AFRL-AFOSR-VA-TR-2016-0077

IMAGING AN DNON IMAGING POLARIMETRIC METHODS FOR REMOTE SENSING

R. John Koshel
ARIZONA UNIV BOARD OF REGENTS TUCSON

02/09/2016
Final Report

Air Force Research Laboratory
AF Office Of Scientific Research (AFOSR)/ RTB1
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
04-02-2016	Final Technical	15-04-10 - 15-12-15

4. TITLE AND SUBTITLE
Imaging and Non-Imaging Polarimetric Methods for Remote Sensing

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-10-1-0114

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
J. Scott Tyo

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Arizona
888 N. EUCLID AVENUE
TUCSON, AZ 85722-3308 US

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Dr. Julie Moses
AFOSR/NE
875 N
Randolph St
Arlington, VA

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

During the lifetime of this project, the research team signi˝cantly advanced the state of understanding of modulated polarimeter systems for active and passive, imaging- and non-imaging polarimetry. Modulated polarimeters infer the polarization state of light by creating a set of polarization-dependent carriers that are modulated by the intensity signal. These carriers can be created in any independent domain, such as time, space, wavenumber, angle of incidence, and they can be created in combinations of domains simultaneously. The work supported on this project has solidified the theory behind such instruments, allowing for new design philosophies that improve state-of-the-art instruments. When the project began five years ago, only a cursory understanding of modulated instruments existed, and the data reduction matrix was the primary means of processing polarization data. In this project, the DRM has been expanded to include a full functional formalism, allowing for a range of new tools in polarimeter design to be brought to bear. During the course of this project, five PhD students, seven MS students, and two undergraduates were trained. Four PhD dissertations were primarily supported by funding from this project

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(Include area code)*

Reset

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATES COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report, e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5d. PROJECT NUMBER. Enter all project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

Imaging and Non-Imaging Polarimetric Methods for Remote Sensing

J. Scott Tyo
College of Optical Sciences
University of Arizona

Grant # FA9550-10-1-0114
Period of Performance: April 15, 2010 – December 15, 2015

ABSTRACT

During the lifetime of this project, the research team significantly advanced the state of understanding of modulated polarimeter systems for active and passive, imaging- and non-imaging polarimetry. Modulated polarimeters infer the polarization state of light by creating a set of polarization-dependent carriers that are modulated by the intensity signal. These carriers can be created in any independent domain, such as time, space, wavenumber, angle of incidence, and they can be created in combinations of domains simultaneously. The work supported on this project has solidified the theory behind such instruments, allowing for new design philosophies that improve state-of-the-art instruments. When the project began five years ago, only a cursory understanding of modulated instruments existed, and the data reduction matrix was the primary means of processing polarization data. In this project, the DRM has been expanded to include a full functional formalism, allowing for a range of new tools in polarimeter design to be brought to bear. During the course of this project, five PhD students, seven MS students, and two undergraduates were trained. Four PhD dissertations were primarily supported by funding from this project.

1. EXECUTIVE SUMMARY

1.1 Principal Accomplishments

This project was carried out in the Advanced Sensing Laboratory of the College of Optical Sciences at the University of Arizona. The original project dates were April 15, 2010 - December 15, 2015. The was initiated based on work that the PI had done under previous USAF funding with microgrid imaging polarimeters.[1–3] While the main motivation at the beginning of the project was to improve the performance of microgrid instruments, the project expanded signficantly to include all forms of modulated polarimeters such as time modulated instruments,[4] prismatic polarimeters,[5] spectrally channeled polarimeters,[6] and even instruments modulated in multiple domains.[7,8] It is important to note that much of the progress was achieved in collaboration with Professor Russell Chipman of the University of Arizona. Prof. Chipman did not receive any direct funding, from this project, but he co-supervised one of the students and was an active collaborator on several of the projects.

The progress supported on this grant can be broken down into the following areas:

- *Fundamental understanding of modulated instruments*: Dr. Charlie LaCasse worked on a basic theory of modoluated instruments for his PhD dissertation. That work extended our earlier work with microgrids and coupled it with Prof. Chipman's work on temporally modulated instruments.[9,10] Dr. LaCasse extended this work to cover random signal processing and the associated transfer function formalisms.[11]
- *Scene-Based Nonuniformity Correction for Focal Plane Arrays*: Dr. Wiley Black worked on methods to correct residual calibration errors in the FPA due to nonuniformity in the array response. Much of his work turned out to be generally applicable to all imagers,[12,13] but he also considered its specific application to infrared microgrid instruments[14]

J. Scott Tyo: E-mail: tyo@ieee.org

- *Coherence Manipulation for Sensing*: Dr. Oscar Rodrìguez-Herrera developed strategies to be able to manipulate the coherence of light in order to indirectly sense BRDF from monostatic measurements.[15] This was follow-on work from preliminary results obtained under earlier AFOSR funding,[16] and continues in current work.

- *Generalized Channeled Polarimetry*: Dr. Andrey Alenin developed a general theoretical framework that allows channeled polarimeters to be designed from *a priori* information.[17] Previous design methods were *ad hoc*, and generally resulted in needlessly sub-optimal systems. This work was a direct outgrowth of Dr. LaCasse's research (mentioned above).

- *Partial Mueller Matrix Polarimeters (pMMPs)*: Dr. Alenin also worked on the development of pMMPs.[18] These devices are active instruments that intentionally omit certain Mueller elements in order to improve bandwidth and were pioneered by the PI and Brian Hoover.[19,20] He also developed methods to apply his generalized channeled polarimeter theory to pMMPs.[21] Results from this work were implemented in an instrument built at AFRL/RYJT by a student from the PI's group.[22]

- *Hybrid Modulated Polarimeters*: As part of his research, Dr. LaCasse explored polarimeters modulated in both space and time.[7] Dr. Israel Vaughn extended this work to include the design and optimization of active spatiotemporally modulated systems,[23] and he built a portable polarimeter based on these principles.[24] These techniques have also been extended by us to hybrid wavelength and time modulated instruments in collaboration with Frans Snik and colleagues in the Netherlands for astronomical applications.[8]

- *Other Collaborative Works*: The PI and Drs. Vaughn and Rodrìguez-Herrera had an ongoing collaboration with Prof. Toshitaka Wakayama from Saitama Medical University in Japan. The collaboration began in 2012 when Dr. Wakayama visited the laboratory for three months. Our group worked with Dr. Wakayama to develop axial waveplates in the terahertz[25] and to apply these instruments to angle-modulated sensing.[26]

1.2 Archival Publications

1. J. S. Tyo, Z. Wang, S. J. Johnson, and B. G. Hoover, "Design, analysis, and optimization of partial Mueller matrix polarimeters," *Appl. Opt.* **49**, pp. 2326 – 2333, 2010.

2. C. F. LaCasse, J. S. Tyo, and R. A. Chipman, "Spatio-temporally modulated polarimetry," in *Proc. SPIE vol. 8160: Polarization Science and Remote Sensing V*, J. A. Shaw and J. S. Tyo, eds., p. 816020, SPIE, (Bellingham, WA), 2011.

3. C. F. LaCasse, , R. A. Chipman, and J. S. Tyo, "Band-limited reconstruction in modulated polarimeters," *Opt. Express* **19**, pp. 14976 – 14989, 2011.

4. A. Alenin, L. Morrison, C. Curiel, and J. S. Tyo, "Hyperspectral measurement of the scattering of polarized light by skin," *Proc. SPIE vol 8160: Polarization Science and Remote Sensing V*, J. A. Shaw and J. S. Tyo, Eds., pp. 816014 (SPIE, Bellingham, 2011)

5. W. T. Black and J. S. Tyo, "Frequency-domain scene-based non-uniformity correction and application to microgrid polarimeters," in *Proc. SPIE vol 8160: Polarization Science and Remote Sensing V*, p. 816003, SPIE, (Bellingham, WA), 2011.

6. O. G. R. Herrera and J. S. Tyo, "Generalized van Cittert-Zernike theorem for the cross-spectral density matrix of quasi-homogeneous planar electromagnetic sources," **29**, pp. 1939 – 1947, 2012.

7. C. F. LaCasse, J. S. Tyo, and R. A. Chipman, "The role of the null space in modulated polarimeters," *Opt. Lett.* **37**, pp. 1097 – 1099, 2012.

8. A. S. Alenin and J. S. Tyo, "Task-Specific Snapshot Mueller Matrix Channeled Spectropolarimeter Optimization," in *Proc. SPIE vol. 8364: Polarization Measurement, Analysis, and Remote Sensing IX*, D. B. Chenault and D. L. Goldstein, Eds., pp. 836402 (SPIE, Bellingham, WA, 2012)

9. C. F. LaCasse and J. S. Tyo, "Modulated polarimeter operators in the presence of stochastic signals," *Proc. SPIE vol. 8364: Polarization Measurement, Analysis, and Remote Sensing IX*, D. B. Chenault and D. L. Goldstein, Eds., pp. 836402 (SPIE, Bellingham, WA, 2012)

10. I. Vaughn, B. G. Hoover, and J. S. Tyo, "Classification using active polarimetry," *SPIE Polarization Measurement, Analysis, and Remote Sensing IX*, D. B. Chenault and D. L Goldstein, Eds., p. 83640S (SPIE, Bellingham, WA, 2012)

11. D. A. LeMaster, A. Mahamat, B. M. Ratliff, A. Alenin, and J. S. Tyo, "SWIR active polarization imaging," in *Proc. SPIE vol. 8873: Polarization Science and Remote Sensing VI*, D. Lemaster and J. A. Shaw, eds., p. 88730O, SPIE, (Bellingham, WA), 2013.

