from a triangle function in the Fourier domain indicate errors in the data reduction method. Reproduced from LaCasse *et al* [43].

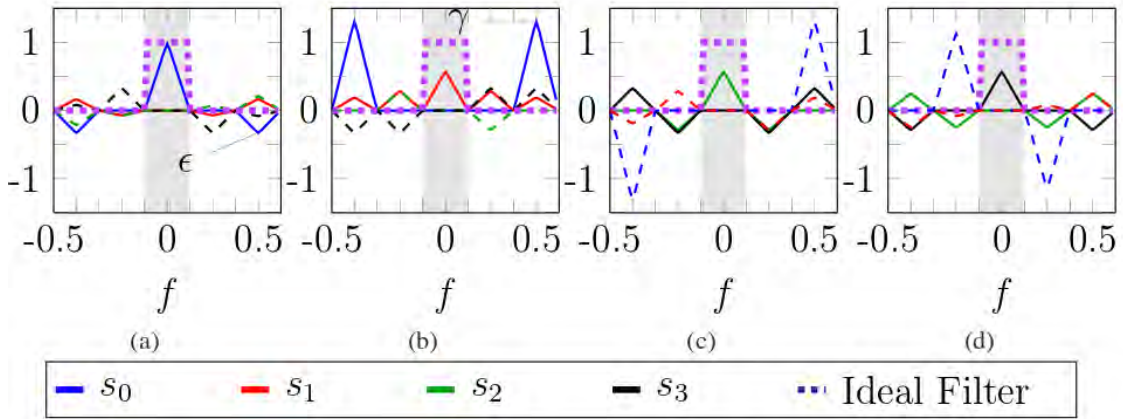


Figure 2.6: The four components of the estimated Stokes parameters given by  $\mathcal{F}\{\mathbf{s}(t)\}_{t \rightarrow f} = \mathcal{F}\{\mathbf{Z}\mathbf{A}(t)I(t)\}_{t \rightarrow f}$  prior to low pass filtering with  $w(t)$  according to Equation (2.4.4). Also shown in the dotted line is the rectangular low pass filter that ideally reconstructs the correct individual Stokes parameters. The marking  $\gamma$  is an example of self-error, while the marking  $\epsilon$  is an example of cross-error. If the low pass filter  $w(t)$  does not reject these frequencies outside of the shaded base band artifacts will arise due to these error terms. Reproduced from LaCasse *et al* [43].

typical DRM scheme has inherent cross talk issues as shown in Figure 2.5. For the specific example in Figures 2.5 and 2.6, the ideal band-limited filter is [43]

$$\mathcal{F}\{w(t)\} = \text{rect}\left(\frac{5f}{f_s}\right) \quad (2.4.6)$$

which is a sinc function in the time domain, *not* a rect function which is implied by the DRM method. The filter selection is intrinsically dependent on *object bandwidth*, and this work began the systematic analysis of bandwidth and filter design in channeled polarimetric systems.

Design of polarimetric systems optimized for bandwidth in multi-domain channeled systems is lacking in the literature, although LaCasse *et al* presented preliminary conference results on a spatio-temporally modulated Stokes polarimeter [33]. A hybrid system with both a rotating retarder + micropolarizer array was presented, shown in Figure 2.7. The last design, (d), is bandwidth optimal but requires a custom microanalyzer array given by [33]

$$\left[1 \quad \frac{1}{\sqrt{3}} \cos \pi m \quad \frac{1}{\sqrt{3}} \cos \pi n \quad \frac{1}{\sqrt{3}} \cos \pi(m+n)\right]^T \quad (2.4.7)$$

where  $(m, n)$  is the pixel location and  $m, n \in \mathbb{N}$ . This type of array cannot currently be purchased in the commercial market. Additionally, constructing such an array would be more difficult than the currently available linear polarizer arrays since each pixel location must analyze either

$$\begin{bmatrix} 1 \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \text{ or } \begin{bmatrix} 1 \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{bmatrix} \text{ or } \begin{bmatrix} 1 \\ -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \text{ or } \begin{bmatrix} 1 \\ \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{bmatrix}, \quad (2.4.8)$$

which necessitates manufacturing microanalyzer elements which are more complicated than linear polarizers rotated to specific angles. Portions of the work in this dissertation expand and extend the hybrid spatio-temporal ideas introduced by LaCasse *et al* applied to active polarimetric instruments.

## 2.4.2 Notation

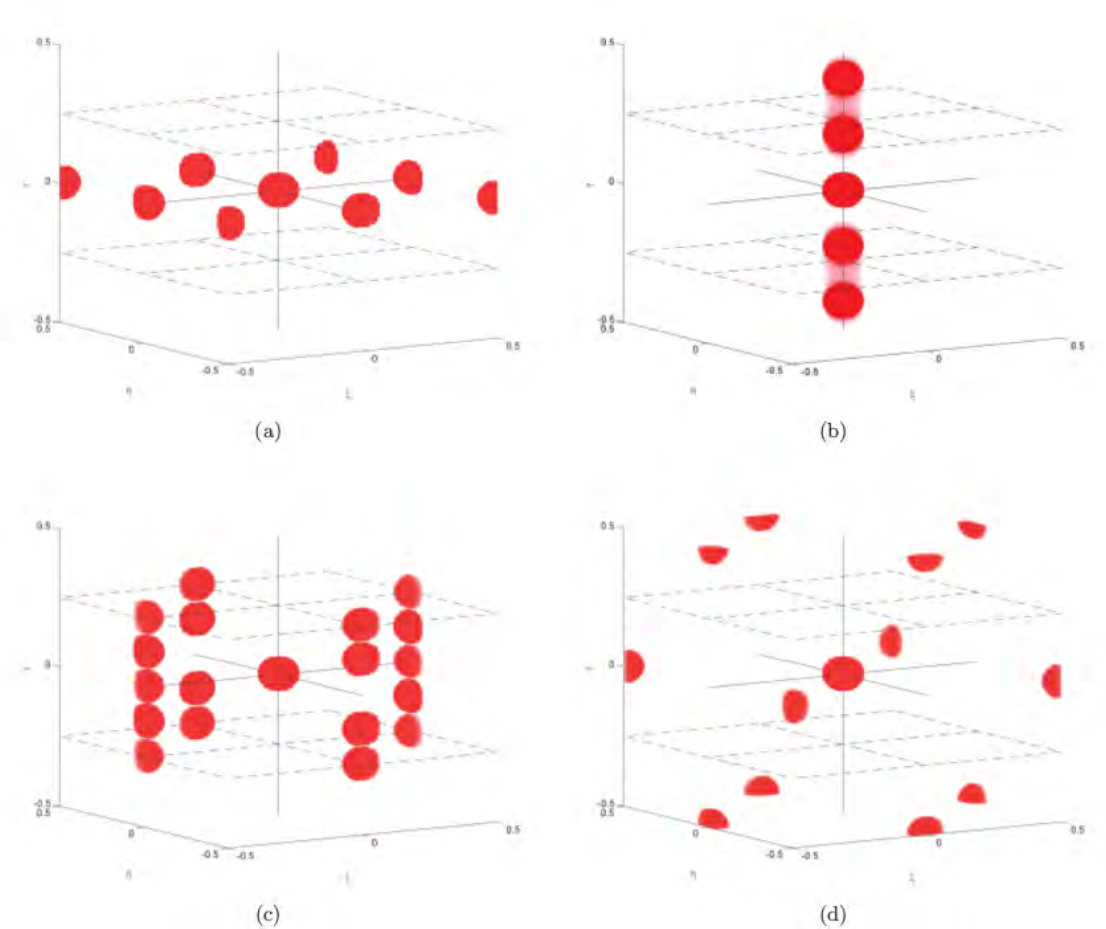






Figure 2.7: Locations of modulated side bands in the frequency domain for four imaging polarimeters: (a) a full micropolarizer array polarimeter, (b) a rotating retarder polarimeter, (c) a polarimeter consisting of a rotating retarder followed by a micropolarizer array polarimeter, (d) and a half wave retarder rotating  $90^\circ$  per temporal sample in front of a micropolarizer array, which creates a polarimeter with maximum spatio-temporal bandwidth. Reproduced from LaCasse *et al* [33].

Table 2.1: Notation for channels.

	positive	negative
real		
imag.		

Visualization of system channel structures can be accomplished by graphing the sets of channels ( $\delta$ -functions). A  $\delta$ -function can be characterized by its position, and its complex magnitude. Table 2.1 outlines the graphical notation that will be used, the blue triangles represent the real part of the magnitude, red triangles represent the imaginary part of the magnitude, the directions that the triangles point represent whether the magnitude is positive or negative, and the size of each triangle represents the absolute value of the real part or the imaginary part. Only channels for a single Mueller matrix element will be displayed for each visualization, with the locations for channels for all other Mueller matrix elements being indicated by light gray-blue circles. Many channels, arising from several Mueller matrix elements, end up being added together at the locations shown. The relationship between the channels of different Mueller matrices is *additive*. An example of the channel structure for a spatio-temporally modulated system corresponding to  $M_{23}(\boldsymbol{\rho})$  is shown in Figure 3.1.

There is linear mixing between channels of the different elements of  $\mathbf{M}_{\text{obj}}$ , described by a matrix,  $\mathbf{Q}$ , [34] so the channels must be unmixed via the pseudoinverse matrix,  $\mathbf{Q}^+$  [4,5] prior to inverting the Fourier transform. The  $\mathbf{Q}$  matrix characterizes noise propagation and channels can be chosen to minimize noise [4, 5, 34]. For the channeled measurements, we can take the Fourier transform of each  $m_{ij}(\mathbf{x})$ :

$$M_{ij}(\boldsymbol{\rho}) = \mathcal{F} \{m_{ij}(\mathbf{x})\}_{\mathbf{x} \rightarrow \boldsymbol{\rho}} \quad (2.4.9)$$

then the channeled system will convolve each  $M_{ij}(\boldsymbol{\rho})$  with the set of  $\delta$ -functions in the channel structure, resulting in a *mixture of data* in the Fourier domain. The exact mixing is characterized by the  $\mathbf{Q}$  matrix [34], and we can unmix by using the pseudoinverse,  $\mathbf{Q}^+$ .

Some of the examples presented will be normalized to a temporal frequency range of  $[-1, 1]$ . Since there is always a maximum sampling rate in practice, the relative bandwidth with respect to a maximum absolute frequency of 1 over the domain(s) of interest (this forms a hypercube) is optimized. Different norms and metrics could be used to normalize for optimization, and result in different optimization

outcomes [5]. Due to the assumption of separable modulation functions for space and time and Mueller matrix physicality conditions, channels are fixed to travel along constrained paths in the Fourier domain. The details of separability are discussed in Chapter 3.

## 2.5 Applications

Stokes polarimeters and partial Stokes polarimeters have been used in a wide variety of applications. Some examples include remote sensing [14], satellite imaging and estimations for weather, aerosols, and vegetation [27, 28, 73–75]. Ocean radiance is also measured both to estimate ocean water column constituents and to verify ocean atmosphere radiance models [76–80].

Active polarimeters, however, have seen less use overall, especially for remote sensing. This is primarily due to the increase in instrument complexity, and the issues of active illumination at range (primarily the problem of irradiance on the object decreasing as range increases, and/or sunlight overpowering the active source for outdoor remote sensing). This dissertation will focus on active polarimetric imagers, so Stokes applications will not be discussed in depth.

Active polarimetric instruments have primarily been used in the lab and as ellipsometers. Ellipsometry is a sub-field of its own, and will not be discussed here. Although there are a dearth of public results on utilization of active polarimetry for remote sensing, there have been some unpublished attempts made by defense agencies and their contractors to build and utilize active polarimeters for remote sensing tasks. Most active instruments which have been built are partial polarimeters which don't measure the complete Mueller matrix. LeMaster *et al* have built a partial Mueller matrix polarimeter (pMMP) for the SWIR wavelength range [81], and Hoover has also apparently built an active polarimeter for remote sensing purposes [70], however there are no details about the instrument available, so it is not currently known if it is a partial or full Mueller matrix polarimeter. This dissertation will develop a channeled framework for spatio-temporally modulated Mueller matrix imaging polarimeters, optimize a specific system for bandwidth, simulate an example of such a system, design and implement an actual instrument which uses the design, and

finally present some data from the actual instrument. The instrument is designed to be an *active, portable, full Mueller matrix imager*.

## CHAPTER 3

### Bandwidth in spatio-temporal channeled systems

#### 3.1 Introduction

Channeled Mueller matrix polarimeters and the concept of using these channels was first introduced by Azzam [46]. Azzam published a very specific case, where

- (a) a specific temporal framework was analyzed,
- (b) an implicit assumption about the object was made: *the object had no temporal bandwidth*, i.e., the object was constant in time.

This ensures a simpler description of the channels and the data, since the data is constant it will be the same in both the temporal and the frequency domains. Prior to the work by LaCasse *et al* [33, 43, 44], bandwidth in channeled polarimetric systems had not been addressed, or only addressed as a consequence of instrumental “error.” Additionally, prior to Alenin and Tyo [34] channeled systems were designed in an *ad-hoc* manner. In this chapter bandwidth is addressed using the *systematic* design tools introduced by Alenin and Tyo [34] for a hybrid spatio-temporally modulated channeled active polarimetric system.

#### 3.2 Spatio-temporal channels

This chapter will discuss a spatio-temporal system, with carriers and modulation in the domain

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ t \end{bmatrix} \quad (3.2.1)$$

and examples will primarily be from a quad-retarder + micropolarizer array system [4]. First convolution with Dirac delta functions will be briefly reviewed. Some



component	spatial carrier generation	temporal carrier generation
polarizer		✓
retarder		✓
variable LC retarder		✓
variable Pockel cell retarder		✓
Faraday rotator		✓
micropolarizer array	✓	
microretarder array	✓	
spatial light modulator	✓	✓

Table 3.1: Table of COTS polarization components.

unknown quantity,  $m(t)$ , can modulate a sinusoidal function. Without loss of generality, cosine is chosen:

$$f(t) = \cos 2\pi\nu_0 t \cdot m(t) \quad (3.2.2)$$

Taking the Fourier transform of  $f(t)$  gives

$$\mathcal{F}\{f(t)\}_{t \rightarrow \nu} = \mathcal{F}\{\cos 2\pi\nu_0 t \cdot m(t)\}_{t \rightarrow \nu} \quad (3.2.3)$$

$$\Rightarrow F(\nu) = \mathcal{F}\{\cos 2\pi\nu_0 t\}_{t \rightarrow \nu} * \mathcal{F}\{m(t)\}_{t \rightarrow \nu} \quad (3.2.4)$$

$$\Rightarrow F(\nu) = \frac{1}{2} \left[ \delta(\nu - \nu_0) + \delta(\nu + \nu_0) \right] * M(\nu) \quad (3.2.5)$$

$$\Rightarrow F(\nu) = \frac{M(\nu - \nu_0)}{2} + \frac{M(\nu + \nu_0)}{2} \quad (3.2.6)$$

where  $\mathcal{F}\{\cdot\}$  is the Fourier transform,  $*$  is convolution, and  $\delta(\nu)$  is the Dirac delta function. The last line is due to the property of convolution with delta functions, a function will be “copied” to the location of a  $\delta$ -function when convolved with that  $\delta$ -function. This gives us some tools for conceptual descriptions for the forward problem and hence the bandwidth.

A spatio-temporally channeled system allows for visualization of a 3-dimensional space for insight, and also allows robust design to occur from common off the shelf components (COTS). Many polarimetric design elements generate carriers either spatially, temporally, or both. Table 3.1 illustrates some examples of ways to generate carriers, with many polarization elements being temporally changed by rotation

or some other movement to change the Mueller matrix with respect to (w.r.t.) the lab frame. Different channel schemes can be combined to yield more complex hybrid schemes like the spatio-temporal channel structures discussed here.

In this chapter it is assumed that for a hybrid spatio-temporal system, the spatial and temporal carriers are *separable*, i.e.,

$$f(x, y, t) = h(x, y) \cdot g(t). \quad (3.2.7)$$

This enforces certain conditions in the channel space, primarily the multiplication between “orthogonal” functions of  $(\xi, \eta)$  and  $\nu$  in the channel space, since

$$\mathcal{F}\{f(x, y, t)\} = \mathcal{F}\{h(x, y)\} \cdot \mathcal{F}\{g(t)\}. \quad (3.2.8)$$

which implies that there is no channel splitting between  $\mathcal{F}\{h(x, y)\}$  and  $\mathcal{F}\{g(t)\}$ , only location copying. This enforces the copying of all of the temporal channels at the locations of the spatial channels. There can be some linear combinations of temporal channels, but arbitrary channel locations on the upper hemisphere (reflected by Hermiticity to the lower hemisphere) of the 3-dimensional frequency space are restricted due to the separability assumption. Although some preliminary work has been carried out on the non-separable condition, it will not be discussed here. It is not completely clear if physical limitations on polarimetric instruments will allow for non-separable carriers, although I hypothesize that something like rotating a focal plane array with a micropolarizer array attached, or pushbroom sensors which inherently couple space and time, will yield non-separable carriers.

Single channels, other than the zero frequency (DC, or baseband), cannot exist in the frequency space. This is due to the fact that carriers are real valued, and therefore the Fourier transform must be Hermitian:

$$f(x, y, t) \in \mathbb{R}^3 = f_{\text{even}}(x, y, t) + f_{\text{odd}}(x, y, t) \quad (3.2.9)$$

$$\Rightarrow \mathcal{F}\{f(x, y, t)\} = F(\xi, \eta, \nu) = \mathcal{F}\{f_{\text{even}}(x, y, t)\} + \mathcal{F}\{f_{\text{odd}}(x, y, t)\} \quad (3.2.10)$$

$$= F_{\text{real}}(\xi, \eta, \nu) + F_{\text{imag}}(\xi, \eta, \nu) \quad (3.2.11)$$

where  $F_{\text{real}}(\xi, \eta, \nu)$  is a real valued function and  $F_{\text{imag}}(\xi, \eta, \nu)$  is a purely imaginary function. Additionally,  $F_{\text{real}}(\xi, \eta, \nu)$  is even and  $F_{\text{imag}}(\xi, \eta, \nu)$  is odd. This implies that

$$F(\xi, \eta, \nu) = F^*(-\xi, -\eta, -\nu) \quad (3.2.12)$$

the definition of Hermitian, where  $*$  here is the complex conjugate. The fact that  $F_{\text{real}}(\xi, \eta, \nu)$  is even imposes a symmetry condition on the real valued channels, any real valued channel in the frequency domain located away from  $(0, 0, 0)$  at say,  $(\xi_0, \eta_0, \nu_0)$  must also have a channel located at  $(-\xi_0, -\eta_0, -\nu_0)$  with the same magnitude and sign. Similarly since  $F_{\text{imag}}(\xi, \eta, \nu)$  is odd, any imaginary valued channel in the frequency domain located at  $(\xi_0, \eta_0, \nu_0) \neq (0, 0, 0)$  must also have a channel located at  $(-\xi_0, -\eta_0, -\nu_0)$  with the same magnitude and opposite sign. Equation (3.2.12) is the formal description of these symmetry conditions. Additionally, the channel information can be represented in a more compact form since if  $F(\xi, \eta, \nu)$  is known for  $\xi, \eta, \nu \geq 0$  then  $F(-\xi, -\eta, -\nu)$  can be deduced from Equation (3.2.12). Due to this, only channels for  $\xi, \eta, \nu \geq 0$  have to be stored/manipulated, and the other channels for  $\xi, \eta, \nu < 0$  can be constructed.

The above Hermiticity and separability conditions constrain the channel(s) location in the frequency domain. Because the carrier is separable,  $F(\xi, \eta, \nu) = H(\xi, \eta) \cdot G(\nu)$ , and Hermiticity implies that

$$H(\xi, \eta) = H^*(-\xi, -\eta) \quad (3.2.13)$$

$$G(\nu) = G^*(-\nu). \quad (3.2.14)$$

Conceptually, this implies that for each channel in the  $(\xi, \eta)$  plane from  $H(\xi, \eta)$ , channels from  $G(\nu)$  will form in lines, perpendicular to the  $(\xi, \eta)$  plane, at each channel from  $H(\xi, \eta)$ . Figure 3.1 illustrates this fact. There are 5 channels from  $H(\xi, \eta)$  in the  $(\xi, \eta)$  plane, and differing numbers of channels (due to superposition, channel cancellation, and that the  $(0, 0, 0)$  components do not have to be copied symmetrically) along lines perpendicular to the  $(\xi, \eta)$  plane from the channels formed by  $G(\nu)$ . Non-separable carriers would allow for channels to exist along trajectories different from lines perpendicular to the plane, provided they can be physically implemented.

### 3.3 Physical constraints

In addition to the constraints enforced by Hermiticity and separability, physical Mueller matrices and Stokes parameters add additional constraints to the carriers, and therefore to the channel structure. There is also concern about the validity of

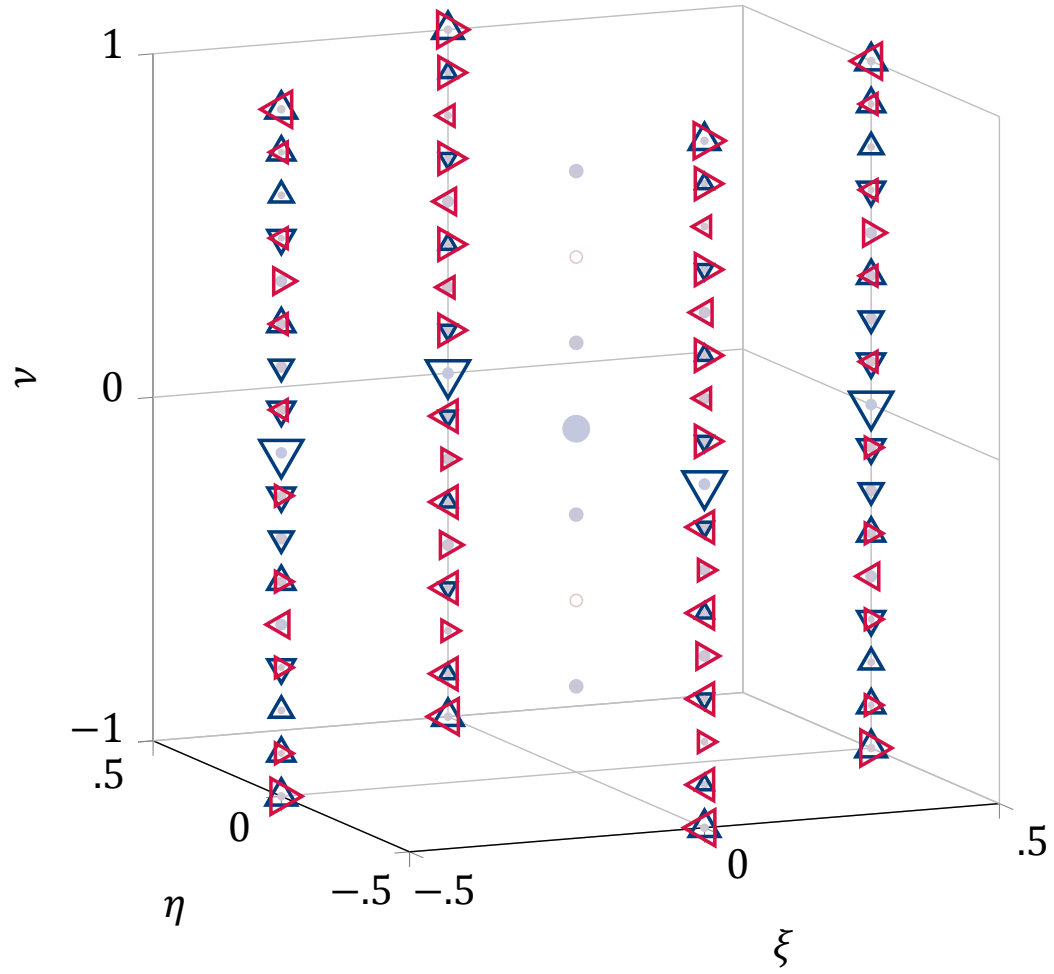


Figure 3.1: Example of a spatio-temporal channel structure with  $\delta$ -functions specific to  $m_{23}$ . The maximum bandwidth corresponds to the minimum distance between two adjacent channels, taken over all possible adjacent channel pairs.

$$\mathbf{H} = \frac{1}{2} \begin{bmatrix} m_{00} + m_{01} + m_{10} + m_{11} & m_{02} + im_{03} + m_{12} + im_{13} & m_{20} + m_{21} - im_{30} - im_{31} & m_{22} + im_{23} - im_{32} + m_{33} \\ m_{02} - im_{03} + m_{12} - im_{13} & m_{00} - m_{01} + m_{10} - m_{11} & m_{22} - im_{23} - im_{32} - m_{33} & m_{20} - m_{21} - im_{30} + im_{31} \\ m_{20} + m_{21} + im_{30} + im_{31} & m_{22} + im_{23} + im_{32} - m_{33} & m_{00} + m_{01} - m_{10} - m_{11} & m_{02} + im_{03} - m_{12} - im_{13} \\ m_{22} - im_{23} + im_{32} + m_{33} & m_{20} - m_{21} + im_{30} - im_{31} & m_{02} - im_{03} - m_{12} + im_{13} & m_{00} - m_{01} - m_{10} + m_{11} \end{bmatrix} \quad (3.3.1)$$

treating micropolarizer arrays as sampled sinusoidal functions ,i.e., the micropolarizers can be modeled as samples of continuously rotating polarizers, with the angle variable being dependent on the spatial image plane coordinate  $(x, y)$ . This assumption is valid under certain conditions but will not be addressed here.

Gil [62] showed the simplest necessary and sufficient conditions for a Mueller matrix to be physical; the coherency matrix,  $\mathbf{H}$ , corresponding to the Mueller matrix  $\mathbf{M}$ , must be Hermitian positive semi-definite.  $\mathbf{H}$  is defined in Equation (3.3.1) [82], where  $m_{jk}$ ,  $0 \leq j, k \leq 3$  are the elements of the associated Mueller matrix  $\mathbf{M}$  and  $i$  is the imaginary unit. The constraints on  $\mathbf{H}$  are easy to understand, but they are not easy to translate into constraints on channels. This definition for physical Mueller matrices can be used to generate test matrices computationally, by computing  $\mathbf{H}$  matrices and then transforming them to Mueller matrices  $\mathbf{M}$ . The two conditions on  $\mathbf{H}$  are

- $\mathbf{H} = \mathbf{H}^\dagger$  where  $\dagger$  here denotes conjugate transpose.
- positive semi-definiteness,  $\mathbf{x}^\dagger \mathbf{H} \mathbf{x} \geq 0$ ,  $\forall \mathbf{x} \in \mathbb{C}^4$ , where  $\dagger$  denotes conjugate transpose.

Translating these conditions to the Mueller-Stokes formalism, i.e., the constraints on  $\mathbf{M}$ , results in a very complicated set of inequality conditions and other conditions [82–86]. The car-

rier equations result from the Mueller-Stokes equations, and the carriers are constrained by the physicality conditions. It is not yet clear what constraints are imposed *in the channel space* by the physicality conditions.

The physicality conditions on Stokes parameters are  $s_0 > 0, s_0^2 \geq s_1^2 + s_2^2 + s_3^2$ , which is imposed on  $\mathbf{s}_G(\mathbf{x})$  defined in Chapter 2. If the assumption is made that the carrier amplitudes are equal for  $s_1(\mathbf{x}), s_2(\mathbf{x}), s_3(\mathbf{x})$  then the physicality condition implies that the maximum carrier amplitude for each Stokes parameter is

$$\frac{s_0}{\sqrt{3}} \quad (3.3.2)$$

for the case when maximum amplitude may be reached by  $s_1(\mathbf{x}), s_2(\mathbf{x}), s_3(\mathbf{x})$  simultaneously. However, the carrier could possibly be designed so that the maximum amplitude of  $s_1(\mathbf{x}), s_2(\mathbf{x}), s_3(\mathbf{x})$ , respectively, are spaced in such a manner that they do not simultaneously reach a maximum. For this case, there will be periodic maximum(minimum) values for each  $s_1(\mathbf{x}), s_2(\mathbf{x}), s_3(\mathbf{x})$ , denoted  $\mathbf{x}_{\max,1}, \mathbf{x}_{\max,2}, \mathbf{x}_{\max,3}$  respectively. Then

$$s_0^2 \geq s_1^2(\mathbf{x}_{\max,1}) + s_2^2(\mathbf{x}_{\max,1}) + s_3^2(\mathbf{x}_{\max,1}) \quad (3.3.3)$$

$$s_0^2 \geq s_1^2(\mathbf{x}_{\max,2}) + s_2^2(\mathbf{x}_{\max,2}) + s_3^2(\mathbf{x}_{\max,2}) \quad (3.3.4)$$

$$s_0^2 \geq s_1^2(\mathbf{x}_{\max,3}) + s_2^2(\mathbf{x}_{\max,3}) + s_3^2(\mathbf{x}_{\max,3}) \quad (3.3.5)$$

and the question arises: If  $s_1(\mathbf{x}), s_2(\mathbf{x}), s_3(\mathbf{x})$  each reach the maximum of  $s_0$  in amplitude, how are the frequencies constrained? A more intuitive analysis may use the Poincaré sphere with radius  $s_0$ . Then, the carrier function can be represented as the movement of the 3-vector

$$\mathbf{s} = s_0 \begin{bmatrix} \cos \theta \sin \phi \\ \sin \theta \sin \phi \\ \cos \phi \end{bmatrix} \quad (3.3.6)$$

around the surface of the Poincaré sphere, where  $\theta$  is the azimuthal angle and  $\phi$  is the elevation or polar angle. For purely a polarized carrier function, this implies that the output is a function of two variables,  $\theta$  and  $\phi$ , which are themselves functions of the modulation domain, i.e.,  $\theta = \theta(\mathbf{x})$  and  $\phi = \phi(\mathbf{x})$ , and of course  $0 \leq \theta < 2\pi$  and  $0 \leq \phi \leq \pi$ . A Fourier transform may then be performed, but there is not a

very useful mathematical form of function composition for a Fourier transform [87] in this case. Rigid rotations on the sphere are characterized by the special orthogonal group in three dimensions,  $\mathbf{SO}(3)$ , a matrix group of rotations. Due to this, quaternion representations may yield some better conceptual insight into the constraints in the channel space, but are not evaluated here.

If the assumption is made, as in Alenin and Tyo [34], that  $s_1(\mathbf{x})$ ,  $s_2(\mathbf{x})$ ,  $s_3(\mathbf{x})$  are finite periodic, that is they can be described by a finite Fourier series or equivalently a finite sum of sinusoidal functions, then the 3-dimensional Fourier series representation, or *Bravais lattice* [88] can be used. The Bravais lattice representation is more general than the one used in Alenin and Tyo [34] and implies

$$s_1(\mathbf{x}) = \cos \theta(\mathbf{x}) \sin \phi(\mathbf{x}) = \sum_{\mathbf{k}_1} F_0(\mathbf{k}_1) e^{2\pi i \mathbf{k}_1 \cdot \mathbf{x}} \quad (3.3.7)$$

$$s_2(\mathbf{x}) = \sin \theta(\mathbf{x}) \sin \phi(\mathbf{x}) = \sum_{\mathbf{k}_2} G_0(\mathbf{k}_2) e^{2\pi i \mathbf{k}_2 \cdot \mathbf{x}} \quad (3.3.8)$$

$$s_3(\mathbf{x}) = \cos \phi(\mathbf{x}) = \sum_{\mathbf{k}_3} H_0(\mathbf{k}_3) e^{2\pi i \mathbf{k}_3 \cdot \mathbf{x}} \quad (3.3.9)$$

where  $F_0(\mathbf{k}_1)$ ,  $G_0(\mathbf{k}_2)$ ,  $H_0(\mathbf{k}_3)$  are complex coefficients and  $\mathbf{k}_1$ ,  $\mathbf{k}_2$ ,  $\mathbf{k}_3$  are the “reciprocal lattice vectors.” Since all Stokes parameters are real,  $F_0(\mathbf{k}_1)$ ,  $G_0(\mathbf{k}_2)$ ,  $H_0(\mathbf{k}_3) = F_0^*(-\mathbf{k}_1)$ ,  $G_0^*(-\mathbf{k}_2)$ ,  $H_0^*(-\mathbf{k}_3)$  respectively where  $*$  denotes the complex conjugate.

All of these equations show general constraints, however translating them into specific conceptual and intuitive constraints on the channels themselves is a subject of future work. Moreover, the analyzer Mueller matrix,  $\mathbf{A}(\mathbf{x})$ , has a more complicated set of constraints, and is also the subject of future work.

### 3.4 Bandwidth

Given Equation (2.4.3) alone, an ideal distribution of delta functions, based on sphere packing optimization already worked out in mathematics [89, 90] could be derived. Optimal sphere packings represent optimal use of volume, i.e., they maximize density for packing spherical objects. This concept is needed in channeled systems design (and communication/information theory uses it extensively), to build bandwidth optimal instruments, *given an upper bound on total system bandwidth*.

The upper bound is typically imposed by engineering or physical constraints. Optimizing sphere packing provides each channel with the greatest possible space when all channels are constrained to some maximum bounding box in the frequency domain. Additionally, sphere packings of unequally sized spheres have also been studied [91, 92]. Unfortunately, the choice of sphere packings of the channels are constrained due to the following requirements which are imposed on Equation (2.2.6) as shown in Section 3.3

- Physical realizability of the  $A_i$  and  $S_j$ , they are not actually made up of independent delta functions.
- Reconstructability when there is delta function co-location (i.e. when some  $M_{ij}$  is actually added to a different  $M_{nm}$  at the same location in the Fourier space).

Alenin's treatment [34] considers reconstructability and optimization based on the equally weighted variance (EWV), which is equivalent to minimization of the *absolute error* [93, p. 24] of a linear system operator. The optimization with respect to EWV is, however, agnostic to bandwidth. I propose to use *error sensitivity* as a constraint or cost parameter in maximizing the bandwidth. The error parameter/constraint can be either condition number (*relative error* [93]) or equally weighted variance, or functions thereof. Conceptually, *absolute error* quantifies the actual maximum error or deviation of a system, while *relative error* quantifies the error relative to the measurements of the system, the two are related but are not the same. For example, an absolute error may have a magnitude of 0.1, but this number means something very different if the magnitudes of the measurements being made are on the order of 0.05 or on the order of  $10^6$ , the former is a poor measurement system and the latter is a good measurement system.

The treatment of bandwidth for channeled systems is mature and well known in

Figure 3.2: Optimal equal sphere packing arrangement for 3 dimensions : hexagonal close packed (hcp). Animated in the electronic version.



information and communications theory [94–96]. The difficulties of using a channeled systems framework for polarimetric instruments are primarily 1) constructing channels in 2 or more dimensions (many systems in communications theory are 1 dimensional), 2) addressing physicality constraints in an analytical way, and 3) addressing the complicated channel mixing behavior that is inherent to polarimetric instruments. Point 3 has mostly been addressed by Alenin and Tyo [34] while point 2 is a complicated subject and details are elucidated in Section 3.3. All of these constraints must be enforced when optimizing for some cost function.

There will exist a set of channels ( $\delta$ -functions) for each Mueller matrix element  $M_{ij}(\boldsymbol{\rho})$  in the Fourier domain. For each channel in that set,  $M_{ij}(\boldsymbol{\rho})$  will be copied at that channel's location with  $M_{ij}(\mathbf{0})$  being located precisely where the delta function is located. The *bandwidth* of  $M_{ij}$  for some threshold  $\epsilon_c \geq 0$  can be defined as the values of  $\boldsymbol{\rho}$  where  $|M_{ij}(\boldsymbol{\rho})| > \epsilon_c$ . Fig.3.3 clarifies these concepts.

The polarimetric system channel structure restrains bandwidth because there is a finite distance between channels as shown in Figure 3.1. The channel structure determines the bandwidth available for reconstruction. When the bandwidth of the data becomes greater than the available bandwidth, then channel crosstalk occurs.

### 3.4.1 Crosstalk

Crosstalk is similar to aliasing, but not the same phenomenon. Crosstalk is the result of limited channel bandwidth, and information (convolutions of data) in the channel exceeding the bandwidth of that channel and "spilling or bleeding" over to an adjacent channel. Crosstalk is a result of the choice of channel structure, as opposed to the sampling rate (aliasing), even a continuously sampled channeled (unaliased) system can have crosstalk. An example of crosstalk is shown in Figure 3.4. For polarimetric instruments *the physical channel structure determines the crosstalk bandwidth limits*, irrespective of the sampling used. That is, the set of physical carriers chosen by the system designer impose a bandwidth limit which is *independent of the sampling rate*.

Aside from and independent of noise, channel crosstalk is the greatest limitation to accurate Mueller matrix reconstruction. If the majority of the bandwidth of the object is greater than the no crosstalk bandwidth limit, then the object cannot be

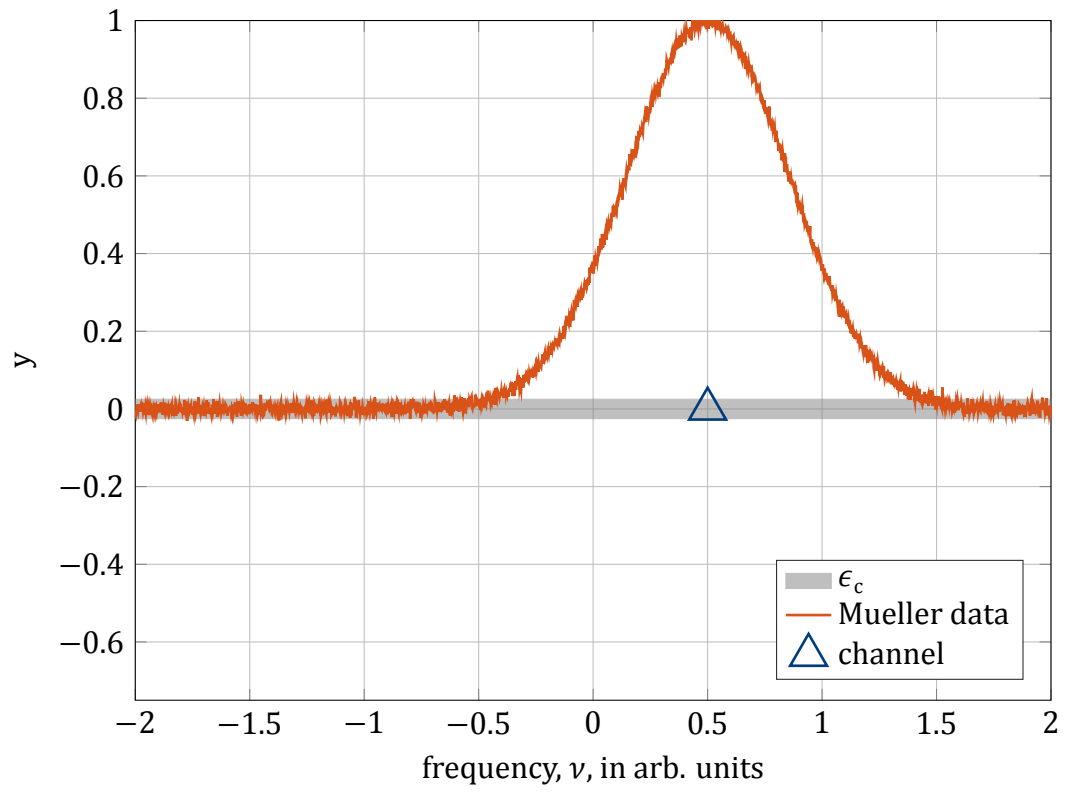


Figure 3.3: An example of convolution of data with a channel. The gray band represents the range of  $\epsilon_c$ , resulting in data being outside of the  $\epsilon_c$  range for  $-0.5 < \nu < 1.5$ . This implies a bandwidth of 2 arb. units for the Mueller data.

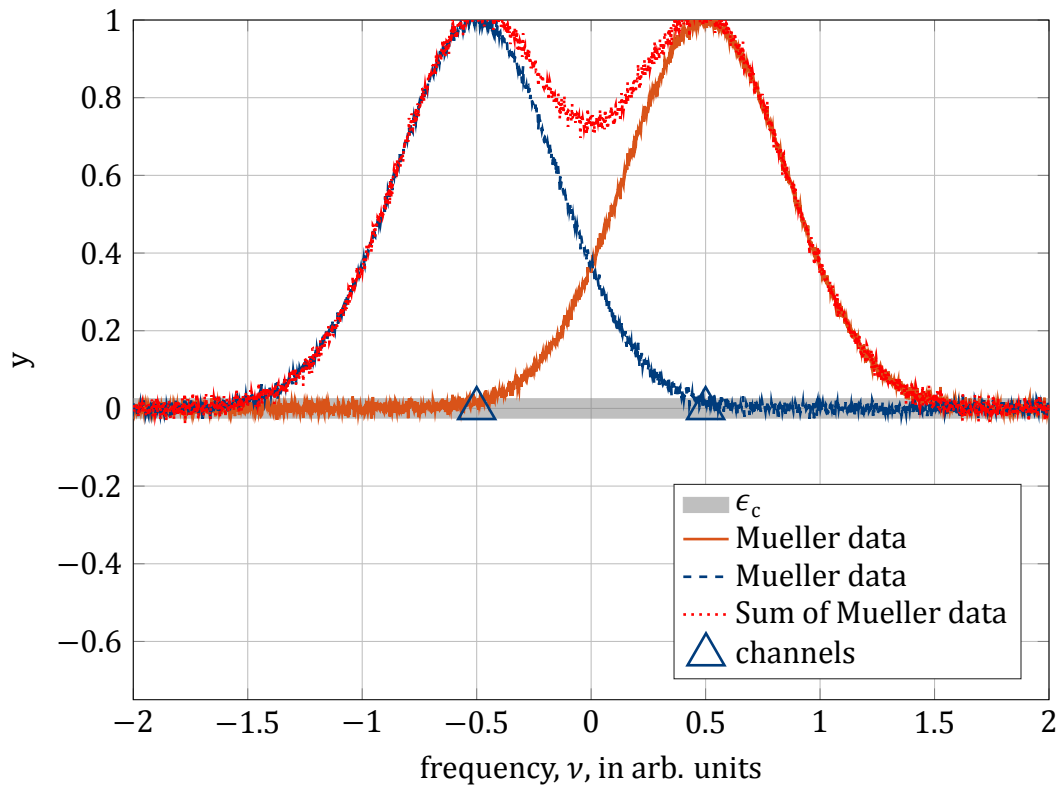


Figure 3.4: An example of channel crosstalk. Mueller data is placed at two channels, with the distance (bandwidth) between them less than the bandwidth of the Mueller data. When added, the Mueller data from different channels adds together, leaving no remedy to differentiate data between channels in the region of bandwidth crossover (the neighborhood of 0) when given arbitrary Mueller data.

accurately reconstructed or estimated. For a Mueller matrix polarimeter system, a general remote sensing instrument should minimize the crosstalk (equivalently maximize the bandwidth) between channels equally. However, if the bandwidth distribution is known or assumed *a priori* for objects of interest, and the bandwidths are different for different Mueller matrix elements, then it may be possible to design systems with unequal bandwidth between channels. This unequal distribution will not be addressed in depth here, and I hope to analyze it in future work. Some theoretical examples of unequal channel bandwidth systems are shown in the results section.