12. T. Wakayama, K. Komaki, I. Vaughn, J. S. Tyo, Y. Otani, T. Yoshizawa, "Evaluation of Mueller matrix of achromatic axially symmetric waveplate," *SPIE Polarization Science and Remote Sensing VI* D. A. LeMaster and J. A. Shaw, Eds., pp. 88730P (SPIE, Bellingham, WA, 2013)

13. W. T. Black and J. S. Tyo, "Frequency-independent scene cancellation nonuniformity correction technique," *J. Electronic Imaging* **23**, p. 023005, 2014.

14. W. T. Black and J. S. Tyo, "Improving feedback-integrated scene cancellation nonuniformity correction through optimal selection of available camera motion," *J. Electronic Imaging* **23**, p. 053014, 2014.

15. A. S. Alenin and J. S. Tyo, "Generalized channeled polarimetry," *J. Opt. Soc. Am. A* **31**, pp. 1013–1022, May 2014.

16. A. S. Alenin and J. S. Tyo, "Structured decomposition design of partial mueller matrix polarimeters," *J. Opt. Soc. Am. A* **32**, pp. 1302–1312, Jul 2015.

17. F. Snik, G. can Harten, J. S. Tyo, A. S. Alenin, and I. Vaughn, "A multi-domain full-stokes polarization modulator that is efficient for 300-2500 nm spectropolarimetry," in *SPIE vol. 9613: Polarization Science and Remote Sensing VII*, J. A. Shaw and D. LeMaster, eds., p. 961315, SPIE, (Bellingham, WA), 2015.

18. C. F. LaCasse, O. G. Rodrìguez-Herrera, R. A. Chipman, and J. S. Tyo, "Spectral density response functions for modulated polarimeters," *Appl. Opt.* **54**, pp. 9490 – 9499, 2015.

19. A. S. Alenin and J. S. Tyo, "Channeled partial mueller matrix polarimetry," in *Proc. SPIE vol. 9613: Polarization Science and Remote Sensing VII*, J. A. Shaw and D. LeMaster, eds., p. 96130M, SPIE, (Bellingham, WA), 2015.

20. I. J. Vaughn, O. G. Rodriguez-Herrera, M. Xu, and J. S. Tyo, "A portable imaging mueller matrix polarimeter based on a spatio-temporal modulation approach: theory and implementation," in *Proc. SPIE vol. 9613: Polarization Science and Remote Sensing VII*, J. A. Shaw and D. LeMaster, eds., p. 961304, SPIE, (Bellingham, WA), 2015.

21. I. J. Vaughn, M. Xu, and J. S. Tyo, "A portable imaging mueller matrix polarimeter based on a spatio-temporal modulation approach: theory and implementation," in *Proc. SPIE vol. 9613: Polarization Science and Remote Sensing VII*, J. A. Shaw and D. LeMaster, eds., p. 961337, SPIE, (Bellingham, WA), 2015.

22. T. Wakayama, T. Higashiguchi, H. Oikawa, K. Sakaue, M. Washio, M. Yonemura, T. Yoshizawa, J. S. Tyo, and Y. Otani, "Determination of the polarization states of an arbitrary polarized terahertz beam: Vectorial vortex analysis," *Scientific Reports* **5**, p. 9416, 2015.

1.3 Researchers Trained on the Project

1.3.1 Postdoctoral Fellows

1. Oscar G. Rodrìguez-Herrera (2011 - 2015)

1.3.2 PhD Students

1. Charles F. LaCasse, 2010 - 2013, Modulated Imaging Polarimetry. Now employed at Sandia National Laboratories.
2. Gabriel C. Birch, 2011 - 2012, Computational Imaging Methods (secondary support). Now employed at Sandia National Laboratories.
3. Wiley T. Black, 2010 - 2014, In-Situ Calibration Of Non-Uniformity In Infrared Staring Systems.
4. Andrey S. Alenin, 2010 - 2015, Matrix structure for information-driven polarimeter design. Now with the University of New South Wales Canberra.
5. Israel J. Vaughn, 2010 - 2015, Bandwidth and noise in spatio-temporally modulated Mueller matrix polarimeters (will be available soon from UA Open Repository). Now with the University of New South Wales Canberra.

1.3.3 MS Students

1. Adoum Mahamat, 2010 - 2014, Now with NAVAIR.

1.3.4 Undergraduates

1. Rafael Rojas, 2012 - 2014.
2. Jeffrey Wilhite, 2014 - 2015. Now pursuing MS Degree at Optical Sciences.

1.4 Related projects funded by other agencies that leveraged AFOSR results

1. J. S. Tyo (PI), "Analysis of Polarization and Coherence Properties of Polarization Tags," Sandia National Laboratories, $37,000, August 2013 – May 2015
2. J. S. Tyo (PI), "Active SWIR Polarimetry," SAIC (sub to AFRL/RYJT), March 2012 – December 2012 $39,985
3. J. S. Tyo (PI), "Polarimetric Data Processing for Seeker Applications," Polaris Sensor Technologies, (Sub to US Army SMDC SBIR Phase II program), $50,000, March 2011 – March 2013
4. J. S. Tyo, "Visible/NIR/SWIR Imaging Polarimeter for Phenomenology," FY2011 DURIP, December 2011 – December 2012 ($243,784, December 2011 – November 2012)
5. Brian G. Hoover (PI) and J. S. Tyo, "Optimization of Partial Mueller Polarimeters," US Army NVESD Phase I SBIR Program, $150,000, June 2011 – December 2012
6. J. S. Tyo (PI), "Processing of LWIR Polarimetric Imagery to Identify Human Activity," AFRL/RYJ and General Dynamics, $24,500, June 2010 – September 2010

2. INTRODUCTION

Optical polarization sensing is an example of an indirect imaging or sensing method,[27] since optical detectors are generally insensitive to the polarization state of the incident radiation. Because of that, the nature of the performance of polarimeters – including SNR, accuracy, calibration error, etc. – is intrinsically related to the processing methods used to compute the desired information from the actual measured intensity signals.

A widely used general formalism for describing the polarization state of optical radiation in the partially coherent case is the Stokes-Mueller calculus.[28] in this formalism, the polarization state of light is described by a set of four Stokes parameters that are related to the second statistical moments of the field. These moments

can be related to any radiometric quantity, but for imaging polarimeters, the irradiance is a reasonable quantity and the Stokes parameters can be defined as

$$\mathbf{S} = \begin{bmatrix} I_0 + I_{90} \\ I_0 - I_{90} \\ I_{45} - I_{135} \\ I_L - I_R \end{bmatrix}^{\mathrm{T}}, \tag{2.1}$$

where I_x, I_y, I_{45}, and I_{135} are the observed irradiances through linear polarizers oriented at 0°, 45°, 90°, and 135°, respectively, and I_L and I_R are the irradiances observed through left- and right-circular polarizers. Note that the angles are defined with respect to an arbitrary coordinate system. When light interacts with an object or material, the polarization properties are altered. When the interactions are linear, the changes in the Stokes parameters can be described by a 4×4 real matrix called the Mueller matrix

$$\mathbf{S}'\left(\overrightarrow{\theta}\right) = \underline{\underline{\mathbf{M}}}\left(\overrightarrow{\theta}\right)\mathbf{S}\left(\overrightarrow{\theta}\right) = \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{bmatrix}\mathbf{S}. \tag{2.2}$$

The notation $\left(\overrightarrow{\theta}\right)$ indicates that the polarimetric quantities are functions of some set of independent parameters that may include spatial coordinates, time, wavenumber, angle of incidence, etc.

2.1 Wavefront Division Polarimeters

A Stokes polarimeter is used to estimate a set of Stokes parameters. There are two broad classes of polarimeters: wavefront division polarimeters and modulated polarimeters. Wavefront division polarimeters operate by making multiple copies of the incident wave field and relaying them to separate detectors through different polarization optics. The most common examples of wavefront division polarimeters are Division of Amplitude polarimeters, which use beam splitters to create the copies of the field, and Division of Aperture Polarimeters, which use an array of optics to split the wavefront in an aperture plane.[29] Subsequet to this splitting, both strategies employ independent polarization optics in each of the channels in order to affect the measurement. Each channel is described by a polarization analyzer vector $\mathbf{A}_i$, and the power measured in the channel is

$$x_i = \mathbf{A} \cdot \mathbf{S}. \tag{2.3}$$

The set of measurements can be compiled into a linear system of equations

$$\mathbf{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1{}^{\mathrm{T}} \\ \vdots \\ \mathbf{A}_N{}^{\mathrm{T}} \end{bmatrix} = \underline{\underline{\mathbf{W}}} \cdot \mathbf{S}, \tag{2.4}$$

and the Stokes parameters are estimated as

$$\hat{\mathbf{S}} = \underline{\underline{\mathbf{W}}}^{+} \cdot \mathbf{X}. \tag{2.5}$$

The matrix $\underline{\underline{\mathbf{W}}}^{+}$ is a suitable inverse or pseudoinverse, and is often referred to as the *data reduction matrix.*[28]

2.2 Modulated Polarimeters

A wavefront division polarimeter splits the light into channels that are measured by different detectors (or subsets of detector elements in some DoAP systems[30]). In contrast, a modulated polarimeter uses a single detector or array of detectors to make all measurements, but uses polarization-dependent carriers to create channels that carry the polarization information. Most systems have employed purely periodic carriers, which results in an orthogonal frequency division multiplexed system, though non-periodic carriers can be employed as well.[8]

The detector output of a modulated polarimeter (which is proportional to irradiance or power received) is

$$X(t) = \mathbf{A}(t) \cdot \mathbf{S}(t) = A_0(t)s_0(t) + A_1(t)s_1(t) + A_2(t)s_2(t) + A_3(t)s_3(t). \tag{2.6}$$

In Eq. 2.6, we have only indicated temporal modulation, or $\left(\overrightarrow{\theta}\right) = (t)$. However, the theory is immediately applicable to any combination of modulation domains. When the analyzer vector is periodic, then the Fourier transform of Eq. 2.6 produces a set of channels that are orthogonal in the Fourier domain. The Stokes parameters can then be estimated through a set of demodulation and filtering operations.

2.3 Mueller Polarimeters

It should be noted that the Stokes polarimeter can be generalized into an active Mueller polarimeter.[31] This instrument includes both a polarization state analyzer (PSA), which is a Stokes polarimeter, and an polarization state generator (PSG). The PSG is a PSA operated in reverse that generates a controlled illumination state. The intensity measured for a particular PSG/PSA combination is

$$x_i = \mathbf{A}_i \underline{\underline{\mathbf{M}}} \mathbf{G}_i = \overrightarrow{\mathbf{D}}_i \cdot \overrightarrow{M}. \tag{2.7}$$

The vector $\overrightarrow{\mathbf{D}}_i$ is a 16×1 vector obtained by rearranging the elements of the dyad product $\mathbf{G}\mathbf{A}^{\mathrm{T}}$, and $\overrightarrow{M}$ is obtained by reordering the elements of $\underline{\underline{\mathbf{M}}}$. All of the analyses that follow are equally applicable to both Stokes and Mueller polarimeters.

3. MODULATED POLARIMETER DATA PROCESSING

3.1 Data Reduction Matrix method for DoT Polarimeters

The conventional theory of operation for DoT polarimeters (and for other classes of polarimeter as well) is well known[28] and is briefly reproduced here. The polarimeter operates by using a collection of retarders and analyzers that can be changed from measurement to measurement. The most common example is the rotation of a wave plate in front of a linear analyzer,[32] but other optical configurations can also be used.[33,34] The incident Stokes vector is altered by the Mueller matrix of the polarization optics, and the modified irradiance is measured by a polarization-blind detector. This irradiance is given mathematically by taking the inner product of the first row of the system Mueller matrix with the incident Stokes vector as

$$I(t) = \mathbf{S}_A(t)^{\mathrm{T}} \mathbf{S}_{in}, \tag{3.1}$$

where $S_A(t)$ is the time-varying anaylzer Stokes vector given by the first row of the system Mueller matrix. For the purposes of the present discussion, we will assume that the irradiance is sampled ideally. The effects of an integration time that is comparable to the time variation of $\mathbf{S}_A(t)$ can be included using standard methods. The n^{th} sample of the irradiance is then

$$I[n] = \mathbf{S}_A[n]^{\mathrm{T}} \mathbf{S} = s_{A0}[n]s_0 + s_{A1}[n]s_1 + s_{A2}[n]s_2 + s_{A3}[n]s_3. \tag{3.2}$$

In general practice, it is assumed that the polarimeter modulates the irradiance at rates that are much faster than the temporal variation of the incident Stokes vector,[28,35] and that the Stokes parameters in Eq. 6.1 are constants. When this is the case, we can take a series of measurements and form a system of linear equations

$$\mathbf{I} = \begin{bmatrix} I[1] \\ \vdots \\ I[N] \end{bmatrix} = \begin{bmatrix} \mathbf{S}_A[1]^{\mathrm{T}} \\ \vdots \\ \mathbf{S}_A[N]^{\mathrm{T}} \end{bmatrix} \mathbf{S}_{in} = \underline{\underline{\mathbf{W}}} \mathbf{S}_{in}. \tag{3.3}$$

The incident Stokes vector is determined by matrix inversion (or pseudoinversion) as

$$\hat{\mathbf{S}}_{in} = \underline{\underline{\mathbf{W}}}^{-1} \mathbf{I} = \left(\underline{\underline{\mathbf{W}}}^{\mathrm{T}} \underline{\underline{\mathbf{W}}} \right)^{-1} \underline{\underline{\mathbf{W}}}^{\mathrm{T}} \mathbf{I}, \tag{3.4}$$

where the hat in Eq. 3.4 indicates that the result is an estimate of the true value. The matrix $\underline{\underline{\mathbf{W}}}^{-1}$ is generally referred to as the data reduction matrix (DRM).[28]

While the above discussion approaches the problem from a linear algebra perspective, a frequency domain approach is often more informative. Equation 3.3 is easiest to understand with the simple example of the rotating analyzer linear polarimeter. If we rotate an ideal linear analyzer at a constant angular frequency ω_0 (in units of degrees per sample) we have

$$\mathbf{S}_A[n] = \frac{1}{2}\begin{bmatrix} 1 & \cos(2\omega_0 n) & \sin(2\omega_0 n) & 0 \end{bmatrix}^{\mathrm{T}} \tag{3.5}$$

and Eq. 6.1 becomes

$$I[n] = \frac{1}{2}\left[s_0 + \cos(2\omega_0 n)s_1 + \sin(2\omega_0 n)s_2\right]. \tag{3.6}$$

Equation 3.6 tells us that the irradiance information is carried in the DC term, the s_1 information is carried in the in-phase $2\omega_0$ term, and the s_2 information is carried in the quadrature $2\omega_0$ term. Since this polarimeter does not have a retardance element, it cannot measure s_3. The polarization information could be demodulated using frequency-domain methods, which is done for some systems using photo-elastic modulators operating at high rates ($\sim$10s of kHz).[36] However, the linear algebra formulation of Eq. 3.4 is often preferred since it is easier to incorporate calibration, error analysis, and other pratical considerations. In principle, only four measurements are required to uniquely invert Eq. 3.4. However, it is common to employ more than 4 measurements in order to help improve the accuracy of the polarimeter and reduce the effects of noise and systematic errors.[4,32,37,38]

3.2 Using a Linear Systems Formalism for Processing

With only a few exceptions,[4,39] most papers in the polarimetry literature treat $\mathbf{S}_{in}$ as approximately constant in computing the DRM. Diner, *et al.*, initially make this assumption in their dual-PEM polarimeter. Their device makes 40 time-sequential measurements to reconstruct s_0, s_1 and s_2, resulting in a highly underdetermined system of equations with 17 unconstrained degrees of freedom. After initially assuming that the measured Stokes vector was constant, they then used one of their degrees of freedom to innoculate the DRM to linear gradients in the underlying Stokes vector.

An examination of Eq. 6.1 or Eq. 3.6 reveals that the constant Stokes vector assumption is overly limiting, and even the linearly-varying-in-time assumption of Diner, *et al.*, is still unnecessarily limiting. It is apparent that basic communications systems theory can be employed allowing the Stokes parameters to be functions of time as well, subject to a bandwidth criterion imposed by the reconstruction process. This result is analagous to earlier developments for spatially modulated imaging polarimeters[3,40] and spectrally modulated spectropolarimeters.[41]

To introduce this concept for DoT polarimeters, we take the Fourier transform of the signal $I(t)$ in Eq. 3.1, but now we allow the Stokes parameters to become functions of time as

$$\tilde{I}(\omega) = \tilde{s}_{A0}(\omega) * \tilde{s}_0(\omega) + \tilde{s}_{A1}(\omega) * \tilde{s}_1(\omega) + \tilde{s}_{A2}(\omega) * \tilde{s}_2(\omega) + \tilde{s}_{A3}(\omega) * \tilde{s}_3(\omega), \tag{3.7}$$

where the tilde indicates the Fourier transform and * is the convolution operator.

Equation 3.7 presents a conventional deconvolution problem that can be inverted by careful design of the analyzer Stokes vector $\mathbf{S}_A(t)$. The specific modulation strategy of Eq. 3.6 creates distinct side bands in frequency space that carry the polarization information. Using the expressions of Eq. 3.6 in Eq. 3.7 produces

$$\tilde{I}(\omega) = \frac{1}{2}\left(\tilde{s}_0(\omega) + \frac{1}{2}\left(\tilde{s}_1(\omega - 2\omega_0) + \tilde{s}_1(\omega + 2\omega_0)\right) + \frac{1}{2j}\left(\tilde{s}_2(\omega - 2\omega_0) - \tilde{s}_2(\omega + 2\omega_0)\right)\right). \tag{3.8}$$

We see from Eq. 3.8 that forcing the Stokes vector to be constant is a severe, and unnecessary, restriction. If we assume that s_0 is band limited to frequencies below W_0 and s_1 and s_2 are band limited to frequencies below W_1, we can accomplish error-free reconstruction (in the noise-free case) provided that

$$W_0 + W_1 \leq \omega_s/2, \tag{3.9}$$

where $\omega_s/2\pi$ is the temporal sampling frequency of the polarimeter. Equation 3.9 is the Nyquist sampling criterion for a rotating analyzer polarimeter. It is worth noting that the band limit requirement is placed on W_0 and W_1 together, so it is possible to apportion a given sampling bandwidth among the various channels in question. Other types of DoT polarimeters with different specific modulations will have band limit criteria modified somewhat from Eq. 3.9, but similar critera can be derived for all modulated polarimeters.