### 3.5 Filtering

Typically crosstalk can be “mitigated” by using filters around the channels to suppress or apodize the region where crosstalk occurs. This does not fully mitigate the corruption from crosstalk however because

- Filters which apodize in some way result in smoothing of the data, essentially removing information.
- Similar to the above, apodization or cutoff from the filters reduces the reconstructable bandwidth of the resulting Mueller data.
- Filters won’t help much in the case where a great deal of crosstalk is present.

Filtering is needed, but cannot fully alleviate the crosstalk issue. Filtering will not be addressed in detail in this chapter; the literature on filtering is vast and mature in control theory and electrical engineering. A discussion on the filters specifically used for the examples and physical polarimetric instrument are in Chapter 5.

Filtering facilitates an additional requirement for reconstruction, however; to provide data suitable for inversion by the  $\mathbf{Q}$  matrix to obtain Mueller matrix estimates. See Section 2.4.2 for the discussion of the  $\mathbf{Q}$  matrix. The data at each channel is a mixture of corresponding Mueller matrix elements and must be unmixed via the pseudoinverse of  $\mathbf{Q}$ , denoted  $\mathbf{Q}^+$  [34]. In order to separate out the “balls” of data around each channel location, filters must be used. The filtered “ball” of data will have a center at the channel location, which can then be moved to  $(0, 0, 0)$ . An ex-

ample of a specific inversion algorithm, including the  $\mathbf{Q}$  matrix and the filtering, is shown in Chapter 5.

### 3.6 Maximizing bandwidth

Increasing relative bandwidth can, of course, reduce crosstalk, and subsequently increase the system resolution or speed for spatio-temporally channeled active polarimetric systems. In order to maximize the relative bandwidth, we must think about the system in a way which addresses *efficiency*, otherwise an optimizer will increase the maximum frequency (and hence the relative frequency distance between channels) *ad infinitum* until a specification is met. Additionally, instruments with arbitrary measurement bandwidth do not exist. In order to constrain the bandwidth maximization to relative frequencies, all of the channels are normalized to be contained in a cube (or rectangular prism in certain cases) where the maximum frequency is normalized to be some fixed value. Different norms can be chosen to accomplish this as long as the norm used is consistent. The two simplest methods are 1) normalize in a 2-norm way, that is your maximum frequency is taken as a vector and normalized by its 2-norm length, and all other channel locations are also normalized by this same length, or 2) normalize in an  $\infty$ -norm way, that is normalize each frequency domain coordinate by the respective maximum frequency channel domain coordinate. The  $\infty$ -norm normalization should be used for separable, or independent domains, while the 2-norm normalization should be used for general, or non-separable domains. To clarify with an example, suppose that our maximum frequency channel is located at  $[0.5, 0.5, 60]^T$ , then the normalization factors would be

$$n_{2\text{-norm}} = \sqrt{0.5^2 + 0.5^2 + 60^2} \quad (3.6.1)$$

$$n_{\infty\text{-norm},0} = 0.5 \quad (3.6.2)$$

$$n_{\infty\text{-norm},1} = 0.5 \quad (3.6.3)$$

$$n_{\infty\text{-norm},2} = 60 \quad (3.6.4)$$

$$(3.6.5)$$

Figure 3.5: An example of how varying retardance parameters change channels in the channel space. Animated in the electronic version, use the controls displayed to view.

Figure 3.6: An example of how varying starting offset parameters change channels in the channel space. Animated in the electronic version, use the controls displayed to view.

and for an arbitrary channel located at  $\mathbf{c}_{\text{arb}} = [\xi_{\text{arb}} \quad \eta_{\text{arb}} \quad \nu_{\text{arb}}]^T$  the two normalizations would be

$$\mathbf{c}_{\text{arb},2\text{-norm}} = \frac{\mathbf{c}_{\text{arb}}}{\sqrt{0.5^2 + 0.5^2 + 60^2}} \quad \text{and} \quad \mathbf{c}_{\text{arb},\infty\text{-norm}} = \begin{bmatrix} \frac{\xi_{\text{arb}}}{0.5} \\ \frac{\eta_{\text{arb}}}{0.5} \\ \frac{\nu_{\text{arb}}}{60} \end{bmatrix} \quad (3.6.6)$$

In the examples shown here, the locations of frequency corresponding to the temporal domain,  $\nu$ , are normalized but the spatial frequencies are not because the examples assume a fixed micropolarizer array which cannot be changed [4]. The spatial frequency coordinates have no effect on the analysis when they are fixed, but in general, if optimization over spatial frequency is an option, the spatial frequency channel coordinates would also need to be normalized. Normalization ensures that for a relative bandwidth optimization an *oranges to oranges* comparison is made as channel location changes. The use of different normalization types will lead to different optimization outcomes.

### 3.7 Optimization

Optimization must account for *bandwidth*, *reconstructability*, and *system conditioning*. Bandwidth for polarimetric instruments has already been introduced earlier in this chapter. Reconstructability defines what elements of the Mueller matrix can be reconstructed from a specific channeled system, i.e., an active instrument could reconstruct the full Mueller matrix, or only  $m_{00}$  (a normal camera paired with “flashlight” illumination) or any partial Mueller matrix polarimeter (pMMP) in between. There are constraints on *which* pMMPs can be realized as channeled systems, and Alenin and Tyo have recently worked out some of these classes of pMMP systems [97]. This chapter will focus on full Mueller matrix systems, but some examples of pMMPs will be presented near the end. The bandwidth optimization process is similar across all channeled system types. System conditioning can be described via the condition number, CN, or equally weighted variance, EWW [98].



Optimization strategy is a field in its own right, and there are a plethora of methods to minimize or maximize some figure of merit or cost function. Often, the difficulty in optimization problems lies in 1) defining an appropriate cost function and 2) applying the correct optimization search strategy given the structure of the cost function. Cost functions can often map multidimensional spaces to other spaces, and may have many local minima or maxima, or they may be well behaved and suitable for more efficient optimization strategies. Two broad types of optimizations exist:

- convex, which implies that a unique local minimum/maximum exists and is the global minimum/maximum.
- non-convex, which implies that many local minima/maxima exist.

Of course, the convexity is with respect to the cost function, and convex cost functions are considered to be “easier” to optimize in practice, however there is no mathematical proof which show that convex problems are necessarily easier to optimize in general. If a cost function is convex, there are often fast and efficient methods to find the global minimum/maximum. When the cost function is non-convex, there are methods which can find *good* but not necessarily *optimal* solutions. Non-convex optimizations include simulated annealing, genetic algorithms, multi-point gradient descent (and associated quasi-Newton methods), conjugate gradient, etc. If possible, cost functions for bandwidth maximization should be designed to be convex.

A general cost function can be defined which accounts for both noise and bandwidth:

$$\mathcal{O}[\mathbf{A}_c(\boldsymbol{\rho}), \mathbf{G}_c(\boldsymbol{\rho})] = \frac{\text{dist}[\mathbf{A}_c(\boldsymbol{\rho}), \mathbf{G}_c(\boldsymbol{\rho})]}{(CN[\mathbf{A}_c(\boldsymbol{\rho}), \mathbf{G}_c(\boldsymbol{\rho})])^n} \quad \text{or} \quad \frac{\text{dist}[\mathbf{A}_c(\boldsymbol{\rho}), \mathbf{G}_c(\boldsymbol{\rho})]}{(EWV[\mathbf{A}_c(\boldsymbol{\rho}), \mathbf{G}_c(\boldsymbol{\rho})])^n} \quad (3.7.1)$$

where  $\mathbf{G}_c(\boldsymbol{\rho})$  are the sets of generator channels from the PSG and  $\mathbf{A}_c(\boldsymbol{\rho})$  are the sets of analyzer channels from the PSA, both are dependent on  $\boldsymbol{\rho}$ , the Fourier domain dual to the modulation domain.  $CN$  and  $EWV$  both grow as systems become ill-conditioned and  $n$  controls the preference of system conditioning over bandwidth in the cost function, as  $n$  becomes larger the cost function will weight system conditioning more heavily. The  $\text{dist}[\mathbf{A}_c(\boldsymbol{\rho}), \mathbf{G}_c(\boldsymbol{\rho})]$  function characterizes, in some way, the bandwidth between channels or the distance between channels. This could be an average distance, a minimum distance, etc.

It must be emphasized that generally, channels from different Mueller matrix elements will be mixed together in the channel space *at the same locations*. This means that the  $\text{dist}[\mathbf{A}_c(\boldsymbol{\rho}), \mathbf{G}_c(\boldsymbol{\rho})]$  must compute the distance between *unique* channel locations in the channel space. This is illustrated in Figures 3.5-3.7, 3.10 and 3.12, the channels corresponding to different Mueller matrix elements are often added to other elements at identical locations in the channel space. The unique channel locations and the distances between them determine bandwidth, while the  $\mathbf{Q}$  matrix characterizes the system conditioning. The  $\text{dist}[\mathbf{A}_c(\boldsymbol{\rho}), \mathbf{G}_c(\boldsymbol{\rho})]$  function is typically smooth if channel cancellation (discussed below) does not occur. For separable systems, however, channel cancellation is explicitly used to improve bandwidth, resulting in discontinuities in  $\text{dist}[\mathbf{A}_c(\boldsymbol{\rho}), \mathbf{G}_c(\boldsymbol{\rho})]$ .

### 3.8 Channel cancellation

For a separable channel structure, only channel cancellation/combination or reduction of overall channels may be used to increase the relative bandwidth between channels. Figure 3.7 shows channel combination as relative retarder frequency is changed for a quad-retarder + micropolarizer array system. At certain relative frequencies, channels combine or cancel depending on their magnitudes, providing larger distance (bandwidth) between channels. Running an optimizer over a cost function,  $\mathcal{O}[\mathbf{A}_c(\boldsymbol{\rho}), \mathbf{G}_c(\boldsymbol{\rho})]$  can then find a good, or possibly even optimal, channel structure. An example of an optimal (to the best of our knowledge) channel structure for the quad retarder + micropolarizer array system is shown in Figure 3.10.

Channel cancellation poses a problem, however, the cost function becomes *discontinuous* when channels cancel. This implies that the cost function is not only a non-convex type, but also requires *derivative free* optimization methods, potentially increasing the computational time. Examples of derivative free methods are simulated annealing and genetic algorithms. As shown in Figure 3.7, the  $\text{dist}[\mathbf{A}_c(\boldsymbol{\rho}), \mathbf{G}_c(\boldsymbol{\rho})]$  part of the cost function can decrease, then jump discontinuously to a higher value when channels cancel or combine. One approach to mitigate this problem is to systematically compute the set of discontinuities in some way [99,100], since heuristically it appears that many of the local bandwidth maxima occur at

Figure 3.7: An example of varying frequency parameters to move channels around in the channel space. Note the cancellation/addition at certain parameter values which opens up more bandwidth between the channels. Animated in the electronic version, use the controls displayed to view.

the discontinuities themselves. Another approach is to compute some definite integral of the  $\text{dist}[\mathbf{A}_c(\boldsymbol{\rho}), \mathbf{G}_c(\boldsymbol{\rho})]$  function, which will lead to minima of the integral function near the discontinuities. A computational approach would be to allow the  $\text{dist}[\mathbf{A}_c(\boldsymbol{\rho}), \mathbf{G}_c(\boldsymbol{\rho})]$  to contain zeros initially (if we allow overlapping channels to mean zero distance), then to find all the zeros using an optimizer and save the set of zeros for later re-evaluation. This approach still has the issues of multiple local minima however, and the associated computational cost.

Optimization for a quad retarder + micropolarizer array system for the examples presented here consisted of a hybrid process of inspection, optimizing a multi-nodal objective function, and manually inputting starting points for hill climbing for local maxima evaluation. Simulated annealing was also used. Sets of maxima found were saved and iterated over many times to find the optimal systems presented in this chapter, see Figure 3.10.

### 3.9 Bandwidth Optimization Discussion

In general a channeled system can be described by  $\mathbf{A}_c(\boldsymbol{\rho}), \mathbf{G}_c(\boldsymbol{\rho})$ , the sets of channels for the PSA and PSG respectively, in practice, however, there is typically a design parameter space which  $\mathbf{A}_c(\boldsymbol{\rho}), \mathbf{G}_c(\boldsymbol{\rho})$  depend upon. During the engineering and designing process for polarimetric systems, other constraints come into play, and a designer will typically be left with some set of free parameters to optimize the channel structure over. These free parameters will be denoted by  $\mathbf{p}$ .

Once the channeled system framework is understood, and the free parameters,  $\mathbf{p}$  of a spatio-temporally channeled system are specified or known, then it is straightforward to design a cost function and run an optimizer over that function to optimize for bandwidth or jointly for bandwidth, noise, and other constraints. The most difficult part of directly optimizing in the channel space is not, however, running an optimization against the cost function. The difficult task is designing a model which properly describes the channel structure itself, with proper physical constraints, from which an appropriate cost function may be derived. I have designed a model for the specific case of a quad-retarder + micropolarizer array system (see Chapter 5), and I hope to adapt the current model to generate generic spatio-temporally channeled systems in

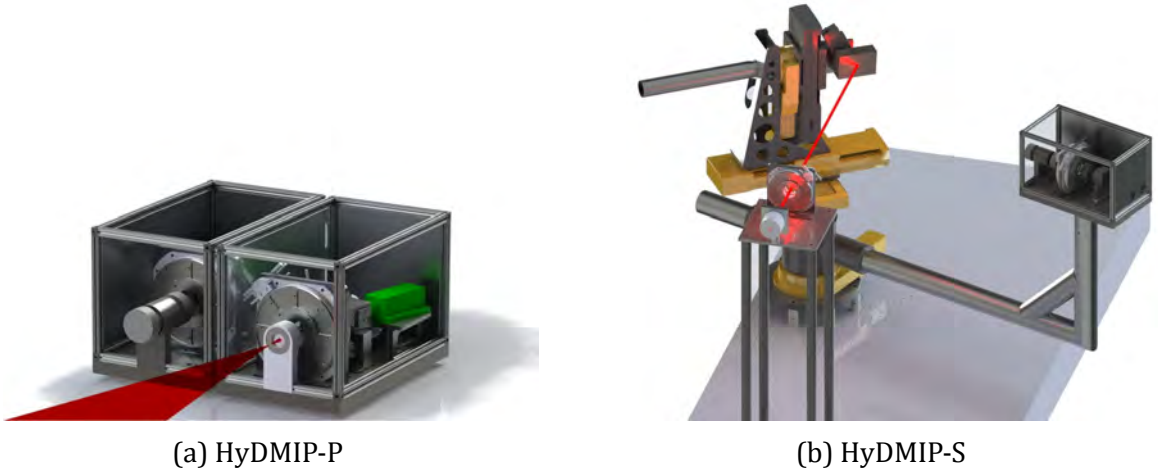


Figure 3.8: The Hybrid Domain Modulated Imaging Polarimeter (HyDMIP). (a) shows the portable part of the testbed, HyDMIP-P; and (b) shows the full lab scatterometer, HyDMIP-S. (a) is the primary focus here.

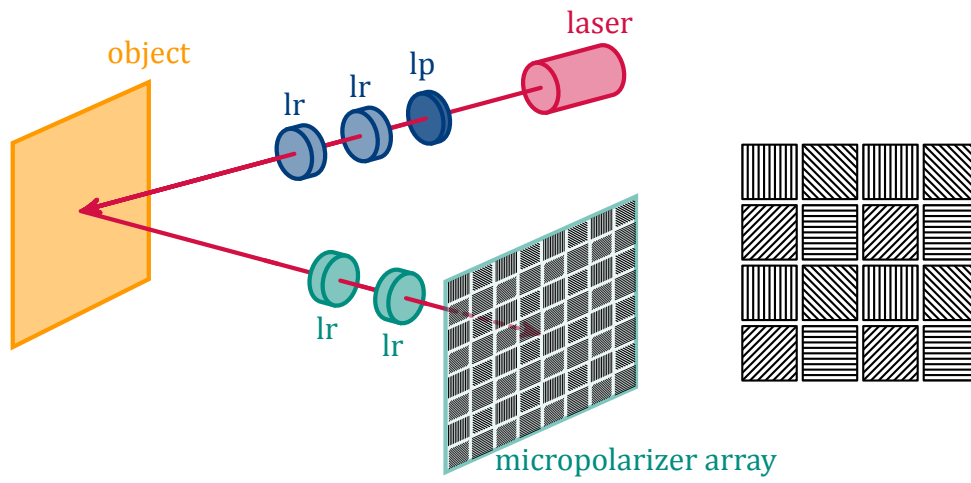
the near future.

Additionally, if the statistics of an object or set of objects are known, then non-uniform bandwidth can be maximized. Most of the examples shown in this dissertation optimize for an equal channel bandwidth between all channels. In certain cases, more bandwidth may be wanted for certain sets of channels over other sets of channels and for certain Mueller matrix elements. This can all be accomplished by using the appropriate cost function, but is non-trivial due to the channel mixing which occurs between Mueller matrix elements.

### 3.10 An Example

Recently, hardware which I designed was acquired for the Advanced Sensing Lab (ASL) at the University of Arizona College of Optical Sciences. The hardware platform was designed to test a broad range of polarimetric techniques and imaging operators, particularly channeled operators. This system is named the Hybrid Domain Modulated Imaging Polarimeter (HyDMIP) shown in Figure 3.8.

Here I will focus on the portable instrument, HyDMIP-P, and derive the channel structure and the free parameters which can be used for the optimization of the bandwidth using a specific objective function  $\mathcal{O}(\mathbf{p})$ , where  $\mathbf{p}$  is the set of free parameters



(a) layout

(b) micropolarizer tiling

Figure 3.9: Spatio-temporally channeled polarimeter schematic, **lp**=linear polarizer, **lr**=linear retarder, blue components denote the polarization state generator (PSG), green components denote the polarization state analyzer (PSA). The micropolarizer array is the conventional tiling, shown in (b). (a) was derived and modified from a figure created by Andrey Alenin.

which can be optimized over. The HyDMIP-P currently has a typical temporal carrier generating polarization state generator, except that additional degrees of freedom are introduced by using two retarders instead of the usual single retarder. The receiver uses spatio-temporal carriers. The temporal carriers for the receiver are also generated via two retarders, and the spatial carriers occur via a micropolarizer array aligned to a silicon sensor. See Chapter 5 for details about the instrument.

### 3.10.1 Instrument layout

The PSG of the instrument is similar to a typical rotating retarder polarimeter design, the only difference is two retarders are used instead of a single retarder to increase the available degrees of freedom for optimization. The PSA is where the hybrid spatio-temporal carrier generation occurs and consists of two rotating retarders coupled with a micropolarizer array attached to the silicon focal plane array (FPA). Figure 3.9 shows a schematic of the system layout. The micropolarizer array is as-

sumed to be the conventional  $0^\circ, 45^\circ, -45^\circ, 90^\circ$  superpixel layout. Other layouts are possible [35, 101], but not commercially available at the moment. Results on optimization for other micropolarizer layouts are not presented here, but the same analysis will generalize to arbitrary micropolarizer layouts. The analysis presented here shows the physical channels available for the system, and sampling is not addressed in this chapter. The methods are general, but the channel structures presented here are ideal physical channels specific to the instrument.

### 3.10.2 Derivation of analyzer and generator equations

Without the spatio-temporal dependencies denoted explicitly, the system equation for HyDMIP-P becomes

$$\mathbf{s}_{\text{out}} = \mathbf{P} \cdot \mathbf{R}_{\delta_4} \cdot \mathbf{R}_{\delta_3} \cdot \mathbf{M} \cdot \mathbf{R}_{\delta_2} \cdot \mathbf{R}_{\delta_1} \mathbf{s}_{\text{in}} \quad (3.10.1)$$

Where  $\mathbf{s}_{\text{in}} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ ,  $\mathbf{R}_{\delta_j}$  indicates a linear retarder with a retardance of  $\delta_j$ ,  $\mathbf{M}_{\text{obj}}$  is the

Mueller matrix which we desire to measure, and  $\mathbf{P}$  is the micropolarizer array. The explicit dependence on  $\mathbf{x}$  or  $\mathbf{p}$  has been dropped here for brevity.

Using the exact same procedure as in Equation (2.2.14), the following inner product can be obtained:

$$s_{0,\text{out}} = (\mathbf{M}^T \cdot \mathbf{R}_{\delta_3}^T \cdot \mathbf{R}_{\delta_4}^T \mathbf{p} \ , \ \mathbf{R}_{\delta_2} \cdot \mathbf{R}_{\delta_1} \mathbf{s}_{\text{in}}) \quad (3.10.2)$$

where  $\mathbf{p}$  is the first row of  $\mathbf{P}$  and  $\mathbf{A}^T$  is the transpose of  $\mathbf{A}$ . Given Equation (3.10.2), we only need to compute  $\mathbf{R}_{\delta_3}^T \cdot \mathbf{R}_{\delta_4}^T \mathbf{p}$  and  $\mathbf{R}_{\delta_2} \cdot \mathbf{R}_{\delta_1} \mathbf{s}_{\text{in}}$  to proceed. For a conventional micropolarizer array, in image plane coordinates [33]:

$$\mathbf{p} = \frac{1}{2} \begin{bmatrix} 1 \\ \frac{\cos \pi x + \cos \pi y}{2} \\ \frac{\cos \pi x - \cos \pi y}{2} \\ 0 \end{bmatrix}, \quad (3.10.3)$$

but  $\mathbf{p}$  could be any spatial periodic arrangement of micropolarizers. A general linear retarder,  $\mathbf{R}_\delta(\theta(t))$  where  $\delta$  is the retardance in radians, and  $\theta(t)$  is the rotation angle as a function of time, is represented by the Mueller matrix [102] (after some trigonometric manipulation):

$$\mathbf{R}_\delta(\theta(t)) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 \frac{\delta}{2} + \sin^2 \frac{\delta}{2} \cos(4\theta(t)) & \sin^2 \frac{\delta}{2} \sin(4\theta(t)) & -\sin \delta \sin(2\theta(t)) \\ 0 & \sin^2 \frac{\delta}{2} \sin(4\theta(t)) & \cos^2 \frac{\delta}{2} - \sin^2 \frac{\delta}{2} \cos(4\theta(t)) & \sin \delta \cos(2\theta(t)) \\ 0 & \sin \delta \sin(2\theta(t)) & -\sin \delta \cos(2\theta(t)) & \cos \delta \end{bmatrix} \quad (3.10.4)$$

The notation for the  $\theta(t)$  terms for each retarder can be simplified. Here, we make the assumption that each retarder is rotating at a constant angular velocity,  $2\pi\nu_j t$  for each  $j$ . Furthermore, we assume that there is a maximum possible velocity for any of the retarders;  $2\pi\nu_0 t$ , i.e.,  $\nu_1, \nu_2, \nu_3, \nu_4 \leq \nu_0$ . These assumptions allow us to write the  $\theta(t)$ s as the following linear relationships:

$$\theta_1(t) = 2\pi(C_1\nu_0 t + \epsilon_1) \quad (3.10.5)$$

$$\theta_2(t) = 2\pi(C_2\nu_0 t + \epsilon_2) \quad (3.10.6)$$

$$\theta_3(t) = 2\pi(C_3\nu_0 t + \epsilon_3) \quad (3.10.7)$$

$$\theta_4(t) = 2\pi(C_4\nu_0 t + \epsilon_4) \quad (3.10.8)$$

where

$$\nu_1 = C_1\nu_0 \quad (3.10.9)$$

$$\nu_2 = C_2\nu_0 \quad (3.10.10)$$

$$\nu_3 = C_3\nu_0 \quad (3.10.11)$$

$$\nu_4 = C_4\nu_0 \quad (3.10.12)$$



and  $\theta_1(t), \theta_2(t), \theta_3(t), \theta_4(t)$  represent the rotation angles of  $\mathbf{R}_{\delta_1}, \mathbf{R}_{\delta_2}, \mathbf{R}_{\delta_3}, \mathbf{R}_{\delta_4}$  respectively,  $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$  are constant offsets in normalized radians from the coordinates defined by the PSG polarizer, and  $0 \leq C_1, C_2, C_3, C_4 \leq 1$  represent the respective angular frequency rates with respect to the maximum rate  $\nu_0$ .

Two linear retarders are multiplied together to obtain the Mueller matrix  $\mathbf{RR}(\nu_i, \nu_j, \epsilon_i, \epsilon_j; \delta_i, \delta_j) = \mathbf{R}_{\delta_j} \cdot \mathbf{R}_{\delta_i}$ . Note that elements of  $\mathbf{RR}$  are denoted by  $RR_{kl}$ . See Appendix A for additional detail.

Now  $\mathbf{R}_{\delta_2} \cdot \mathbf{R}_{\delta_1} \mathbf{s}_{\text{in}}$  is derived from Equation (3.10.2):

$$\begin{bmatrix} 1 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 1 \\ RR_{11}(\nu_1, \nu_2, \epsilon_1, \epsilon_2; \delta_1, \delta_2) \\ RR_{21}(\nu_1, \nu_2, \epsilon_1, \epsilon_2; \delta_1, \delta_2) \\ RR_{31}(\nu_1, \nu_2, \epsilon_1, \epsilon_2; \delta_1, \delta_2) \end{bmatrix}. \quad (3.10.13)$$

The  $RR_{kl}$  definitions can be found in appendix A.  $\mathbf{R}_{\delta_3}^t \cdot \mathbf{R}_{\delta_4}^t \mathbf{p}$ , the other portion of the inner product in Equation (3.10.2), is:

$$\frac{1}{2} \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 2 \\ \cos \pi x (RR_{11} + RR_{21}) + \cos \pi y (RR_{11} - RR_{21}) \\ \cos \pi x (RR_{12} + RR_{22}) + \cos \pi y (RR_{12} - RR_{22}) \\ \cos \pi x (RR_{13} + RR_{23}) + \cos \pi y (RR_{13} - RR_{23}) \end{bmatrix} \quad (3.10.14)$$

where each  $RR_{kl} = RR_{kl}(\nu_3, \nu_4, \epsilon_3, \epsilon_4; \delta_3, \delta_4)$  above due to space constraints.

### 3.10.3 Fourier transform of analyzer and generator equations

The channel structure can now be extracted from the equations that we derived in the last subsection. This is simple conceptually, we take the three dimensional Fourier transforms of  $1, s_1, s_2, s_3$  and  $1, a_1, a_2, a_3$  and then convolve them as prescribed in Equation (2.2.6). In practice, the computation is somewhat tedious, and for brevity many of the steps are omitted here.

First, the Fourier transform of each of  $1, s_1, s_2, s_3$  and  $1, a_1, a_2, a_3$  are computed (here the factor of  $\frac{1}{2}$  in Equation (3.10.14) can be ignored since it multiplies all channels in the final channel structure) and simplified. Then, without explicitly comput-

ing the delta functions from appendix A we have:

$$\mathcal{F}\{\mathbf{R}_{\delta_2} \cdot \mathbf{R}_{\delta_1} \mathbf{s}_{\text{in}}\}_{[x,y,t] \rightarrow [\xi,\eta,v]} = \delta(\xi)\delta(\eta) \begin{bmatrix} \delta(v) \\ \mathcal{F}\{RR_{11}(v_1, v_2, \epsilon_1, \epsilon_2; \delta_1, \delta_2)\}_{t \rightarrow v} \\ \mathcal{F}\{RR_{21}(v_1, v_2, \epsilon_1, \epsilon_2; \delta_1, \delta_2)\}_{t \rightarrow v} \\ \mathcal{F}\{RR_{31}(v_1, v_2, \epsilon_1, \epsilon_2; \delta_1, \delta_2)\}_{t \rightarrow v} \end{bmatrix}. \quad (3.10.15)$$

and, given the following Fourier transform pairs:

$$\cos \pi x \leftrightarrow \frac{1}{2} \left[ \delta\left(\xi + \frac{1}{2}\right) + \delta\left(\xi - \frac{1}{2}\right) \right]$$

$$\cos \pi y \leftrightarrow \frac{1}{2} \left[ \delta\left(\eta + \frac{1}{2}\right) + \delta\left(\eta - \frac{1}{2}\right) \right]$$

then

$$2\mathcal{F}\{\mathbf{R}_{\delta_3}^t \cdot \mathbf{R}_{\delta_4}^t \mathbf{p}\}_{[x,y,t] \rightarrow [\xi,\eta,v]} \quad (3.10.16)$$

$$= \frac{1}{4} \begin{bmatrix} 4\delta(\xi)\delta(\eta)\delta(v) \\ \left[ \delta\left(\xi + \frac{1}{2}\right) + \delta\left(\xi - \frac{1}{2}\right) \right] \mathcal{F}\{RR_{11} + RR_{21}\}_{t \rightarrow v} + \left[ \delta\left(\eta + \frac{1}{2}\right) + \delta\left(\eta - \frac{1}{2}\right) \right] \mathcal{F}\{RR_{11} - RR_{21}\}_{t \rightarrow v} \\ \left[ \delta\left(\xi + \frac{1}{2}\right) + \delta\left(\xi - \frac{1}{2}\right) \right] \mathcal{F}\{RR_{12} + RR_{22}\}_{t \rightarrow v} + \left[ \delta\left(\eta + \frac{1}{2}\right) + \delta\left(\eta - \frac{1}{2}\right) \right] \mathcal{F}\{RR_{12} - RR_{22}\}_{t \rightarrow v} \\ \left[ \delta\left(\xi + \frac{1}{2}\right) + \delta\left(\xi - \frac{1}{2}\right) \right] \mathcal{F}\{RR_{13} + RR_{23}\}_{t \rightarrow v} + \left[ \delta\left(\eta + \frac{1}{2}\right) + \delta\left(\eta - \frac{1}{2}\right) \right] \mathcal{F}\{RR_{13} - RR_{23}\}_{t \rightarrow v} \end{bmatrix}.$$

This can be expanded into a linear combination of three separate vectors:

$$\begin{bmatrix} \delta(\xi)\delta(\eta)\delta(\nu) \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (3.10.17)$$

$$+\frac{1}{4} \begin{bmatrix} 0 \\ \left[ \delta\left(\xi + \frac{1}{2}\right) + \delta\left(\xi - \frac{1}{2}\right) \right] \mathcal{F}\{RR_{11} + RR_{21}\}_{t \rightarrow \nu} \\ \left[ \delta\left(\xi + \frac{1}{2}\right) + \delta\left(\xi - \frac{1}{2}\right) \right] \mathcal{F}\{RR_{12} + RR_{22}\}_{t \rightarrow \nu} \\ \left[ \delta\left(\xi + \frac{1}{2}\right) + \delta\left(\xi - \frac{1}{2}\right) \right] \mathcal{F}\{RR_{13} + RR_{23}\}_{t \rightarrow \nu} \end{bmatrix} \quad (3.10.18)$$

$$+\frac{1}{4} \begin{bmatrix} 0 \\ \left[ \delta\left(\eta + \frac{1}{2}\right) + \delta\left(\eta - \frac{1}{2}\right) \right] \mathcal{F}\{RR_{11} - RR_{21}\}_{t \rightarrow \nu} \\ \left[ \delta\left(\eta + \frac{1}{2}\right) + \delta\left(\eta - \frac{1}{2}\right) \right] \mathcal{F}\{RR_{12} - RR_{22}\}_{t \rightarrow \nu} \\ \left[ \delta\left(\eta + \frac{1}{2}\right) + \delta\left(\eta - \frac{1}{2}\right) \right] \mathcal{F}\{RR_{13} - RR_{23}\}_{t \rightarrow \nu} \end{bmatrix} \quad (3.10.19)$$

where, again, the parameter dependence,  $(\nu_3, \nu_4, \epsilon_3, \epsilon_4; \delta_3, \delta_4)$ , is omitted for brevity. Appendix A contains detailed Fourier transform results for  $RR_{11}, RR_{12}, RR_{13}, RR_{21}, RR_{22}, RR_{23}, RR_{31}$ . Upon inspection of these results, the following equalities hold:

$$\mathcal{F}\{RR_{21}\}_{t \rightarrow \nu} = -i\mathcal{F}\{RR_{11}\}_{t \rightarrow \nu} + i\cos^2 \frac{\delta_4}{2} \cos^2 \frac{\delta_3}{2} \delta(\nu) \quad (3.10.20)$$

$$\mathcal{F}\{RR_{22}\}_{t \rightarrow \nu} = i\mathcal{F}\{RR_{12}\}_{t \rightarrow \nu} + \cos^2 \frac{\delta_4}{2} \cos^2 \frac{\delta_3}{2} \delta(\nu) \quad (3.10.21)$$

$$\mathcal{F}\{RR_{23}\}_{t \rightarrow \nu} = i\mathcal{F}\{RR_{13}\}_{t \rightarrow \nu}. \quad (3.10.22)$$

These may then be substituted into Equation (3.10.16) to obtain

$$\begin{bmatrix} \delta(\xi)\delta(\eta)\delta(\nu) \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (3.10.23)$$

$$+ \frac{1}{4} \begin{bmatrix} 0 \\ \left[ \delta\left(\xi + \frac{1}{2}\right) + \delta\left(\xi - \frac{1}{2}\right) \right] \left[ (1-i)\mathcal{F}\{RR_{11}\}_{t \rightarrow \nu} + i \cos^2 \frac{\delta_4}{2} \cos^2 \frac{\delta_3}{2} \delta(\nu) \right] \\ \left[ \delta\left(\xi + \frac{1}{2}\right) + \delta\left(\xi - \frac{1}{2}\right) \right] \left[ (1+i)\mathcal{F}\{RR_{12}\}_{t \rightarrow \nu} + \cos^2 \frac{\delta_4}{2} \cos^2 \frac{\delta_3}{2} \delta(\nu) \right] \\ \left[ \delta\left(\xi + \frac{1}{2}\right) + \delta\left(\xi - \frac{1}{2}\right) \right] (1+i)\mathcal{F}\{RR_{13}\}_{t \rightarrow \nu} \end{bmatrix} \quad (3.10.24)$$

$$+ \frac{1}{4} \begin{bmatrix} 0 \\ \left[ \delta\left(\eta + \frac{1}{2}\right) + \delta\left(\eta - \frac{1}{2}\right) \right] \left[ (1+i)\mathcal{F}\{RR_{11}\}_{t \rightarrow \nu} - i \cos^2 \frac{\delta_4}{2} \cos^2 \frac{\delta_3}{2} \delta(\nu) \right] \\ \left[ \delta\left(\eta + \frac{1}{2}\right) + \delta\left(\eta - \frac{1}{2}\right) \right] \left[ (1-i)\mathcal{F}\{RR_{12}\}_{t \rightarrow \nu} - \cos^2 \frac{\delta_4}{2} \cos^2 \frac{\delta_3}{2} \delta(\nu) \right] \\ \left[ \delta\left(\eta + \frac{1}{2}\right) + \delta\left(\eta - \frac{1}{2}\right) \right] (1-i)\mathcal{F}\{RR_{13}\}_{t \rightarrow \nu} \end{bmatrix} \quad (3.10.25)$$

We can then proceed to directly compute the convolutions, and then optimize over the parameters, but this may become extremely tedious and prone to mistakes. Instead, I have exploited some of the symmetries of the convolution process and the resultant delta functions and have written software which can optimize over the delta function combinations and their parameters, with the convolution step taken *after* a parameter evaluation step.

#### 3.10.4 Discussion of parameters

By examining the equations in the Fourier domain derived above, we can see that there are sets of delta functions, which are then convolved with one another to obtain the final channel structure for the HyDMIP-P polarimeter. The effect of the offsets of the waveplates with respect to the reference frame,  $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$  is to scale the delta functions between the real and imaginary domains in Fourier space as shown in Figure 3.6. The effect of  $C_1, C_2, C_3, C_4$  is to move the delta functions "up and down" perpendicular to the  $\xi - \eta$  plane in Fourier space as shown in Figure 3.7. The location on the  $\xi - \eta$  plane is fixed by the layout of the micropolarizer array.

#### 3.11 HyDMIP Results

After the free parameters for a quad-retarder + micropolarizer array system were determined, a cost function was specified and optimized. As specified in the previous section, each linear retarder consists of 3 free parameters to optimize over; retardance,  $\delta_j$ ; starting position or offset,  $\epsilon_j$ ; and frequency or speed,  $\nu_j$ . Hence, the total number of free parameters for the system comes to 12;  $\mathbf{p} = (\delta_1, \dots, \delta_4, \epsilon_1, \dots, \epsilon_4, \nu_1, \dots, \nu_4)$ . The starting offsets,  $\epsilon_j$ s, however, do not have an effect on the ideal system bandwidth performance for a balanced system, reducing the number of effective free parameters to only 8. A *balanced system* is a system with equal bandwidth for all channels. It appears that the offsets can be used to improve bandwidth in unbalanced systems. Additionally in the chapter on noise and systematic error the  $\epsilon_j$ s can be used to re-optimize the channeled system for the case of non-ideal retardances, and are useful for reducing systematic error.

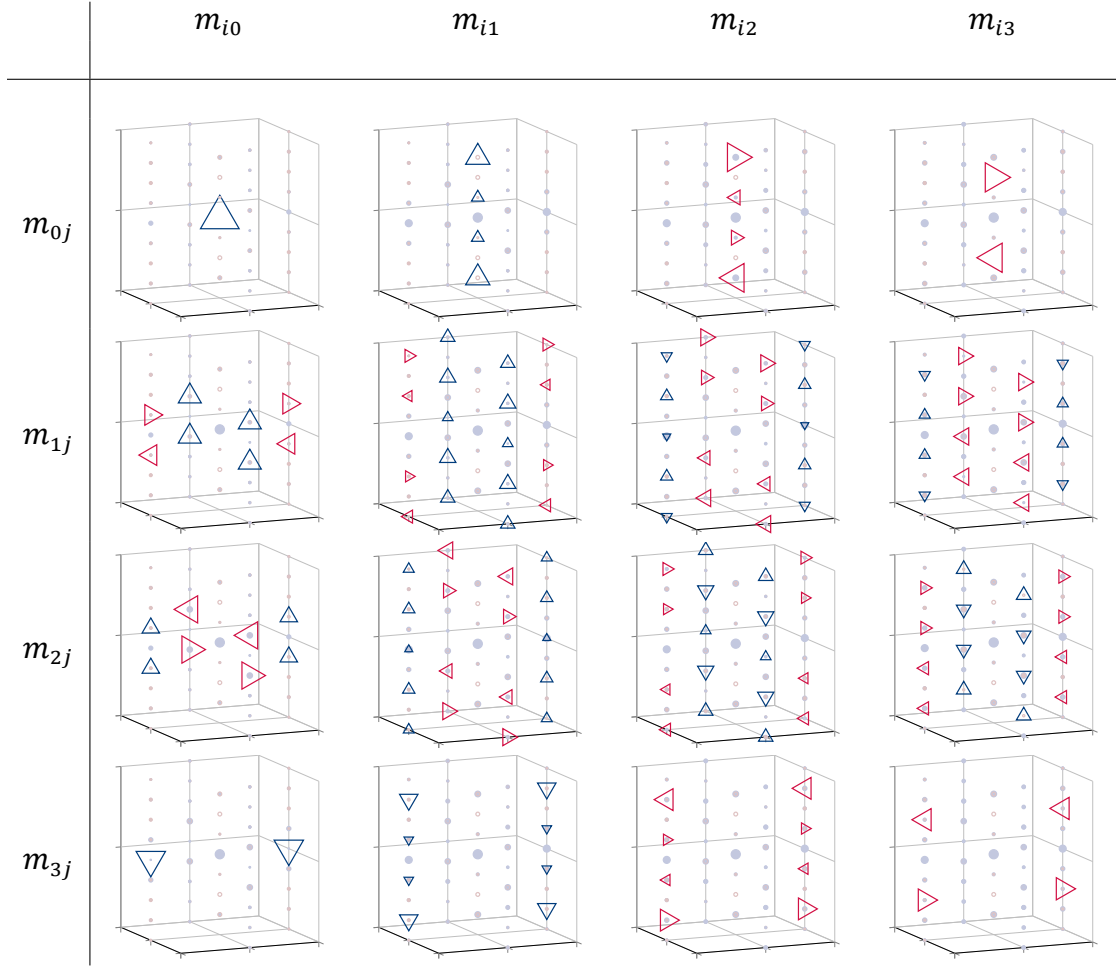


Figure 3.10: Optimal channel structure for a specific quad retarder + micropolarizer system. The system is balanced (equal channel bandwidth).

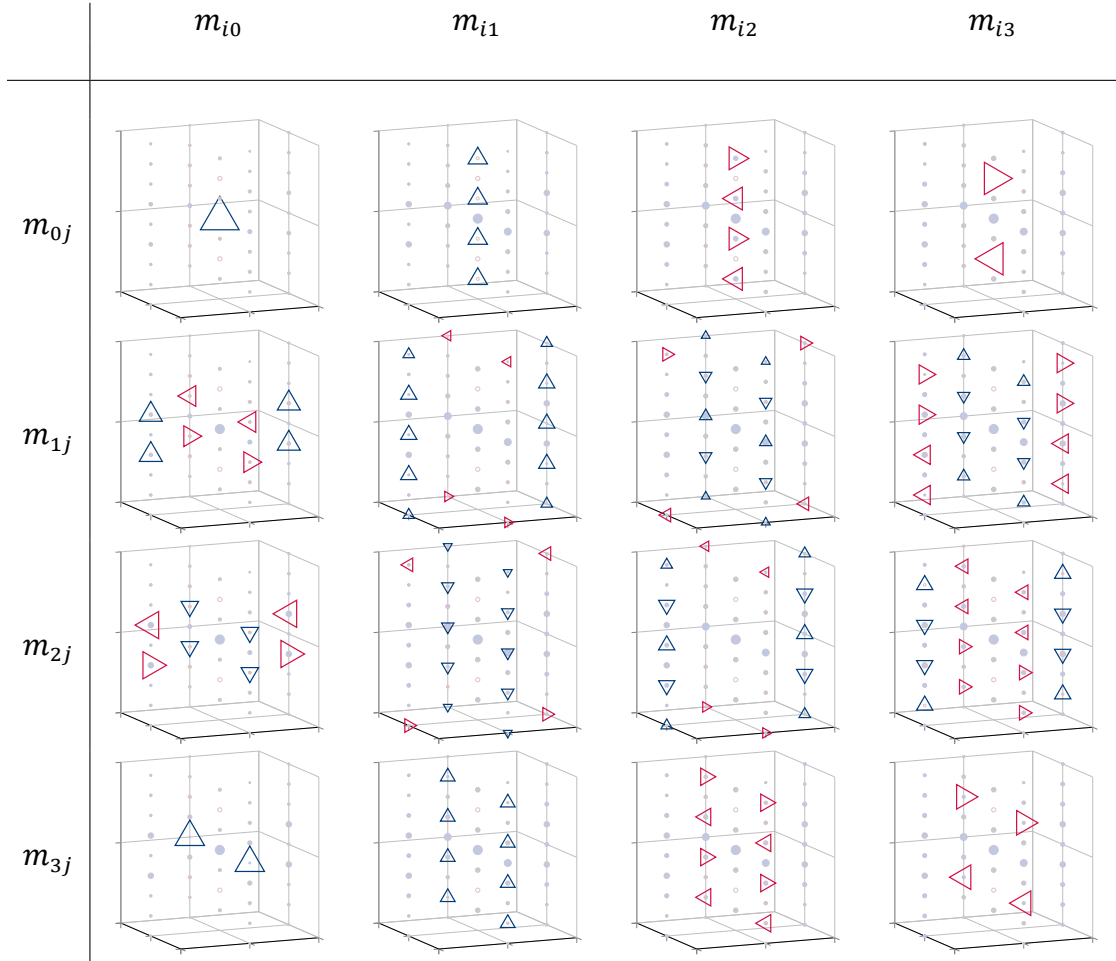


Figure 3.11: Close to optimal channel structure for a specific quad retarder + micropolarizer system. The system is balanced (equal channel bandwidth). The axes are identical to those in Figure 3.1.  $\delta_2 = \pi$  here instead of the optimized  $\delta_2 = \pi - \cos^{-1} \frac{1}{\sqrt{3}}$ , this retardance is what was available to test these types of systems in our physical instrument.

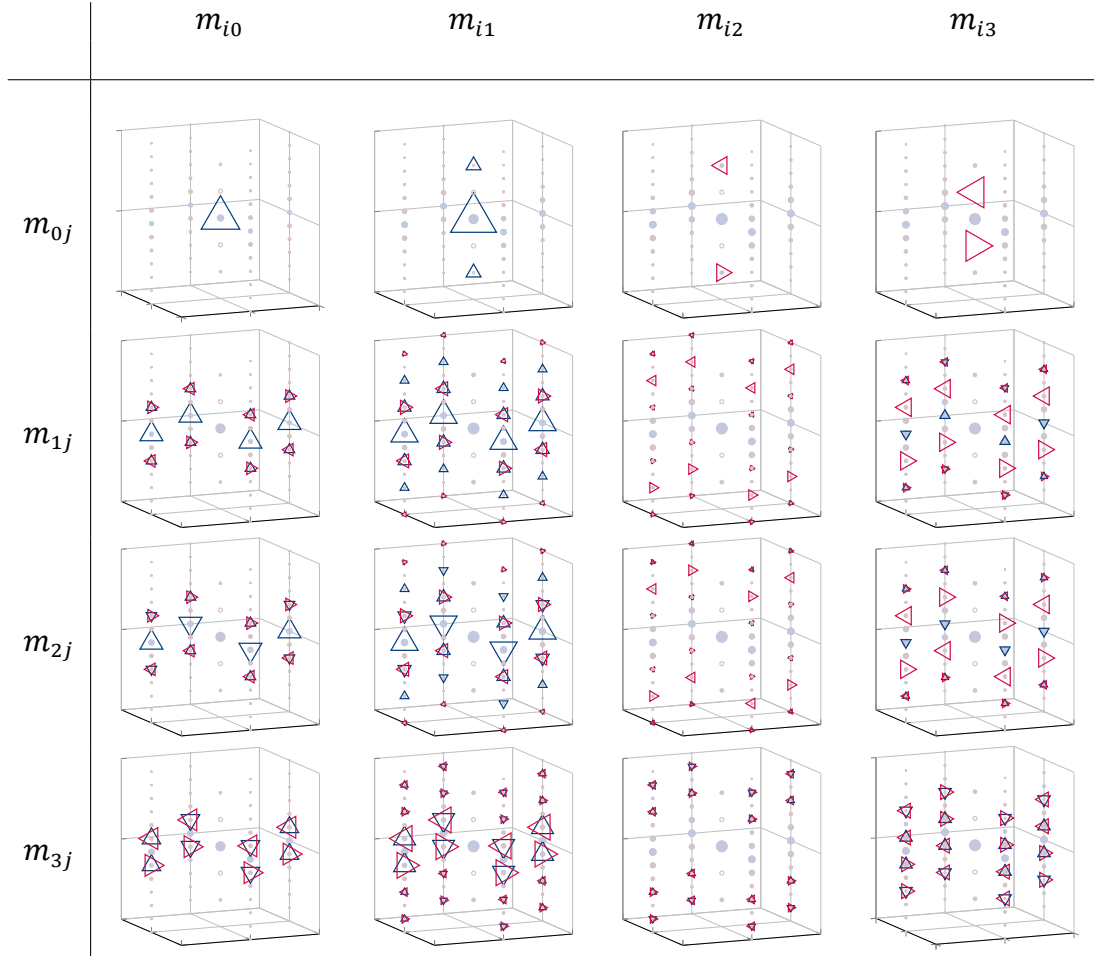


Figure 3.12: Optimized channel structure for a specific dual retarder + micropolarizer system. The system is balanced (equal channel bandwidth). The axes are identical to those in Figure 3.1.



retarder number, $j$	retardance, $\delta_j$	offset, $\epsilon_j$	frequency magnitude, $C_j$
1	$\pi$	0	1/2
2	$\pi - \arccos\left[\frac{1}{\sqrt{3}}\right]$	0	-1
3	$\pi - \arccos\left[\frac{1}{\sqrt{3}}\right]$	0	1
4	$\pi$	0	1/2

Table 3.2: Specifications for optimized quad retarder + micropolarizer system.

The cost function for the results was defined as

$$\mathcal{O}(\mathbf{p}) = \frac{\text{dist}(\mathbf{p})}{[CN(\mathbf{p})]^{\frac{1}{2}}} \quad (3.11.1)$$

where  $CN(\mathbf{p})$  is the condition number of the resulting  $\mathbf{Q}$  matrix [34] and  $\text{dist}(\mathbf{p})$  is balanced temporal channel bandwidth, i.e. the distance between the channels in the  $\nu$  direction is equally spaced. Some unbalanced examples are also presented later in this section. Running the cost function through both simulated annealing and multi-point iterative hill climbing optimizers resulting in the best channel structure shown in Figure 3.10, and an similar structure that we could physically build is shown in Figure 3.11. This channel structure has the specifications shown in Table 3.2. The condition number of  $\mathbf{Q}$  for this system is 2.53.

I also evaluated a dual rotating retarder system over the 6 optimization parameters  $(\delta_1, \delta_2, \epsilon_1, \epsilon_2, \nu_1, \nu_2)$ . The best channel structure found for the dual retarder system is shown in Figure 3.12. This channel structure has the specifications shown in Table 3.3. The condition number of  $\mathbf{Q}$  for this system is 3.4. The quad retarder + micropolarizer array system results in a significant temporal bandwidth improvement over other instruments presented in the current literature. The dual retarder system also realizes some bandwidth gains. The gains are shown in Table 3.4. This specific

retarder number, $j$	retardance, $\delta_j$	offset, $\epsilon_j$	frequency magnitude, $C_j$
1	$2\pi/e$	0	1
2	$2\pi/e$	0	-1/2

Table 3.3: Specifications for optimized dual retarder + micropolarizer system.  $e$  in  $2\pi/e$  here denotes the natural number.

system type	balanced?	temporal bandwidth	sensor resolution
typical dual rotating retarder	yes	1/12	1
quad retarder + micropolarizer array	yes	1/4	1/4
dual retarder + micropolarizer array	yes	1/6	1/4

Table 3.4: Performance gains and tradeoffs for various system types. Note that temporal bandwidth here is given as a fraction of some maximum frequency set to  $1/s$ . When sampling and Nyquist is taken into account, the actual framerate of the instrument will be  $1/2$  of temporal bandwidth listed in the table, multiplied by the base camera framerate. Sensor resolution is the reconstructable resolution of the Mueller matrix images, as a fraction of the base full sensor size.

system type trades spatial bandwidth for temporal bandwidth. These systems only have  $1/4$  of the available sensor resolution in the Mueller matrix image estimates. The quad retarder + micropolarizer array channeled system achieves a 300% temporal bandwidth improvement over a typical dual rotating retarder Mueller matrix system.

### 3.11.1 Other spatio-temporal systems

A few other systems were briefly evaluated, including a full Stokes system. These systems are not necessarily balanced, i.e., they may have unequal bandwidths between channel groupings. One full Stokes system, shown in Figure 3.13 is of particular interest, it is unbalanced, but has the ability to reconstruct the full Stokes parameters at the *full framerate* of the base camera. This optimization utilized retarder offsets to achieve this particular channel structure.

An unbalanced system was also found for the full Mueller case, utilizing the Stokes system described above as the PSA. The system achieves a temporal bandwidth fraction of  $2/7$  and a condition number of 4. The specifications are  $\delta_1 = \pi, \delta_2 = 2\pi/3, \delta_3 = 3\pi/2, \delta_4 = \pi, \epsilon_1 = \epsilon_2 = \epsilon_3 = 0, \epsilon_4 = -0.03125, v_1 = v_0/4, v_2 = v_0, v_3 = v_0/4, v_4 = v_0/8$ . The channels are shown in Figure 3.14. If we change  $v_2$  to  $v_2 = v_0/2$  in this system, then a rank 10 pMMP results with a temporal bandwidth fraction of  $2/3$  and a condition number of 2, for the 10 reconstructable linear combinations of Mueller matrix elements. The channel structure is shown in Figure 3.15

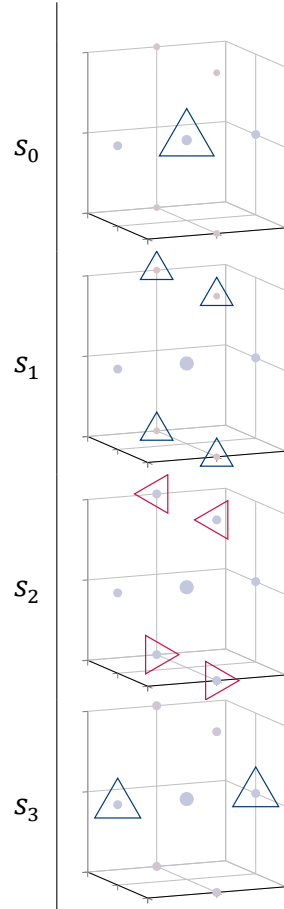


Figure 3.13: Optimized channel structure for a dual retarder + micropolarizer array Stokes system.  $\delta_1 = 3\pi/2$ ,  $\delta_2 = \pi$ ,  $\epsilon_1 = 0$ ,  $\epsilon_2 = -0.03125$ ,  $v_1 = v_0$ ,  $v_2 = v_0/2$ . The axes are identical to those in Figure 3.1.

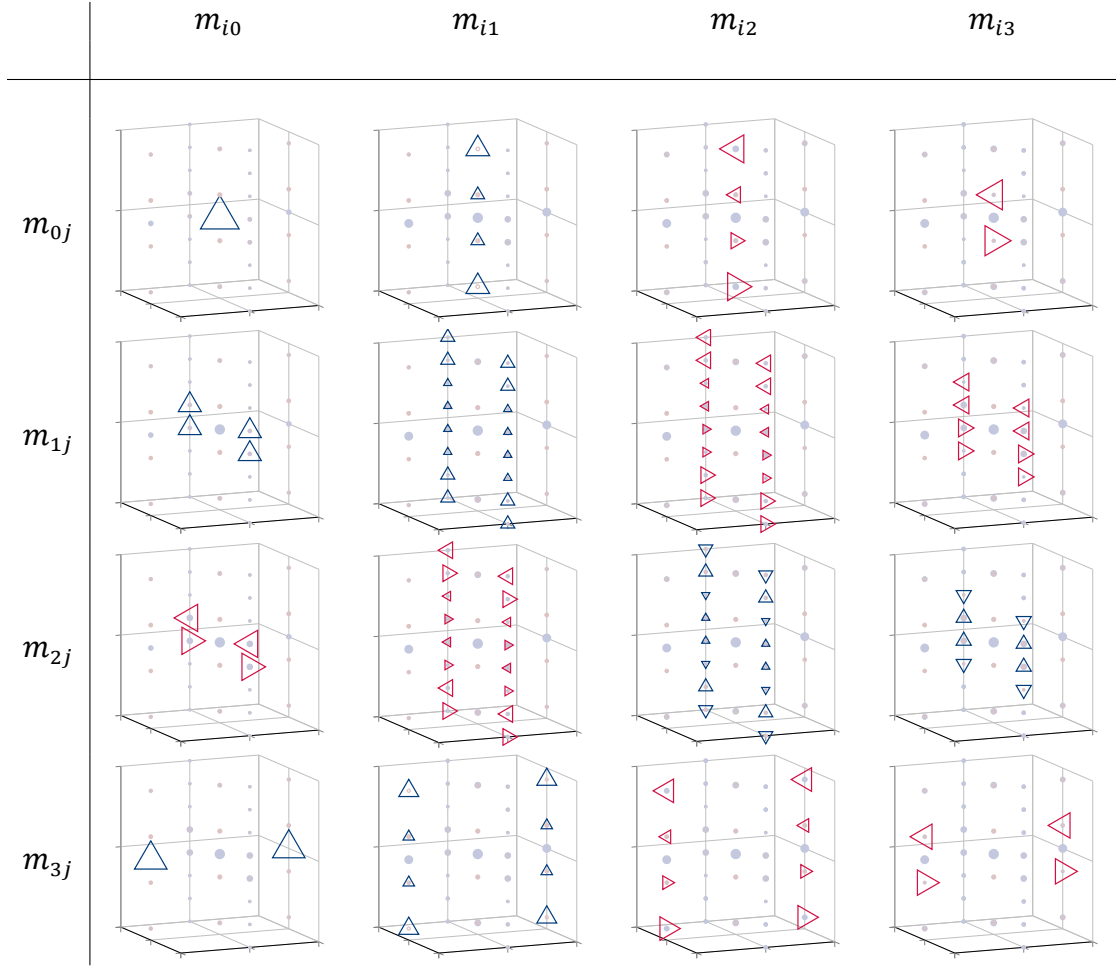


Figure 3.14: Optimized channel structure for a specific unbalanced quad retarder + micropolarizer system. The axes are identical to those in Figure 3.1.

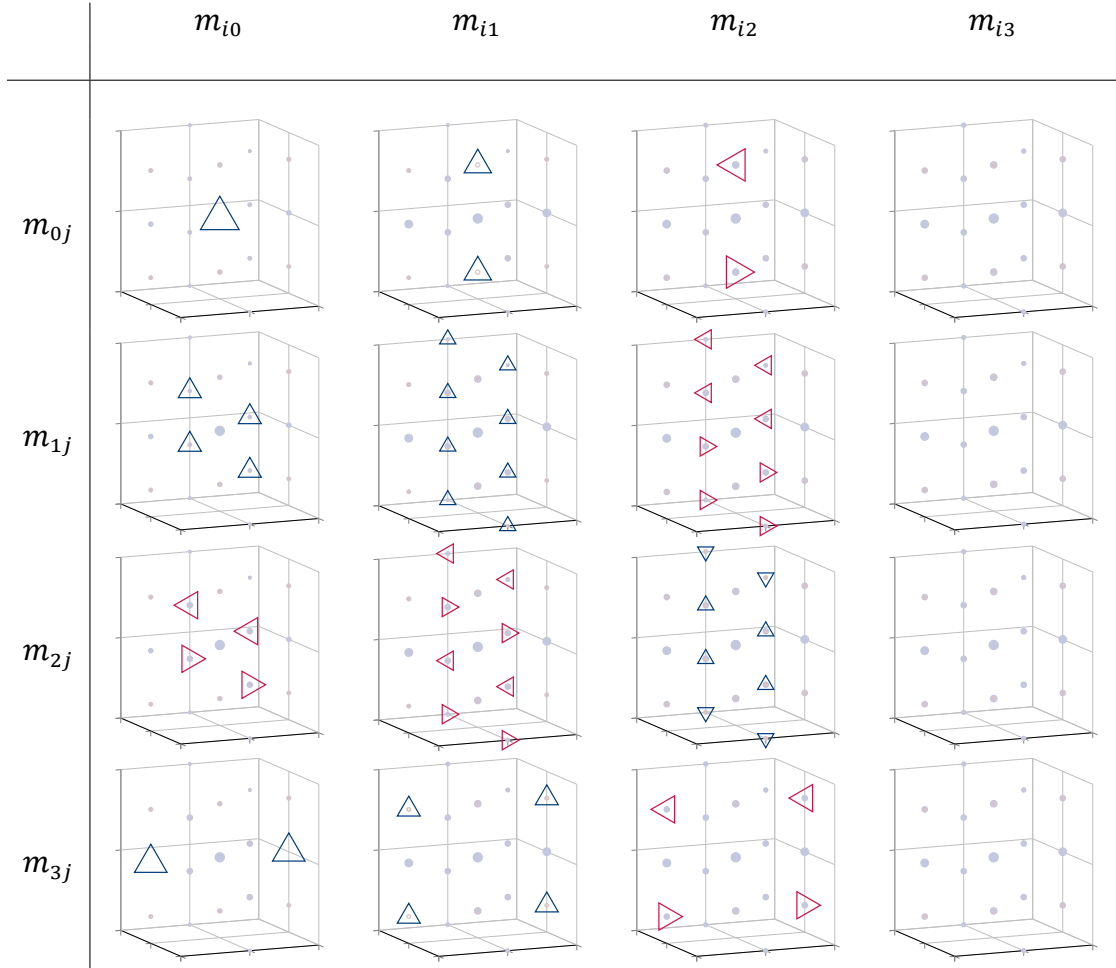


Figure 3.15: Optimized channel structure for a specific unbalanced quad retarder + micropolarizer rank 10 pMMP system. The axes are identical to those in Figure 3.1.

### 3.12 Conclusions

Fast active polarimetric imagers are required for usable, deployable, polarimetric remote sensing applications. There have been a dearth of such imagers presented in the literature to date, for both theoretical and actual system designs. I have shown a *systematic* approach for bandwidth optimization using polarimetric linear systems theory [34], and have shown examples of various systems and the bandwidth improvement which can theoretically be realized. What remains is to analyze system sensitivity to deviations for the ideal theoretical systems laid out in this chapter, and to verify that the system descriptions as presented here are correct. In the next chapter, noise and systematic error are addressed, and in Chapter 5 an actual instrument and results from that instrument are presented.

The theoretical results so far imply that an active Mueller matrix imaging system may be built with as much as a 343% improvement over the typical dual rotating retarder system as shown in Figure 3.14, however even better results may be obtained in the future. Various pMMP and Stokes polarimetric imaging schemes also benefit from the bandwidth analysis shown here, with a full Stokes polarimeter at the full camera framerate presented in Figure 3.13. I have shown that for a base camera framerate of  $30fps$ , it is possible to

- Obtain full Mueller matrix images at  $3.75fps$ , and perhaps even at  $4.3fps$  if the system shown in Figure 3.14 can be built,
- Obtain full Stokes parameter images at  $30fps$ ,
- Obtain specific pMMP elements (linear combinations of Mueller elements) at higher framerates, for the pMMP system shown in Figure 3.15, the 10 elements could be imaged at  $10fps$ .

The design paradigm and bandwidth optimization derived in this chapter is general, and can be applied to other spatio-temporal systems, and with model changes, can be applied to arbitrary systems. Future work includes providing an arbitrary bandwidth optimization framework, which only depends on *the modulation domain* and *the physical constraints* of a system.

## CHAPTER 4

### Noise and Systematic Error in Channeled Polarimeters

#### 4.1 Introduction

The bandwidth optimized channeled systems shown in Chapter 3 resulted from optimizations given an ideal model. Real systems differ from the ideal systems; the primary factors of deviation from the ideal systems are random noise, systematic deviation (tolerances), and other systematic effects like lens aberration and interference effects. Noise and systematic deviations affect channeled systems and the available bandwidth of those systems. The performance of real systems must be characterized in order to build robust active polarimetric imagers.

Noise and systematic error in traditional polarimetric systems is well studied in the literature [14, 63, 103–121], however not much has published on the effects of noise and systematic error in channeled systems [34, 45, 122, 123]. Noise and systematic error can significantly affect polarimetric systems, and must be addressed in the channeled systems paradigm in order for channeled systems design methodologies to be adopted by practitioners in the field.

In this chapter noise and systematic error in channeled systems are discussed. The impact of noise on bandwidth is also presented, and some mitigation strategies are reported.

#### 4.2 Noise and Error

Two primary phenomena contribute to invalid estimation and reconstruction from polarimetric data: random noise and systematic error. Random noise is something probabilistic, is unpredictable, and tends to change from measurement to measurement. Systematic errors are consistent, yet unknown, deviations of a physical instrument from what a user expects, or from some instrument model. These two error

sources drive calibration methodologies and impact reconstruction/estimation accuracies, and calibration can sometimes mitigate systematic errors.

#### 4.2.1 Systematic Error

Systematic error must always be addressed in polarimetric instruments, otherwise consistently inaccurate measurements result. Typical calibration methods require measuring known samples, and subsequently solving a matrix equation (or an eigenvalue problem) to derive a *system measurement matrix* which quantifies the systematic errors [46, 60, 71, 124, 125]. These same methods are relevant to channeled polarimeters, and provide system matrices which can be used for calibration. The way the calibration is used for channeled systems, however, differs from the standard approach.

Sources of systematic error include deviations in retardances of waveplates, non-normal raypaths through crystalline retarders, polarizers which are only partial diattenuators, geometric projections of certain diattenuators (i.e. either tilt in the element or non-normal ray angles), dispersion for polarization gratings, etc. Mechanical sources of error can include positioning bias; total error motion, axial error motion, and tilt error motion for mechanical rotation stages; wedge angle in optical elements (non-parallel entrance and exit faces); micro-element positioning error (e.g. micropolarizer angle errors); etc. These systematic errors tend to introduce repeatable *bias* into the measurements which, with proper analysis and measurements, can be removed.

Calibration is the process of taking experimental measurements which can be used to deduce the systematic errors. The measured errors are really differences between the model (and the inverse) used for the system and the physical system, the measured errors can be used to update the system model to better reflect reality. This process can be done carefully once, or iteratively over time, or a combination of the two. Typical calibration for polarimetric systems is carried out under the assumption that the measured objects have no bandwidth (often for temporally channeled systems) or that the measured objects are band-limited (for spatially or spectrally channeled systems). In the channeled systems approach, this assumption must be lifted.



#### 4.2.2 Random Noise

Random noise is the result of some process which 1) the operator cannot control and 2) is not predictable but can be modeled by as some random variable. All measurement systems will have random noise, and the impact of random noise can be minimized via system design. The electrical engineering literature is quite robust in the general description of random noise and mitigation strategies.

For polarimetric imaging systems, random noise sources include sensor noise (both Johnson-Nyquist noise and shot noise), random (uncontrollable and unpredictable) mechanical and opto-mechanical errors, illumination phenomena (e.g. laser speckle), electrical voltage noise on inputs for voltage controlled polarization optical carrier generation elements, etc. Random noise is usually modeled as some sampled random variable pulled from an adequate distribution or statistical description.

Many commercially available pixelated sensors used to detect light in the visible range (the polarimetric instrument designed in this dissertation operates in the visible wavelength range) consist of either a charge coupled device (CCD) technology or complementary metal-oxide-semiconductor (CMOS) technology. Historically performance has been better for CCD, but recently, scientific CMOS (sCMOS) has begun to overtake CCD devices in terms of performance [126], which includes noise, speed, and dynamic range. Typical random noise sources for digital optical sensors include inherent quantum effects (Poisson distributed noise), thermal noise (Gaussian distributed noise), readout circuitry effects (distributions vary, often Gaussian), and  $1/f$  noise, which has not yet been fully explained [88, 126, 127]. Fixed pattern “noise”, or spatially varying defects and other effects, are often addressed as noise, but these effects are systematic errors that should be addressed as part of the system model. Most imaging camera devices are operated at frequencies high enough that  $1/f$  noise becomes negligible for most imaging purposes [88], so  $1/f$  noise will not be included here. The primary noise sources are then quantum (photon) noise, thermal noise, and readout noise. An additional factor in the noise analysis is the efficiency of the detector to convert photons to detectable signals. In total, the signal

to noise ratio ( $SNR$ ) equation for sensors is typically written as [126]

$$SNR^2 = \frac{\mu^2}{\sigma_d^2 + \sigma_r^2 + \sigma_n^2} \quad (4.2.1)$$

where  $\mu$  is the mean signal,  $\sigma_d^2$  is the dark or thermal noise variance,  $\sigma_r^2$  is the readout noise variance, and  $\sigma_n^2$  is the shot or quantum noise variance. Thermal noise typically has a normal/Gaussian distribution, shot noise typically has a Poisson distribution, and the distribution for readout noise is sensor dependent. For CCD sensors, readout noise is typically Gaussian, but for sCMOS sensors it often non-Gaussian. Depending on the author, occasionally  $SNR$  is substituted for  $SNR^2$  in Equation (4.2.1).

When a CCD or sCMOS sensor is operating in a moderate, but unsaturated, irradiance range, typically the dominant noise component is the shot noise, with variance  $\sigma_n^2$ . Occasionally, the readout noise can also have noticeable effects in moderate irradiance conditions, but usually only impacts low light imaging. Thermal noise only has significant effects when operating in low light, low irradiance conditions.

#### 4.3 Random noise in channeled systems

Random noise in channeled polarimetric systems should be addressed in the channel, or Fourier, domain. Measurements taken in the modulation domain can have noise sufficiently modeled by techniques in the literature and from the general field of electrical engineering [14, 63, 103–121]. The question then arises: what happens to the noise when transformed into the Fourier domain?

Noise in the modulation domain can be modeled via the typical imaging equation

$$\mathbf{g} = \mathcal{H}\mathbf{f}(\mathbf{x}) + \mathbf{n} \quad (4.3.1)$$

where  $\mathbf{g}$  is a finite vector of length  $j$ ,  $\mathcal{H}$  is some general imaging operator,  $\mathbf{f}(\mathbf{x})$  is some object as function of the domain  $\mathbf{x}$ , and  $\mathbf{n}$  represents the noise, also of length  $j$  [88]. Generally, the noise is not necessarily additive even though the above equation implies additive noise. For channeled polarimetric systems, we have

$$\mathcal{F}\{\mathbf{g}\} = \mathcal{F}\{\mathcal{H}\mathbf{f}(\mathbf{x}) + \mathbf{n}\} \quad (4.3.2)$$

$$= \mathcal{F}\{\mathcal{H}\mathbf{f}(\mathbf{x})\} + \mathcal{F}\{\mathbf{n}\} \text{ in the case the noise is additive} \quad (4.3.3)$$

In real systems, we have discrete measurements, and use discrete Fourier transforms (DFTs), typically implemented as fast Fourier transforms (FFTs). Fortunately, Schoukens and Renneboog have already worked out this case for DFTs and additive noise [128]. The result is quite important for channeled polarimetric systems, *no matter the noise distribution in the modulation domain, the noise is approximately normal or Gaussian in the Fourier domain*. Additionally, the covariance matrix of the noise among the Fourier coefficients is approximately diagonal even for colored noise in the modulation domain. There are conditions which must be met: 1) the number of samples must be “large enough” and 2) the noise in the modulation domain must be *statistically stationary*. For a 1-dimensional noise vector, the Fourier coefficients are given by [128]

$$a_j = \sum_{k=0}^{N-1} n_k e^{-2\pi i k j / N}, \quad 0 \leq j \leq N-1 \quad (4.3.4)$$

where  $n_k$  is the  $k$ th element of the noise vector  $\mathbf{n}$ . Note that  $n_k$  can be represented as a random variable. Since  $n_k$  is a random variable in Equation (4.3.4), by the central limit theorem the  $a_j$ s will approach a normal/Gaussian distribution as  $N$  increases [128]. The only issue is that the distribution parameters may differ between the real and imaginary parts of the distribution of the  $a_j$ s.

Schoukens and Renneboog also show that the correlation matrix between Fourier coefficients is usually approximately diagonal, which implies independence of the Fourier coefficients of the noise. The conditions are defined by [128]

$$|\rho_{ij}^{RR}| < \frac{4p}{N(1-p^2) - 4p} \quad (4.3.5)$$

$$|\rho_{ij}^{II}| < \frac{4p^2}{N(1-p^2)} \quad (4.3.6)$$

$$|\rho_{ij}^{RI}| < \frac{2p}{N(1-p)} \quad (4.3.7)$$

where  $\rho_{ij}^{RR}$  is the real-real expectation part of the  $ij$ th element of the correlation matrix,  $\rho_{ij}^{II}$  is the imaginary-imaginary expectation part of the  $ij$ th element of the correlation matrix, and  $\rho_{ij}^{RI}$  is the real-imaginary part of the  $ij$ th element of the correlation matrix, and  $i \neq j$ . Schoukens and Renneboog define the correlation matrix in the

standard way:

$$R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}} \quad (4.3.8)$$

where  $R_{ij}$  is an element of the correlation matrix and  $C_{ij}$  is an element of the standard covariance matrix [128]. The correlation matrix here is taken amongst the  $a_j$ s, the coefficients of the DFT of the noise.  $p$  in the above equations is a specific quantity related to the color of the noise in  $\mathbf{n}$  [128].  $N$  can be adjusted, provided the  $p$  for a specific noise type is known, until all of the off diagonal  $\rho_{ij}$ s are “small enough.”

Schoukens’ and Renneboog’s analysis only addresses the 1-dimensional DFT case however, for spatio-temporally channeled systems, we really have the 3-dimensional imaging equation

$$\mathbf{G} = \mathcal{H}\mathbf{f}(\mathbf{x}) + \mathbf{N} \quad (4.3.9)$$

where  $\mathbf{G}, \mathbf{N}$  are 3-dimensional arrays with elements  $g_{k_1 k_2 k_3}, n_{k_1 k_2 k_3}$  respectively. The 3-dimensional DFT is then used resulting in :

$$a_{j_1 j_2 j_3} = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \sum_{k_3=0}^{N_3-1} n_{k_1 k_2 k_3} e^{-2\pi i k_1 j_1 / N_1} e^{-2\pi i k_2 j_2 / N_2} e^{-2\pi i k_3 j_3 / N_3} \quad (4.3.10)$$

It is clear that the same central limit theorem argument applies here, except that the multidimensional central limit theorem is used, and the result is approximately the multi-dimensional normal distribution. What is not entirely clear is the behavior of the covariance matrix for the multidimensional case. It should be noted that noise amplification from the inversion of the  $\mathbf{Q}$  matrix has been addressed by Alenin and Tyo [34].

#### 4.3.1 Sensor noise

An analytical result for the 3-dimensional case will not be worked out here, but some simulations of the noise using typical distributions for CCD sensors will be presented. The typical unit used for noise descriptions for commercially available sensors is the photoelectron,  $e^-$ . Noise values are often given in root mean square (*RMS*). For example, the camera in a spatio-temporally channeled polarimeter which I built (Chapter 5) has a thermal noise of  $50e^-$  *RMS*. For additive white Gaussian noise, this

corresponds to a variance of  $2500e^{-2}$ . Figure 4.1 shows the covariance matrix between Fourier transformed noises on a 3-dimensional space consisting of an  $8 \times 8$  pixel array and 4 time points. The covariance matrix is the result of 100,000 simulated noise realizations in the space time domain. The maximum magnitude of any off-diagonal covariance matrix entry is 0.01. The read noise and thermal noise distributions were both assumed to have standard deviation values of  $75e^{-}$  and means of  $75e^{-}$ . No incident light was assumed, i.e., Poisson noise was zero to isolate the effects of the Gaussian noise for analysis which is equivalent to subtracting off a deterministic mean illumination. This is not something which would occur in a real sensor since sensor outputs cannot be negative. The maximum magnitude of any off-diagonal covariance matrix entry was 0.011. These specific illumination simulations assumed uniform, non-spatially varying, illumination however. These simulations indicate that the covariance derivation of Schoukens and Renneboog [128] likely holds in the multi-dimensional DFT case for spatially uniform illumination.

In an active polarimeter, we often have a varying illumination pattern over the field of view (FOV). This will cause a spatial noise variation in practice for medium to large illumination, when the dominate noise source is represented as a Poisson distribution. The variance, and the SNR, are then spatially varying parameters, which combines aspects of random noise and systematic error and *breaks the stationarity assumption* of Schoukens and Renneboog. The instrument which I designed (Chapter 5) uses a laser for the source, which has a 2-dimensional spatial Gaussian irradiance profile in the object plane. This implies that the shot noise will have a spatial dependence over the sensor. The noise result for this spatial dependence when transferred to the Fourier domain are shown in Figure 4.4. This figure was created with the same read and thermal noise parameters as in Figure 4.1, but with Poisson distributed noise for a spatial Gaussian profile with a peak value of  $25,000e^{-}$ , shown in Figure 4.2. The space-time domain noise volume was also a  $8 \times 8$  pixel array and 4 time points, and Poisson noise was generated by sampling a Poisson distribution with the correct mean at each spatial point, then subtracting the mean and adding back 4 standard deviations. The last operation was implemented to keep the noise positive. For  $N$  photoelectrons a sample is

$$n_{P,k_1k_2k_3} = P(e^{-a(x_{k_1}^2+y_{k_2}^2)}N) - e^{-a(x_{k_1}^2+y_{k_2}^2)}N + 4\sqrt{e^{-a(x_{k_1}^2+y_{k_2}^2)}N} \quad (4.3.11)$$

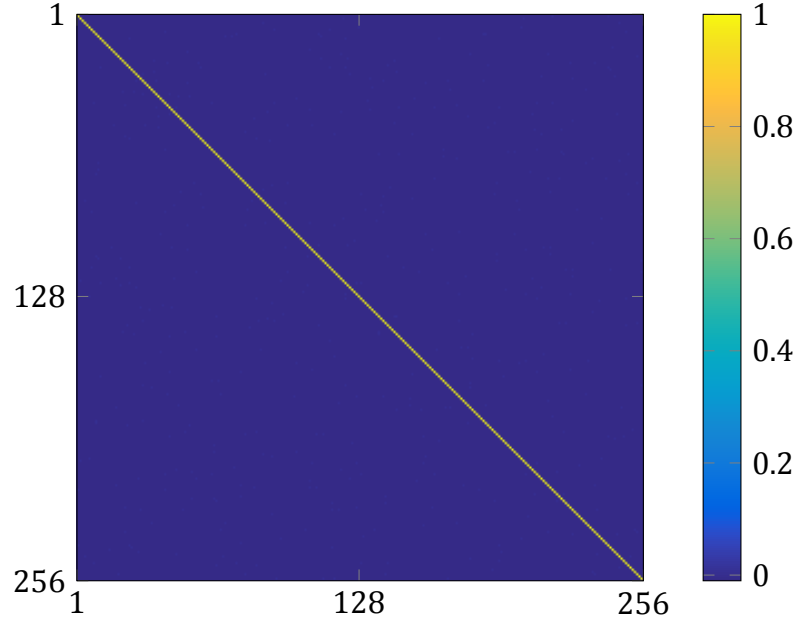


Figure 4.1:  $8 \times 8$  pixel array and 4 time point (256 total measurements) noise covariance matrix for thermal and read noise only (no shot noise). 100,000 realizations were Fourier transformed and the subsequent covariance matrix was computed between all Fourier coefficients. The imaginary part is not shown, since all imaginary parts of the matrix elements were less than 0.01. Both the thermal and read noise were assumed to be sampled from Gaussian distributions with means and standard deviations of  $75e-$ .

where  $n_{P,k_1k_2k_3}$  is the noise sampled from the spatially dependent Poisson distribution  $P(e^{-a(x_{k_1}^2+y_{k_2}^2)}N)$ .  $-1 \leq x_{k_1}, y_{k_2} \leq 1$  are coordinates dependent on the spatial pixel location and  $a$  changes the “spread” of the Gaussian. In the examples here  $a = 4$ . An example of a noise realization is shown in Figure 4.3. This analysis shows that for a spatially varying irradiance profile in the object plane, the sensor noise in the channel space *becomes correlated* to some degree.

#### 4.3.2 Other noise

Noise sources other than sensor noise are also present, including thermal fluctuations in the optical elements, random rotation errors for rotating carrier generation elements, random electrical signal noise for electrically driven carrier generation elements, etc. Generally high frequency thermal fluctuations have negligible impact,

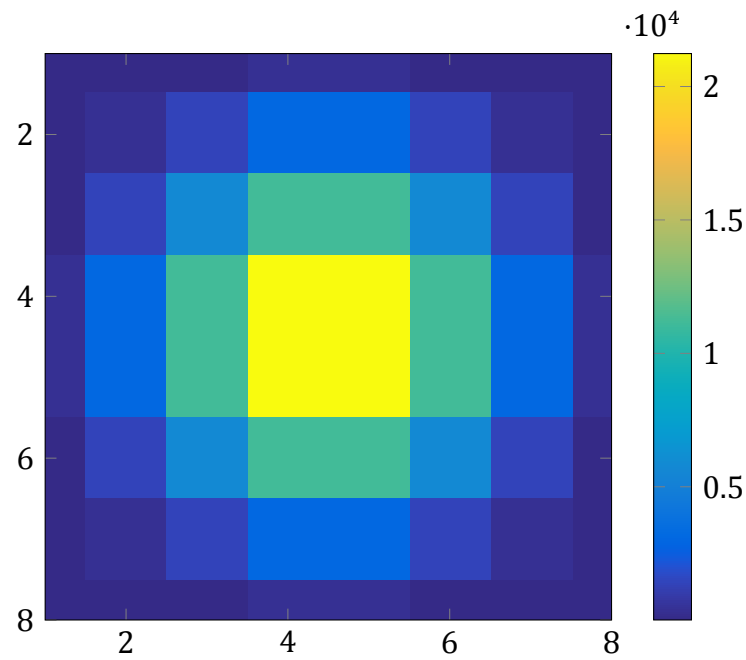


Figure 4.2:  $8 \times 8$  pixel array spatial Gaussian irradiance profile. The peak value is  $25,000e^-$ .

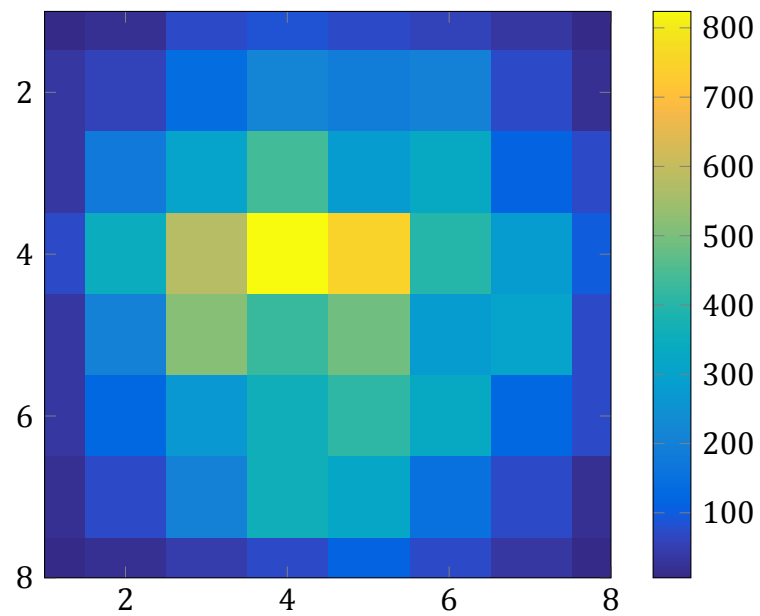


Figure 4.3:  $8 \times 8$  pixel array Poisson noise realization from a spatial Gaussian profile in photoelectrons  $e^-$ .

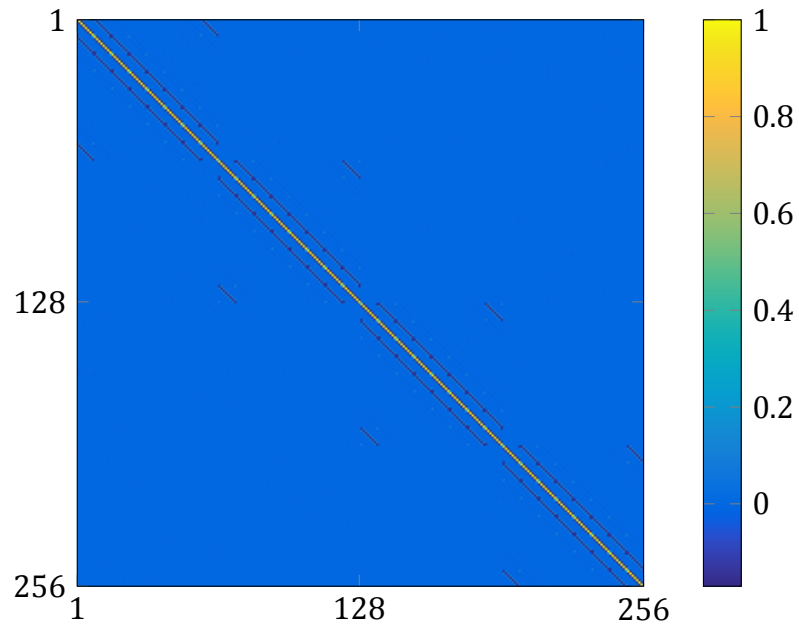


Figure 4.4:  $8 \times 8$  pixel array and 4 time point (256 total measurements) noise covariance matrix for thermal, read noise, and shot noise with a Gaussian irradiance profile illuminating the detector. 100,000 realizations were Fourier transformed and the subsequent covariance matrix was computed between all Fourier coefficients. The imaginary part is not shown, since all imaginary parts of the matrix elements are small. The largest off off-diagonal magnitudes are about 0.16.



and slowly varying thermal effects can be mitigated by systematic error strategies, e.g. measuring a waveplate's retardance over temperature and then compensating a measurement by including a temperature reading at the same time the measurement is acquired. The primary noise sources that a system designer must include are random carrier generation errors, introduced by rotation or electrical noise. Noise also originates from the illumination source for the PSG, but high frequency small magnitude source fluctuations are indistinguishable from sensor noise and can be lumped together with sensor noises, and low frequency variation can be mitigated in a systematic way by either using well controlled sources or measuring the source radiance/irradiance directly and compensating for source variation.