3.3 Unifying the DRM and linear systems methods

In this section we generalize the linear systems formalism to encompass the DRM method. Doing this shows an immediate drawback of the DRM method that leads to unncessary polarimetric aliasing that can easily be remedied by choosing a different weighting scheme in formulating the DRM.

Consider a Stokes vector signal to be measured that is a function of space, time, and wavelength $\mathbf{S}_{in}(x, y, t, \lambda)$. The system has an impulse response function that is described by $h(x, y, t, \lambda)$, a detector with a sampling function $d(x, y, t, \lambda)$, and a polarimetric modultation described by $\mathbf{S}_A(x, y, t, \lambda)$. We also consider a uniform sampling process in space, time, and wavelength with sampling intervals X, Y, T, and L*. Nonuniform sampling is possible as well, but complicates the analysis unnecessarily. With these definitions, the modulated irradiance is given as

$$I(x, y, t, \lambda) = \left[\left[\mathbf{S}_A(x, y, t, \lambda)^{\mathrm{T}} \left(h(x, y, t, \lambda) * \mathbf{S}_{in}(x, y, t, \lambda)\right)\right] * d(x, y, t, \lambda)\right] \operatorname{comb}\left(\frac{x}{X}, \frac{y}{Y}, \frac{t}{T}, \frac{\lambda}{L}\right). \tag{3.10}$$

In Eq. 3.10 the operator $*$ defines convolution; however, the various functions are not necessarily linear, shift-invariant (LSI) functions in general. When the system is not LSI, the convolution integrals have to take the more general form

$$g_1(x) = \int_{-\infty}^{\infty} f(\alpha) h(x, \alpha) d\alpha \tag{3.11}$$

rather than the more familiar form for LSI systems

$$g_2(x) = \int_{-\infty}^{\infty} f(\alpha) h(x - \alpha) d\alpha, \tag{3.12}$$

where in both cases the input is $f(x)$ and the impulse response is $h(x)$. Furthermore, the point spread function $h(x, y, t, \lambda)$ is assumed to be scalar. In reality, the system is really described by a polarimetric impulse response matrix that describes how the optical system alters the polarization state between object and image plane before it is ever sampled.[42] For convenience we will define our units so that $X = Y = 1$, $T = 1$, and $L = 1$, so we will be working in normalized frequency space (cycles per sample).

Referring back to Eq. 3.4, the DRM formalism computes the pseudoinverse as

$$\underline{\underline{\mathbf{W}}}^{-1} = \left(\underline{\underline{\mathbf{W}}}^{\mathrm{T}} \underline{\underline{\mathbf{W}}}\right)^{-1} \underline{\underline{\mathbf{W}}}^{\mathrm{T}}, \tag{3.13}$$

and the irradiance $I(x, y, t, \lambda)$ is rearranged into a column vector that $\underline{\underline{\mathbf{W}}}^{-1}$ operates on. The first part of the pseudoinverse in Eq. 3.13 is

$$\underline{\underline{\mathbf{Z}}} = \left(\underline{\underline{\mathbf{W}}}^{\mathrm{T}} \underline{\underline{\mathbf{W}}}\right)^{-1} = \left(\sum_{n=1}^{N} \mathbf{S}_A(n) \mathbf{S}_A(n)^{\mathrm{T}}\right)^{-1}. \tag{3.14}$$

In this notation, the polarimeter uses N measurements to form the measurement matrix. For example, a rotating retarder polarimeter might make $N = 16$ measurements as the retarder is rotated from 0° to 360°, or a DoFP polarimeter might be decomposed into 4-element (2×2) superpixels. In the linear systems formalism, we propose a similar formation of $\underline{\underline{\mathbf{Z}}}$

$$\underline{\underline{\mathbf{Z}}} \equiv \left(\sum_{x,y,t,\lambda} \mathbf{S}_A(x, y, t, \lambda) \mathbf{S}_A(x, y, t, \lambda)^{\mathrm{T}}\right)^{-1}. \tag{3.15}$$

The $\mathbf{W}^T\mathbf{W}$ part of the inverse is the integral of the product of the modulation functions contained in $\mathbf{S}_A$, so the inversion of this quantity describes how to separate the polarization signals if their modulation is not completely orthogonal. For modulation schemes that are orthogonal over the integral the quantity $\mathbf{W}^T\mathbf{W}$ and its inversion will simply be a diagonal matrix. We refer to $\underline{\underline{\mathbf{Z}}}$ as The modulator inner product inversion matrix.

*The system could equivalently be uniformly sampled in wavenumber $1/\lambda$, which is also common.

If the measurement has been taken such that there is an integer number of periods of modulation in all of the functions contained in $\mathbf{a}$, this quantity will be a constant with respect to the initial phase of the modulation. However, in cases where the polarimeter varies in time or in space, as would be the case when there is drift in the absolute angular position of a rotating retarder or in the absolute phase of oscillation of the dual-PEM system,[4] then it is possible that $\underline{\underline{\mathbf{Z}}}$ might need to be computed locally.[43]

Next we turn to the $\underline{\underline{\mathbf{W}}}^{\mathrm{T}}$ term in Eq. 3.14. To understand this element it is useful to examine how this term operates on the modulated irradiance in the standard DRM formalism. We have

$$\underline{\underline{\mathbf{W}}}^{\mathrm{T}}\mathbf{I} = \underline{\underline{\mathbf{W}}}^{\mathrm{T}}\underline{\underline{\mathbf{W}}}\,\mathbf{S}_{in}. \tag{3.16}$$

Examining this equation closely in the context of the linear systems formalism reveals that the matrix $\underline{\underline{\mathbf{W}}}^{\mathrm{T}}$ plays two roles. The first role is the homodyne in the demodulation process. The matrix $\underline{\underline{\mathbf{W}}}$ includes the modulation strategy of $\mathbf{S}_A(x,y,t,\lambda)$, multiplying the input signal by $\underline{\underline{\mathbf{W}}}$ then $\underline{\underline{\mathbf{W}}}^{\mathrm{T}}$ is equivalent to mixing with a carrier frequency once to move the base band signal up to the side bands and a second time to create a copy at base band along with spurious copies at higher frequencies. In communications theory, the next step is to low pass filter. This low pass filter is implicitly included in the matrix multiplication $\underline{\underline{\mathbf{W}}}^{\mathrm{T}}\underline{\underline{\mathbf{W}}}$. However, in the matrix multiplication the low pass filter has a rectangular footprint (in time, space, wavelength). This rectangular footprint results in a sing-function low pass filter that results in leakage of spurious high frequency signals into the reconstruction as we will see below.

In the linear systems formalism we want to leave ourselves the freedom to use an arbitrary low-pass filter. Specifically, we want to use band-limited low pass filters that are designed to only pass the base band signal, thus eliminating polarimetric aliasing. For the linear systems formalism, we accomplish this by separating the homodyne from the low pass filter as

$$\underline{\underline{\mathbf{W}}}^{\mathrm{T}}\mathbf{I} \rightarrow m(x,y,t,\lambda) * \mathbf{S}_A(x,y,t,\lambda)I(x,y,t,\lambda) = w(x,y,t,\lambda) * \mathbf{S}_A(x,y,t,\lambda)\mathbf{S}_A(x,y,t,\lambda)^{\mathrm{T}}\mathbf{S}_{in}, \tag{3.17}$$

where $w(x,y,t,\lambda)$ is the windowing function in the space-time-wavelength domain that corresponds to the desired low pass filter in the corresponding frequency domain.

Putting all of the terms together, we can form the linear systems version of the DRM method. The estimated Stokes vector distribution is

$$\hat{\mathbf{S}} = \underline{\underline{\mathbf{W}}}^{-1}\left\{I(x,y,t,\lambda)\right\} = w(x,y,t,\lambda) * \underline{\underline{\mathbf{Z}}}\mathbf{S}_A(x,y,t,\lambda)I(x,y,t,\lambda). \tag{3.18}$$

3.4 Discussion

In order to understand the polarimetric demodulation operation from the linear systems perspective, we consider the two examples of a rotating retarder polarimeter and a dual-PEM polarimeter.

3.4.1 Rotating Retarder Polarimeter

Consider once again the rotating retarder (RR) polarimeter. We assume that the polarimeter is an ideal linear retarder of retardance δ followed by an analyzer. Since our reconstruction was designed to be insensitive to absolute phase, we take the phase to be 0 with respect to $t = 0$. In this case the modulation term for the k^{th} temporal sample is

$$\mathbf{S}_A[n,m,k,l] = \mathbf{S}_A[k] = \tfrac{1}{2}\begin{bmatrix} 1 \\ \cos^2\frac{4\pi k}{K} + \cos\delta\sin^2\frac{4\pi k}{K} \\ (1-\cos\delta)\cos\frac{4\pi k}{K}\sin\frac{4\pi k}{K} \\ -\sin\delta\sin\frac{4\pi k}{K} \end{bmatrix} = \tfrac{1}{2}\begin{bmatrix} 1 \\ \frac{1+\cos\delta}{2} + \frac{1-\cos\delta}{2}\cos(\frac{8\pi k}{K}) \\ \frac{(1-\cos\delta)}{2}\sin\frac{8\pi k}{K} \\ -\sin\delta\sin\frac{4\pi k}{K} \end{bmatrix}, \tag{3.19}$$

where there are K samples per 360° rotation of the retarder. The modulator inner product inversion matrix is

$$\underline{\underline{\mathbf{Z}}} = \begin{bmatrix} \frac{12\cos^2\delta+8\cos\delta+12}{(\cos\delta-1)^2} & -\frac{16\cos\delta+16}{(\cos\delta-1)^2} & 0 & 0 \\ -\frac{16\cos\delta+16}{(\cos\delta-1)^2} & \frac{32}{(\cos\delta-1)^2} & 0 & 0 \\ 0 & 0 & \frac{32}{(\cos\delta-1)^2} & 0 \\ 0 & 0 & 0 & \frac{8}{\sin^2\delta} \end{bmatrix}. \tag{3.20}$$

We begin by assuming that the excitation has the form $\text{sinc}()^2(t)$ in only one of the four Stokes parameters. Even though this excitation is non-physical because it only includes one Stokes parameter, we can decompose an arbitrary physical excitation into these components in order to understand where the information from each Stokes parameter lies in the side bands. Figure 3.1 shows the measured irradiance for the $\text{sinc}()^2$ signal in each parameter.