Random carrier noise, together with sensor noise, are often the most important sources of noise for the system designer. Carrier noise has the effect of perturbing the carrier with respect to the ideal measurement or sample point(s) on the Poincaré sphere. This can be conceptualized as introducing small amounts of aperiodicity into the carrier generation scheme or function. This noise does not change the channel locations, but reduces the magnitude of the channels ( $\delta$ -functions) as the noise increases. This effect is shown in Figure 4.5.

#### 4.4 Systematic error in channeled systems

Systematic error in channeled systems has not been developed in depth, and I could only find a single article by Dubreuil *et al* which specifically quantifies the effects of systematic errors on a polarization grating channeled polarimeter [123]. In any channeled system, systematic error effects are limited to

- Moving channel locations,
- Changing channel magnitudes (both/either real or imaginary parts),
- Inducing spectral leakage (since real systems are sampled),
- Adding *unwanted/extraneous channels*,
- Corrupting channel ( $\delta$ -function) shape.

Figure 4.5: Channel effects as rotation wander noise increases for a carrier generation element like a retarder. On the left is the carrier function with noise, and on the right are the channels resulting from the carrier. The noise free carrier resulted in channels with an amplitude of 512. Animated in the electronic version.

Systematic errors which move the channel locations change the available bandwidth. Spurious channels also can change the available bandwidth, and contribute to channel crosstalk. Changes in channel magnitudes result in changes to the conditioning of the  $\mathbf{Q}$  matrix, which impacts noise amplification and reconstructability. Spectral leakage can occur when carrier function positions are systematically different from specifications. System errors can be both static or dynamic.

Retardance deviations can cause extraneous channels for rotating retarders as shown in Figure 4.6. Here, channel cancellation was used to obtain a high bandwidth spatio-temporally channeled Mueller matrix system, and the actual retardances of retarders used to build the system created spurious channels. Fortunately, for bandwidth optimization, the effects of spurious channels can be reduced; details will be shown later in the chapter. Another systematic effect which can have significant impacts on the channel structure is *drift*. Drift occurs when some carrier generation element which was designed to be periodic is periodic at a slightly different frequency than the design frequency, which causes the carrier frequency and the sampling frequency to drift apart from correct phasing. Drift can be caused by control system

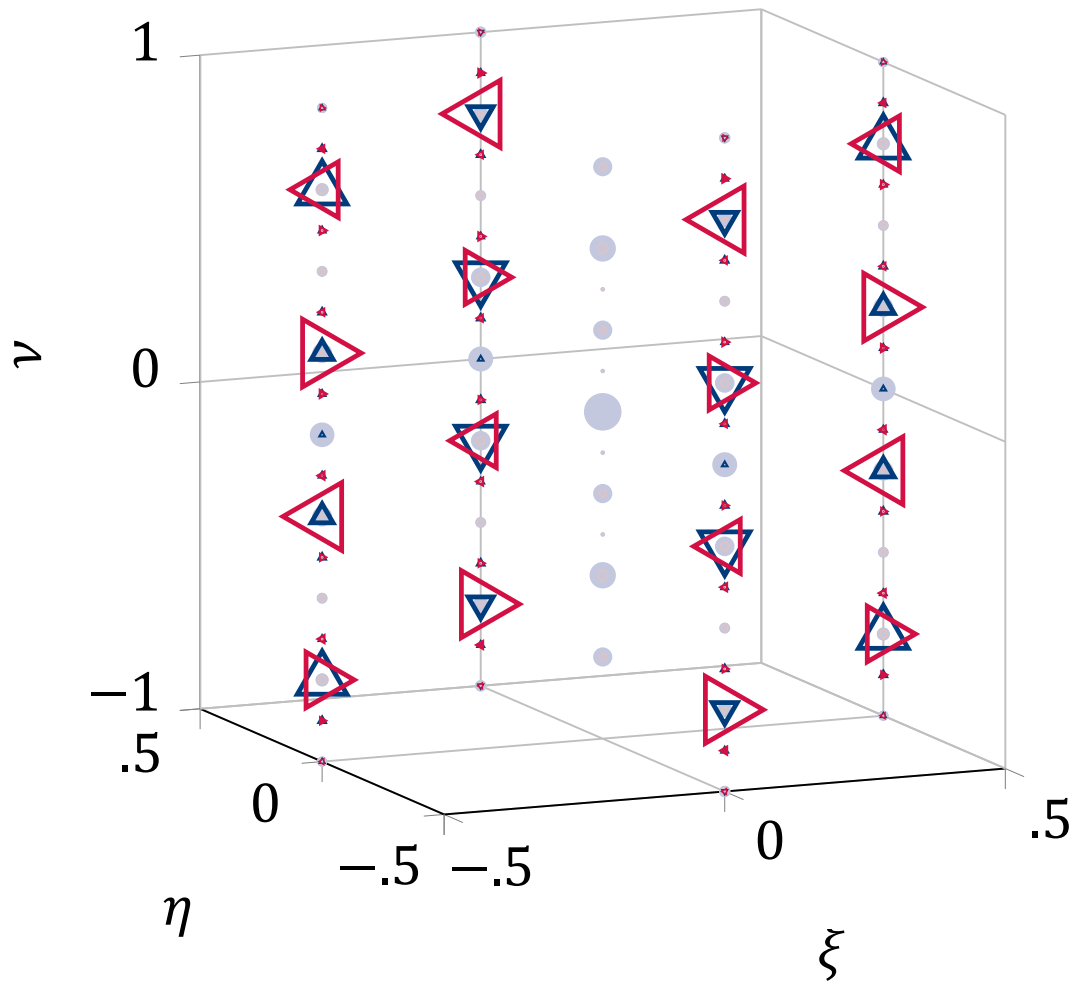
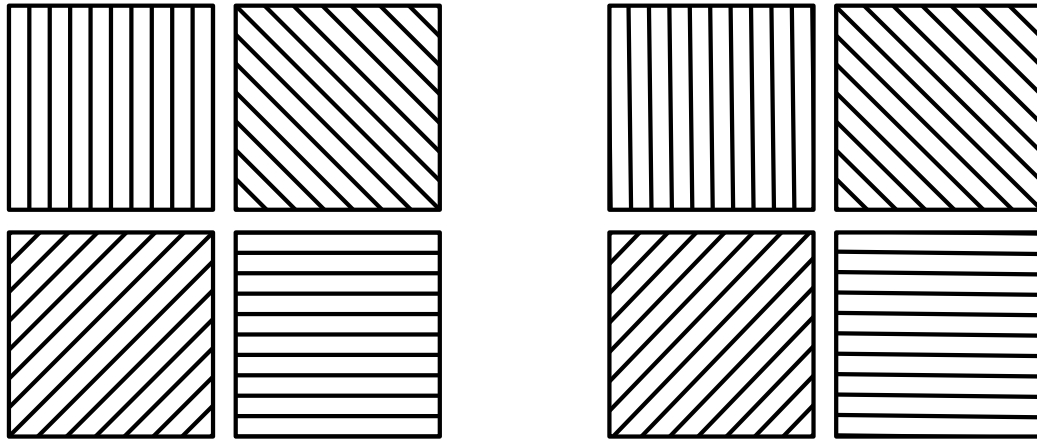


Figure 4.6: An example of systematic error in a real instrument, the small triangles represent extraneous channels which are present due to retardance deviation from the specifications.

Figure 4.7: Channel effects as rotation drift occurs for a carrier generation element like a retarder. On the left is the carrier function with random biased drift applied, and on the right are the channels resulting from the carrier, blue, —, is the real part and red, —, is the imaginary part. Without drift there are two real (blue) channels with no imaginary components. Animated in the electronic version.

bias for rotating elements, or possibly unaccounted for hysteresis in electrical carrier generation elements, among other causes. An example of drift for a cosine carrier is shown in Figure 4.7. Drift can also be caused by a skip or time shift, even when the sampling and carrier frequency have a correct ratio to one another. An example of a skip is a missed sample trigger, which results in an unknown time shift in the carrier function, but the subsequent sampling frequency may remain accurate. Time shifts causing channel structure changes is due to the fact that polarimetric channel structures are *not shift invariant* linear systems. A missed trigger from a real instrument channel structure is shown in Figure 4.8. In a real system, if the number of triggers missed is unknown, then continued measurements are not possible since the channel structure will no longer be known. This type of systematic error can be mitigated by externally tracking the carrier generator position and updating the reconstruction algorithm to handle trigger errors, or by carefully designing a system to have a very low probability of missing triggers.

Figure 4.8: Channel effects of 3 missed trigger events in a physical instrument, likely due to a Microsoft Windows system interrupt. Note the spectral leakage present when the time shift occurs in the DFT window; when the DFT window is before any missed triggers there is little to no spectral leakage, and when the DFT window is after any missed triggers there is also little to no spectral leakage, but the channel structure between the two is inconsistent. Animated in the electronic version.



(a) specified micropolarizer tiling

(b) mean micropolarizer tiling

Figure 4.9: Specified (left) and the mean of the physical (right) micropolarizer array tiling unit cells. Array adapted from a figure by Andrey Alenin.

#### 4.5 Spatio-temporal example

A test and validation system was built as discussed in Chapter 5 and some specific results concerning noise are shown here. The system was designed using channel cancellation/combination to maximize bandwidth, and the system is a separable channel design. For separable channeled systems (see Section 3.2) a potentially serious problem arises; at the locations in the frequency space where the channel cancellation(s) occurred, channels may again be present due to deviations of real components, resulting in a non-ideal physical system. These *extraneous channels* will be convolved with Mueller data, and introduce channel crosstalk.

Systematic errors and noise sources are quantified using a *semi-empirical model*, details are in Chapter 5, and some of the results shown here are from simulations using this model.

##### 4.5.1 Systematic error

Systematic errors in our specific instrument include variations in micropolarizer array orientations from specifications, retardance deviations for waveplates, waveplate wedge, and spatially varying irradiance in the object plane due to laser illumination.

This dissertation will not address the issue of waveplate wedge. The first systematic error that must be managed is the spurious channels which result from the waveplate retardance errors. An example of extraneous channels is shown in Figure 4.6. The small triangles represent the spurious channels.

The second major systematic error is imparted from the manufacturing defects of the micropolarizer array. Defects consist of per pixel orientation error and pixel-to-pixel extinction ratio variation. Each pixel of the micropolarizer array used in the instrument was characterized for extinction ratio and orientation. Details on the specific micropolarizer array used in the instrument are in Chapter 5. Ideal micropolarizer array channels (the Fourier domain structure) were computed using the mean values of the physical array orientations,  $-0.72^\circ$ ,  $+45.33^\circ$ ,  $-43.95^\circ$ ,  $89.34^\circ$  which differs from the original micropolarizer array specifications of  $0^\circ$ ,  $+45^\circ$ ,  $-45^\circ$ ,  $90^\circ$ ; see Figure 4.9 for examples of the micropolarizer array tiling. Graphs of the real parts of the resultant channels for the micropolarizer array by itself are shown in Figures 4.10 and 4.11. Note that because of the magnitude of the  $\delta$ -functions, an inverse hyperbolic sine scaling is used instead of a log scale, because it can handle zero values appropriately [129]. The transformation is

$$f(y, C) = \frac{\log_{10} \left( Cy + \sqrt{C^2 y^2 + 1} \right)}{C}. \quad (4.5.1)$$

The graphs shown in Figures 4.10 and 4.11 use  $C = 1.057$ , which gives approximately log scaling for large values.

In Figure 4.11, the graphs for the ideal and physical cases are shown for the lines along  $\eta = 0$  and  $\xi = 0$ , taken from the 2-dimensional plots shown in Figure 4.10, and they show finer detail since they are not downsampled like the figures in Figure 4.10.

The third primary systematic error is the result of non-uniform illumination, specifically a spatially Gaussian irradiance profile in the object plane. As shown in Figure 4.4 a spatially varying irradiance profile has an effect on the noise in the modulation domain, and it also has a systematic effect on the channel structure. The assumption is made that the illumination at the object plane is purely polarized, which implies that the Stokes parameter vector at each object space pixel differs only by a multiplicative constant, which in our case is a 2-dimensional Gaussian. This implies that the channel spatial shapes will be the Fourier transforms of possibly shifted

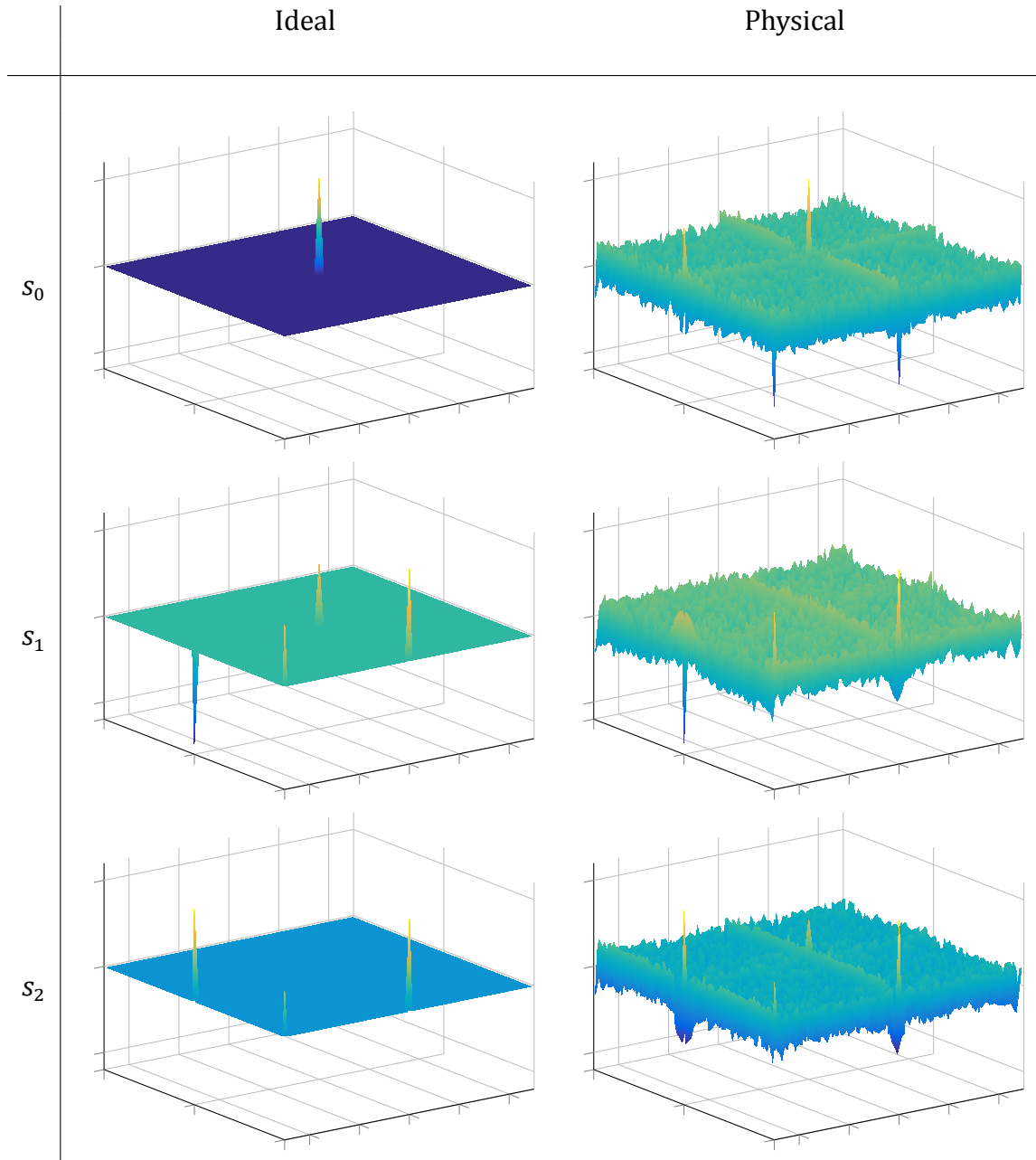


Figure 4.10: Physical micropolarizer array channels vs ideal micropolarizer array channels. Each graph is a 2-dimensional DFT of the physical and ideal micropolarizer arrays. The ideal array parameters used the mean polarizer angle values taken from the physical microgrid. In the figures the axes ranges are  $-1/2 \leq \xi, \eta \leq 1/2$  and  $-6 \leq z \leq 6$ ; the  $\xi$  axis is on the lower right and the  $\eta$  axis is on the lower left. We can observe some extraneous spatial micropolarizer channels which occur at the  $(0, 0)$  and  $(-1/2, -1/2)$  locations due to the mean angles being slightly different from  $0^\circ, +45^\circ, -45^\circ, 90^\circ$ . The vertical  $z$ -axis is inverse hyperbolic sine scaled. The array size was  $980 \times 980$ . Note that the graphs have been downsampled for visualisation purposes, the actual  $\delta$ -functions are much thinner and difficult to view.



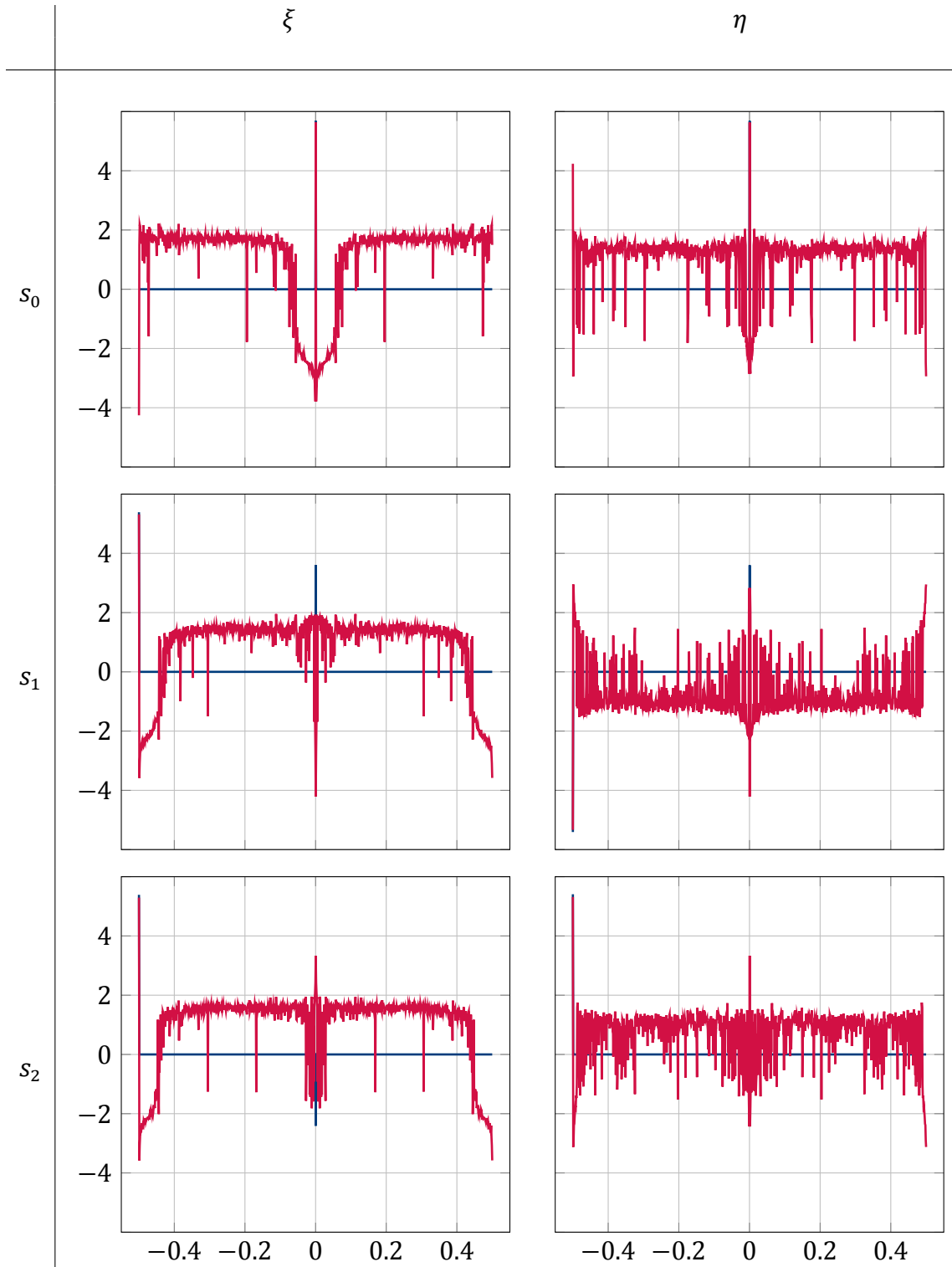


Figure 4.11: Physical micropolarizer array channels vs ideal micropolarizer array channels. Each graph is taken from the 2-dimensional DFT (Figure 4.10) along  $\eta = 0$  and  $\xi = 0$  respectively, since all the relevant channels lie on these lines. The vertical axis is inverse hyperbolic sine scaled. The array size was  $980 \times 980$ . Blue, —, denotes the ideal array and red, —, denotes the physical array.

Figure 4.12: An example of the effect of Gaussian illumination on the channel structure. Left pane shows the real part of the channel structure with Gaussian illumination in the  $\nu = 0$  plane for the system shown in Figure 3.10 and the channels for  $m_{00}$ ,  $m_{11}$ ,  $m_{12}$ , with the magnitude inversely hyperbolic sine scaled. The right pane is the spatial Gaussian irradiance profile at the object plane. Note that a narrow Gaussian profile was used to exaggerate the effect on the channels. Animated in the electronic version, with position of the Gaussian profile varying.

Gaussians. It could be argued that this systematic error is not necessarily an error, as the illumination has a spatial Gaussian profile, and the reconstruction will reflect this fact. For remote sensing purposes, it may be beneficial to remove the Gaussian profile, either prior to reconstruction, e.g., deconvolving the known Gaussian in the channel space, or in post processing, e.g., normalizing the Mueller matrix image. The former would likely be well suited to the use of a 2-dimensional Wiener filter [130] done in the spatial domain (since the Gaussian profile can be viewed as a multiplicative operation in space) prior to Fourier transforming to the Fourier domain. This operation, however, assumes that the Gaussian illumination profile is well characterized. If the Gaussian profile is removed in an infinite SNR (ideal) system case then the channel structure becomes the same as in the uniform illumination case, so the effects of the spatially varying noise must be evaluated first. Figure 4.12 graphically illustrates the channel domain effect of a spatial Gaussian illumination profile.

#### 4.5.2 Noise

For our physical instrument, random errors primarily consist of sensor noise, wave-plate rotation wander, and laser speckle. Speckle will not be discussed, as our instrument has a low enough  $f/\#$  that  $\sim 21$  speckles are incident on each pixel when imaging with the stop(s) open. In the example system discussed here the effect of rotation wander error is the same on the system channels as the simpler channels shown in Figure 4.5, i.e., rotation wander noise decreases the magnitude of the channels. Since rotation wander was adequately addressed in Section 4.3 it will not be discussed further here.

The noise source remaining for analysis is the sensor noise, and a CCD was used in our instrument. As previously stated, the noise sources for a CCD sensor are thermal noise, with a Gaussian/normal distribution, readout noise, also with a Gaussian/normal distribution, and photon/quantum noise, with a Poisson distribution. Simulations will be presented here which show the effects of the various sensor noise types on the channel structure of the system described in Figure 3.10.

Both readout noise and thermal noise are approximately Gaussian in CCD sensors and hence the analysis for both noise types is identical. The Gaussian noise effects on channel structure is shown as a function of SNR in Figure 4.13, for Gaussian additive noise in the modulation domain. The results are simulated, but show that the channel structure itself is quite robust to thermal and readout detector noise. Here we switch to a different view of the channels to better view the noise effects. Each line of channels in the  $\nu$  direction is plotted in Figure 4.13 and Gaussian noise is added in a simulation to the final irradiance for this result. The SNR is shown in the figure, and it appears that the channel structure is stable for an SNR above somewhere between 1 and 2.

The final noise source is the Poisson distributed noise, and the additional effects of the Poisson noise being spatially varying. Spatially varying Poisson noise was generated as a function of irradiance on the detector from the spatial distribution as well as the irradiance fluctuation from the spatio-temporal carriers at simulated time-points. These requirements for the noise modeling (the joint dependence on sensor irradiance and the spatial distribution) make it more difficult to simulate than the readout and thermal noises. Poisson distributions become expensive to compute as

### Gaussian detector noise

Figure 4.13: Gaussian detector noise effects on channel structure. Note that these are channels similar to the channels as shown in Fig. 3.10, but plotted along the  $\nu$  lines where the channels are located. They are not exactly the same channels because this model incorporated many of the systematic errors of the real instrument, resulting in different channel magnitudes. The left column is the real part, the right column is the imaginary part, there are 16 colors in each graph representing each Mueller matrix element. The first row is  $\nu$  at  $\xi = 0, \eta = -0.5$ , the second row is  $\nu$  at  $\xi = -0.5, \eta = 0$  and the third row is  $\nu$  at  $\xi = 0, \eta = 0$ . Animated in the electronic version. The animation is a spectrogram with window length 64 samples.

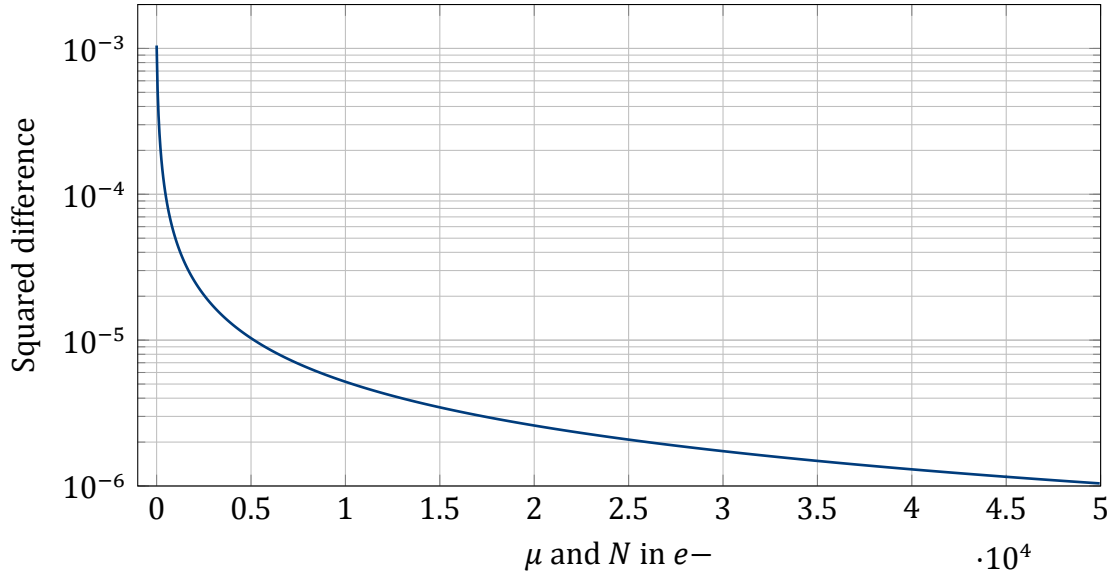


Figure 4.14: Normalized error between a Poisson distribution with mean  $N$ , denoted  $pdf_{P(N)}$ , and the approximate Gaussian distribution with mean  $\mu = N$  and variance  $N$ , denoted  $pdf_{G(N,N)}$ , as a function of the mean in photoelectrons,  $e^-$ .

the mean increases; it is then beneficial to switch to using the Gaussian distribution which approximates a Poisson distribution for large  $N$ . Figure 4.14 shows a graph of the normalized error between the Poisson and approximate Gaussian distributions as a function of the mean  $N$ . The normalized error calculation is

$$E(N) = \sum_{j=1}^{100,000} \frac{[pdf_{P(N)}(j) - pdf_{G(N,N)}(j)]^2}{pdf_{P(N)}^2(j)}, \quad (4.5.2)$$

i.e., the normalized square “area” where  $j$  corresponds to the integer number of photoelectrons on the  $x$  axis for the distributions. My Matlab code for generating Poisson distributed noise switches over to using the corresponding Gaussian distribution when the number of mean photoelectrons is above 1000. Poisson distributed noise for a uniform spatial irradiance was also computed and compared to the spatial Gaussian profile. Figure 4.15 shows the  $m_{00}$  channels for both irradiance profiles, other channels have similar differences. The difference is as expected, a Gaussian multiplied by a sinusoidal function due to the spatial Gaussian irradiance.

The question remains, what effects does the spatially dependent noise have on the channel structure?

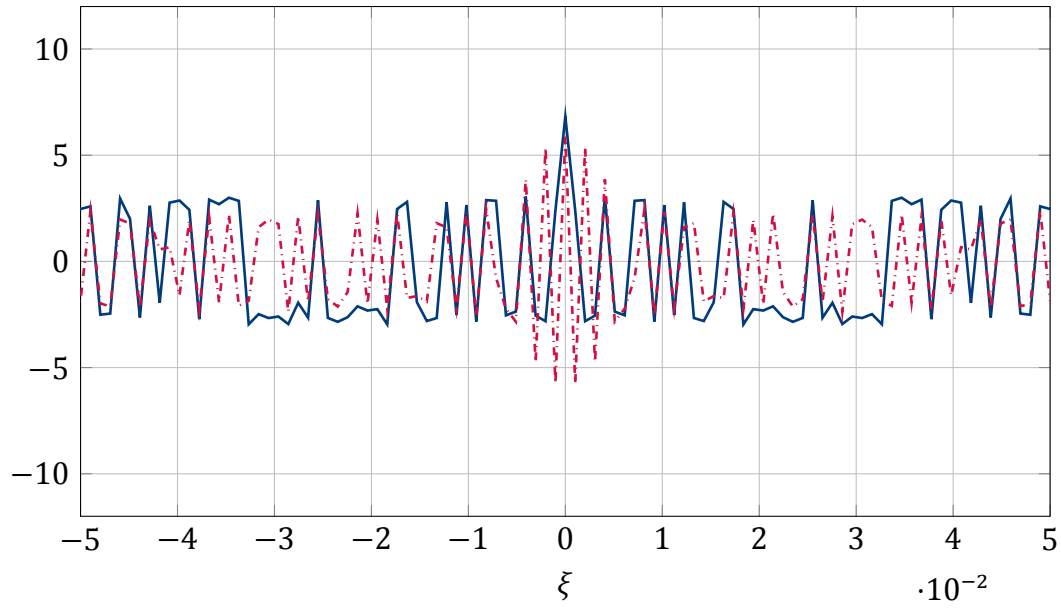
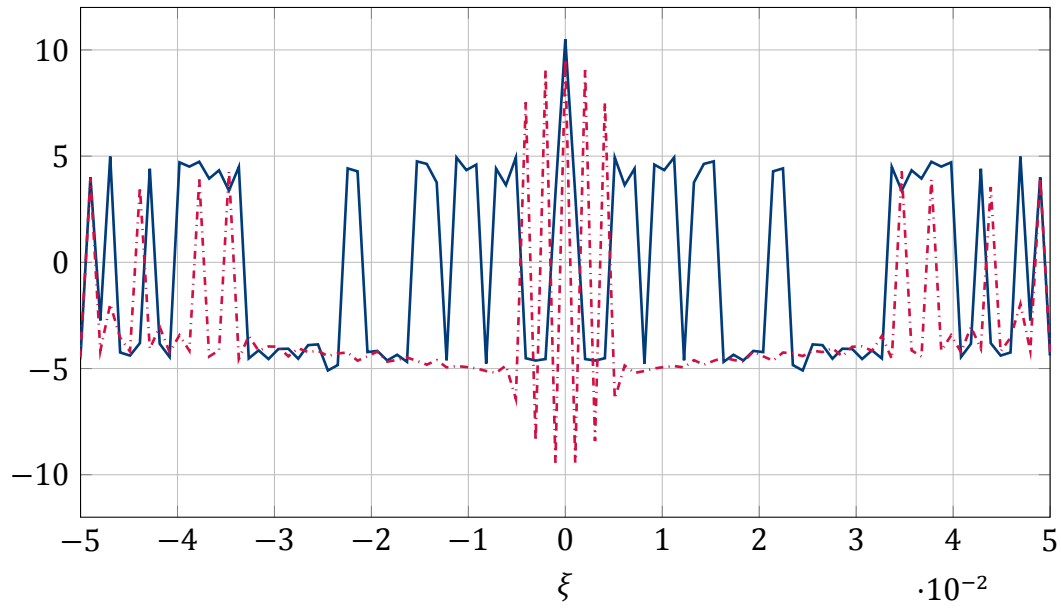
(a)  $6e-$ (b)  $49,006e-$ 

Figure 4.15: Uniform irradiance profile with Poisson noise vs Gaussian irradiance profile with Poisson noise for the  $m_{00}$  channel. The  $y$ -axis is inversely hyperbolic sine scaled. Other channels have similar differences. All curves are taken in the  $\nu = 0$  plane, along  $\xi$ , i.e., the  $\eta = 0$  line. (a) shows the case where the peak pixel irradiance corresponds to  $6e-$  detected per pixel, and (b) shows the case where the peak pixel irradiance corresponds to  $40,006e-$  detected per pixel.

#### 4.6 Optimization

For some types of systematic error, it is possible to mitigate at least some of the impacts on system performance by re-optimizing the channel system. For example, the bandwidth of the system shown in Figure 3.10 does not depend on the initial retarder starting positions *in the ideal case*. This leaves 4 parameters to utilize in the real system for mitigation of systematic errors. Specifically the increased crosstalk from extraneous channels can be minimized to some degree by optimization over the retarder starting positions.

The effects of the extraneous channels for the quad retarder + micropolarizer system can be minimized by re-optimizing over the system channels using the remaining free parameters available while fixing the ones constrained by the physical instrument components. For the quad-retarder + micropolarizer system example (Figure 3.10), once the actual retardances are fixed, the system can be re-optimized using the starting position of each retarder and adding a parameter to the cost function which characterized the magnitudes of the spurious channels compared with the magnitudes of the adjacent channels. For general spatio-temporal systems this kind of method will work after some of the physical components are specified, if there are any free parameters left to optimize over. The secondary optimization function for the quad-retarder + micropolarizer system is

$$\max_{\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4} \min_{c_j, S_{k_j}} \frac{\text{mag}[c_j]}{\text{mag}[S_{k_j}]} \quad (4.6.1)$$

where  $\epsilon_i$  are the offset/starting positions of the retarders,  $c_j$  are the list of specified channels identified by location,  $S_{k_j}$  are lists of each  $c_j$ 's *adjacent extraneous channels* and  $\text{mag}[\cdot]$  is the magnitude of a channel. The adjacent spurious channels, for this specific optimization, are only computed *temporally*, that is each  $c_j$  will have at most 2 adjacent spurious channels. In the general case, there may be spurious channels located arbitrarily (subject to physicality constraints) in the channel domain, and some partition would need to be made to define which spurious channels are “adjacent” and which are not. A simple way to accomplish this would be to define a norm dependent (hyper)-sphere centered on each channel  $c_j$  and define anything inside of the sphere to be adjacent. Implementing this is a subject of future work. For

Figure 4.16: An example of varying the starting position of a retarder,  $\epsilon_2$  to optimize the impact of extraneous channels on the bandwidth optimized system shown in Figure 3.10 for  $m_{22}$ .



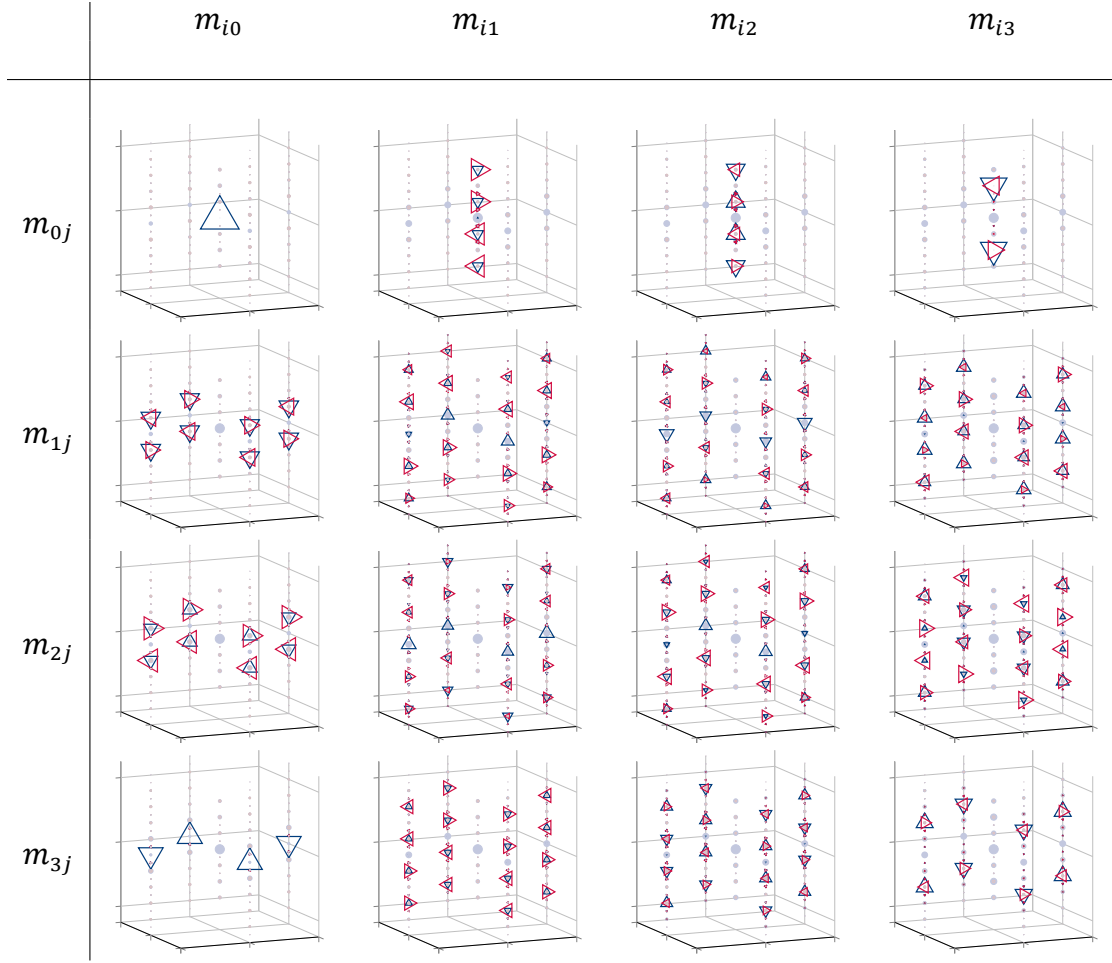


Figure 4.17: Re-optimized channel structure for a specific quad retarder + micropolarizer system with extraneous channels. The axes are identical to those in Figure 3.1.

temporal carrier generation we may use the simpler 1-dimensional adjacency definition. The optimization defined here maximizes the ratio of the magnitude of the channels we want,  $c_j$  to the spurious channels  $S_{k_j}$ . Figure 4.16 shows the effects of varying the starting position of the 2nd retarder,  $\epsilon_2$ , for a bandwidth optimized quad retarder + micropolarizer array system. An optimized system for actual retardances  $\delta_1 = 3.116, \delta_2 = 1.566, \delta_3 = 2.173, \delta_4 = 3.148$  is shown in Figure 4.17.

Other optimizations to mitigate systematic error are possible, for instance in spectropolarimeters [31, 45, 53, 122] it may be possible to re-optimize the retardances after each manufactured retarder is characterized to enforce the specified channel structure, as is done in the modulation domain by Snik *et al* [131] to enforce

a different constraint (system conditioning minimization over a broad spectrum). For the balanced quad retarder + micropolarizer array system shown in figure Figure 3.10 channel cancellation is dependent on the two outside retarders having  $\pi$  retardance, so extraneous channels cannot be mitigated through retardance adjustment of the other waveplates after manufacturing specifications are known. However, it may be possible to adjust the retardances to mitigate systematic error for other systems, like the unbalanced system shown in Figure 3.14.

If the system designer has the time, the best approach is an iterative one with a single component being sourced at a time, the component being measured and characterized, then the system being re-optimized to reset the other components specifications to minimize the systematic errors, with the process being repeated [131] until the system is complete.

#### 4.7 Conclusion

The system designer must account for systematic errors and random noise in polarimetric channeled system designs. Most random noise becomes approximately Gaussian distributed and approximately independent in the channel space. However, spatially dependent illumination creates an exception to noise independence in the channel domain for spatio-temporally channeled systems, and generally non-statistically stationary noise in the modulation domain can create noise correlations in the channel domain.

I have shown examples of these noise profiles and systematic errors for a spatio-temporally channeled polarimeter and how both noise and systematic errors can affect the channel structure. I have also shown how to mitigate some forms of systematic error using free optimization parameters which may remain for the system designer to use for this purpose.

## CHAPTER 5

### Hybrid domain modulated imaging polarimeter (HyDMIP)

#### 5.1 Introduction

Previously, Chapters 3 and 4 introduced channel structures for spatio-temporally modulated polarimeters and analyzed the noise effects in those channeled systems. In this chapter I present an instrument which utilizes these concepts, including the design, the physical instrument, and results from the instrument. Discussions with Andrey Alenin, Charles LaCasse, and Scott Tyo led to the following general design requirements for the portable HyDMIP, HyDMIP-P:

- Ability to generate and analyze arbitrary Stokes parameters, with full coverage on the Poincaré sphere.
- Visible wavelength range, preferably in the 635-830nm wavelength range, due to preliminary skin cancer detection indications from Alenin.
- Have imaging capability.
- Have the ability to test some of the hybrid modulation theory developed by our research group.
- Be portable.
- Acquire data for remote sensing classification tasks.
- Be able to validate classification/discrimination algorithms, including multi-class algorithms.

The first requirement, to access arbitrary states of the Poincaré sphere, lead to initial designs utilizing two retarders in both the PSG and PSA respectively. A half wave linear retarder in combination with another linear retarder allows for the generation of (and analysis of) arbitrary Stokes parameters. This led to the decision to use at most

4 rotating retarders in the system design, which results in temporally modulation. The decision to test hybrid modulation theory developed by LaCasse and Tyo [33,43] determined the requirement to use a micropolarizer array bonded to a silicon detector array. Finally, portability and visible wavelengths were specified due to a number of applications which interested our group at the time, namely skin cancer detection, general outdoor/indoor remote sensing tasks and material discrimination including fruit defect detection, and trained classifier algorithm performance evaluation for polarimetric detection tasks.

The work in this dissertation grew from the need to maximize the temporal bandwidth of the portable instrument, given the other design constraints. An end to end system design and the implementation of that design will be shown here, including mechanical and opto-mechanical design, control system and software implementations, troubleshooting/implementation issues, instrument validation, and results from the instrument. The physical instrument results imply that the underlying theory from Chapters 3 and 4 is correct. The portable instrument is a component of a larger instrument system, the hybrid domain modulated imaging polarimeter (HyDMIP) which consists of a lab instrument and a portable instrument as shown in Figure 3.8. This chapter will only discuss the portable instrument.

## 5.2 Physical design

The instrument design is based on the initial layout composed of a polarizer state generator (PSG) consisting of a linearly polarized source which generates temporal carriers via two linear rotating retarders and a polarization state analyzer (PSA) consisting of two rotating linear retarders followed by a micropolarizer array at the focal plane of the camera. The basic layout is shown in Figure 3.9.

### 5.2.1 Polarization state generator

The PSG is a typical rotating retarder design [46], with the polarization components placed in a collimated optical space, the only difference is that we use two rotating retarders instead of one. Figure 5.3 shows the detailed source design. The initial design was specified to be spectrally broadband from  $635 - 830nm$ . The design in-

cluded broadband retarders and a 2400W (635W radiant power) Xenon (Xe) arc lamp source. The Xe source would have allowed for low sunlight imaging, near dawn or dusk, however costs for the Xe source were determined to be too high. An alternative design was developed, an inexpensive yet powerful solid state 671nm laser source was chosen, and the design was refactored for this source in the the polarization state generator (PSG). This allows for the PSA to be operated as a broadband portable full Stokes polarimeter by itself, or as a narrowband active Mueller matrix polarimeter when combined with the 671nm PSG.

The PSG with the laser source in the final design cannot image in sunlight, this is because near the 671nm wavelength sunlight has about  $1.4W/m^2$  per nm of irradiance at air mass 1.5 (AM1.5) [132]. Even during sunset conditions, the irradiance on average drops to about 1/3 of the AM1.5 value, which can be computed using the formula [133].

$$I = 1.1 \cdot I_0 \cdot (0.7)^{AM^{0.678}} \quad (5.2.1)$$

where  $I_0 = 1353W/m^2$  for the AM0 curve integrated over wavelength. The AM can be calculated from the approximation

$$AM = \frac{1}{\cos \theta} \quad (5.2.2)$$

where  $\theta$  is the sun zenith angle. According to the National Oceanic and Atmospheric Administration (NOAA), in Tucson, AZ on August 1, 2015 at 6:30pm,  $\theta = 80.85^\circ$  corresponding to  $AM = 6.29$ . When this value is plugged into Equation (5.2.1) we obtain a value about 0.35 of the AM1.5 value. Substituting into the per wavelength values, we then obtain about  $0.5W/m^2$  per nm of irradiance near 671nm. Assuming a 5nm wide filter on our PSA optics, a total of  $2.5W/m^2$  of irradiance from sunlight will be incident on our object plane (of course this depends on the plane angle with respect to the sun, but this is an order of magnitude calculation). To put polarized irradiance in a 1-to-1 ratio with the sun irradiance would then require large laser powers in the 1 – 25W range, depending on range and optics. Lasers in this power range were cost prohibitive. This constrained us to computing a requirement for laser power for night imaging.

Prior to computing the required laser power, sensor specifications need to be known. The sensor for the instrument was sourced from 4D Technologies, Inc. A

QE	Pixel Size	# of Pixels	Fill Factor	Full Well ( $e^-$ )	Total Noise ( $e^-$ )
0.2	$9\mu m$	$1 \times 10^6$	0.55	50,000	72

Table 5.1: KAI-1010 sensor specifications for  $671nm$ .

Moxtek™ micropolarizer array was bonded to the focal plane of an Illunis RMV-1010 camera with a Kodak KAI-1010 CCD focal plane array. The sensor specifications are shown in Table 5.1. Each photon impinging on a detector pixel will have energy  $hc/\lambda$  where  $\lambda$  is the wavelength,  $h$  is Planck's constant, and  $c$  is the speed of light. The pixel area then determines the number of photons irradiating it, and the quantum efficiency determines how many of those photons are converted to photoelectrons. The number of photoelectrons for our specific sensor is approximately

$$N_{e-} = 0.55 \cdot N_p \eta = 0.11 N_p \quad (5.2.3)$$

where  $N_p = N \cdot A_p$ ,  $N$  is the number of incident photons per unit area,  $A_p$  is the pixel area,  $\eta$  is the quantum efficiency (QE), and the 0.55 constant is due to the fill factor. This implies a ratio of 9.09 photons to each photoelectron. The energy per photon at  $671nm$  is

$$E_p = \frac{hc}{\lambda} \approx \frac{2.0 \cdot 10^{-16} J \cdot nm}{671nm} \approx 3.0 \cdot 10^{-19} J. \quad (5.2.4)$$

Instead of each individual pixel the computation can utilize the entire sensor area, it does not change the above equations. Furthermore, in an imaging system, the approximation can be made that our source beam spread is matched to our optical field of view (FOV) which implies that for planar objects nearly perpendicular to our instrument's optical axis, we can assume that nearly all of the reflected light is collected, with the reflection characterized by the albedo of the object. Typical dark land materials (soil, etc.) have an albedo of 0.1. If we would like a uniform surface with an albedo of 0.1 to, on average, produce around 20,000 photoelectrons per pixel per image, then a  $1MP$  sensor requires a total of  $2 \cdot 10^{10}$  photoelectrons per image. This corresponds to an energy of

$$E_{p,i} = 9.09 \cdot 2 \cdot 10^{10} \cdot 3.0 \cdot 10^{-19} J = 5.4 \cdot 10^{-8} J \quad (5.2.5)$$

per image. If we assume that the integration time is  $1/60$  seconds ( $16.67ms$ ), then

for a continuous wave (CW) laser, this corresponds to

$$\frac{5.4 \cdot 10^{-8} J}{1s/60} = 3.2 \cdot 10^{-6} W. \quad (5.2.6)$$

When the albedo is taken into account, a factor of 10 increase is required;  $3.2 \cdot 10^{-5} W$ . Finally, our specific optical system has  $f/\# = 2.4$  which corresponds to the light collection ability, and remains a slow enough optical system that the approximation

$$\Omega = \frac{\pi}{4f/\#^2} = 0.14 \quad (5.2.7)$$

holds, where  $\Omega$  is the solid angle. A Lambertian surface reflects radiance equally into the hemisphere of  $2\pi$  steradians, so the collection factor is

$$\frac{0.14}{2\pi} \approx 0.02. \quad (5.2.8)$$

The total power is then divided by this value to get the source power required;  $3.2 \cdot 10^{-5} W / 0.02 = 1.5 \cdot 10^{-3} W$ . When the micropolarizer array elements are taken into account (a reduction of about half in light received on average) then the requirement becomes  $3.0 \cdot 10^{-3} W$ . If the exposure time of the sensor is assumed to be  $1ms$  instead of  $16.67ms$ , then the source power requirement becomes  $5.0 \cdot 10^{-2} W$  or about  $50mW$  which corresponds well to what we see with the actual instrument. The source which was procured is rated for a maximum of  $200mW$  CW, and is adjustable in output from about  $5mW$  to  $200mW$ . The source has a vertically linearly polarized output beam with respect to the mounting plane of the laser. The source is a Laserglow<sup>TM</sup> LRS-0671-PFM-00100-05 LabSpec  $671nm$  DPSS Laser System with a beam output power variation of  $< 5\%$  RMS over 4 hours. This laser source has a  $1/e^2$  beam diameter of  $2mm$  and a quality factor of  $M^2 < 1.2$  [134].

The laser beam is subsequently expanded via a Thorlabs<sup>TM</sup> BE04R/M reflective beam expander. A reflective beam expander was chosen for 1) broadband wavelength capability as compared with a refractive beam expander, 2) compactness for weight savings elsewhere. The broadband capability of the reflective beam expander allows for flexibility to install a laser source and polarization components at wavelengths other than  $671nm$  if needed in the future. The compactness allows for a smaller breadboard footprint, saving weight. A beam expander is needed in the design to mitigate systematic errors which could result from spatial variation in the rotating waveplates, i.e the full clear aperture of the waveplate should be illuminated to

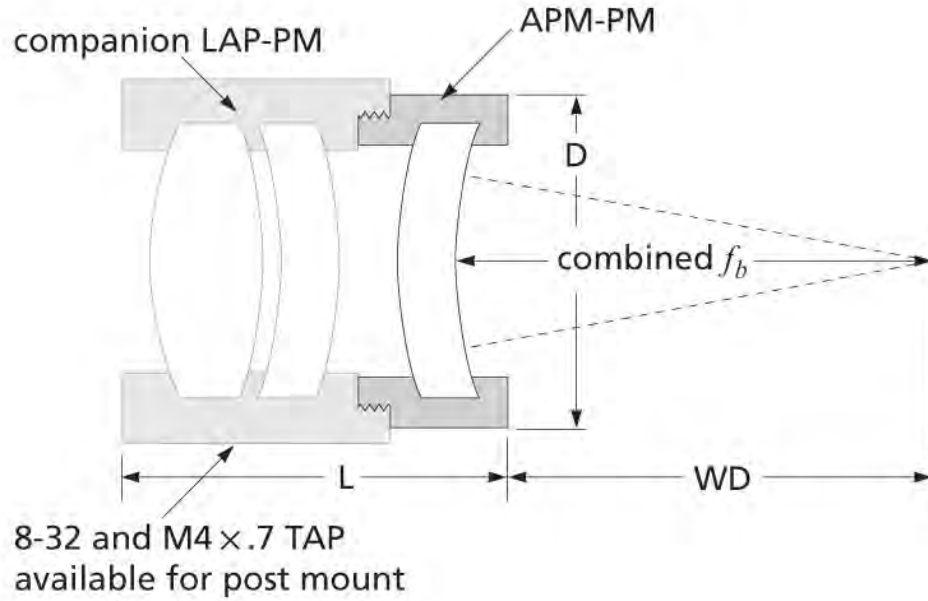


Figure 5.1: Aplanatic lens system. The specified lens has  $f = 51mm$ ,  $D = 31.7mm$ ,  $L = 25.4mm$ ,  $WD = 35mm$ . These specifications are for  $633nm$ , so the focal length will change slightly for  $671nm$  light. Courtesy of CVI Melles Griot.

eliminate the effects of spatial variation in the retarders. A Thorlabs™ GL10-B Glan Laser polarizer was mounted to the front output port of the beam expander to clean up the laser polarization. The extinction ratio is approximately 100,000 : 1 for this polarizer. The output beam is collimated.

The expanded linearly polarized beam then passes through two retarders. The retarders are Meadowlark™ precision polymer coated for  $670nm$  with a  $25.4mm$  diameter. The first retarder is a half wave ( $\pi$ ) retarder, and the second is a quarter wave retarder ( $\pi/2$ ). See Figure 5.3 for opto-mechanical details. These retarders were ordered prior to the optimization work from Chapter 3, otherwise the second retarder would have been specified with a retardance of  $\pi - \cos^{-1}(1/\sqrt{3})$  which is approximately 0.348 waves. The rotation of the retarders modulates the polarization state of the beam, generating carriers. To ensure minimal systematic errors from



rotation errors, a position accuracy and repeatability requirement of 30 arcseconds was specified for the rotation stages. Additionally, a rotation speed of 1000rpm was specified to allow for future speed increases, and to allow for experiments on various system types and speeds to be robust. IntelLiDrives™ built custom low profile rotation stages ( model number ACR-55UT) for this purpose with a 55mm aperture to install waveplates or polarizers into. Stage repeatability was tested via a fixture consisting of a telescope system mounted to the rotation stages. The telescope system was used to magnify the angular errors of the rotation stages. A laser was mounted so that the beam passed through the lens system on the stage, resulting in a spot on the opposite wall. The stages were then rotated 360° 100 times and the spot was imaged with a camera. Since the spot was approximately Gaussian, a centroiding algorithm was used to give good approximate centers on the wall from the images taken with the camera. Measurements were taken between all components, and an angular error was computed, which resulted in a single pixel shift on the camera corresponding to about 2.5 arcseconds of angular change of the rotary stage. 6 stages were ordered, and all but one were well within specifications, with one being right at the repeatability specification of 30 arcseconds. The 4 stages with the best repeatability were selected to be assembled into the HyDMIP-P, all with < 10 arcsecond repeatability. The repeatability test was designed with help from Rafael Rojas, and Rafael Rojas conducted the tests.

After the Stokes parameter generation from the retarders, the beam then passes through a lens system with a focal length designed to ensure that the beam spread in the far field approximately matches the FOV at distances greater than 50m. The lens system is a CVI Melles Griot APM-75.0-15.0-633 aplanat. It is designed for 633nm; we did not solicit a custom 671nm aplanat due to cost constraints (CVI laser quoted a cost of ~ \$6000 for the custom element). The details of the lens design are shown in Figure 5.1. The clear aperture is 15mm, and the beam diameter entering the beam expander corresponds to 2mm [134]. The beam expander then expands the beam diameter to about 8mm. We can use [135]

$$\omega_0 = \frac{M^2 \lambda f}{\pi r_{in}} = \frac{1.2 \cdot 671nm \cdot 50}{\pi \cdot 4} \approx 3.2\mu m \quad (5.2.9)$$

to compute the focal spot size at the beam waist after the lens where  $\omega_0$  is the beam



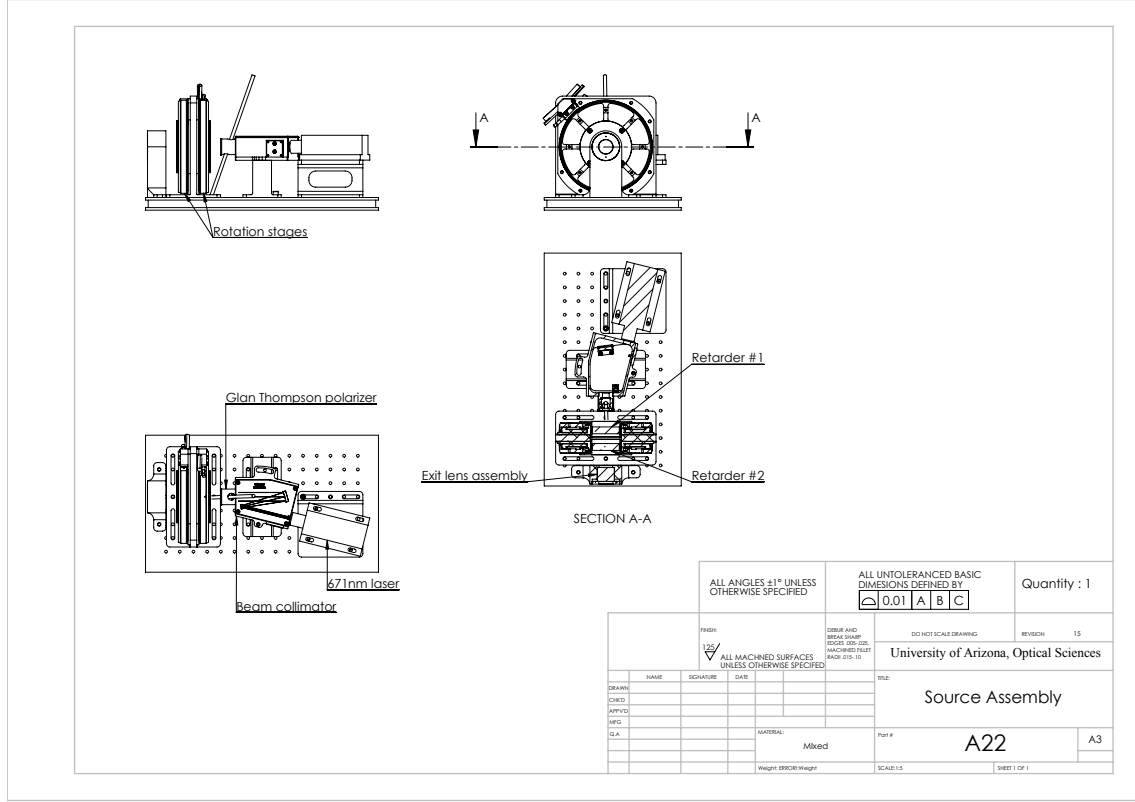


Figure 5.3: Source assembly, zoom in for detail. The laser beam is expanded to about a  $8mm$  beam diameter prior to passing through the retarders.

waist radius at the focus and  $r_{in}$  is the radius of the beam impinging on the lens system. This equation is really only accurate to first order. A focal length of  $50mm$  was assumed for  $671nm$  light due to lack of information from CVI Melles Griot regarding the focal length at  $671nm$  or alternatively the lens prescription, which would have allowed me to compute the  $671nm$  focal length myself. The far field divergence angle for the beam is then

$$\theta_d = \frac{\lambda M^2}{\pi \omega} \quad (5.2.10)$$

where  $\omega$  is the beam waist radius. Substituting in  $\omega_0$  from above yields

$$\theta_d = \frac{0.671\mu m(1.2)}{\pi 3.2\mu m} = 0.08 \approx 4.6^\circ \quad (5.2.11)$$

The Kodak sensor is approximately  $9.1mm \times 9.1mm$  in size, which is circumscribed by a circle of diameter  $12.87mm$ . The focal length for the PSA optical imaging system

is  $\sim 103\text{mm}$ . At an object plane  $50\text{m}$  from the lens this corresponds to a circular field of view (FOV) with diameter  $6.25\text{m}$  which corresponds to an subtended half-angle of about  $3.6^\circ$  from the observer. We are near infinity focus at  $50\text{m}$  so the  $3.6^\circ$  is approximately the same FOV subtended angle for any object plane  $> 50\text{m}$ . This is within the far field angular beam spread for the source of  $4.6^\circ$ . A custom lens assembly mount was designed to work with the other components, shown in Figure 5.2. The complete source assembly is shown in Figure 5.3.

### 5.2.2 Polarization state analyzer

The polarization state analyzer (PSA) is a spatio-temporal hybrid channeled system. The micropolarizer array provides spatial carrier generation [33, 35, 43], and two rotating retarders provide temporal carrier generation. After the beam from the PSG reflects from the object, the light is collected by the PSA. The design provides for the rotating retarders to be placed in a collimated optical space via an afocal element. The imaging optics assembly was designed with the following constraints

- Minimal cost,
- Flat contrast and resolution limits across the image (i.e. MTF is fixed across FOV),
- Changeable front lens,
- Variable focus,
- Variable aperture stop.
- Applicable to the  $9\text{mm} \times 9\text{mm}$  sensor size.

These constraints led to commercial off the shelf lenses being selected due to the performance versus cost advantages that they have over custom designs or semi-custom designs (e.g. Edmund Optics® or Thorlabs® lens assemblies). For the visible range, Pentax®, Canon®, Nikon®, and others make high performance, low cost camera lenses. After reviewing MTFs curves from different manufacturers for various lenses, Pentax® lenses were chosen due to their consistency across the FOV. Pentax

had, to the the best of our knowledge from the technical specifications we had access to, the most consistent MTF, contrast, and vignetting profiles across the field of view. Attempts were made to obtain lens prescriptions or Zemax black boxes to facilitate optical design, but none of the companies were willing to provide us with this information. The flange focal distance ( $FFD$ ) is, however, a public specification, and for Pentax® K-mount camera lenses  $FFD = 45.46mm$ . The K-mount bayonet system allows for quick front lens changes on the system if desired. Another benefit to the Pentax® lenses was the ability to install a computer controlled autofocus system in the future. The lenses contain a small screw like element which is spun by the camera body to autofocus, and would allow us to use a motor to control the lens focus. Other brands have autofocus built into the lens through an electronic communication protocol, making it infeasible to utilize in our polarimeter since this protocol is not public.

The lens assembly design required that the rotating retarders be located in a nominally collimated optical space. Additionally, the Pentax® lenses were designed for a film format sensor plane size of  $36mm \times 24mm$ . This created a requirement for minimification to the  $9mm \times 9mm$  actual sensor size of our micropolarizer array camera. A fairly simple design was implemented, an initial lens forms an intermediate image, then a lens is placed so that the intermediate image is at the focal length, forming an afocal system. This results in rays passing through the retarders being collimated, and the lens assembly is also designed to minify the intermediate image. The design is shown in Figure 5.4 modeled via thin lenses given the known focal lengths. The assumption is made that the first lens forms an image at an intermediate image plane, the next lens images the intermediate image to infinity, and finally the third lens images from infinity to the sensor. Lenses are notated  $L_1, L_2, L_3$  in order of light propagation.  $L_1$  is designed to be a changeable lens, and for the full active system a Pentax SMCP-FA 77mm f/1.8 Limited was used,  $L_2$  is a Pentax 70mm f/2.4 DA Limited Lens and  $L_3$  is a Pentax SMCP-FA 43mm f/1.9 Limited Lens attached to the sensor body with a K-mount to C-mount converter. The system  $f/\# = 2.4$  with the 77mm focal length lens for  $L_1$ .

The first two lenses  $L_1$  and  $L_2$  had to be affixed in an afocal configuration and required custom lens mounts. In order to mitigate the time intensive reverse engineer-

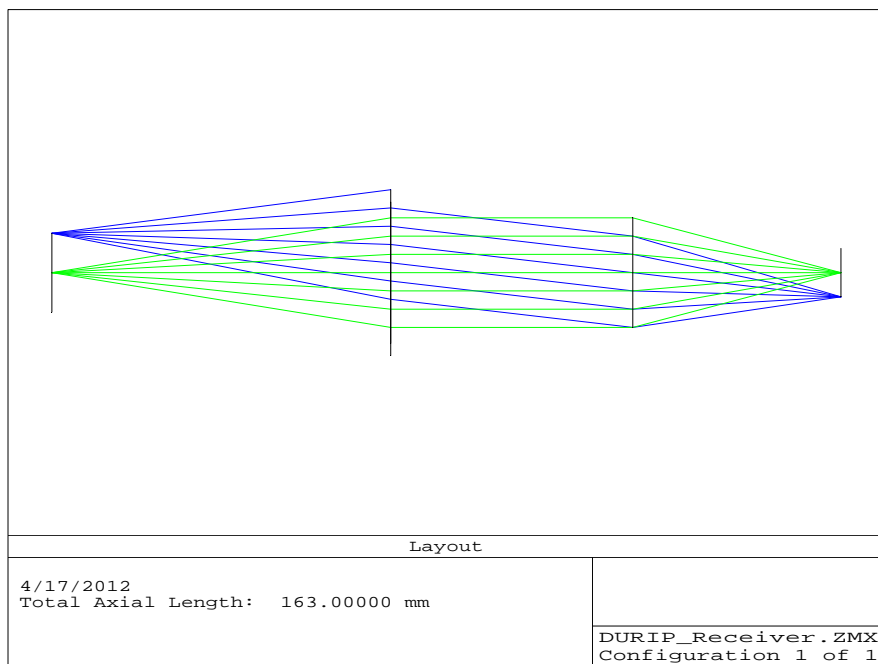


Figure 5.4: Receiver lens design. The left plane is the intermediate image with a total height of  $24\text{mm}$ , the second plane is  $L_2$  as a thin lens, the third plane is  $L_3$  as a thin lens, and the fourth plane is the sensor plane with a total height of  $\sim 10.5\text{mm}$ . The actual sensor is slightly smaller with a height of  $9.1\text{mm}$ , but this was the closest match available with the focal lengths available from Pentax.

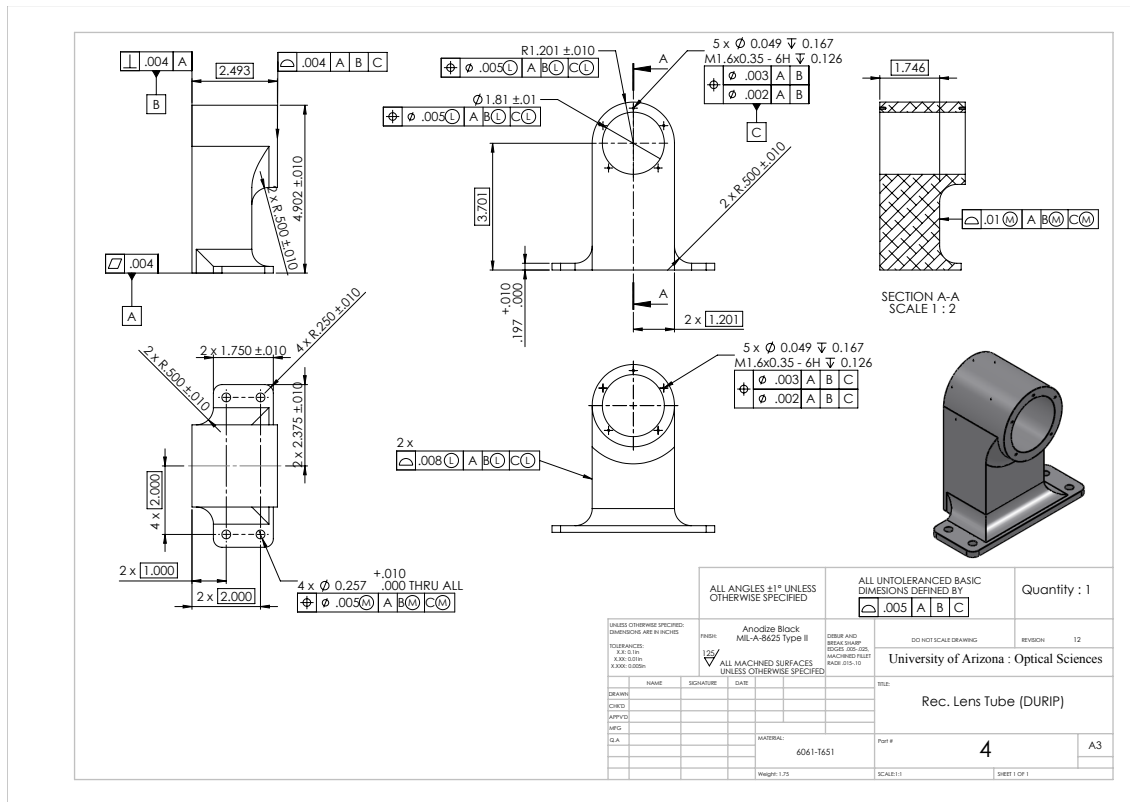


Figure 5.5: Receiver afocal lens mount body design. The Samsung adapters are not shown, but attach to either side. Zoom in for detail.

ing of the mechanical K-mount, Samsung MA9NXK adapters were purchased to be incorporated into the afocal lens mount. These adapters are used to convert K-mount lens to a Samsung NX10 mount. Rafael Rojas helped to measure the adapters and build Solid Works models of the adapters for use in the assembly design. The Samsung MA9NXK adapters were modified to be bolted to the lens mount body shown in Figure 5.5. This required machining, drilling, and reassembly. The complete lens mount assembly is shown in Figure 5.6. Because the  $FFD$  is known for the lenses, the requirement for the afocal lens mount assembly is that the flange to flange distance be  $2FFD = 90.92mm$ . The afocal assembly provides (nearly) collimated rays through the rotating retarders and the final lens then re-images onto the micropolarizer focal plane array sensor. The mounting plates for the rotation stages were also custom designed as shown in Appendix B, as well as other opto-mechanical details.

The sensor is a Kodak 1010 CCD array with a micropolarizer array bonded di-

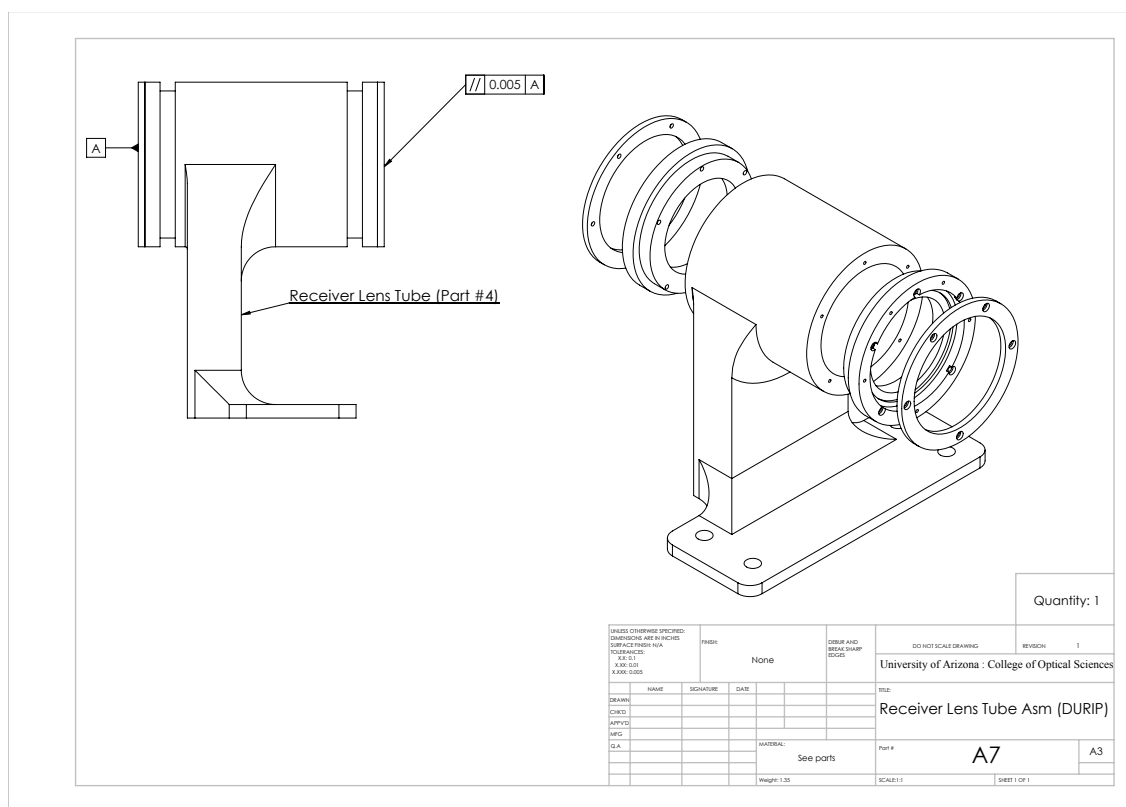


Figure 5.6: Receiver afocal lens mount assembly. The modified Samsung adapters are shown exploded in the bottom right view. Zoom in for detail.



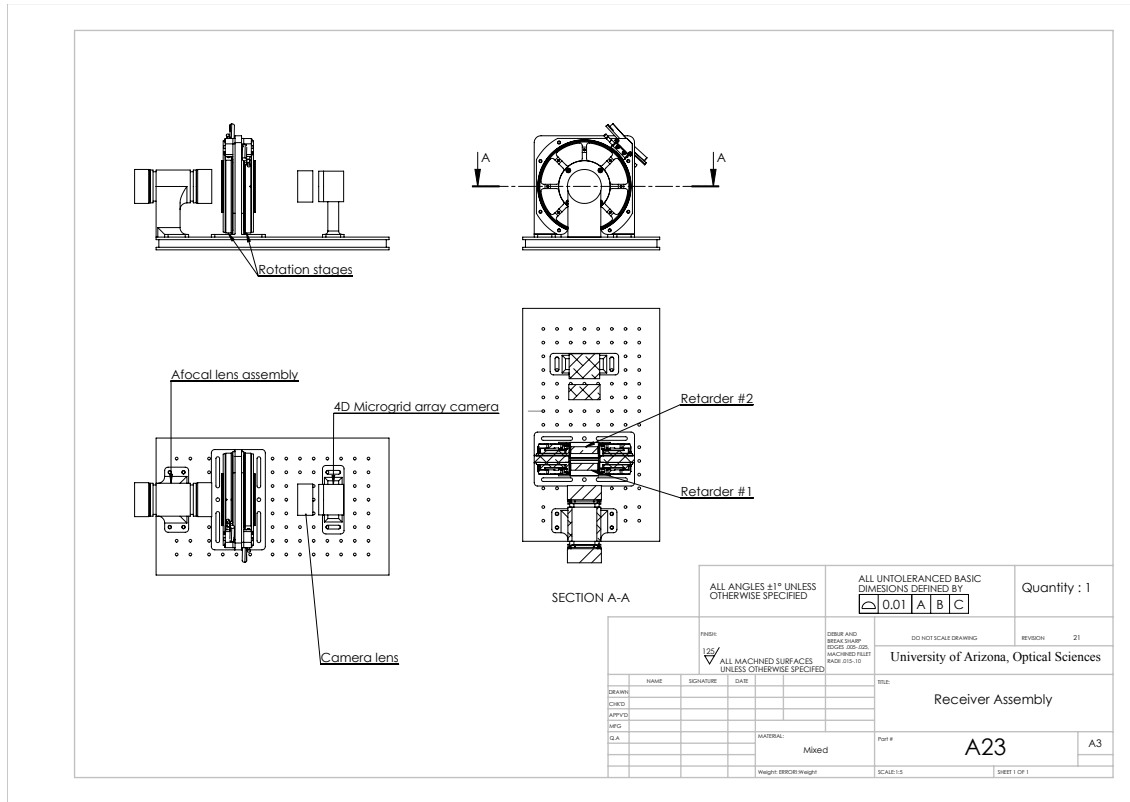


Figure 5.7: Receiver (PSA) assembly, zoom in for detail.

rectly to the focal plane by 4D Technologies<sup>TM</sup>. The camera can operate at up to  $28.4\text{fps}$  when in a triggered mode. The camera uses a cameralink data connection. The complete PSA assembly is shown in Figure 5.7.

### 5.2.3 Triggering

Timing is critical for channeled polarimeters since the systems are not shift invariant. This requires a triggered camera, with triggers coinciding to specific retarder positions. The controllers for the rotation stages unfortunately cannot be programmed to output pulse trains which are some multiple of the position, or of some multiple of the encoder pulse. The only output available from the motor controllers is the digital encoder pulse train, with 540,000 pulses per rotation. Sampling right at the Nyquist frequency of the system shown in Figure 3.11 requires 32 samples per rotation of the slow motors and 16 samples per rotation of the fast motors. This results in 16,875

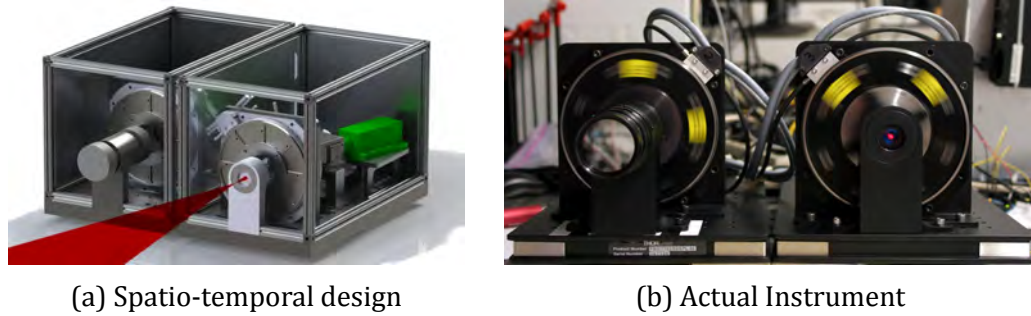


Figure 5.8: Hybrid Domain Modulated Imaging Polarimeter - Portable (HyDMIP-P)

and 33,750 pulses between triggers respectively. I had access to an Arduino Uno microcontroller and decided use the Arduino to downsample the pulses and output camera triggers. Several iterations of Arduino control software were required, due to the pulse width of the encoder output pulses being  $\sim 1\mu s$ . This narrow pulse width caused the first few iterations of the Arduino software to miss pulses. The final version of the Arduino software is shown in Appendix B.

#### 5.2.4 Overview

The physical system consists of 4 rotating retarders and a micropolarizer array. The retarders are rotated by precision motors with a maximum rotation speed of  $1000rpm$ . Both the PSG and PSA were designed so that the retarders reside in a collimated optical space, and the source beam spread was matched to the PSG field of view. A rendering of the complete instrument and a photograph of the physical instrument are shown in Figure 5.8. A summary of the instrument specifications are shown Table 5.2

### 5.3 Control system

The HyDMIP polarimeter requires a control system and the associated software to procure and process images. The control software design architecture consists of 6 main modules;

- a motor control module,
- a camera control module,

Table 5.2: System specifications

	Component	Type	Detail
PSG	source	671nm laser	200mW
PSG	retarder(s)	671nm	1in dia.
PSG	stages(s)	up to 1000rpm	< 15arc sec accuracy
PSG	optics	aplanatic @ 633nm	matched to PSA FOV
PSA	sensor	silicon 1MP array @ 28.4 fps	9 $\mu$ m pixels, micropolarizer array
PSA	retarder(s)	630 – 835nm	2in dia.
PSA	stages(s)	up to 1000rpm	< 15arc sec accuracy
PSA	optics	afocal section for retarders	f/2.4

- an image aggregator and pre-processor module,
- a reconstruction module,
- a classification module,
- a user interface module.

The classification module is only partially built and will not be discussed in detail; however it is designed as a modularized architecture where different classification algorithm modules can be plugged in for different classification or estimation tasks from the polarimetric data. The user interface module will also not be discussed. The reconstruction module is currently templated in Matlab, but is not implemented in final form, which will be in C++ and CUDA for real-time reconstruction capability. CUDA is a C-like language which is compiled for Nvidia graphical processing units (GPUs). GPUs offer large and inexpensive computational speed ups for linear algebra operations, and are ideal for parallelizable computations. The motor control module and the camera control module are complete and written to the C++14 standard. They are multi-threaded and can take advantage of modern hardware. The code is also written in a platform independent way and can be compiled on both Microsoft Windows® and Linux operating systems. All modules were designed to be as general

as possible, for example the motor control module is designed to accept any number of motors and even different motor types. The camera module has an interface allowing for any number of camera hardware implementations to be used if necessary. Figure 5.9 shows a flowchart of the overall control system design.

### 5.3.1 Stage control

The IntelLiDrive rotation stages are driven and controlled by Copley Controls™ Xenus XTL amplifiers. These amplifiers have two interfaces for computer control, RS-232 serial and CANOpen. CANOpen control requires a CiA compatible CANOpen software stack combined with a CAN bus card installed in a computer. CAN is an acronym for Controller Area Network and is a physical communication bus used in industrial and automation applications. CANOpen is a communications protocol which uses a CAN bus. The CANOpen protocol is much faster and more flexible than the RS-232 serial communication; however it is more difficult to use. The open source CANOpen stack CANFestival 3.0 was used.

The CANOpen protocol consists of two types of messages, the service data object (SDO) and the process data object (PDO). SDOs are usually used for initialization, network protocol setup, etc. PDOs are used once the system moves to an event driven and/or active stage. SDOs require more network overhead and are therefore slower. The CANOpen syntax and semantics are esoteric and use a hexadecimal addressing scheme and many of the data received or sent require exercises in bit flipping. To alleviate this error prone syntax I wrote a complete semantic translation of the CANOpen objects and functions supported by typical automation devices and specifically the Xenus XTL. For example to read the position via an SDO now only requires

```
1 sdoCommunicator.Read(i, CANOpenFunctions::PositionActualValue, data)
```

in C++, where *i* is the CANid corresponding to a specific stage and *data* is a structure where the position data will be placed. The `CANOpenFunctions` enumerated class translates semantic meaning to the hexadecimal addresses.

The stage control software consists of 5 sub-modules, 1) a master controller, 2) a master settings object, 3) an SDO communicator object, 4) a PDO communicator

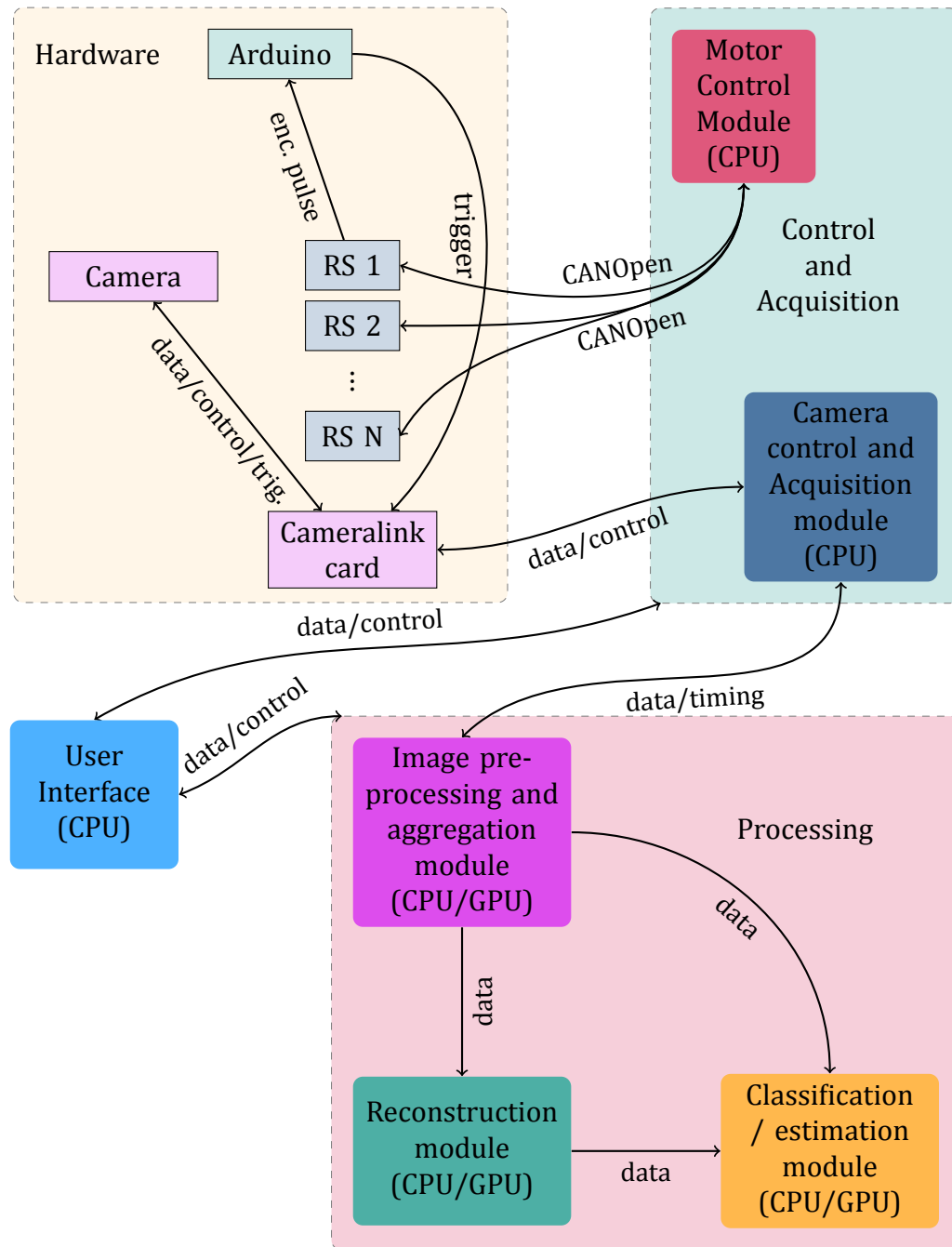


Figure 5.9: Control system design overview flowchart. RS denotes rotation stage. The motor control and camera control modules are fully implemented. The remaining modules are only partially or not implemented at this time. The reconstruction module is implemented in Matlab, but cannot operate in real time at the moment.

object, 5) a per stage information/settings object, and 6) a data storage object. These sub-modules work together to make using the stages straightforward. The master settings object is used to set all of the network and motor parameters needed, then passed to the constructor of the master controller. The master controller initializes all stages, homes them, and sets them to the appropriate starting positions. It then waits for a start function call, which starts the rotation of all motors specified in the master settings object. Once started the master automatically records all specified PDO information, including velocity, position, and time elapsed since the last update, for each stage. This requires setting the update speed, which is currently every  $667\mu s$  for each stage.