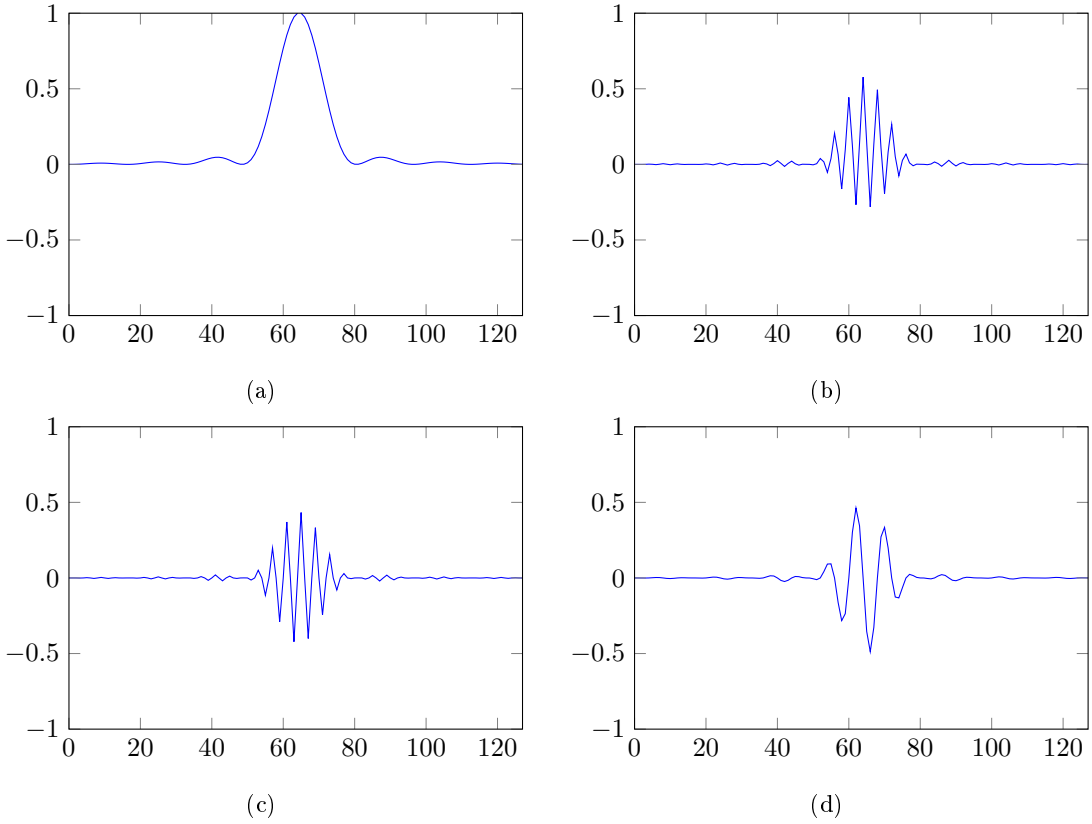
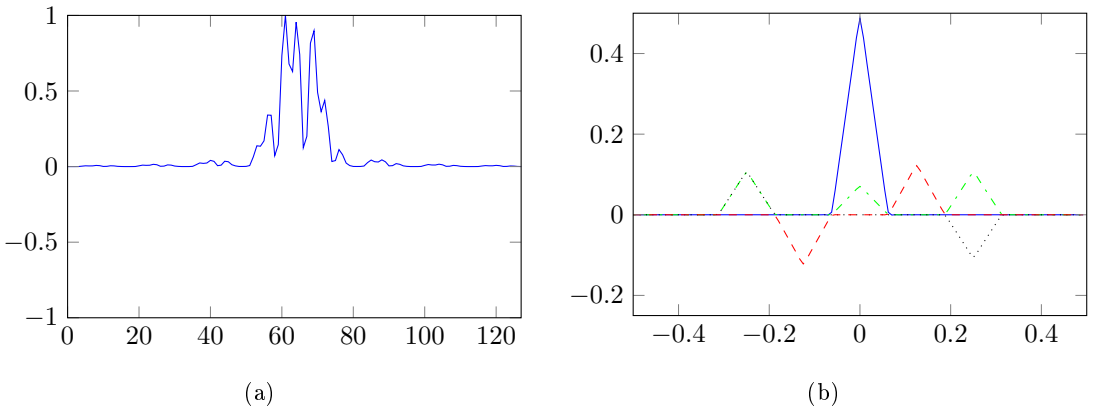


Figure 3.1: Shows the components of the irradiance due to S_0(a), S_1(b), S_2(c), and S_3(d)

Figure 3.2 shows the measured irradiance and its Fourier transform for the fully polarized signal

$$\mathbf{S} = \text{sinc}(t)^2 \begin{bmatrix} \sqrt{3} & 1 & 1 & 1 \end{bmatrix}^{\mathrm{T}}. \tag{3.21}$$

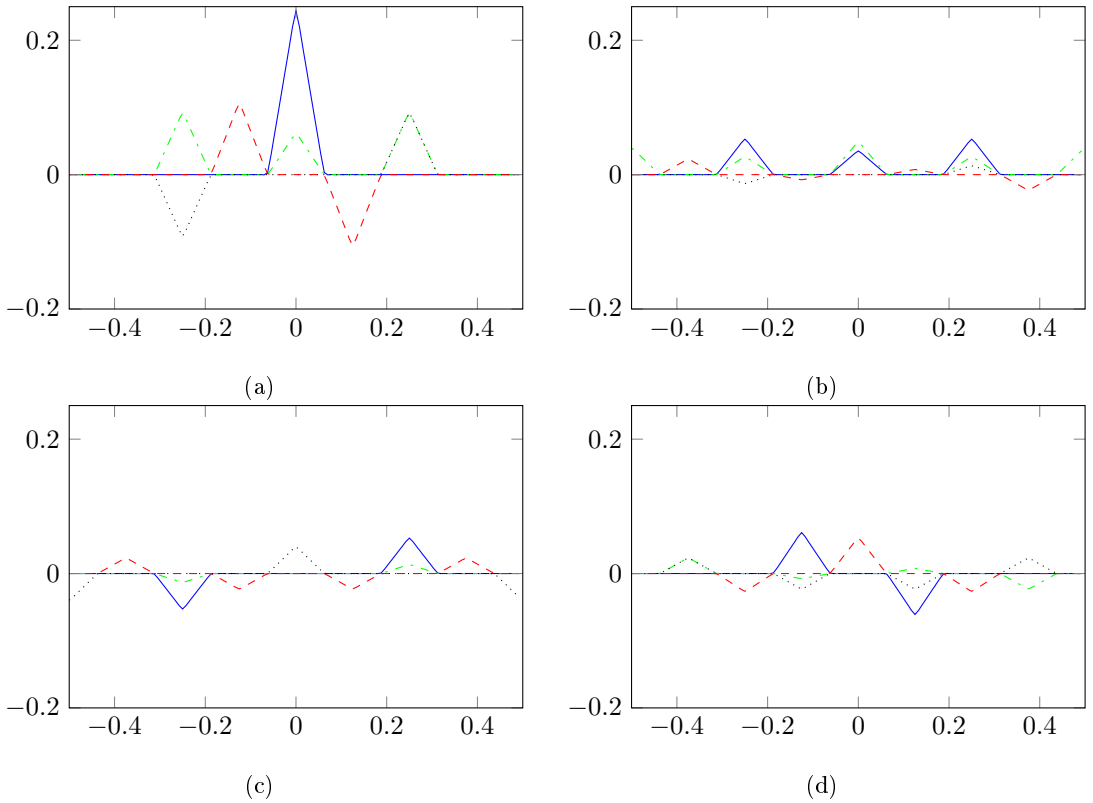


(a) (b)

Figure 3.2: (a) shows the measured irradiance for $\mathbf{S} = [\sqrt{3}111]$ in the temporal domain.and (b) shows the Fourier Transform of the measured irradiance. **Need to explain colors, and real/imag**

As expected, the irradiance signal shown in Fig. 3.2a is now all non-negative. Figure 3.2b shows the Fourier transform of the measured irradiance decomposed into the components that are due to each of the signals from Fig. 3.1.

The first step in the reconstruction process it to multiply the measured irradiance with the functions given in Eq. 3.19 in order to move the information for the desired Stokes parameters down to base band.