### 5.3.2 Camera control

The camera control module is written in C++14 and has interfaces abstracted from specific implementations of camera types. It has multi-threaded threadsafe queuing for storing images until they are retrieved. It is general and robust enough to handle any camera hardware implementation, provided new camera implementation code is written. The interface abstraction allows any code which uses the camera(s) to be written for the interface application programming interface (API), and the interface will not change for a change in camera implementation. This allows for new camera hardware or different cameras to be used and installed in the system without rewriting the software stack. The only software which has to be written for a camera change is a specific implementation module which inherits the general interface.

There are two primary interfaces in the camera control module, `AbstractSingleCameraWrapper` and `MultipleCameraCaptureControllerInterface`. `AbstractSingleCameraWrapper` provides an interface for any single camera. A dummy camera instance, `DummySingleCameraWrapper`, has been built for extensive testing and evaluation of the single camera interface and programs relying upon it. `MultipleCameraCaptureControllerInterface` is an interface which can control any group of instances derived from `AbstractSingleCameraWrapper`. This allows for multiple camera control for generic cameras.

Finally, images are stored in the `Image` class, which automatically stores images in the correct format depending on the actual implementation of the camera hardware

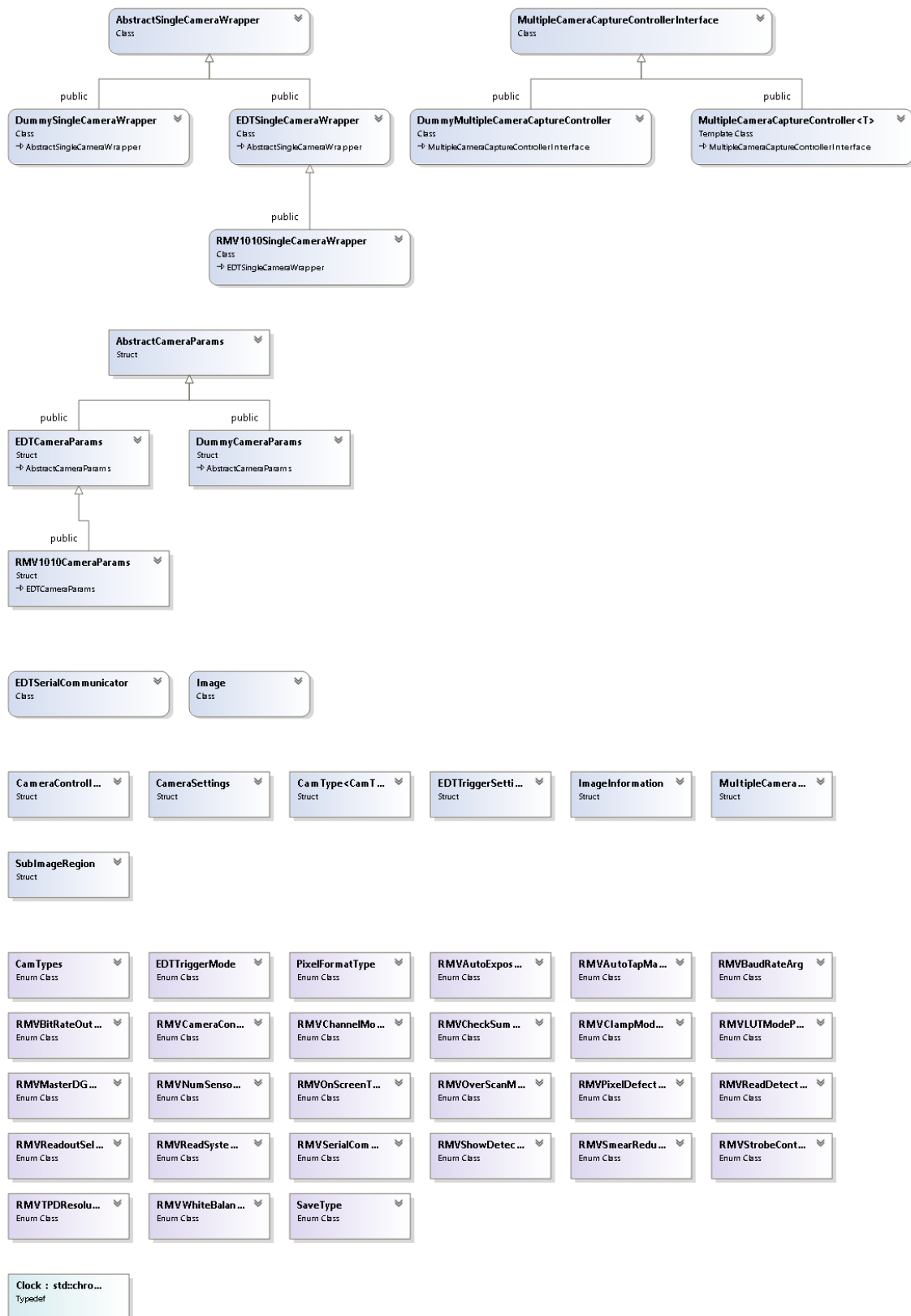


Figure 5.10: Class diagram for the camera module. The lines with arrows denote inheritance.

which provided the image. This includes parameters like color/monochrome, bits per pixel, image size, etc. The Image class also automatically converts and stores raw data to floating point data, for graphical processing unit (GPU) computations when/if needed. If a region of interest is needed, the Image class provides this capability. The Image class also provides an image id field for each image.

The camera control interfaces provide a user with the ability to obtain images or queues (stacks) of images in a threadsafe manner. Any implementation is required to stamp images with unique image ids, so that any lost or dropped image can be detected by a user of the interfaces. Once a user reads an image, then the user has responsibility for that image, i.e., the camera interfaces are designed to delete images once they are read by a user, they are not persistent after a read in camera control implementations. See Figure 5.10 for the class diagram of the camera control module.

### 5.3.3 Pre-processing

The pre-processing module collates and packages the images, and if needed, queues images or image stacks onto the GPU for processing. This module is also designed to perform simple pre-processing image operations like thresholding, masking, ROI, or normalization. It is written in C++14 and is thread-safe. This module is not yet complete, and in the future I would like to add morphological operations and fast convolutions to the pre-processing capabilities.

Aggregation capabilities include the ability to apply a general "shuffling" or permutation map when images are collected from multiple cameras, including shuffling of ROIs within the same image if needed. This capability was not needed for the specific polarimeter discussed here since we have a single camera.

### 5.3.4 Reconstruction

The reconstruction module is planned to be implemented in C++14 and CUDA, but is currently implemented in Matlab as a proof of concept. The reconstruction module takes an image stack, a  $\mathbf{Q}$  matrix (as described in Chapter 3), a filter set, and a channel location set to produce reconstructed Mueller matrix images. The filtering will be discussed in detail in a later section. Each filter produces a spatio-temporal



data cube. After the filters are applied, the resulting data cubes are linearly transformed according to  $\mathbf{Q}^+$ , the pseudoinverse of  $\mathbf{Q}$ . Recall that this transformation is applied in the channel, or Fourier, space. The transformation results in a set of 16 data cubes (in the  $(\xi, \eta, \nu)$  domain) corresponding to each Mueller matrix element. These data cubes are then each individually inverse Fourier transformed to obtain a spatio-temporal data cube (in the  $(x, y, t)$  domain) for each Mueller matrix element, which can then be displayed or processed further.

## 5.4 Calibration

Actual instruments differ from the ideal models we derived in Chapter 3. Our channel structure optimization assumed perfect polarizers, retarders, etc. This gave us a good design to first order, but calibration is needed to compute a proper  $\mathbf{Q}$  matrix [34] and compute the physical channel structure. The  $\mathbf{Q}$  matrix is computed with the semi-empirical model described below in the next section. This section describes the calibration of the instrument, with the majority of the calibration effort devoted to micropolarizer array characterization. Retardance calibration was not done, retardance specifications were provided with each retarder by the manufacturer which were used in the modeling.

### 5.4.1 Micropolarizer array

The first step of calibration was calibrating the micropolarizer array. This calibration was accomplished in a non-imaging setup by 1) placing white paper in front of a rotatable polarizer as a diffuser, which was then placed in front of the micropolarizer array and masked off for stray light, 2) rotating the polarizer with  $1^\circ$  increments and recording an image frame at each increment, 3) fitting a Malus' like function to each micropolarizer pixel and recording the parameters. The Malus' like function is

$$f(a, b, \theta') = a + b \cdot \cos^2(\theta + \theta') \quad (5.4.1)$$

where  $\theta$  is the rotation angle of the reference polarizer and  $\theta'$  is the angle of the micropolarizer of the pixel with respect to the lab reference frame (set by the reference polarizer). The extinction ratio can then be estimated by either using the actual data

for each pixel, or by

$$ER = \frac{a}{a + b} \quad (5.4.2)$$

note that the above doesn't account for dark current. If bad pixels are eliminated, our extinction ratios range from about 1 : 10 to 1 : 50, with the mean at  $\sim 1 : 17$ . Bad pixels are pixels which are either "hot" or "dark," i.e., they don't respond to variations in irradiance and output a very high digital number or a low digital number respectively. These ratios were calculated directly from the data. The Malus' like law parameter fits for each pixel are then used as inputs for Mueller matrices of diattenuators. The calibration data was recorded with help from Rafael Rojas and the initial fits to the Malus' like function were performed by Oscar G. Rodríguez-Herrera, however I re-ran the fits using a more stringent stopping criteria and a multi-threaded optimization routine in Matlab. Initially there appeared to be a much larger angular variance among the nominally  $0^\circ$  micropolarizers as compared with the other elements, however this appears to be an artifact from the first optimization procedure; the subsequent more stringent optimization procedure appears to have eliminated this variance difference.

The systematic effects on the channel structure of the physical micropolarizer array are addressed in Section 4.5

#### 5.4.2 PSA and PSG reference frame

Another issue with the instrument is the in plane, i.e., around the optical axis, rotation between the PSG and the PSA. In the ideal system, the components are all aligned to some fixed reference frame, with the Stokes parameters of the input into the modulation optics of the PSG being the horizontally linearly polarized state

$$\mathbf{s}_{\text{in}} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad (5.4.3)$$

and the micropolarizer array and camera in the PSA are oriented so that the  $0^\circ$  micropolarizer elements are aligned to the same  $xy$ -axis as the Stokes parameters after

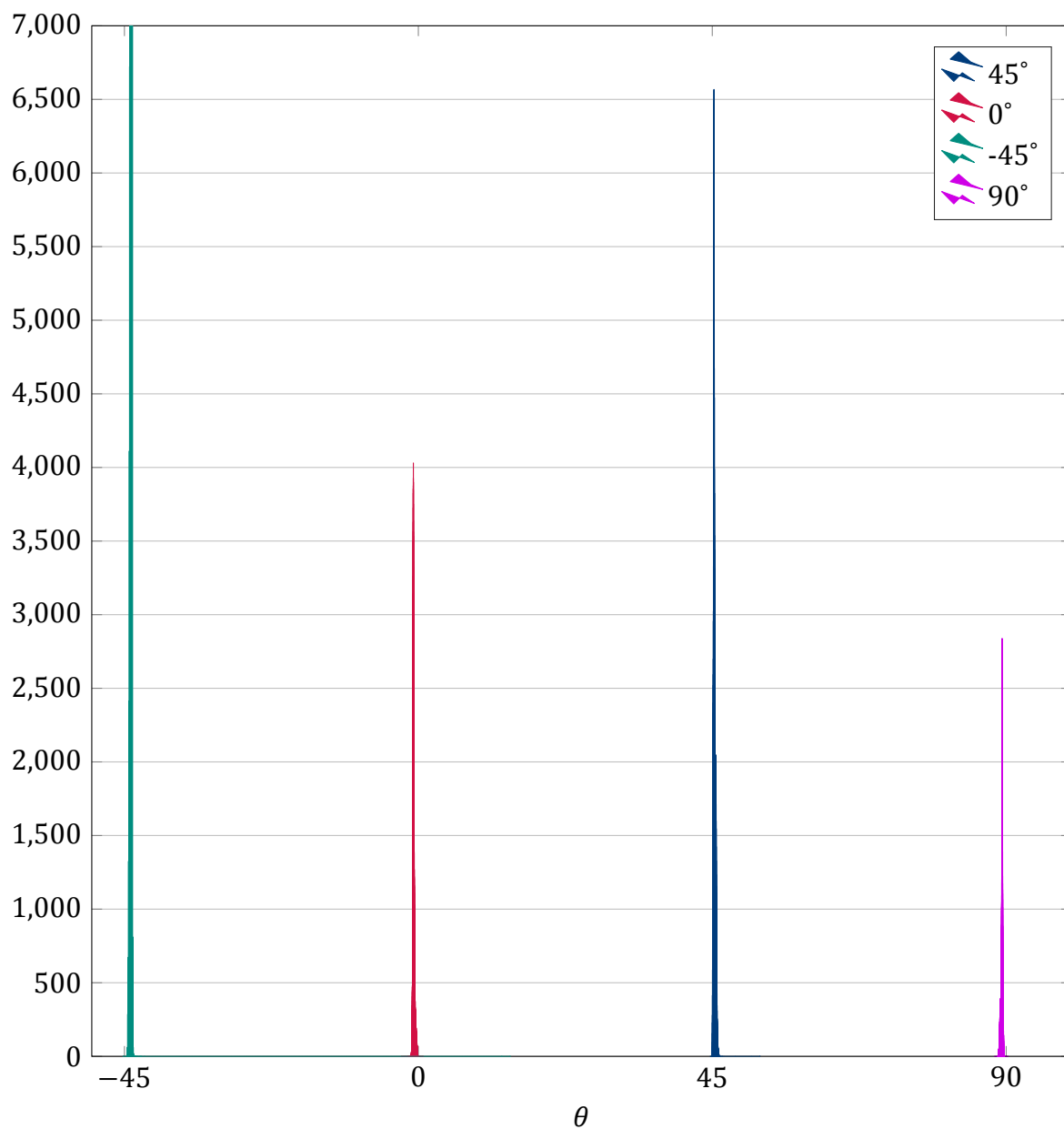


Figure 5.11: Histograms of angles of the micropolarizer elements. The  $-45^\circ$  histogram is much narrower than the rest, with the center bin reaching around 70,000 counts, exceeding the range of the graph. Each histogram shows the angular spread from the specified values over the pixels.

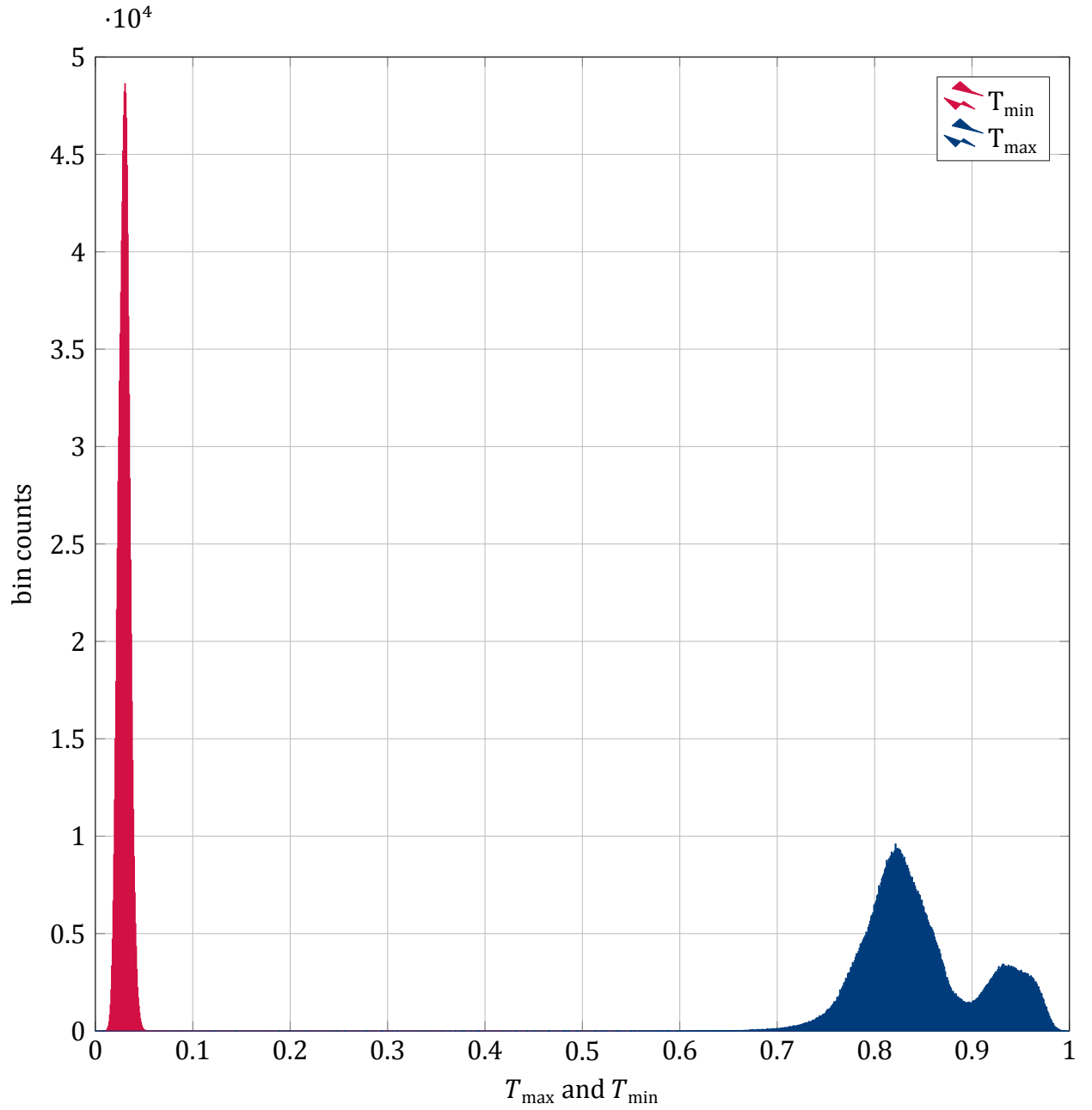


Figure 5.12: Histograms of the entire sensor with  $T_{\min} = a$  from Equation (5.4.1) in red and  $T_{\max} = a + b$  from Equation (5.4.1) in blue. The odd shape of the  $T_{\max}$  histogram is explained in Figure 5.13.

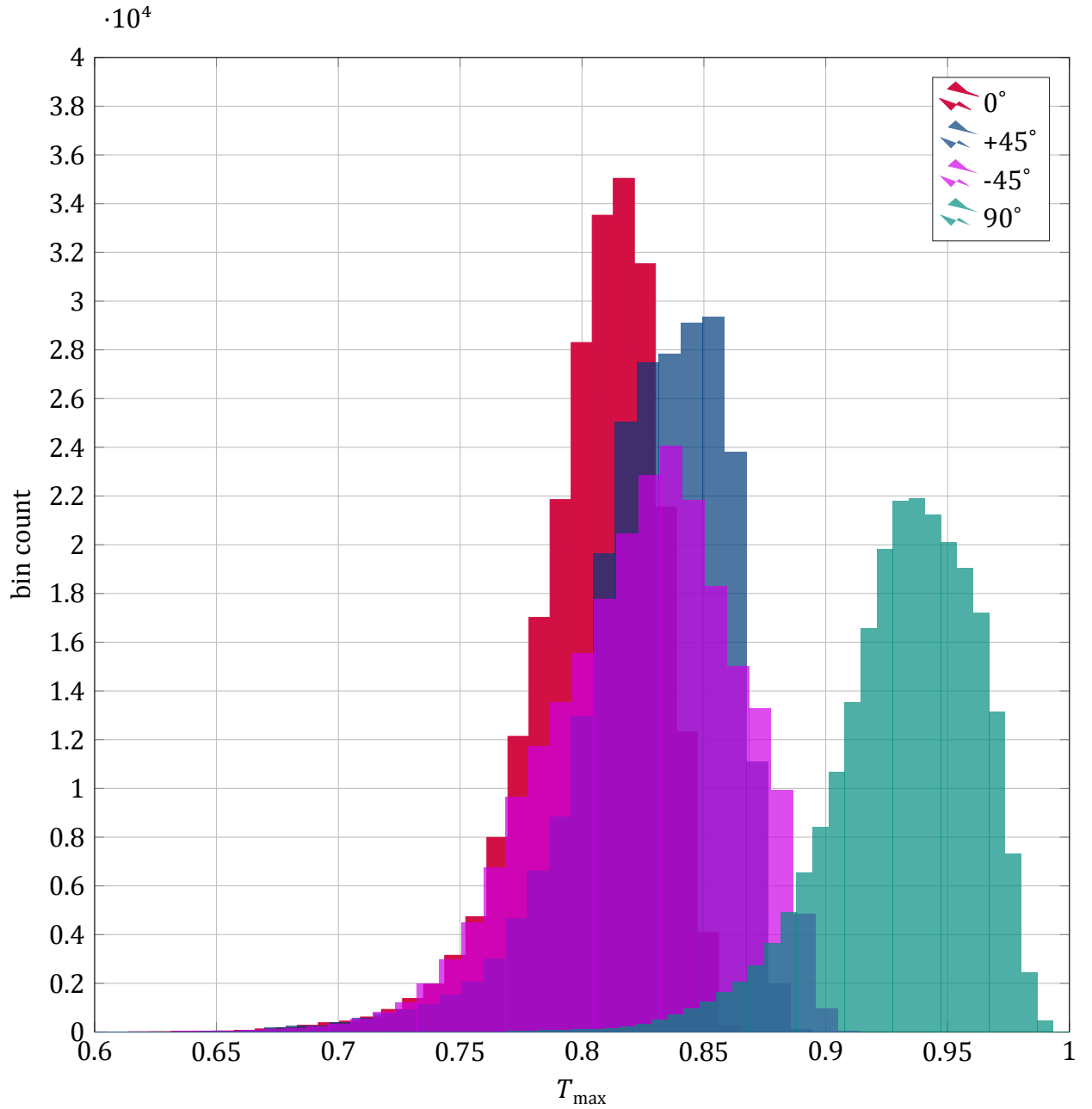


Figure 5.13: Histograms of  $T_{\max} = a + b$  from Equation (5.4.1) for each micropolarizer orientation. Notice the much higher overall  $T_{\max}$  for the  $90^\circ$  elements. This is a result of the micropolarizer etching and manufacturing process [136].

the polarizer in the PSG. However, this is not the case in the physical system. First, the laser output linear polarization state is nearly vertical, not horizontal, and there is a rotation between that vertical state and the  $xy$ -coordinate frame attached to the mean  $0^\circ$  micropolarizer element. The mean  $0^\circ$  element here denotes the literal mean of all  $0^\circ$  elements of the micropolarizer array. Furthermore, once a reference is defined, the fast axis of each retarder in the system must be aligned in a systematic way so that the fast axis position is known with respect to the defined reference frame.

As shown in Chapter 3, the rotation of the retarders has no effect on the ideal system, but it can affect the systematic errors of the real system. The simplest way to find and set the retarder orientations is to use the element by element calibration method. The laser is polarized to a high quality linear state after the Glan-Laser polarizer, and it is used as the reference axis. That is, the fast axes of the retarders are oriented  $90^\circ$  from linear polarization axis of the PSG laser. Note that the fast axis of each retarder is visually marked by the manufacturer. Retarder orientation calibration can be accomplished in the following way:

- (a) prior to any retarders being inserted into the beam path, either orient the PSA and PSG so that the beam path goes directly from the PSG through the PSA, or orient the PSA and PSG in a reflective configuration where a mirror is placed far away in a quasi-monostatic orientation and the beam is reflected through the apertures for the retarders.
- (b) after the last retarder aperture on the PSA, place another rotatable linear reference polarizer and rotate it until maximum extinction is reached, i.e., until the PSG polarizer and the reference polarizer are crossed.
- (c) place the first retarder (in optical propagation order) into the stage mount.
- (d) rotate the stage until the visible mark is approximately horizontal. Then rotate the stage until maximum extinction is reached after the reference polarizer. Record the stage position. Repeat this 10 to 20 times. Take an average of the recorded stage positions. This position is the position for which the slow axis of the retarder is aligned to the linear polarization of the PSG. Lock the position of the stage for this particular retarder and leave the retarder in place.

- (e) repeat the above step for each subsequent retarder.

The above procedure is not currently automated, but I intend to automate in the future.

The remaining orientation to calibrate is the rotation angle between the PSG and the micropolarizer array. This can be accomplished via a procedure similar to the retarder calibration procedure. First insert a reference polarizer and maximize the extinction of the PSG light and record the angular position. This polarizer must be fixed in a location where the micropolarizer array camera can image through it. I used a Meadowlight 50mm Versalight wire grid polarizer in an accurate rotatable mount with rotation scale markings. Next, insert the camera into the final position and securely mount it. When imaging through the polarizer (into something with enough illumination, I used a white wall), the darkest set of pixels will correspond to the micropolarizer elements which are linearly polarized in nominally the same direction as the PSG light. Finally, rotate the reference polarizer until this dark set of pixels is minimized in irradiance (maximized in extinction). Record the difference between the original angular position of the reference polarizer and the current position. This is the angular rotation offset between the micropolarizer array and the PSG and PSA retarders. This angular difference is used in the semi-empirical model described in the next section. The measured rotation offset was  $\sim -3.9^\circ$ .

## 5.5 Semi-empirical model

Reconstruction requires accurate knowledge of the actual channels of the physical instrument, which dictates the entries of the  $\mathbf{Q}$  matrix for Mueller matrix reconstruction. Since it is difficult to probe the entire system matrix dynamically, a semi-empirical model was developed to characterize the physical system. All calibrations and known retardances were used in the model. Additionally, the model was designed to support potential systematic errors and noise sources. The actual bulk retardances of each retarder are used in the model, which is a valid assumption since the retarders are in collimated optical spaces and therefore the retardance doesn't have any significant spatial variation at the object or image planes. Finally, sampling at the correct framerate is implemented into the model. The model was designed to

address Johnson (Gaussian like) detector noise, Poisson noise, spatially varying illumination, and the remaining systematic error of retarder position wander during rotation. The model also allows for any Mueller matrix object to be inserted, and assumes that the object has a spatial size equivalent to the camera image frame size. The model does not account for the systematic errors induced from the imaging optics, speckle, lens aberrations, etc. This is a forward model, given a set of system parameters, noise parameters, and time resolved Mueller matrix objects, it computes a simulated spatio-temporal data cube for the imaging camera, which can then be Fourier transformed into the channel space.

The model uses all system, calibration and noise information available to compute the Stokes parameters at each pixel,  $(x_k, y_l)$  at each time step,  $t_m$  as follows:

$$\begin{aligned} \mathbf{s}_{\text{out}}(x_k, y_l, t_m) = & \mathbf{P}(x_k, y_l) \cdot \mathbf{R}([v_4(t_m) + n_{v_4}], \epsilon_4, \delta_4) \cdot \mathbf{R}([v_3(t_m) + n_{v_3}], \epsilon_3, \delta_3) \\ & \cdot \mathbf{M}_{\text{test},I}(t_m) \cdot \mathbf{R}([v_2(t_m) + n_{v_2}], \epsilon_2, \delta_2) \cdot \mathbf{R}([v_1(t_m) + n_{v_1}], \epsilon_1, \delta_1) \mathbf{s}_{\text{in}} \end{aligned} \quad (5.5.1)$$

where

$$\mathbf{s}_{\text{in}} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \text{ due to the laser polarization orientation} \quad (5.5.2)$$

and  $\mathbf{M}_{\text{test},I}(t_m)$  is a Mueller matrix object multiplied spatially by the illumination pattern  $I$  and can change in time  $t_m$ . The  $n_{v_k}$  are noise realizations sampled from whatever rotation wander distribution is chosen.  $\mathbf{P}(x_k, y_l)$  represents the calibrated and properly rotated micropolarizer array at pixel  $(k, l)$ , and is a linear polarizer matrix assumed to be [102]

$$\frac{1}{2} \begin{bmatrix} A & B \cos 2\theta & B \sin 2\theta & 0 \\ B \cos 2\theta & A \cos^2 2\theta + C \sin^2 2\theta & (A - C) \cos 2\theta \sin 2\theta & 0 \\ B \sin 2\theta & (A - C) \cos 2\theta \sin 2\theta & C \cos^2 2\theta + A \sin^2 2\theta & 0 \\ 0 & 0 & 0 & C \end{bmatrix} \quad (5.5.3)$$

$$\text{where} \quad (5.5.4)$$

$$A = T_{\text{max}} + T_{\text{min}}, \quad (5.5.5)$$



$$B = T_{\max} - T_{\min}, \quad (5.5.6)$$

$$C = 2\sqrt{T_{\max}T_{\min}} \quad (5.5.7)$$

at each pixel, where  $T_{\max}$  is the maximum transmission corresponding to  $a + b$  from Equation (5.4.1) and  $T_{\min}$  is the minimum transmission corresponding to  $a$  from Equation (5.4.1). The system channels can be derived by making  $\mathbf{M}_{\text{test},l}(t_m)$  constant over time and with one element at  $0 \leq i, j \leq 3$  equal to 1 and all other elements zero, then evaluating the system over all  $i, j$ . This gives the channel response of each specific Mueller matrix element. Finally, the  $s_0$  component of the final Stokes parameters are taken, and Gaussian (corresponding to read noise and thermal noise) and Poisson distribution samples are applied to simulate the CCD sensor noise. Noise strengths can be changed for the Gaussian component, and the Poisson component is computed per pixel and is dependent on the  $s_0$  magnitude at that pixel. This implicitly captures illumination variations and the effects of these variations on Poisson noise. The available noise types for rotation wander are Gaussian and uniform distributions. Each response for each time step and each pixel is then assembled into a 3-dimensional data cube and Fourier transformed. The channel weights can then be read off to assemble the  $\mathbf{Q}$  matrix.

The forward model was used with a region of interest of size  $980 \times 980$  pixels on the focal plane. On a computer with 32GB of RAM blocks of up to 72 temporal frames could be processed through the model in reasonable times of less than an hour. For this size,  $980 \times 980 \times 72$ , 32GB of RAM was necessary, attempts on computers with less RAM could not complete even when allowed to run for multiple days. The semi-empirical model was used to generate many of the simulated results in this dissertation, and was used to generate the  $\mathbf{Q}$  matrix required for reconstruction of Mueller matrix images from actual data shown later in this chapter.

## 5.6 Filtering and sampling

Prior to inversion to obtain the Mueller matrix images, the data *must be filtered in the Fourier domain* in order to obtain data "cubes" to apply the  $\mathbf{Q}$  matrix to. The unique channel locations are determined by the design of the system channel structure, and the sampling scheme. The micropolarizer array samples spatially right at

the Nyquist frequency, and the instrument is also designed to sample temporally right at the Nyquist frequency as well. This is done primarily to mitigate *spectral leakage*.

The filtering process described here implicitly sets the center of the filtered data to zero frequency (moved to “base band”), which is denoted *demodulation* in the communications literature, i.e. the data is both filtered and demodulated before unmixing.

### 5.6.1 Spectral leakage

In the polarimetric system design, we must temporally sample (via our focal plane array, i.e., a camera). We also are forced to use a window of samples, since we don’t have infinite time or computer memory. These windows must be integer periods of our carrier frequencies, otherwise we get spectral leakage. The seminal paper on spectral leakage and mitigation is by Harris [137]. When performing a discrete Fourier transform (DFT, including FFT), if the window length is not an integer multiple of the period, then frequencies from the sharp cutoff become introduced into the DFT.

The camera on our polarimeter has a maximum framerate of 28.4 frames per second (fps); however the following analysis addresses differences between sampling at 30fps and 32fps. The balanced bandwidth optimal system design for this particular type of system has sets of 9 temporal channels, including the DC or 0 channel, corresponding to the temporal modulation dimension. Due to Hermiticity, this implies that there are 4 unique temporal frequencies other than 0 for this design, and they are evenly spaced (balanced). The maximum frequency is limited by the physical rotation speed limitations of a retarder, and the sample rate of the camera (due to Nyquist).

The above puts constraints on the windows, frequencies, and sampling rates that we can use without spectral leakage. It may be desirable to sample at slightly above Nyquist; however in our case spectral leakage does not warrant doing so. Our window must be an integer number of periods of all 4 frequencies. Practically this means that we do not want to increase the sampling rate to something slightly over Nyquist, because of the interplay between Nyquist, spectral leakage, and channel frequency spacing.

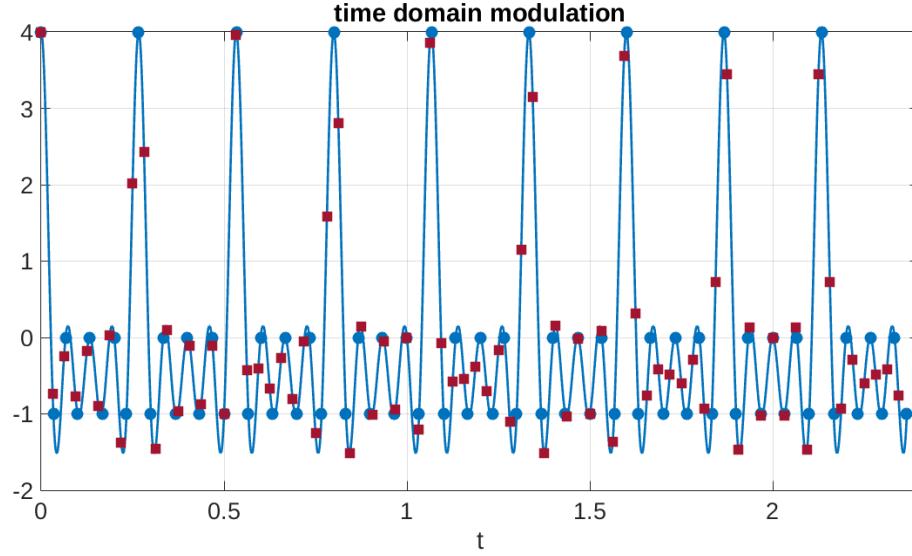


Figure 5.14: Sum of cosines curve sampled at the Nyquist frequency (blue dots) and slightly higher than the Nyquist frequency (red squares).

Figure 5.14 is a graph of

$$f(t) = \cos 2\pi(3.75t) + \cos 2\pi(7.5t) + \cos 2\pi(11.25t) + \cos 2\pi(15t). \quad (5.6.1)$$

The blue dots denote sampling at 30 fps (or once every  $\frac{1}{30}s$ ), and the red squares denote sampling at 32 fps.

When we take the Fourier transform  $\mathcal{F}\{f(t)\}_{t \rightarrow \nu}$  for the 30fps sampled case we obtain the graph in blue in Figure 5.15. This is the correct representation, there is no imaginary part, and the discrete  $\delta$ -functions are located precisely at the frequencies expected from the corresponding real (non-discrete) Fourier transform results. The red graphs show the spectral leakage issue, the imaginary part becomes non-zero, and the  $\delta$ -functions "split" into positive and negative parts. This creates problems when trying to reconstruct data after filtering. Note that we have the leftmost  $\delta$ -function at twice the magnitude of the others for the blue, 30fps, case and a missing  $\delta$ -function on the right side. This is due to the channel being sampled right at Nyquist. There are ways to mitigate spectral leakage when we reconstruct our data from the channel(s), but this makes writing the reconstruction code more difficult, more bug prone, etc. Suppose that we denote our sampling rate

$$f_s = n_s \frac{1}{s}, \quad n_s \in \mathbb{N}, \quad (5.6.2)$$

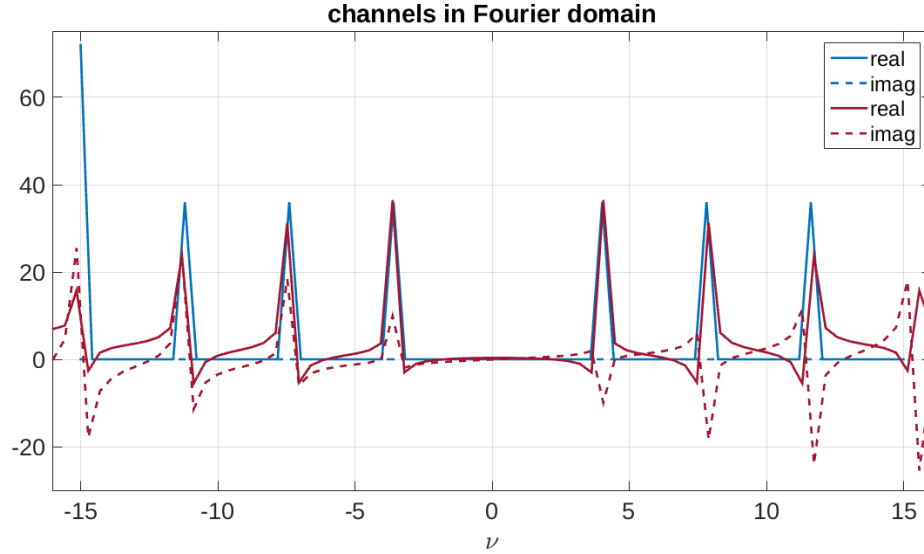


Figure 5.15: Fast Fourier transforms of the sampled sum of cosines curve, blue corresponds to the Nyquist sampling rate, and red corresponds to the slightly greater than Nyquist rate.

i.e.,  $n_s$  evenly spaced samples per second, where  $n_s$  is an integer. Now suppose that we also have defined an ordered list of channel frequencies,

$$\nu_0 \leq \nu_1 \leq \dots \leq \nu_k \leq \frac{f_s}{2}, \quad (5.6.3)$$

where the last frequency is necessarily less than or equal to the Nyquist frequency (caveat : when sampling at Nyquist there exist functions which cannot be reconstructed, the Nyquist condition is strictly less than, but we typically may use the equality in engineering applications). Given these assumptions, we now want to find the minimum window length which encloses an integer number of sampled periods, for each frequency.

The sample increment is

$$\Delta t = \frac{1}{f_s}, \quad (5.6.4)$$

so we need to find  $m$  such that

$$f_0(0) = f_0(m\Delta t) \quad (5.6.5)$$

$$f_1(0) = f_1(m\Delta t) \quad (5.6.6)$$

$$\vdots \quad (5.6.7)$$

$$f_k(0) = f_k(m\Delta t) \quad (5.6.8)$$

simultaneously, where  $f_j(t) = \cos 2\pi(\nu_j t)$  or  $f_j(t) = \sin 2\pi(\nu_j t)$ . This implies that for an arbitrary  $j$ ,

$$1 = \cos 2\pi(\nu_j m\Delta t) \quad \text{or} \quad 0 = \sin 2\pi(\nu_j m\Delta t) \quad (5.6.9)$$

which then leads to the fact

$$\nu_j m\Delta t = l_j, \quad l_j \in \mathbb{N} \quad (5.6.10)$$

$$\Rightarrow m = \frac{l_j}{\nu_j \Delta t}, \forall j \quad (5.6.11)$$

$$\Rightarrow m = \frac{l_j f_s}{\nu_j}, \forall j \quad (5.6.12)$$

This last equation reveals a key piece of information;  $f_s/\nu_j$  must be a rational number.

The last line above can then be rearranged to solve for  $l_j$ ,

$$l_j = \frac{m\nu_j}{f_s}, \quad (5.6.13)$$

since  $m$  is fixed and both  $\nu_j$  and  $f_s$  can be specified as shown previously, for example

$$\cos 2\pi(3.75t) + \cos 2\pi(7.5t) + \cos 2\pi(11.25t) + \cos 2\pi(15t), \quad (5.6.14)$$

sampled at 30fps. For our example, we have the following set of equations

$$l_0 = \frac{3.75m}{30} = \frac{m}{8} \quad (5.6.15)$$

$$(5.6.16)$$

$$l_1 = \frac{7.5m}{30} = \frac{m}{4} \quad (5.6.17)$$

$$(5.6.18)$$

$$l_2 = \frac{11.25m}{30} = \frac{3m}{8} \quad (5.6.19)$$

$$(5.6.20)$$

$$l_3 = \frac{15m}{8} = \frac{m}{2}. \quad (5.6.21)$$

By inspection we can see that  $m \geq 8$ , and plugging in  $m = 8$  does in fact satisfy all of the equations, resulting in  $l_0 = 1, l_1 = 2, l_2 = 3, l_3 = 4$ . For this example, the smallest window is 8 samples.

A spectral leakage free channel construction requires:

$$m = \frac{l_j f_s}{\nu_j}, \forall j, \quad (5.6.22)$$

where  $m \in \mathbb{N}$  is the minimum number of samples,  $f_s$  is the sampling rate,  $\nu_j$  is each of the channel frequencies,  $l_j$  corresponds to the number of periods in the window for each respective  $\nu_j$ , and finally  $f_s/\nu_j$  must be rational. Notice that, typically  $f_s$  is fixed, which implies that there can never be an irrational ratio between any of the channel frequencies if we want the channels to remain free of spectral leakage.

### 5.6.2 Filter types

Filters must be applied in the channel space around each unique channel ( $\delta$ -function) location to both suppress crosstalk and to isolate the Mueller object data which is convolved with that channel. If crosstalk is large, then the suppression effect of the filters is limited. If the spatio-temporal frequency distribution of specific objects or classes of objects is known then filters can be designed which reject frequencies not in the distribution. Recent work by LaCasse *et al* has shown a way to compute spectral density response functions for channeled polarimetric systems [138], which quantifies how the power spectral density of Stokes parameters is transformed in the channeled system. This work shows how the power spectral density transforms *for a given system* and then shows how to compute optimal filters. However, further work is needed to *design an optimal system* give the power spectral density.