(a) (b) (c) (d)

Figure 3.3: Pre modulation of the measured irradiance in the Fourier Domain by a_0(a), a_1(b), a_2(c), a_3(d)

Figure 3.3 shows the Fourier transform of the modulated irradiance using each of the four modulation functions in Eq. 3.19. As we can see, the first and second modulation functions move the s_0 and s_1 information to base band. The third row moves the s_2 information to base band, while the fourth row moves the s_3 information to base band. In order to separate s_0 and s_1, it is still necessary to multiply by the modulator inner product inversion matrix $\underline{\underline{\mathbf{Z}}}$, and the result of that multiplication is shown in Fig. 3.4.

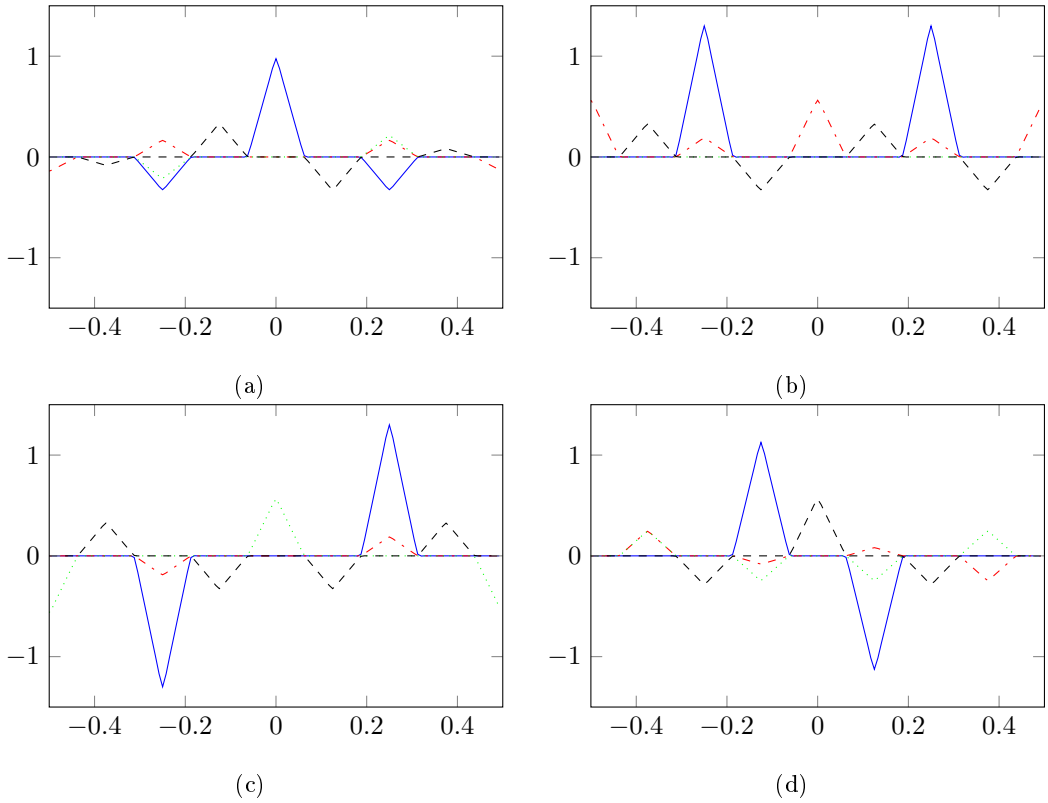


Figure 3.4: Pre filtered S_r for S_{r0}(a), S_{r1}(b), S_{r2}(c), S_{r3}(d)

At this point we can examine Fig. 3.4 and see that the band limited information has been moved to base band independently for each of the four Stokes paramters. Now the only step left is to select a suitable low-pass filter to extract the data. This is the step where virtually all division of time polarimeter reconstruction strategies make the crucial mistake of using a time-limited reconstruction window $w(x, y, t, \lambda)$ rather than a band-limited one. Figure 3.5 shows the effect of this when using a standard, 16-element rectangular window for $w(x, y, t, \lambda)$. Each of the four panels shows the Fourier transform of one of the Stokes paramters decomposed into the portions that arise from each of the inputs in Fig. 3.1.

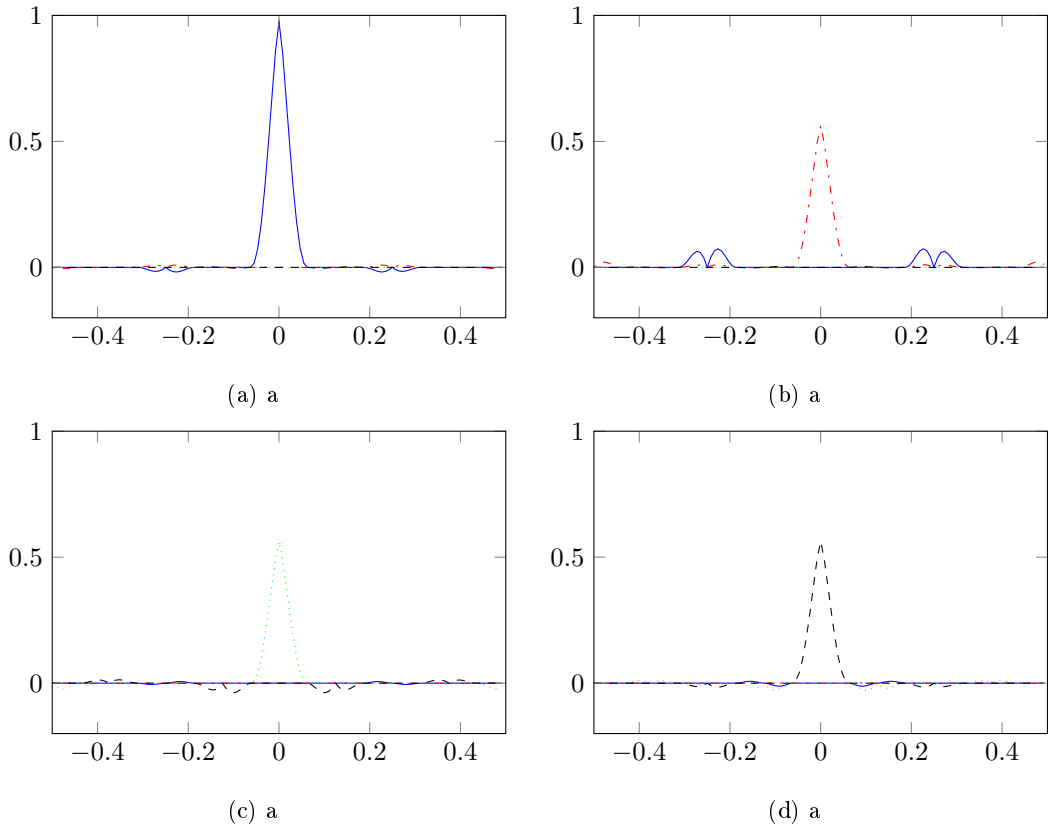


Figure 3.5: Result of multiplication of a filter from a 16 element rect window in the Fouerier Domain. S_{r0}(a), S_{r1}(b), S_{r2}(c), S_{r3}(d)

We can see that the proper signal is reconstructed at base band in each case, but there is also a form of polarimetric aliasing error. This error manifests both as self-error (e.g., baseband s_0 information aliasing up to higher frequencies in s_0 through the demodulation process) and cross error (e.g., base band s_0 information showing up as high frequency error in the s_1 signal). Figure 3.6 shows the reconstructed Stokes parameters in the time domain, where the high frequency reconstruction error is evident.

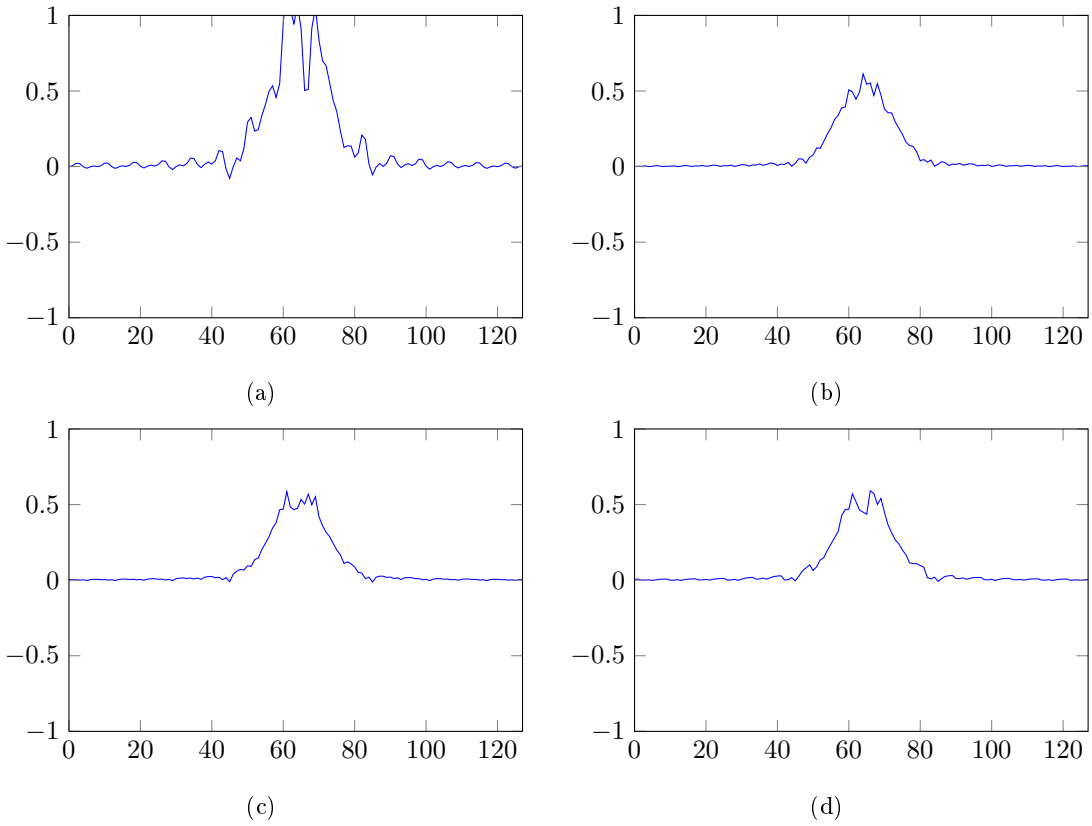


Figure 3.6: Result of multiplication of a filter from a 16 element rect window in the temporal domain. S_{r0}(a), S_{r1}(b), S_{r2}(c), S_{r3}(d)