For a general instrument, few assumptions can be made about the statistical properties (the power spectral densities) of the objects being polarimetrically imaged. We know that spatial frequencies are limited by the imaging optics, and that unaliased temporal frequencies are limited by the CCD sensor temporal bandwidth and integration times. As discussed in Chapter 3 the channel structure then limits the reconstructable spatio-temporal bandwidth. Many natural scenes have bandwidth power grouped around low frequencies in both space and time, so this assumption

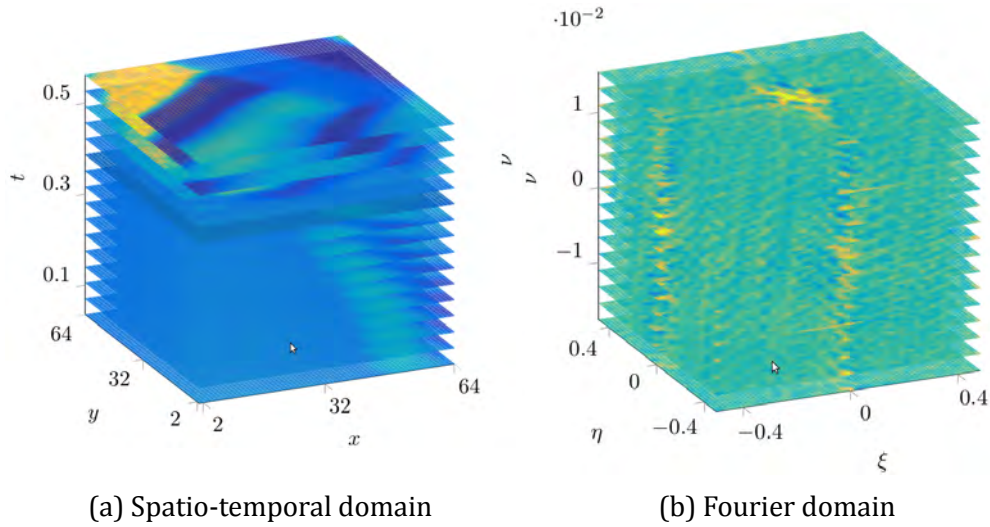


Figure 5.16: Space-time  $(x, y, t)$  modulated data "cube" transformed into Fourier domain.

will be made. Note that a different analysis is needed for something like a system which scans (pans) across a scene at some known rate, as this will change the assumption about the temporal frequencies being clustered around the zero frequency.

An example of the data prior to filtering is shown in Figure 5.16. A data "cube" of a spatial region of interest of my hand moving while holding a piece of ground glass is shown on the left, and the Fourier transform of those data is shown on the right. Since we have a general instrument, the filters we use should pass as much of the data as possible, up to the band limit. There are a variety of filters which accomplish this, from the simple rectangular filter, Gaussian and Gaussian like filters, Hann and Hamming windows, and various flat top filters such as the Tukey, etc. I chose a flat top filter which I came across in mathematical analysis: the Planck-taper window [139]. I chose Planck-taper filters because

- They are functions with compact support.
- They are  $C^\infty$ , i.e. infinitely differentiable.
- They have a flat top and can approach the *rect()* function
- The transition from 0 to 1 is controlled by a single parameter,  $\epsilon$ .

The Planck-taper filter is defined in 1 dimension over the interval  $[-T, T]$  by

$$f(t, \epsilon) = \begin{cases} 0 & t \leq -T \\ \frac{1}{\exp\left(\frac{T-T(1-2\epsilon)}{t-T} + \frac{T-T(1-2\epsilon)}{t+T(1-2\epsilon)}\right)+1} & -T < t < -T(1-2\epsilon) \\ 1 & -T(1-2\epsilon) \leq t \leq T(1-2\epsilon) \\ \frac{1}{\exp\left(\frac{T(1-2\epsilon)-T}{t-T(1-2\epsilon)} + \frac{T(1-2\epsilon)-T}{t-T}\right)+1} & T(1-2\epsilon) < t < T \\ 0 & T \leq t \end{cases} \quad (5.6.23)$$

which is easily adapted to 3 dimensions by substituting a radius from each channel location for  $t$ . I use a weighted radius to form ellipsoidal filters for the spatio-temporal system since the channel space is not equally scaled. The data presented later in this chapter were filtered with 3 dimensional Planck taper filters with  $\epsilon = 0.25$ .

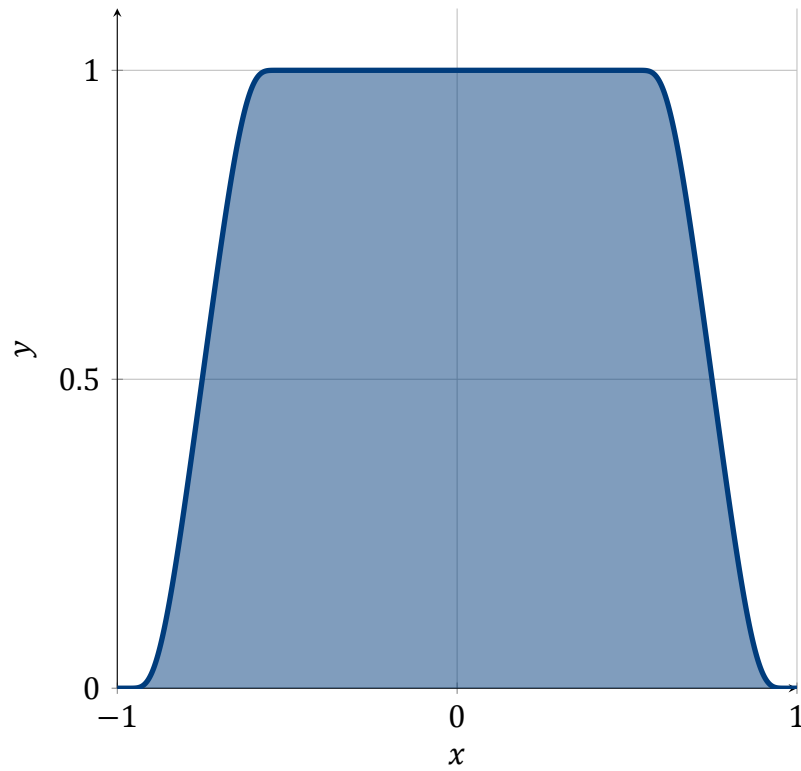
### 5.6.3 Implementation

Difficulties arose during the implementation of filtering and application of the  $\mathbf{Q}^+$  matrix. As shown in Figure 5.17 the 3 dimension filters are only complete for the center column of the data in the spatio-temporal channel space. Due to sampling right at the Nyquist frequency, the resulting data and filters only cover half (and for two channel locations, a quarter) of the space. To apply  $\mathbf{Q}^+$  properly then requires reconstruction of the other half (or other three quarters of the space) using the Hermiticity relations. This reconstruction is straight forward in theory, but is error prone in practice, tedious, and subject to errors when writing the filtering code. After extensive testing the Hermitian extension code was verified to be correct and produced results which are correct when using simulated data from the semi-empirical model. Filtered data of an entire  $980 \times 980 \times 72$  for  $\xi, \eta, \nu$  data "cube" is shown in Figure 5.18.

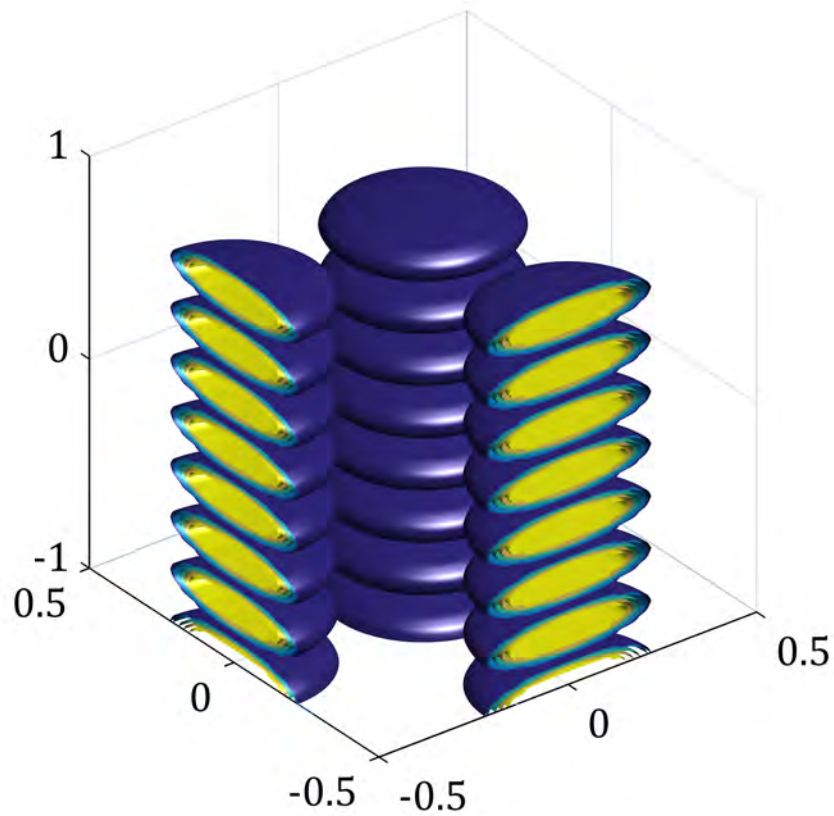
## 5.7 Validation

After calibration was finished, a small validation procedure was undertaken. This validation procedure is not exhaustive and a full procedure will need to be designed in the future. A mirror was placed across the room from the polarimeter and light was directed so that it was imaged onto the focal plane. A defocus was introduced





(a) 1-dimensional



(b) 3-dimensional

Figure 5.17: Planck-taper filters used for our system. In (b) yellow represents a magnitude of 1 and blue represents a magnitude close to, but greater than, 0.

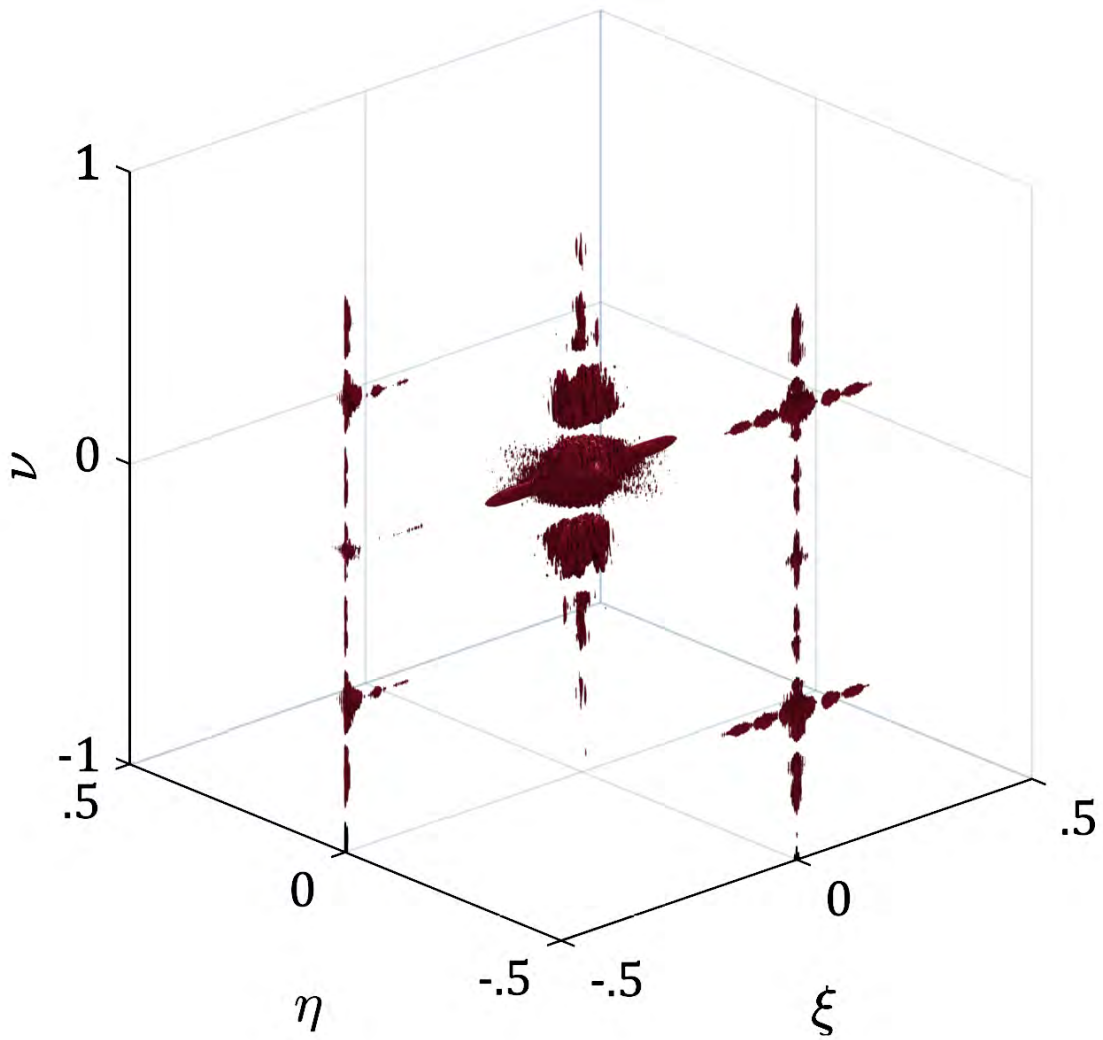


Figure 5.18: Filtered data in the Fourier domain, filtered via Planck taper filters. The figure depicts an isosurface (surface of equal magnitude) to show the data.

via the lens focusing mechanism on the PSA in order to minimize detector pixel saturation, and the laser was turned to minimum power, about  $4mW$ . A polarizer was introduced into the beam path, and rotated until it was at maximum extinction with the source, this ensured that the polarizers transmission axis was nominally aligned with the  $x$ -axis of the instruments polarization reference frame. Raw data was then collected and compared to data generated by the semi-empirical model for an ideal linear polarizer Mueller matrix. The comparison was done in the channel space. Both sets of channel data were normalized in magnitude by the zero frequency component magnitude in order to make them comparable.

As shown in Figures 5.19 and 5.20 the channel structures are close to one another, providing some validation of the instruments actual channel structure. There remain some differences which need to be analyzed further in the future, however the channel structure generated by the semi-empirical model can be used to attempt to reconstruct data, and is used for the results in the next section.

## 5.8 Imaging results

After calibration was finished and the semi-empirical model verified, all of the hardware and software components were in place to take data and compute Mueller matrix images. The procedure for Mueller matrix image reconstruction is exactly the same as for the reconstruction of the semi-empirical model output: collect a data stream of 72 images, Fourier transform the 3 dimensional data "cube," filter the data through the Planck taper filters, and recombine the data linearly as prescribed by the  $Q^+$  matrix generated by the semi-empirical model. As stated previously, this reconstruction cannot yet be accomplished in real time, the reconstruction algorithm is written in Matlab which is not performant enough to accomplish real time speeds.

The data were collected was at a distance of approximately  $4 - 5m$  and the polarimeter was stationary (i.e. not panned or rotated while the images were taken). After collection began I moved my hand in the frame while holding a piece of ground glass. Laser power was set to  $\sim 10mW$  and integration time was  $30ms$ . Position data were carefully checked to ensure that there were no missed camera triggers for the data used for reconstruction.

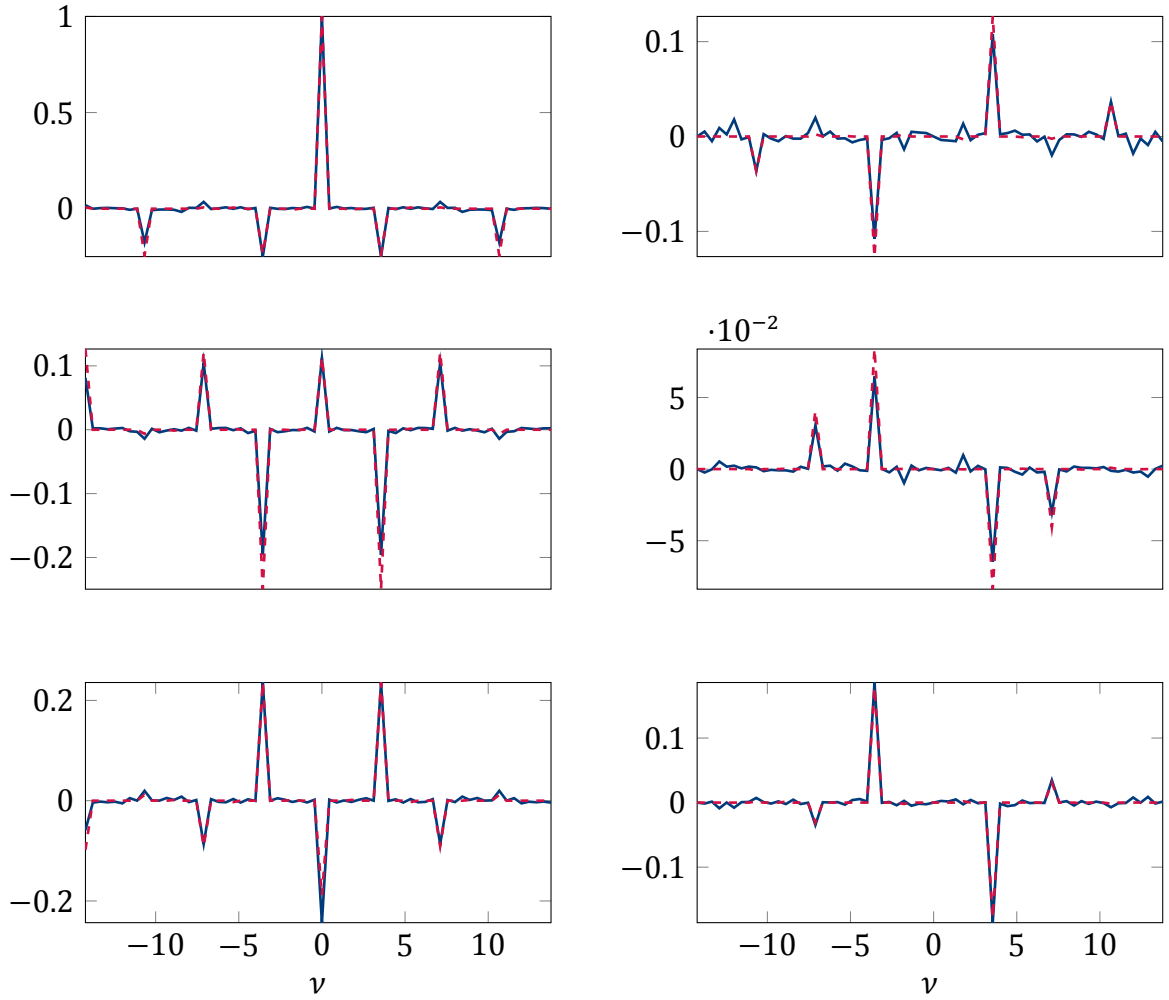


Figure 5.19: Polarizer channel response from the semi-empirical model (red, dashed) versus the physical instrument (blue, solid). The left column are the real parts and the right column are the imaginary parts. The first row is taken along the center column line,  $\nu$ , of the channel structure where  $\xi = \eta = 0$ , the second row is taken along the line  $\xi = 0, \eta = -0.5$  and the third row is taken along the line  $\xi = -0.5, \eta = 0$ . The channels match well, but there are some differences.

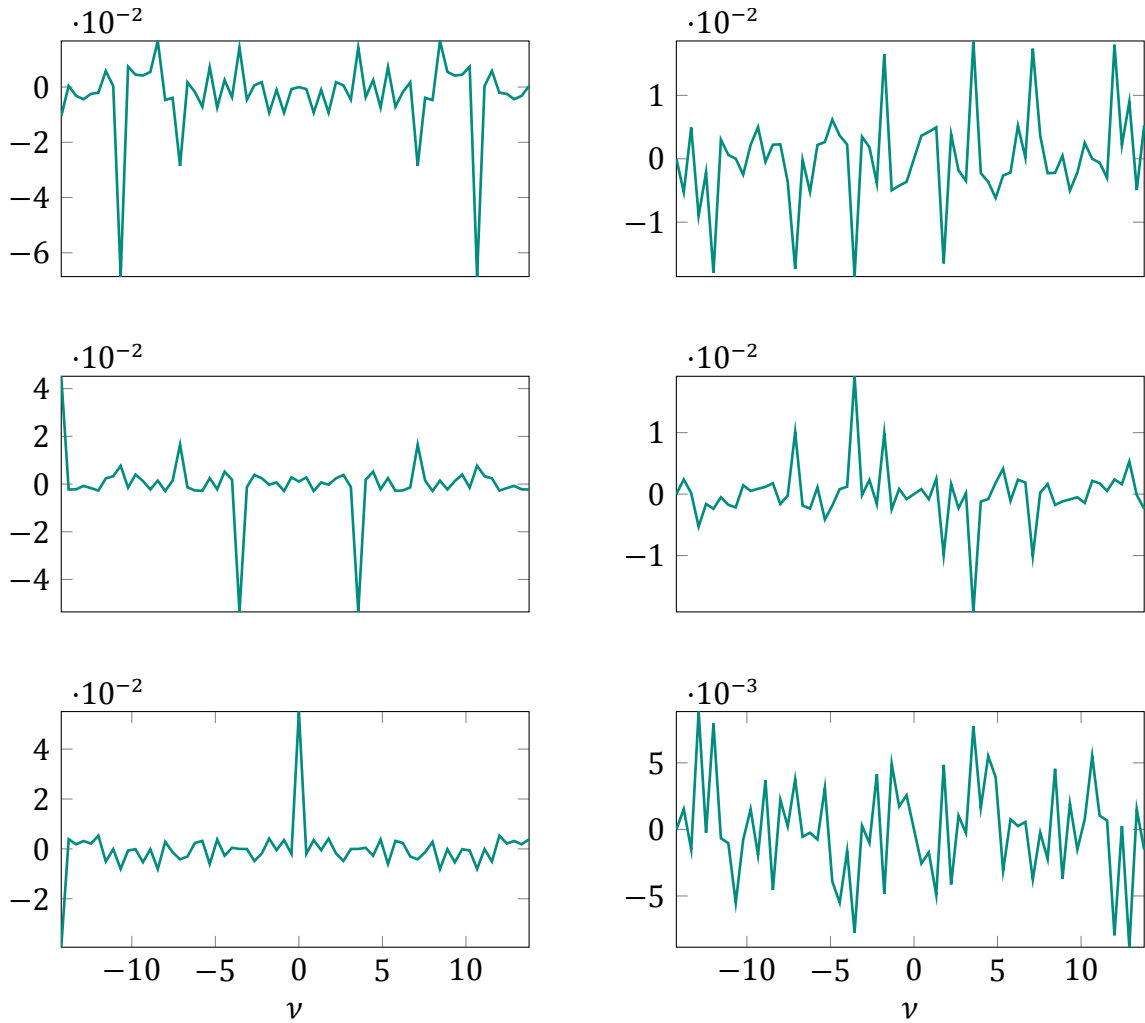


Figure 5.20: Polarizer channel response differences between the semi-empirical (SE) model and the physical instrument, the physical instrument data was subtracted from the SE model data. The left column are the real parts and the right column are the imaginary parts. The first row is taken along the center column line,  $\nu$ , of the channel structure where  $\xi = \eta = 0$ , the second row is taken along the line  $\xi = 0, \eta = -0.5$  and the third row is taken along the line  $\xi = -0.5, \eta = 0$ . The channels match well, but there are some differences.

Figure 5.21: Moving hand Mueller matrix reconstruction. All elements except for  $m_{00}$  are normalized by  $m_{00}$ . Animated in the electronic version.

Figure 5.22: Moving hand reconstruction. On the left is the reconstruction of  $m_{01}$  and on the right is a downsampled raw camera image. Animated in the electronic version.

The complete Mueller reconstruction output is shown in Figure 5.21, and in the electronic version is animated at the actual framerate of  $3.5fps$ . A comparison of the reconstructed  $m_{01}$  and a downsampled raw image is shown in Figure 5.22 and the animation is at the raw camera frame rate of  $28.4fps$ . When my hand movements exceed the temporal bandwidth, the effects of crosstalk can be seen as “ringing” edges in the reconstructed Mueller images.

There appear to be some inconsistencies in the reconstructed images that are not physically or intuitively correct. For instance, for a scattering surface like my hand and ground glass, it is unlikely that my hand is actually diattenuating like the first row shown in Figure 5.21 suggests. I have not shown it to be the case, but I strongly suspect that this artifact is due to the possible strong crosstalk into the  $m_{01}, m_{02}, m_{03}$  channels by the zero frequency component. The zero frequency component has a large magnitude and so could produce more crosstalk into the temporally adjacent channels, which consist of the channels for the  $m_{01}, m_{02}, m_{03}$  elements. See the first row of the system in Figure 3.11 for the details about this potential issue.

Mitigation of this crosstalk could be eliminated if an independent estimate of  $m_{00}$

could be obtained. A first attempt can be made by summing over an ideal micropolarizer “superpixel.” Recall

$$s_0(\mathbf{x}) = \sum_{i=0}^3 \sum_{j=0}^3 a_{0i}(\mathbf{x}) s_j(\mathbf{x}) m_{ij}(\mathbf{x})$$

which implies for an ideal micropolarizer “superpixel”

$$\mathbf{a}_{0^\circ} = \begin{bmatrix} 1 \\ RR_{11} \\ RR_{12} \\ RR_{13} \end{bmatrix}, \mathbf{a}_{45^\circ} = \begin{bmatrix} 1 \\ RR_{21} \\ RR_{22} \\ RR_{23} \end{bmatrix}, \mathbf{a}_{-45^\circ} = \begin{bmatrix} 1 \\ -RR_{21} \\ -RR_{22} \\ -RR_{23} \end{bmatrix}, \mathbf{a}_{90^\circ} = \begin{bmatrix} 1 \\ -RR_{11} \\ -RR_{12} \\ -RR_{13} \end{bmatrix}. \quad (5.8.1)$$

explicitly. If the pixels of the “superpixel” are then added together we obtain:

$$\begin{aligned} & \sum_{j=0}^3 s_j(t) m_{0j}(\mathbf{x}_0) + RR_{11}(t) \sum_{j=0}^3 s_j(t) m_{1j}(\mathbf{x}_0) \\ & \quad + RR_{12}(t) \sum_{j=0}^3 s_j(t) m_{2j}(\mathbf{x}_0) + RR_{13}(t) \sum_{j=0}^3 s_j(t) m_{3j}(\mathbf{x}_0) \\ & + \sum_{j=0}^3 s_j(t) m_{0j}(\mathbf{x}_1) + RR_{21}(t) \sum_{j=0}^3 s_j(t) m_{1j}(\mathbf{x}_1) \\ & \quad + RR_{22}(t) \sum_{j=0}^3 s_j(t) m_{2j}(\mathbf{x}_1) + RR_{23}(t) \sum_{j=0}^3 s_j(t) m_{3j}(\mathbf{x}_1) \\ & + \sum_{j=0}^3 s_j(t) m_{0j}(\mathbf{x}_2) - RR_{21}(t) \sum_{j=0}^3 s_j(t) m_{1j}(\mathbf{x}_2) \\ & \quad - RR_{22}(t) \sum_{j=0}^3 s_j(t) m_{2j}(\mathbf{x}_2) - RR_{23}(t) \sum_{j=0}^3 s_j(t) m_{3j}(\mathbf{x}_2) \\ & + \sum_{j=0}^3 s_j(t) m_{0j}(\mathbf{x}_3) - RR_{11}(t) \sum_{j=0}^3 s_j(t) m_{1j}(\mathbf{x}_3) \\ & \quad - RR_{12}(t) \sum_{j=0}^3 s_j(t) m_{2j}(\mathbf{x}_3) - RR_{13}(t) \sum_{j=0}^3 s_j(t) m_{3j}(\mathbf{x}_3), \quad (5.8.2) \end{aligned}$$

where  $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$  represent the same time points  $t$ , but different positions within the “superpixel.” If the assumption that  $\mathbf{x}_0 \approx \mathbf{x}_1 \approx \mathbf{x}_2 \approx \mathbf{x}_3 \approx \bar{\mathbf{x}}$  is made then



Equation (5.8.2) can be simplified to

$$4 \sum_{j=0}^3 s_j(t) m_{0j}(\bar{\mathbf{x}}). \quad (5.8.3)$$

This equation does not give  $m_{00}$ , but instead the inner product of the input carrier vector with the first row of the Mueller matrix object. Further analysis is needed to extract  $m_{00}$  alone.

## 5.9 Discussion

I designed, built, and calibrated an instrument to test one of the theoretical systems developed in Chapter 3. A physical design was implemented and custom parts specified and assembled. A partial software stack was developed with rotation stage and camera control modules being completely implemented. The reconstruction module is templated and working in Matlab, but needs to be transferred to a C++/CUDA implementation to allow the instrument to reconstruct Mueller matrix images in real or reasonable times. The overall software architecture is finished, but the GUI and classification modules need to be finished in the future.

Some preliminary validation showed the the system design, system modeling, and the physical system nominally agree with one another. More robust validation is needed in the future, however. Additionally, the microgrid array calibration was accomplished via a diffuse piece of paper, a proper Lambertian source, e.g. an integrating sphere, needs to be used to recalibrate the microgrid array with greater accuracy. This is planned in the near future.

Finally, reconstruction of data from the instrument shows that crosstalk remains an issue that polarimetric system designers must contend with, and there are also some artifacts which remain to be resolved, i.e. the first row of the reconstructed data in Figure 5.21 does not appear to be physically correct and further validation and analysis is needed.

## CHAPTER 6

### Closing Remarks

Using the tools of linear systems for polarimetric systems allows both the physical system limitations to become lucid and to design polarimetric systems in a systematic way. In this dissertation I have shown how to optimize spatio-temporal systems for bandwidth using the inherent geometry present in channeled systems. I have shown the results of this optimization, which allow for as much as a 343% improvement in temporal bandwidth over conventional Mueller matrix polarimeters.

In real systems, noise and systematic error must be addressed, and can cause reconstruction issues for polarimetric instruments. There is not much in the literature addressing the specific effects of noise on polarimetric channeled systems, and I begin to address these effects in Chapter 4. Noise effects include noise from spatially non-uniform illumination sources like lasers. Furthermore, systematic error in channeled systems has only been addressed previously for a specific system in a single journal article. In Chapter 4 I address systematic errors for channeled systems in a more general way, and enumerate the ways specific types of systematic errors affect channeled systems. This analysis has both general contributions, and contributions specific to spatio-temporal systems.

Lastly, there is a dearth of Mueller matrix polarimeters being used for remote sensing purposes, which is mainly due to the difficulty of manufacturing fast instruments. I have designed and built a fast, portable, Mueller matrix polarimeter to test the theoretical ideas outlined in Chapter 3. This system achieves  $3.5fps$  when a camera at  $28.4fps$  is used as the base camera. The system validates the theory in Chapter 3, and shows that the bandwidth optimizations result in faster Mueller matrix polarimeters. There remain issues, however. The  $m_{00}$  channel has a very large magnitude, resulting in large crosstalk into the channels adjacent to  $m_{00}$  along  $\nu$ , which corrupts the resulting Mueller matrix image reconstruction. This crosstalk issue must be mitigated in order for these types of instruments to be viable for the field.

Continuation of both channeled system design and bandwidth optimization in the

future will include a general design framework whereby a system design can choose channels directly in the channel space in an arbitrary domain, and the framework will produce polarization elements which can produce those channels. The instrument needs further testing and more development on the control systems, including implementing real time reconstruction.

## APPENDIX A

## Derivations of Quad-retarder + micropolarizer array equations

## A.1 Two retarder Mueller matrix

The two retarder Mueller matrix  $\mathbf{RR}(\nu_k, \nu_l, \epsilon_k, \epsilon_l; \delta_k, \delta_l) = \mathbf{R}_{\delta_l} \cdot \mathbf{R}_{\delta_k}$  has the following form :

$$RR_{00} = 1, \quad RR_{01} = RR_{02} = RR_{03} = RR_{10} = RR_{20} = RR_{30} = 0 \quad (\text{A.1.1})$$

$$\begin{aligned} RR_{11} = & \left[ \cos^2 \frac{\delta_l}{2} + \sin^2 \frac{\delta_l}{2} \cos 8\pi(\nu_l t + \epsilon_l) \right] \left[ \cos^2 \frac{\delta_k}{2} + \sin^2 \frac{\delta_k}{2} \cos 8\pi(\nu_k t + \epsilon_k) \right] \\ & + \sin^2 \frac{\delta_l}{2} \sin^2 \frac{\delta_k}{2} \sin 8\pi(\nu_l t + \epsilon_l) \sin 8\pi(\nu_k t + \epsilon_k) \\ & - \sin \delta_l \sin \delta_k \sin 4\pi(\nu_l t + \epsilon_l) \sin 4\pi(\nu_k t + \epsilon_k) \end{aligned} \quad (\text{A.1.2})$$

$$\begin{aligned} RR_{12} = & \left[ \cos^2 \frac{\delta_l}{2} + \sin^2 \frac{\delta_l}{2} \cos 8\pi(\nu_l t + \epsilon_l) \right] \sin^2 \frac{\delta_k}{2} \sin 8\pi(\nu_k t + \epsilon_k) \\ & + \sin^2 \frac{\delta_l}{2} \sin 8\pi(\nu_l t + \epsilon_l) \left[ \cos^2 \frac{\delta_k}{2} - \sin^2 \frac{\delta_k}{2} \cos 8\pi(\nu_k t + \epsilon_k) \right] \\ & + \sin \delta_l \sin \delta_k \sin 4\pi(\nu_l t + \epsilon_l) \cos 4\pi(\nu_k t + \epsilon_k) \end{aligned} \quad (\text{A.1.3})$$

$$\begin{aligned} RR_{13} = & - \left[ \cos^2 \frac{\delta_l}{2} + \sin^2 \frac{\delta_l}{2} \cos 8\pi(\nu_l t + \epsilon_l) \right] \sin \delta_k \sin 4\pi(\nu_k t + \epsilon_k) \\ & + \sin^2 \frac{\delta_l}{2} \sin \delta_k \sin 8\pi(\nu_l t + \epsilon_l) \cos 4\pi(\nu_k t + \epsilon_k) \\ & - \sin \delta_l \cos \delta_k \sin 4\pi(\nu_l t + \epsilon_l) \end{aligned} \quad (\text{A.1.4})$$

$$\begin{aligned} RR_{21} = & \sin^2 \frac{\delta_l}{2} \sin 8\pi(\nu_l t + \epsilon_l) \left[ \cos^2 \frac{\delta_k}{2} + \sin^2 \frac{\delta_k}{2} \cos 8\pi(\nu_k t + \epsilon_k) \right] \\ & + \left[ \cos^2 \frac{\delta_l}{2} - \sin^2 \frac{\delta_l}{2} \cos 8\pi(\nu_l t + \epsilon_l) \right] \sin^2 \frac{\delta_k}{2} \sin 8\pi(\nu_k t + \epsilon_k) \\ & + \sin \delta_l \sin \delta_k \cos 4\pi(\nu_l t + \epsilon_l) \sin 4\pi(\nu_k t + \epsilon_k) \end{aligned} \quad (\text{A.1.5})$$

$$\begin{aligned}
RR_{22} = & \sin^2 \frac{\delta_l}{2} \sin^2 \frac{\delta_k}{2} \sin 8\pi(\nu_l t + \epsilon_l) \sin 8\pi(\nu_k t + \epsilon_k) \\
& + \left[ \cos^2 \frac{\delta_l}{2} - \sin^2 \frac{\delta_l}{2} \cos 8\pi(\nu_l t + \epsilon_l) \right] \left[ \cos^2 \frac{\delta_k}{2} - \sin^2 \frac{\delta_k}{2} \cos 8\pi(\nu_k t + \epsilon_k) \right] \\
& - \sin \delta_l \sin \delta_k \cos 4\pi(\nu_l t + \epsilon_l) \cos 4\pi(\nu_k t + \epsilon_k)
\end{aligned} \tag{A.1.6}$$

$$\begin{aligned}
RR_{23} = & -\sin^2 \frac{\delta_l}{2} \sin \delta_k \sin 8\pi(\nu_l t + \epsilon_l) \sin 4\pi(\nu_k t + \epsilon_k) \\
& + \left[ \cos^2 \frac{\delta_l}{2} - \sin^2 \frac{\delta_l}{2} \cos 8\pi(\nu_l t + \epsilon_l) \right] \sin \delta_k \cos 4\pi(\nu_k t + \epsilon_k) \\
& + \sin \delta_l \cos \delta_k \cos 4\pi(\nu_l t + \epsilon_l)
\end{aligned} \tag{A.1.7}$$

$$\begin{aligned}
RR_{31} = & \sin \delta_l \sin 4\pi(\nu_l t + \epsilon_l) \left[ \cos^2 \frac{\delta_k}{2} + \sin^2 \frac{\delta_k}{2} \cos 8\pi(\nu_k t + \epsilon_k) \right] \\
& - \sin \delta_l \sin^2 \frac{\delta_k}{2} \cos 4\pi(\nu_l t + \epsilon_l) \sin 8\pi(\nu_k t + \epsilon_k) \\
& + \cos \delta_l \sin \delta_k \sin 4\pi(\nu_k t + \epsilon_k)
\end{aligned} \tag{A.1.8}$$

$$\begin{aligned}
RR_{32} = & \sin \delta_l \sin^2 \frac{\delta_k}{2} \sin 4\pi(\nu_l t + \epsilon_l) \sin 8\pi(\nu_k t + \epsilon_k) \\
& - \sin \delta_l \cos 4\pi(\nu_l t + \epsilon_l) \left[ \cos^2 \frac{\delta_k}{2} - \sin^2 \frac{\delta_k}{2} \cos 8\pi(\nu_k t + \epsilon_k) \right] \\
& - \cos \delta_l \sin \delta_k \cos 4\pi(\nu_k t + \epsilon_k)
\end{aligned} \tag{A.1.9}$$

$$\begin{aligned}
RR_{33} = & -\sin \delta_l \sin \delta_k \sin 4\pi(\nu_l t + \epsilon_l) \sin 4\pi(\nu_k t + \epsilon_k) \\
& - \sin \delta_l \sin \delta_k \cos 4\pi(\nu_l t + \epsilon_l) \cos 4\pi(\nu_k t + \epsilon_k) \\
& + \cos \delta_l \cos \delta_k
\end{aligned} \tag{A.1.10}$$

where  $RR_{kl}$  are elements of  $\mathbf{RR}$  and we have dropped the arguments for brevity.