3.4.2 Performance Issues

The theory presented in section 3.3 along with the example of the RR polarimeter in section 3.4.1 demonstrate some issues associate with traditional DRM processing of modulated polarimeter data. The conventional DRM formulation is created assuming that the input Stokes vector is constant in the modulated dimension. For that reason a standard matrix formalism is obvious that equally weights all of the observations in an N-element window in performing the inversion. Since the excitation in this case is assumed DC, the analysis only guarantees proper reconstruction for DC signals. In this paper, we have demonstrated that there is a fundamental band limit criterion that must be satisfied for a polarimeter that is related to both the physical sampling frequency *and* the modulation frequency that introduces the polarization side bands. So long as Eq. 3.9 is satisfied, error-free polarimetric reconstruction is possible. This result relaxes the requirement for extremely high modulation frequencies that has been used recently to improve error in DoT polarimeters.[35,44]

The analysis of Section 3.3 used an analysis of the DRM formalism to illuminate the linear systems viewpoint. However, it is equally possible to to use the knowledge that band limited reconstructions exists to create a DRM picture that can be used in the traditional linear algebra sense. The two viewpoints must be identical, but some readers might find one easier to implement than the other for their purposes.

In order to construct band-limited DRMs that can reconstruct signals that obey Eq. 3.9, it is necessary to augment the rows of the instrument matrix with vectors that represent the Fourier components in the base band used to reconstruct the Stokes parameters.

The theory presented here was focused on DoT polarimeters. However, this scheme applies equally to all forms of modulated polarimeters. Similar linear systems analyses have been performed for channeled spectropolarimeters,[3,45] but this paper represents the first attempt to unify the two theories. Furthermore, in this and

other analyses, the modulation is always assumed to be in only one domain, i.e. temporal modulation in a DoT polarimeter, spatial modulation in a DoFP polarimeter, or spectral modulation in a channeled spectropolarimeter. Equation 3.10 anticipates the case where the polarimetric modulation represented by $\mathbf{S}_A(x, y, t, \lambda)$ actually modulates across dimensions in a way that allows more accurate reconstruction and/or greater bandwidth in particular applications.

4. SPECTRAL DENSITY RESPONSE FUNCTIONS FOR MODULATED POLARIMETERS

4.1 Background

Conventional imaging systems are often described by their point spread function (PSF) or optical transfer function (OTF) in space and their temporal impulse response. These functions can be used to objectively compare systems based on criteria that are determined by the specific imaging problem. When comparing modulated polarimeters, transfer functions in their conventional form cannot be employed. This is because modulated polarimeters multiplex signals in the frequency domain; a single frequency sinusoidal object will result in an image with multiple frequencies. This ruins the linear shift invariant assumption required for the transfer function formalism to be valid, since a measurement at a given frequency could have come from a multiplicity of input signals. A response function for modulated polarimeters is developed here to perform an analogous objective comparison for polarimetric imaging tasks that the OTF/PSF formalism allows for conventional imaging systems.

The band-limited filter algorithm presented in section 3 is to establish a method to objectively compare modulation schemes (spatial, temporal, or some combination) to determine the method that performs "best" on a polarized object that has finite bandwidth both in the temporal and spatial domain. This tool is termed the *spectral density response function*. The tool is used in Section 4.3 to determine an object-specific, optimal Wiener filter as an example of the application of the theory.

Power spectral densities (PSDs) are used to describe objects and systems when deterministic properties about the measurement process are not known, but statistical properties are known or can be reasonably well modeled. Methods using PSDs describe the performance of the system on average, and any optimization of a system using such techniques may not be optimal for a specific instantiation (i.e., a matched filter would be best for an instantiation where the object and any interfering signals are deterministically known).

In remote sensing applications, polarimeters are often used to image large areas of similar terrain, and they are optimized for measurement of the entire data set as opposed to one particular view. The fact that these imagers measure large quantities of data, and have prior knowledge of signal statistics for given applications, makes comparison based on the PSDs a reasonable choice. For instance, in satellite imaging applications, an existing large pool of data allows estimation of the nominal intensity statistics for a variety of wavelengths, even if the polarization PSDs are not known.

For Stokes polarimeters, the object to be measured can be represented as a vector, i.e., an image of three (linear polarimeters) or four (complete polarimeters) Stokes parameters. The power spectrum must therefore be represented as a spectral density matrix. The spectral density matrix (SDM) is defined as the expected value of the outer product of the Stokes parameters ($\mathbf{S}(f)$) as a function of frequency,

$$\mathbf{SDM}(f) = \left\langle \mathbf{S}(f)\mathbf{S}^{\dagger}(f) \right\rangle = \left\langle \|\mathbf{S}(f)\|^2 \right\rangle. \tag{4.1}$$

The SDM contains information about the Fourier transforms of the auto- and cross-correlation functions among the Stokes parameters.

The goal of this effort is to develop a spectral density response function, analogous to the transfer function of conventional optics. The spectral density response function relates the SDM of the object Stokes parameters to the spectral densities of the image Stokes parameters. However, the measurement process for modulated polarimeters is not linear shift invariant (LSI), so the resulting spectral density response function must be able to characterize the magnitude of response of all output frequencies for a given input frequency, which is somewhat more complicated than for a LSI system. As an example, consider a pushbroom polarimeters such as the Multiangle Spectro-Polarimetric Imager (MSPI),[4,46] which builds up an image one line at a time as the

platform moves across the scene. For these instruments, spatial variation is translated into the temporal response of the detector, which causes spatial variations and temporal noise to combine with the temporal modulation strategy. The formalism presented here can be applied to such imagers.

4.2 Spectral density response

4.2.1 Derivation

Consider designing an imaging polarimeter for an application where the Stokes parameter SDM is known. In this situation it is desirable to predict the PSDs of the output Stokes parameters and modify the polarimeter design to optimize the output based on criteria such as the signal to noise ratio (SNR). The polarimeter imaging operator is defined as

$$\hat{\mathbf{S}}[n_x, n_y, n_t] = \mathcal{P}\left\{\mathbf{S}(x, y, t)\right\}, \tag{4.2}$$

where $\mathbf{S}(x, y, t)$ is a continuous incident set of object Stokes parameters and $\hat{\mathbf{S}}[n_x, n_y, n_t]$ is a discrete estimate of the incident polarization quantities. The polarimeter operator includes both the physical measurement of the incident fields and the reconstruction from that measurement into the final estimation of the scene. This operator is useful because it provides a description of the entire system from measurement to estimation at once, instead of first choosing a measurement scheme and then the estimation algorithm. Optimizing the polarization operator for an imaging task will potentially lead to different solutions from conventional optimization techniques.[27]

The operator $\mathcal{P}$ is developed from a basic description of the system. This discussion will address time modulated polarimeters, since such instruments are single variate. However, the approach can be readily applied to more dimensions, such as spatially modulated polarimeters (microgrids)[3] or a system modulated in both space and time,[7] and is directly applicable to all periodically modulated instruments. Generalization to non-periodic polarimeter modulation schemes[47] is left to future work. The optics will be assumed to have an ideal (flat) transfer function (infinite aperture diffraction limited optics), but the detector impulse response is explicitly included and modeled as a finite integrator:

$$h(t) = \text{rect}\left(\frac{t}{T}\right), \tag{4.3}$$

where $\text{rect}(t) = 1$ for $|t| < 1/2$, $\text{rect}(t) = 1/2$ for $|t| = 1/2$, and 0 otherwise. The sampling function is defined as

$$\text{Ш}(n, t) = \delta(t - nT), \tag{4.4}$$

which assumes periodic sampling with a sampling interval of T and sample index n. The flux that passes through a time-varying analyzer is given by

$$P(t) = \mathbf{a}^{\mathrm{T}}(t)\mathbf{S}(t), \tag{4.5}$$

where $\mathbf{a}(t)$ is the analyzer vector (the first row of the Mueller Matrix of the polarimeter optics.[28] This flux is then sampled by the detector

$$I[n] = \int_{-\infty}^{\infty} \left(h(t) * P(t)\right) \text{Ш}(n, t)\mathrm{d}t, \tag{4.6}$$

where $*$ is the convolution operator. The band-limited data reduction algorithm proposed in[9] is described by

$$\hat{\mathbf{S}}[n] = w[n] * \underline{\underline{\mathbf{Z}}}^{-1}[n]\mathbf{a}_s[n]I[n], \tag{4.7}$$

where $\mathbf{a}_s[n]$ is a sampled version of the analyzer vector, $w[n]$ is the reconstruction window (which will be discussed in detail below), and the analyzer inversion matrix is

$$\underline{\underline{\mathbf{Z}}}^{-1}[n] = \left(w[n] * \mathbf{a}_s^{\mathrm{T}}[n]\mathbf{a}_s[n]\right)^{-1}. \tag{4.8}$$

Equations 4.6, 4.7, and 4.8 constitute the full mathematical model of the polarimeter operator using band-limited reconstruction. This operator can now be used to compute the spectral density response. This analysis will be most easily accomplished in the Fourier domain for the periodically modulated polarimeters considered here.