## A.2 Fourier transform of two retarder Mueller matrix elements

We only explicitly calculate the Fourier transform for  $RR_{11}, RR_{12}, RR_{13}, RR_{21}, RR_{22}, RR_{23}, RR_{31}$  since they are the relevant elements to

our specific application. The calculations are tedious, but straightforward.

$$\begin{aligned}
\mathcal{F}\{RR_{11}\}_{t \rightarrow \nu} &= \cos^2 \frac{\delta_l}{2} \cos^2 \frac{\delta_k}{2} \delta(\nu) \\
&+ \frac{e^{2\pi i \nu \epsilon_k}}{2} \cos^2 \frac{\delta_l}{2} \sin^2 \frac{\delta_k}{2} [\delta(\nu - 4\nu_k) + \delta(\nu + 4\nu_k)] \\
&+ \frac{e^{2\pi i \nu \epsilon_l}}{2} \sin^2 \frac{\delta_l}{2} \cos^2 \frac{\delta_k}{2} [\delta(\nu - 4\nu_l) + \delta(\nu + 4\nu_l)] \\
&+ \frac{e^{2\pi i \nu (\epsilon_l - \epsilon_k)}}{2} \sin^2 \frac{\delta_l}{2} \sin^2 \frac{\delta_k}{2} [\delta(\nu - 4[\nu_l - \nu_k]) + \delta(\nu + 4[\nu_l - \nu_k])] \\
&- \frac{e^{2\pi i \nu (\epsilon_l - \epsilon_k)}}{4} \sin \delta_l \sin \delta_k [\delta(\nu - 2[\nu_l - \nu_k]) + \delta(\nu + 2[\nu_l - \nu_k])] \\
&+ \frac{e^{2\pi i \nu (\epsilon_l + \epsilon_k)}}{4} \sin \delta_l \sin \delta_k [\delta(\nu - 2[\nu_l + \nu_k]) + \delta(\nu + 2[\nu_l + \nu_k])]
\end{aligned} \tag{A.2.1}$$

$$\begin{aligned}
\mathcal{F}\{RR_{21}\}_{t \rightarrow \nu} &= -\frac{ie^{2\pi i \nu \epsilon_k}}{2} \cos^2 \frac{\delta_l}{2} \sin^2 \frac{\delta_k}{2} [\delta(\nu - 4\nu_k) - \delta(\nu + 4\nu_k)] \\
&- \frac{ie^{2\pi i \nu \epsilon_l}}{2} \sin^2 \frac{\delta_l}{2} \cos^2 \frac{\delta_k}{2} [\delta(\nu - 4\nu_l) - \delta(\nu + 4\nu_l)] \\
&- \frac{ie^{2\pi i \nu (\epsilon_l - \epsilon_k)}}{2} \sin^2 \frac{\delta_l}{2} \sin^2 \frac{\delta_k}{2} [\delta(\nu - 4[\nu_l - \nu_k]) - \delta(\nu + 4[\nu_l - \nu_k])] \\
&- \frac{ie^{2\pi i \nu (\epsilon_l - \epsilon_k)}}{4} \sin \delta_l \sin \delta_k [\delta(\nu - 2[\nu_l - \nu_k]) - \delta(\nu + 2[\nu_l - \nu_k])] \\
&+ \frac{ie^{2\pi i \nu (\epsilon_l + \epsilon_k)}}{4} \sin \delta_l \sin \delta_k [\delta(\nu - 2[\nu_l + \nu_k]) - \delta(\nu + 2[\nu_l + \nu_k])]
\end{aligned} \tag{A.2.2}$$

$$\begin{aligned}
\mathcal{F}\{RR_{12}\}_{t \rightarrow \nu} &= \frac{ie^{2\pi i \nu \epsilon_k}}{2} \cos^2 \frac{\delta_l}{2} \sin^2 \frac{\delta_k}{2} [\delta(\nu - 4\nu_k) - \delta(\nu + 4\nu_k)] \\
&+ \frac{ie^{2\pi i \nu \epsilon_l}}{2} \sin^2 \frac{\delta_l}{2} \cos^2 \frac{\delta_k}{2} [\delta(\nu - 4\nu_l) - \delta(\nu + 4\nu_l)] \\
&- \frac{ie^{2\pi i \nu (\epsilon_l - \epsilon_k)}}{2} \sin^2 \frac{\delta_l}{2} \sin^2 \frac{\delta_k}{2} [\delta(\nu - 4[\nu_l - \nu_k]) - \delta(\nu + 4[\nu_l - \nu_k])] \\
&+ \frac{ie^{2\pi i \nu (\epsilon_l - \epsilon_k)}}{4} \sin \delta_l \sin \delta_k [\delta(\nu - 2[\nu_l - \nu_k]) - \delta(\nu + 2[\nu_l - \nu_k])] \\
&+ \frac{ie^{2\pi i \nu (\epsilon_l + \epsilon_k)}}{4} \sin \delta_l \sin \delta_k [\delta(\nu - 2[\nu_l + \nu_k]) - \delta(\nu + 2[\nu_l + \nu_k])]
\end{aligned}$$

$$\begin{aligned}
\mathcal{F}\{RR_{22}\}_{t \rightarrow \nu} &= \cos^2 \frac{\delta_l}{2} \cos^2 \frac{\delta_k}{2} \delta(\nu) \\
&- \frac{e^{2\pi i \nu \epsilon_k}}{2} \cos^2 \frac{\delta_l}{2} \sin^2 \frac{\delta_k}{2} [\delta(\nu - 4\nu_k) + \delta(\nu + 4\nu_k)]
\end{aligned} \tag{A.2.3}$$

$$\begin{aligned}
& -\frac{e^{2\pi i v \epsilon_l}}{2} \sin^2 \frac{\delta_l}{2} \cos^2 \frac{\delta_k}{2} \left[ \delta(\nu - 4\nu_l) + \delta(\nu + 4\nu_l) \right] \\
& + \frac{e^{2\pi i v (\epsilon_l - \epsilon_k)}}{2} \sin^2 \frac{\delta_l}{2} \sin^2 \frac{\delta_k}{2} \left[ \delta(\nu - 4[\nu_l - \nu_k]) + \delta(\nu + 4[\nu_l - \nu_k]) \right] \\
& - \frac{e^{2\pi i v (\epsilon_l - \epsilon_k)}}{4} \sin \delta_l \sin \delta_k \left[ \delta(\nu - 2[\nu_l - \nu_k]) + \delta(\nu + 2[\nu_l - \nu_k]) \right] \\
& - \frac{e^{2\pi i v (\epsilon_l + \epsilon_k)}}{4} \sin \delta_l \sin \delta_k \left[ \delta(\nu - 2[\nu_l + \nu_k]) + \delta(\nu + 2[\nu_l + \nu_k]) \right] \\
\mathcal{F} \{RR_{13}\}_{t \rightarrow \nu} &= -\frac{ie^{2\pi i v \epsilon_k}}{2} \cos^2 \frac{\delta_l}{2} \sin \delta_k \left[ \delta(\nu - 2\nu_k) - \delta(\nu + 2\nu_k) \right] \\
& - \frac{ie^{2\pi i v \epsilon_l}}{2} \sin \delta_l \cos \delta_k \left[ \delta(\nu - 2\nu_l) - \delta(\nu + 2\nu_l) \right] \\
& + \frac{ie^{2\pi i v [2\epsilon_l - \epsilon_k]}}{2} \sin^2 \frac{\delta_l}{2} \sin \delta_k \left[ \delta(\nu - 2[2\nu_l - \nu_k]) - \delta(\nu + 2[2\nu_l - \nu_k]) \right]
\end{aligned} \tag{A.2.4}$$

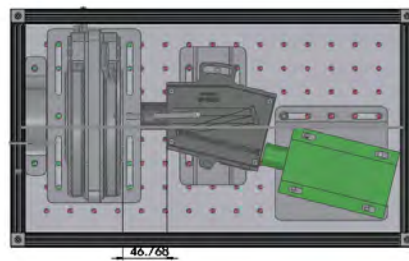
$$\begin{aligned}
\mathcal{F} \{RR_{23}\}_{t \rightarrow \nu} &= \frac{e^{2\pi i v \epsilon_k}}{2} \cos^2 \frac{\delta_l}{2} \sin \delta_k \left[ \delta(\nu - 2\nu_k) + \delta(\nu + 2\nu_k) \right] \\
& + \frac{e^{2\pi i v \epsilon_l}}{2} \sin \delta_l \cos \delta_k \left[ \delta(\nu - 2\nu_l) + \delta(\nu + 2\nu_l) \right] \\
& - \frac{e^{2\pi i v [2\epsilon_l - \epsilon_k]}}{2} \sin^2 \frac{\delta_l}{2} \sin \delta_k \left[ \delta(\nu - 2[2\nu_l - \nu_k]) + \delta(\nu + 2[2\nu_l - \nu_k]) \right]
\end{aligned} \tag{A.2.5}$$

$$\begin{aligned}
\mathcal{F} \{RR_{31}\}_{t \rightarrow \nu} &= -\frac{ie^{2\pi i v \epsilon_k}}{2} \cos \delta_l \sin \delta_k \left[ \delta(\nu - 2\nu_k) - \delta(\nu + 2\nu_k) \right] \\
& - \frac{ie^{2\pi i v \epsilon_l}}{2} \sin \delta_l \cos^2 \frac{\delta_k}{2} \left[ \delta(\nu - 2\nu_l) - \delta(\nu + 2\nu_l) \right] \\
& - \frac{ie^{2\pi i v [\epsilon_l - 2\epsilon_k]}}{2} \sin \delta_l \sin^2 \frac{\delta_k}{2} \left[ \delta(\nu - 2[\nu_l - 2\nu_k]) - \delta(\nu + 2[\nu_l - 2\nu_k]) \right]
\end{aligned} \tag{A.2.6}$$

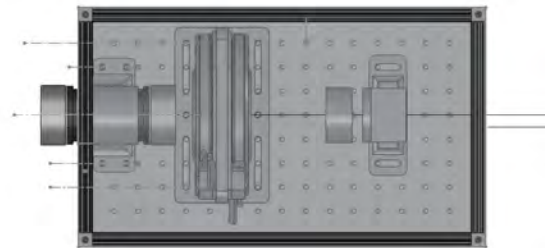
## APPENDIX B

### Instrument details

We include here some opto-mechanical details, the Source Assembly is the PSG, and the Receiver Assembly is the PSA. Fig B.1 shows a top view of the instrument design.



(a) Top view of PSG



(b) Top view of PSA

Figure B.1: Renderings of the source and receiver for HyDMIP.

#### B.0.1 PSG custom parts



[illegible]

Figure B.3: PSG beam expander mount.

### B.0.3 Mounting and display hardware

A  
 B  
 5 x  $\varnothing$  0.07 THRU ALL  
 $\varnothing$  0.14  $\nabla$  0.05  
 $\varnothing$  .0025 A B C  
 $\varnothing$  .0025 A B C  
 This face is machined from existing part.  
 (Existing Holes) 5 x  $\varnothing$  0.05  $\nabla$  0.17  
 M1.6x0.35 - 6H  $\nabla$  0.13  
 2 x .886  
 2 x .512  
 1 .024  
 2 x .724  
 2 x .724

ALL ANGLES $\pm 1^\circ$ UNLESS OTHERWISE SPECIFIED		ALL UNTOLERANCED BASIC DIMENSIONS DEFINED BY $\nabla$ 0.01 A B C		Quantity : 2
UNLESS OTHERWISE SPECIFIED (DIMENSIONS AND FINISHES) TOLERANCES: FRACTIONS DECIMALS ANGLES SURF. FINISH		FINISH: None ALL MACHINED SURFACES UNLESS OTHERWISE SPECIFIED		ORDER AND SERIAL NUMBER (SEE DGS ATTACHED) INCH 201-10 REVISION: 4
NAME: _____ SIGNATURE: _____ DATE: _____		UNIVERSITY OF ARIZONA : Optical Sciences PENTAX ADAPTER MOCKUP (DURIP)		SIZE: 31 A3
MATERIAL: Plastic, Ink Type		Part #		SHEET 1 OF 1

Figure B.6: Modified Samsung K-mount adapter, additional detail and model.

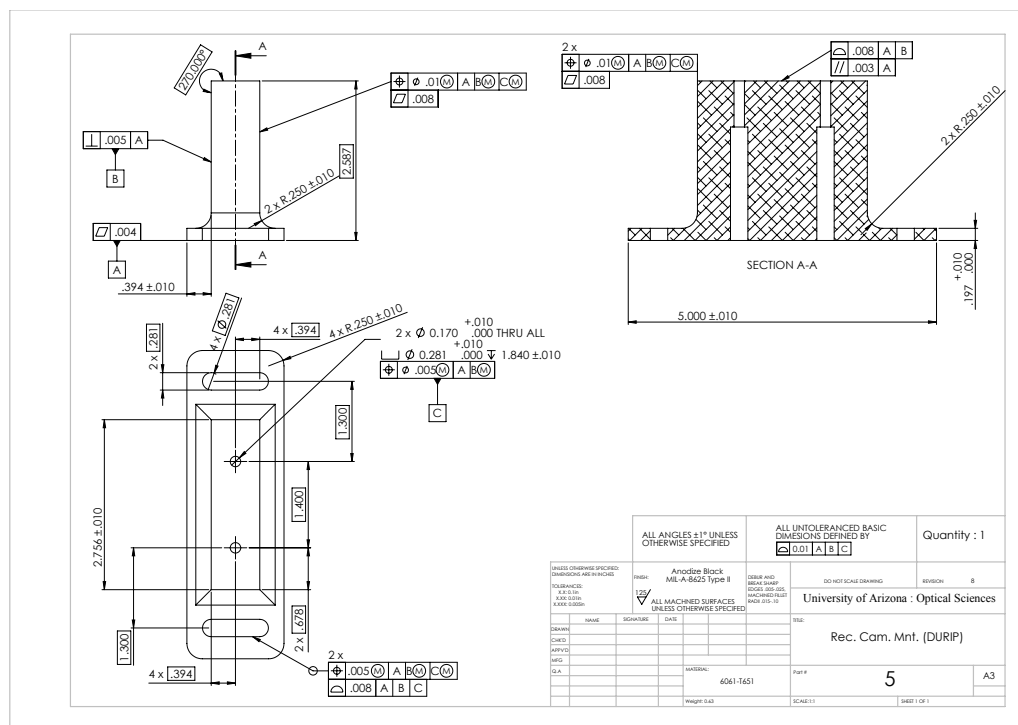


Figure B.7: 4D Technologies™ micropolarizer camera mount.



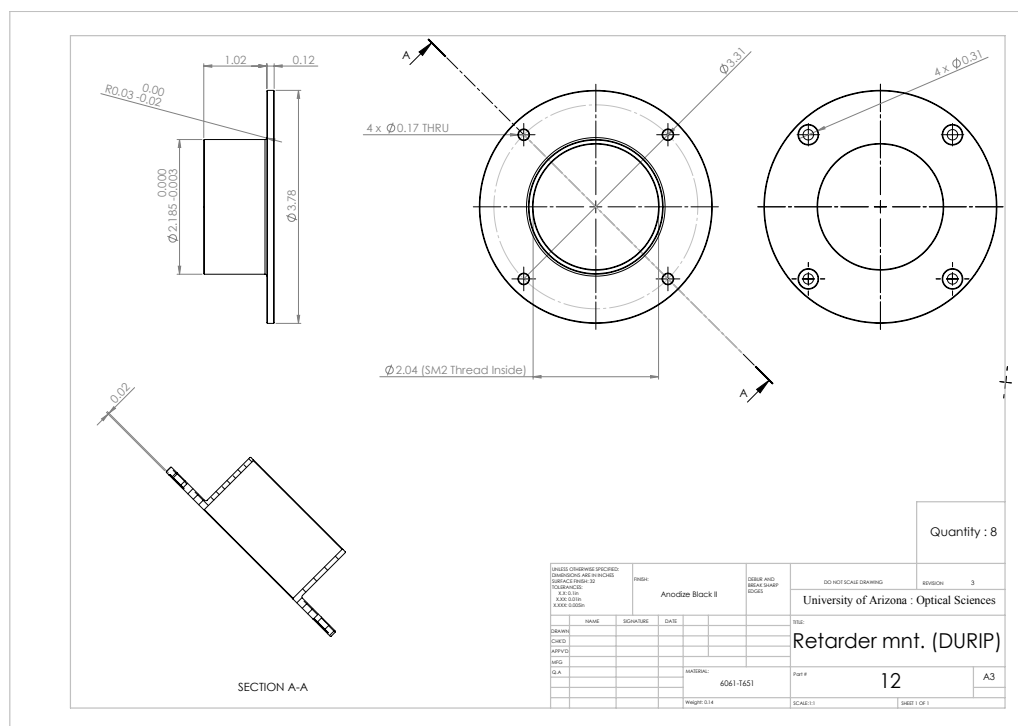


Figure B.9: Retarder mounting fixture. Bolts to IntelliDrives stages. The interior is threaded to ThorLabs SM-2 for the use of COTS retaining rings. Used in both the PSG and PSA.

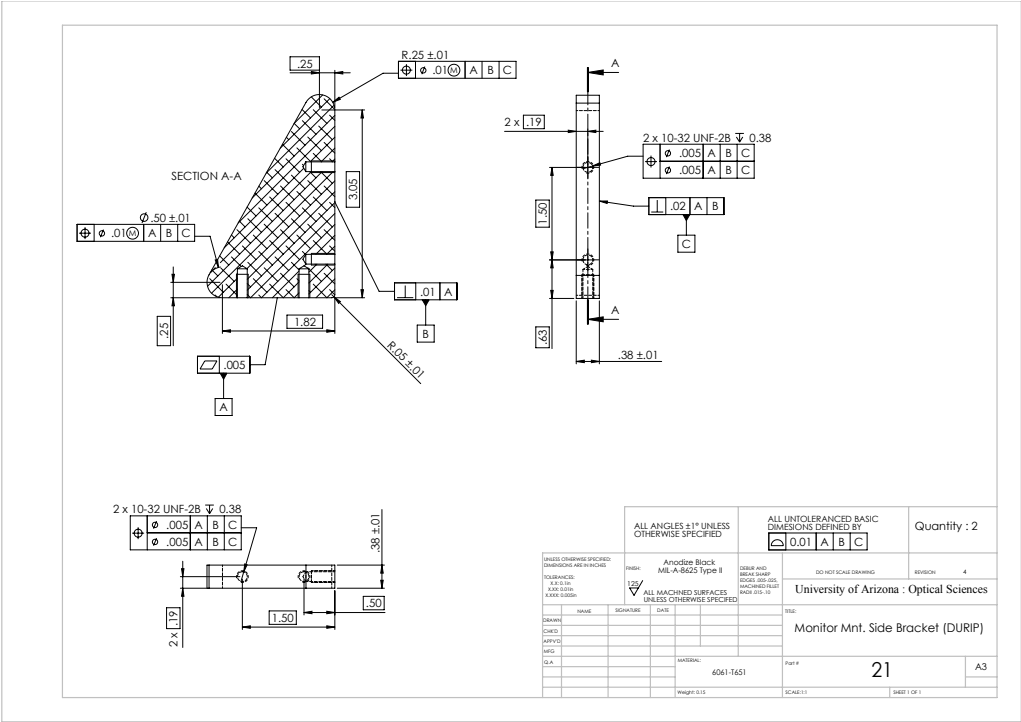


Figure B.10: Angle bracket for monitor mount.







## B.0.4 Code

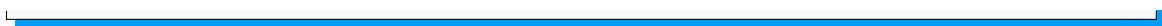
```

2 //Pulse downsampling code for Arduino.
void setup() {
4     pinMode(3,OUTPUT);
    pinMode(12,INPUT);
6 }

//super fast pulse counting loop, works for 480kHz pulse train.
8 void loop() {
    byte toggle;
10    int encoderCount;
    uint8_t sampledPin, lastVal, portDOut;
12    portDOut = B00001000;

14    toggle = 0;
    encoderCount = 0;
16    lastVal = PINB & B00010000;
    noInterrupts();
18    while(true)
    {
20        sampledPin = PINB & B00010000;
        if (lastVal ^ sampledPin)
22        {
            encoderCount++;
24            lastVal = sampledPin;
        }
26        // Each pulse has 2 states, low and high, this
        // with a modulus toggle had to be done for speed instead
28        // of counting pulses it counts state changes.
        if (encoderCount == 8437 + toggle)
30        {
            portDOut ^= B00001000;
32            PORTD = portDOut;
            encoderCount = 0;
34            toggle = (~toggle) & 1;
        }
36    }
}

```



## REFERENCES

- [1] I. J. Vaughn, "The imaging equation for a microgrid linear stokes polarimeter," in "SPIE Optical Engineering+ Applications," (International Society for Optics and Photonics, 2011), pp. 816008–816008.
- [2] I. J. Vaughn, B. G. Hoover, and J. S. Tyo, "Classification using active polarimetry," in "SPIE Defense, Security, and Sensing," (International Society for Optics and Photonics, 2012), pp. 83640S–83640S.
- [3] I. J. Vaughn, "Polarimetric imaging for target and threat identification," *Electronic Military and Defense Magazine* **5**, 8–12 (2015).
- [4] I. J. Vaughn, O. G. Rodríguez-Herrera, M. Xu, and J. S. Tyo, "A portable imaging mueller matrix polarimeter based on a spatio-temporal modulation approach: theory and implementation," (2015), vol. 9613.
- [5] I. J. Vaughn, O. G. Rodríguez-Herrera, M. Xu, and J. S. Tyo, "Bandwidth and crosstalk considerations in a spatio-temporally modulated polarimeter," (2015), vol. 9613.
- [6] I. J. Vaughn, O. G. Rodríguez-Herrera, and J. S. Tyo, "Spatio-temporally channeled mueller matrix polarimeters I : bandwidth," (in preparation) (2015).
- [7] I. J. Vaughn, O. G. Rodríguez-Herrera, and J. S. Tyo, "Spatio-temporally channeled mueller matrix polarimeters II : noise and systematic error," (in preparation) (2016).
- [8] I. J. Vaughn, O. G. Rodríguez-Herrera, and J. S. Tyo, "A fast portable imaging mueller matrix polarimeter," (in preparation) (2016).
- [9] W. Egan, "Optical stokes parameters for farm crop identification," *Remote Sensing of Environment* **1**, 165–180 (1970).
- [10] W.-I. Boerner, H. Mott, and E. Luneburg, "Polarimetry in remote sensing: Basic and applied concepts," in "Geoscience and Remote Sensing, 1997. IGARSS'97. Remote Sensing-A Scientific Vision for Sustainable Development, 1997 IEEE International," , vol. 3 (IEEE, 1997), vol. 3, pp. 1401–1403.
- [11] J. Zallat, P. Grabbling, and Y. Takakura, "Using polarimetric imaging for material classification," in "Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on," , vol. 2 (IEEE, 2003), vol. 2, pp. II–827.

- [12] F. Goudail, P. Terrier, Y. Takakura, L. Bigue, F. Galland, and V. DeVlaminck, "Target detection with a liquid-crystal-based passive stokes polarimeter," *Applied Optics* **43**, 274–282 (2004).
- [13] J. S. Harchanko and D. B. Chenault, "Water-surface object detection and classification using imaging polarimetry," in "Optics & Photonics 2005," (International Society for Optics and Photonics, 2005), pp. 588815–588815.
- [14] J. S. Tyo, D. L. Goldstein, D. B. Chenault, and J. A. Shaw, "Review of passive imaging polarimetry for remote sensing applications," *Appl. Opt.* **45**, 5453–5469 (2006).
- [15] V. Thilak, D. G. Voelz, and C. D. Creusere, "Polarization-based index of refraction and reflection angle estimation for remote sensing applications," *Applied optics* **46**, 7527–7536 (2007).
- [16] D. J. Diner, A. Davis, B. Hancock, G. Gutt, R. A. Chipman, and B. Cairns, "Dual-photoelastic-modulator-based polarimetric imaging concept for aerosol remote sensing," *Appl. Opt.* **46**, 8428–8445 (2007).
- [17] J. R. Schott, *Fundamentals of polarimetric remote sensing*, vol. 81 (Spie Press, 2009).
- [18] F. Waquet, B. Cairns, K. Knobelspiesse, J. Chowdhary, L. D. Travis, B. Schmid, and M. Mishchenko, "Polarimetric remote sensing of aerosols over land," *Journal of Geophysical Research: Atmospheres* (1984–2012) **114** (2009).
- [19] L. J. Cheng, M. Hamilton, C. Mahoney, and G. Reyes, "Analysis of AOTF hyperspectral imaging," in "Algorithms for Multispectral and Hyperspectral Imagery," , vol. 2231 of *Proc. SPIE*, A. Iverson, ed. (1994), vol. 2231 of *Proc. SPIE*, p. 158–16.
- [20] 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).
- [21] M. P. Silverman and W. Strange, "Object delineation within turbid media by backscattering of phase-modulated light," *Optics Communications* **144**, 7 – 11 (1997).
- [22] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, "Polarization-based vision through haze," in "ACM SIGGRAPH ASIA 2008 courses," (ACM, New York, NY, USA, 2008), SIGGRAPH Asia '08, pp. 71:1–71:15.

- [23] J. van der Laan, D. Scrymgeour, S. Kemme, and E. Dereniak, "Detection range enhancement using circularly polarized light in scattering environments for infrared wavelengths," *Applied optics* **54**, 2266–2274 (2015).
- [24] V. Thilak, D. G. Voelz, and C. D. Creusere, "Image segmentation from multi-look passive polarimetric imagery," (SPIE, 2007), vol. 6682, p. 668206.
- [25] K. Sassen, "Polarization in Lidar," in "Lidar," , vol. 102 of *Springer Series in Optical Sciences*, C. Weitkamp, ed. (Springer Berlin / Heidelberg, 2005), pp. 19–42.
- [26] D. W. Tyler, A. M. Phenis, A. B. Tietjen, M. Virgen, J. D. Mudge, J. S. Stryjewski, and J. A. Dank, "First high-resolution passive polarimetric images of boosting rocket exhaust plumes," (SPIE, 2009), vol. 7461, p. 74610J.
- [27] J. R. Dymond and J. D. Shepherd, "Correction of the topographic effect in remote sensing," *IEEE Transactions on Geoscience and Remote Sensing* **37**, 2618–2619 (1999).
- [28] R. N. Treuhaft and S. R. Cloude, "The structure of oriented vegetation from polarimetric interferometry," *IEEE Transactions on Geoscience and Remote Sensing* **37**, 2620–2624 (1999).
- [29] S. L. Blakeney, S. E. Day, and J. N. Stewart, "Determination of unknown input polarisation using a twisted nematic liquid crystal display with fixed components," *Optics Communications* **214**, 1 – 8 (2002).
- [30] J. E. Wolfe and R. A. Chipman, "Polarimetric characterization of liquid-crystal-on-silicon panels," *Appl. Opt.* **45**, 1688–1703 (2006).
- [31] K. Oka, "Singleshot spectroscopic polarimetry using channeled spectrum," in "Photonics Asia 2002," (International Society for Optics and Photonics, 2002), pp. 167–175.
- [32] M. W. Kudenov, M. J. Escuti, N. Hagen, E. L. Dereniak, and K. Oka, "Snapshot imaging mueller matrix polarimeter using polarization gratings," *Optics letters* **37**, 1367–1369 (2012).
- [33] C. F. LaCasse, T. Ririe, R. A. Chipman, and J. S. Tyo, "Spatio-temporal modulated polarimetry," (2011), vol. 8160, pp. 81600K–81600K–11.
- [34] A. S. Alenin and J. S. Tyo, "Generalized channeled polarimetry," *J. Opt. Soc. Am. A* **31**, 1013–1022 (2014).
- [35] D. A. LeMaster and K. Hirakawa, "Improved microgrid arrangement for integrated imaging polarimeters," *Opt. Lett.* **39**, 1811–1814 (2014).

- [36] J. L. Pezzaniti and J. M. Lindberg, "Apparatus and process for the non-invasive measurement of optically active compounds," (1998). US Patent 5,788,632.
- [37] M. H. Smith, P. D. Burke, A. Lompado, E. A. Tanner, and L. W. Hillman, "Mueller matrix imaging polarimetry in dermatology," in "BiOS 2000 The International Symposium on Biomedical Optics," (International Society for Optics and Photonics, 2000), pp. 210–216.
- [38] A. W. Dreher and K. Reiter, "Scanning laser polarimetry of the retinal nerve fiber layer," in "San Diego'92," (International Society for Optics and Photonics, 1992), pp. 34–41.
- [39] J. M. Bueno and P. Artal, "Double-pass imaging polarimetry in the human eye," *Optics Letters* **24**, 64–66 (1999).
- [40] K. Twietmeyer, R. Chipman, A. E. Elsner, Y. Zhao, and D. VanNasdale, "Mueller matrix retinal imager with optimized polarization conditions," *Optics express* **16**, 21339–21354 (2008).
- [41] A. Baldwin, J. Chung, J. Baba, C. Spiegelman, M. Amos, and G. Cote, "Mueller matrix imaging for cancer detection," in "Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE," , vol. 2 (IEEE, 2003), vol. 2, pp. 1027–1030.
- [42] A. S. Alenin, L. Morrison, C. Curiel, and J. S. Tyo, "Hyperspectral measurement of the scattering of polarized light by skin," in "SPIE Optical Engineering+ Applications," (International Society for Optics and Photonics, 2011), pp. 816014–816014.
- [43] C. F. LaCasse, R. A. Chipman, and J. S. Tyo, "Band limited data reconstruction in modulated polarimeters," *Opt. Express* **19**, 14976–14989 (2011).
- [44] C. F. LaCasse, J. S. Tyo, and R. A. Chipman, "Role of the null space of the drmm in the performance of modulated polarimeters," *Optics letters* **37**, 1097–1099 (2012).
- [45] A. S. Alenin and J. S. Tyo, "Task-specific snapshot mueller matrix channeled spectropolarimeter optimization," (2012).
- [46] R. M. A. Azzam, "Photopolarimetric measurement of the mueller matrix by fourier analysis of a single detected signal," *Opt. Lett.* **2**, 148–150 (1978).
- [47] K. Oka and T. Kato, "Spectroscopic polarimetry with a channeled spectrum," *Optics Letters* **24**, 1475–1477 (1999).

- [48] K. Oka and T. Kaneko, "Compact complete imaging polarimeter using birefringent wedge prisms," *Optics express* **11**, 1510–1519 (2003).
- [49] D. Sabatke, A. Locke, E. L. Dereniak, M. Descour, J. Garcia, T. Hamilton, and R. W. McMillan, "Snapshot imaging spectropolarimeter," *Optical engineering* **41**, 1048–1054 (2002).
- [50] M. W. Kudenov, M. J. Escuti, E. L. Dereniak, and K. Oka, "White-light channeled imaging polarimeter using broadband polarization gratings," *Applied optics* **50**, 2283–2293 (2011).
- [51] N. Hagen and M. W. Kudenov, "Review of snapshot spectral imaging technologies," *Optical Engineering* **52**, 090901–090901 (2013).
- [52] M. Dubreuil, S. Rivet, B. Le Jeune, and J. Cariou, "Snapshot mueller matrix polarimeter by wavelength polarization coding," *Optics express* **15**, 13660–13668 (2007).
- [53] N. Hagen and E. L. Dereniak, "Snapshot mueller matrix spectropolarimetry," in "Optical Engineering+ Applications," (International Society for Optics and Photonics, 2007), pp. 668207–668207.
- [54] G. Myhre, W.-L. Hsu, A. Peinado, C. LaCasse, N. Brock, R. A. Chipman, and S. Pau, "Liquid crystal polymer full-stokes division of focal plane polarimeter," *Opt. Express* **20**, 27393–27409 (2012).
- [55] X. Zhao, A. Bermak, F. Boussaid, and V. G. Chigrinov, "Liquid-crystal micropolarimeter array for full stokes polarization imaging in visible spectrum," *Optics express* **18**, 17776–17787 (2010).
- [56] M. Mansuripur, *Field, Force, Energy and Momentum in Classical Electrodynamics* (Bentham Science Publishers, 2011).
- [57] P. D. Lax, *Functional Analysis* (Wiley-Interscience, New York, New York, 2002).
- [58] R. B. Marcus, "The significance of negative frequencies in spectrum analysis," *Electromagnetic Compatibility, IEEE Transactions on* **9**, 123–126 (1967).
- [59] E. Wolf, *Introduction to the Theory of Coherence and the Polarization of Light* (Cambridge University Press, New York, United States, 2007).
- [60] E. Collett, "Polarized light. fundamentals and applications. 1993," .
- [61] R. Chipman, "Handbook of optics: Volume i: Geometrical and physical optics, polarized light, components and instruments," (2009).



- [62] J. J. Gil, "Characteristic properties of mueller matrices," *Journal of the Optical Society of America A* **17**, 328–334 (2000).
- [63] J. Zallat, S. Ainouz, and M. P. Stoll, "Optimal configurations for imaging polarimeters: impact of image noise and systematic errors," *Journal of Optics A: Pure and Applied Optics* **8**, 807–814 (2006).
- [64] B. Le Jeune, J. Cariou, and J. Lotrian, "Underwater imaging by laser: Target discrimination by polarization and speckle characteristics," in "1988 Intl Congress on Optical Science and Engineering," (International Society for Optics and Photonics, 1989), pp. 193–200.
- [65] J. Cariou, B. Le Jeune, J. Lotrian, and Y. Guern, "Polarization effects of seawater and underwater targets," *Applied optics* **29**, 1689–1695 (1990).
- [66] G. W. Kattawar and M. J. Raković, "Virtues of mueller matrix imaging for underwater target detection," *Applied optics* **38**, 6431–6438 (1999).
- [67] J. D. Ellis and A. C. Dogariu, "Remote sensing through reduced mueller matrix elements," in "Aerospace/Defense Sensing, Simulation, and Controls," (International Society for Optics and Photonics, 2001), pp. 509–520.
- [68] G. W. Kattawar and D. J. Gray, "Mueller matrix imaging of targets in turbid media: effect of the volume scattering function," *Applied optics* **42**, 7225–7230 (2003).
- [69] A. Kashani and M. Humayun, "Combined spectral and polarimetry imaging and diagnostics," (2011). US Patent App. 13/021,524.
- [70] B. G. Hoover, "Hoover (Advanced Optical Technologies) active polarimeter," <http://www.advanced-optical.com/featured/newspolarimeterupdate.php>. Accessed: 2015-09-11.
- [71] D. H. Goldstein and R. A. Chipman, "Error analysis of a mueller matrix polarimeter," *JOSA A* **7**, 693–700 (1990).
- [72] B. Ratliff, C. Lacasse, and J. Tyo, "Quantifying ifov error and compensating its effects in dofp polarimeters," *Opt. Express* **17**, 9112–9125 (2009).
- [73] M. I. Mishchenko and L. D. Travis, "Satellite retrieval of aerosol properties over the ocean using polarization as well as intensity of reflected sunlight," *Journal of Geophysical Research: Atmospheres* (1984–2012) **102**, 16989–17013 (1997).

- [74] D. J. Diner, A. Davis, B. Hancock, G. Gutt, R. A. Chipman, and B. Cairns, "Dual-photoelastic-modulator-based polarimetric imaging concept for aerosol remote sensing," *Applied optics* **46**, 8428–8445 (2007).
- [75] D. J. Diner, A. Davis, B. Hancock, S. Geier, B. Rheingans, V. Jovanovic, M. Bull, D. M. Rider, R. A. Chipman, A.-B. Mahler *et al.*, "First results from a dual photoelastic-modulator-based polarimetric camera," *Applied optics* **49**, 2929–2946 (2010).
- [76] H. R. Gordon, T. Du, and T. Zhang, "Atmospheric correction of ocean color sensors: analysis of the effects of residual instrument polarization sensitivity," *Applied optics* **36**, 6938–6948 (1997).
- [77] J. T. Adams, E. Aas, N. K. Højerslev, and B. Lundgren, "Comparison of radiance and polarization values observed in the mediterranean sea and simulated in a monte carlo model," *Applied optics* **41**, 2724–2733 (2002).
- [78] S. Ahmed, A. Gilerson, A. Gill, B. Gross, F. Moshary, and J. Zhou, "Separation of fluorescence and elastic scattering from algae in seawater using polarization discrimination," *Optics communications* **235**, 23–30 (2004).
- [79] A. Gilerson, J. Zhou, M. Oo, J. Chowdhary, B. M. Gross, F. Moshary, and S. Ahmed, "Retrieval of chlorophyll fluorescence from reflectance spectra through polarization discrimination: modeling and experiments," *Applied optics* **45**, 5568–5581 (2006).
- [80] P. Bhandari, K. J. Voss, and L. Logan, "An instrument to measure the downwelling polarized radiance distribution in the ocean," *Optics express* **19**, 17609–17620 (2011).
- [81] D. A. LeMaster, A. H. Mahamat, B. M. Ratliff, A. S. Alenin, J. S. Tyo, and B. M. Koch, "Swir active polarization imaging for material identification," in "SPIE Optical Engineering+ Applications," (International Society for Optics and Photonics, 2013), pp. 887300–887300.
- [82] R. Simon, "The connection between mueller and jones matrices of polarization optics," *Optics Communications* **42**, 293–297 (1982).
- [83] R. Barakat, "Bilinear constraints between elements of the 4 x 4 mueller-jones transfer matrix of polarization theory," *Optics Communications* **38**, 159–161 (1981).
- [84] S. R. Cloude, "Conditions for the physical realisability of matrix operators in polarimetry," **1166**, 177–185 (1989).

- [85] A. B. Kostinski, C. R. Givens, and J. M. Kwiatkowski, "Constraints on mueller matrices of polarization optics," *Applied Optics* **32**, 1646 (1993).
- [86] D. G. M. Anderson and R. Barakat, "Necessary and sufficient conditions for a mueller matrix to be derivable from a jones matrix," *Journal of the Optical Society of America A* **11**, 2305 (1994).
- [87] S. Bergner, T. Möller, D. Weiskopf, and D. J. Muraki, "A spectral analysis of function composition and its implications for sampling in direct volume visualization," *Visualization and Computer Graphics, IEEE Transactions on* **12**, 1353–1360 (2006).
- [88] H. H. Barrett and K. J. Myers, *Foundations of Image Science* (Wiley-Interscience, Hoboken, New Jersey, United States, 2004).
- [89] T. C. Hales, "The sphere packing problem," *Journal of Computational and Applied Mathematics* **44**, 41–76 (1992).
- [90] T. C. Hales, "A proof of the kepler conjecture," *Annals of mathematics* pp. 1065–1185 (2005).
- [91] A. B. Hopkins, F. H. Stillinger, and S. Torquato, "Densest binary sphere packings," *Physical Review E* **85**, 021130 (2012).
- [92] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*, vol. 290 (Springer Science & Business Media, 2013).
- [93] J. Mathews and K. Fink, *Numerical Methods Using MATLAB*, Featured Titles for Numerical Analysis Series (Pearson Prentice Hall, 2004).
- [94] H. Nyquist, "Certain topics in telegraph transmission theory," *American Institute of Electrical Engineers, Transactions of the* **47**, 617–644 (1928).
- [95] D. Gabor, "Theory of communication. part 1: The analysis of information," *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering* **93**, 429–441 (1946).
- [96] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review* **5**, 3–55 (2001).
- [97] A. S. Alenin and J. S. Tyo, "Structured decomposition design of partial mueller matrix polarimeters," *JOSA A* **32**, 1302–1312 (2015).
- [98] D. Sabatke, M. Descour, E. Dereniak, W. Sweatt, S. Kemme, and G. Phipps, "Optimization of retardance for a complete stokes polarimeter," *Optics Letters* **25**, 802–804 (2000).

- [99] T. F. Chan and X.-C. Tai, "Level set and total variation regularization for elliptic inverse problems with discontinuous coefficients," *Journal of Computational Physics* **193**, 40–66 (2004).
- [100] L. Vese and S. Osher, "The level set method links active contours, mumford-shah segmentation, and total variation restoration," *CAM Reoirt* pp. 02–05 (2002).
- [101] A. S. Alenin, "Matrix structure for information-driven polarimeter design," Ph.D. thesis, The University of Arizona. (2015).
- [102] R. Chipman, "Ch. 14 : Mueller matrices," in "Handbook of Optics Volume I: Geometrical and Physical Optics, Polarized Light, Components and Instruments," M. Bass, ed. (McGraw Hill, New York, 2009), 3rd ed.
- [103] D. H. Goldstein and R. A. Chipman, "Error analysis of a mueller matrix polarimeter," *Journal of the Optical Society of America A* **7**, 693 (1990).
- [104] E. Voigtman, "Effect of source  $1/f$  noise on optical polarimeter performance," *Analytical Chemistry* **64**, 2590–2595 (1992).
- [105] A. Ambirajan and D. C. Look, "Optimum angles for a polarimeter: part 1," *Optical Engineering* **34**(6), 1651–1655 (1995).
- [106] D. Sabatke, M. Descour, E. Dereniak, W. Sweatt, S. Kemme, and G. Phipps, "Optimization of retardance for a complete stokes polarimeter," *Optics Letters* **25**, 802–804 (2000).
- [107] J. C. del Toro Iniesta and M. Collados, "Optimum modulation and demodulation matrices for solar polarimetry," *Applied Optics* **39**, 1637 (2000).
- [108] J. S. Tyo, "Noise equalization in stokes parameter images obtained by use of variable-retardance polarimeters," *Optics Letters* **25**, 1198 (2000).
- [109] M. H. Smith, "Optimization of a dual-rotating-retarder mueller matrix polarimeter," *Applied Optics* **41**, 2488 (2002).
- [110] J. S. Tyo, "Design of optimal polarimeters: Maximization of signal-to-noise ratio and minimization of systematic error," *Applied Optics* **41**, 619 (2002).
- [111] V. L. Gamiz and J. F. Belsher, "Performance limitations of a four-channel polarimeter in the presence of detection noise," *Optical Engineering* **41**, 973–980 (2002).
- [112] J. F. Belsher, "Performance limitations of a four-channel polarimeter in the presence of detection noise," *Optical Engineering* **41**, 973 (2002).

- [113] Y. Takakura and J. E. Ahmad, "Noise distribution of mueller matrices retrieved with active rotating polarimeters," *Applied Optics* **46**, 7354 (2007).
- [114] I. J. Vaughn and B. G. Hoover, "Noise reduction in a laser polarimeter based on discrete waveplate rotations," *Optics Express* **16**, 2091 (2008).
- [115] K. M. Twietmeyer and R. A. Chipman, "Optimization of mueller matrix polarimeters in the presence of error sources," *Optics Express* **16**, 11589 (2008).
- [116] A. A. Ramos and M. Collados, "Error propagation in polarimetric demodulation," *Applied Optics* **47**, 2541 (2008).
- [117] J. R. Valenzuela and J. A. Fessler, "Regularized estimation of stokes images from polarimetric measurements," in "Computational Imaging VI," , vol. 6814 of *Computational Imaging VI* (2008), vol. 6814 of *Computational Imaging VI*, pp. 681403–681403–10.
- [118] D. Lara and C. Paterson, "Stokes polarimeter optimization in the presence of shot and gaussian noise," *Optics Express* **17**, 21240 (2009).
- [119] J. R. Valenzuela and J. A. Fessler, "Joint reconstruction of stokes images from polarimetric measurements," *Journal of the Optical Society of America A* **26**, 962 (2009).
- [120] F. Goudail and A. Bènière, "Estimation precision of the degree of linear polarization and of the angle of polarization in the presence of different sources of noise," *Appl. Opt.* **49**, 683–693 (2010).
- [121] M. Kupinski, R. Chipman, and E. Clarkson, "Relating the statistics of the angle of linear polarization to measurement uncertainty of the stokes vector," *Optical Engineering* **53**, 113108 (2014).
- [122] D. Sabatke, A. Locke, E. Dereniak, and R. McMillan, "Linear calibration and reconstruction techniques for channeled spectropolarimetry," *Optics express* **11**, 2940–2952 (2003).
- [123] M. Dubreuil, S. Rivet, B. Le Jeune, and J. Cariou, "Systematic errors specific to a snapshot mueller matrix polarimeter," *Applied Optics* **48**, 1135 (2009).
- [124] P. Hauge, "Mueller matrix ellipsometry with imperfect compensators," *JOSA* **68**, 1519–1528 (1978).
- [125] E. Compain, S. Poirier, and B. Drevillon, "General and self-consistent method for the calibration of polarization modulators, polarimeters, and mueller-matrix ellipsometers," *Applied optics* **38**, 3490–3502 (1999).

- [126] Y. Sabharwal, J. Joubert, and D. Sharma, "Digital camera technologies for scientific bio-imaging. part 4: Sampling and signal," *Microscopy and Analysis* p. 1 (2011).
- [127] Y. Sabharwal, "Digital camera technologies for scientific bio-imaging. part 4: signal-to-noise ratio and image comparison of cameras," *Microscopy and Analysis-UK* p. S4 (2012).
- [128] J. Schoukens and J. Renneboog, "Modeling the noise influence on the fourier coefficients after a discrete fourier transform," *Instrumentation and Measurement, IEEE Transactions on* **1001**, 278–286 (1986).
- [129] J. B. Burbidge, L. Magee, and A. L. Robb, "Alternative transformations to handle extreme values of the dependent variable," *Journal of the American Statistical Association* **83**, 123–127 (1988).
- [130] P. Hoeher, S. Kaiser, and P. Robertson, "Two-dimensional pilot-symbol-aided channel estimation by wiener filtering," in "Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on," , vol. 3 (IEEE, 1997), vol. 3, pp. 1845–1848.
- [131] F. Snik, G. v. Harten, A. S. Alenin, I. J. Vaughn, and J. S. Tyo, "A multi-domain full-stokes polarization modulator that is efficient for 300-2500nm spectropolarimetry," (2015), vol. 9613.
- [132] ASTM G173-03, "Standard tables for reference solar spectral irradiances: Direct normal and hemispherical on 37° tilted surface," ASTM International (2012).
- [133] A. Meinel and M. Meinel, *Applied solar energy: an introduction*, Addison-Wesley series in physics (Addison-Wesley Pub. Co., 1976).
- [134] N. S. Browes, Laser Glow Technologies, personal communication.
- [135] W. Silfvast, *Laser Fundamentals* (Cambridge University Press, 2004).
- [136] N. Brock, 4D Technologies, personal communication at SPIE Optics + Photonics.
- [137] F. J. Harris, "On the use of windows for harmonic analysis with the discrete fourier transform," *Proceedings of the IEEE* **66**, 51–83 (1978).
- [138] C. F. LaCasse, O. G. Rodríguez-Herrera, R. A. Chipman, and J. S. Tyo, "Spectral density response functions for modulated polarimeters," *Appl. Opt.* **54**, 9490–9499 (2015).

- [139] L. W. Tu, “Bump functions and partitions of unity,” *An Introduction to Manifolds* pp. 127–134 (2008).