First we note that the integral over all time in Eq. 4.6 can be computed by evaluating the Fourier transform of the integrand at $f = 0$:

$$I[n] = \left(\tilde{h}(f)\tilde{P}(f)\right) * \tilde{\text{III}}(n, f)\Big|_{f=0}. \tag{4.9}$$

The Fourier transform of $P(t)$ is found by taking the Fourier transform of Eq. 4.5,

$$\tilde{P}(f) = \tilde{\mathbf{a}}^{\mathrm{T}}(f) * \tilde{\mathbf{S}}(f). \tag{4.10}$$

Substituting Eqs. 4.9 and 4.10 into Eq. 4.7 yields

$$\hat{\mathbf{S}}[n] = w[n] * \underline{\underline{\mathbf{Z}}}^{-1}[n]\mathbf{a}_s[n] \left(\left(\tilde{h}(f)\left(\tilde{\mathbf{a}}^{\mathrm{T}}(f) * \tilde{\mathbf{S}}(f)\right)\right) * \tilde{\text{III}}(n, f)\right)\Big|_{f=0}. \tag{4.11}$$

A choice that seems reasonable for the sampled analyzer vector calculation is

$$\mathbf{a}_s[n] = \int_{-\infty}^{\infty} (h(t) * \mathbf{a}(t))\, \text{III}\,(n, t)\, \mathrm{d}t, \tag{4.12}$$

since this samples the analyzer vector the same way the flux is sampled. Another reasonable choice might be sampling the analyzer with ideal point sampling. The point of this discussion is that this parameter is free for the algorithm designer to choose to best fit the underlying measurement. In fact, the sampled analyzer vector can be augmented with additional signals so that the estimation of the Stokes parameters is orthogonal to certain known sources of error; the authors of[4] augmented their conventional data reconstruction matrix (DRM) with linear gradients to force the system estimation to ignore slowly varying components of the signal.

Equation 4.11 becomes more manageable when we make some simplifications to $\tilde{\mathbf{a}}(f)$ and $\tilde{\mathbf{S}}(f)$. The first simplification is to assume that $\mathbf{a}(t)$ has a period B and can be written in terms of its Fourier series[48]

$$\mathbf{a}(t) = \sum_{b=-\infty}^{\infty} \mathbf{c}_b \exp\left(2\pi i \frac{b}{B} t\right) = \sum_{b=-\infty}^{\infty} \mathbf{c}_b \exp\left(2\pi i f_b t\right), \tag{4.13}$$

with $f_b = b/B$. The Fourier transform of this Fourier series is a weighted sum of delta functions

$$\tilde{\mathbf{a}}(f) = \sum_{b=-\infty}^{\infty} \mathbf{c}_b \delta\left(f - \frac{b}{B}\right) = \sum_{b=-\infty}^{\infty} \mathbf{c}_b \delta(f - f_b). \tag{4.14}$$

We will also assume periodicity of the quantity $\underline{\underline{\mathbf{Z}}}^{-1}[n]\mathbf{a}_s[n]$, since we have already assumed periodicity of the analyzer vector. Likewise, the Fourier series of this function is

$$\underline{\underline{\mathbf{Z}}}^{-1}[n]\mathbf{a}_s[n] = \sum_{p=0}^{P-1} \mathbf{d}_p \exp\left(\frac{2\pi i p n}{P}\right) = \sum_{p=0}^{P-1} \mathbf{d}_p \exp\left(2\pi i T f_p n\right), \tag{4.15}$$

where P is the period of $\underline{\underline{\mathbf{Z}}}^{-1}[n]\mathbf{a}_s[n]$, p is an integer, and $f_p = p/PT$. Note that P has no units since it is the period of a discrete set, while B has units of time since it is defined in the continuous domain. The summation only has P elements, i.e., the number of measurements taken over one period of the function in Eq. 4.15. The period T must be included in the definition of f_p so that this quantity has units of frequency. Another simplification made here is the introduction of the quantity

$$\beta(f - f_o) = \tilde{\mathbf{S}}(f)\delta(f - f_o), \tag{4.16}$$

where $\beta(f - f_o)$ is a frequency-domain function that represents a pure sinusoidal excitation in time at a frequency given by f_o with a Fourier domain amplitude described by $\tilde{\mathbf{S}}(f)$. This modification is made so that the response function can be built up by integrating over all f_o to find the total response to the input.

$$\tilde{\mathbf{S}}(f) = \int_{-\infty}^{\infty} \beta(f - f_o)\mathrm{d}f_o. \tag{4.17}$$

The function $\beta(f-f_o)$ will replace $\tilde{\mathbf{S}}(f)$ in the above equations for the derivation of the spectral density response function.

Using the simplifications in Eqs. 4.14 and 4.15, the representation of the sampled flux in Eq. 4.9 changes to

$$I[n] = \left(\tilde{h}(f) \left(\sum_{b=-\infty}^{\infty} \mathbf{c}_b^{\mathrm{T}} \delta(f - f_b) * \tilde{\mathbf{S}}(f) \right) \right) * \tilde{\mathrm{III}}(n, f) \Bigg|_{f=0} . \tag{4.18}$$

The convolution of $\tilde{\mathbf{S}}(f)$ with delta functions simply shifts the argument of $\tilde{\mathbf{S}}(f)$ in frequency, therby creating side bands in the Fourier domain[48]

$$I[n] = \left(\tilde{h}(f) \left(\sum_{b=-\infty}^{\infty} \mathbf{c}_b^{\mathrm{T}} \tilde{\mathbf{S}}(f - f_b) \right) \right) * \tilde{\mathrm{III}}(n, f) \Bigg|_{f=0} . \tag{4.19}$$

When examining the definition of $I[n]$ in Eq. 4.6, we find that only the sampling function $\mathrm{III}(n,t)$ depends on n. The Fourier transform of $\mathrm{III}(n,t)$ over t is

$$\mathcal{F}_{t\to f} \left\{ \mathrm{III}\left(n,t\right) \right\} = \exp\left(2\pi i n T f\right), \tag{4.20}$$

and the discrete Fourier transform (DFT) of the sampling function over index n is

$$\begin{aligned} \mathcal{F}_{n\to k} \left\{ \tilde{\mathrm{III}}\left(n,f\right) \right\} &= \sum_{n=0}^{N-1} \exp\left(2\pi i n \left[Tf - \frac{k}{N} \right] \right) \\ &= \frac{1 - \exp\left(2\pi i N T f\right)}{1 - \exp\left(2\pi i T \left[f - \frac{k}{NT} \right] \right)}. \end{aligned} \tag{4.21}$$

Since k is an integer and does not have units of frequency, a useful variable to define is the sampled frequency f_k:

$$f_k = \frac{k}{NT}, \tag{4.22}$$

which relates the sampled frequencies back to the excitation frequencies. Using Eq. 4.22 in Eq. 4.21 yields

$$\tilde{\mathrm{III}}\left(f_k, f\right) = \frac{1 - \exp\left(2\pi i N T f\right)}{1 - \exp\left(2\pi i T \left[f - f_k \right]\right)}. \tag{4.23}$$

The behavior of the sampling function is not ideal; the Fourier transform of the ideal sampling function would have the form $\delta(f - f_k)$, where non-zero reconstructed values would only occur exactly when the sampling frequency equals the incident frequency. The form of Eq. 4.23 allows frequency leakage from frequencies where $f_k \neq f$. Frequency leakage is a fundamental sampling issue, but is specifically discussed here to fully develop a model of modulated imaging polarimetry. Now we can consider that the sampling delta function is changed to include two shifts introduced by the modulation of the measuring analyzer vector and the modulation of the data processing algorithm. When the integers b and p are zero, the sampling occurs exactly at $f_k = f$ just as with a conventional imager. As b and p change they allow the excitation of a pure sinusoidal input with a frequency f to create a response in the system at a frequency given by $f_k = f + f_b + f_p$. The sampling still occurs at specific, discrete intervals, but now there is ambiguity in the estimation because multiple excitations can create overlapping responses.

There are two aspects to consider when dealing with sampling artifacts in polarimeters: aliasing and frequency leakage. For the examples used in this discussion, the effects of aliasing will be shown in the following section. The response of the real system to frequencies where $f \neq f_k$ is given by Eq. 4.23, which, for f sufficiently close to f_k, can be approximated with the function $\tilde{\mathrm{III}}(f_k, f) = \mathrm{sinc}(T(f - f_k))$, where the sinc function allows leakage from a band in f to a particular value f_k. The behavior when $f = f_k$ is still ideal, but for frequencies $f_{k-1} < f < f_k$ there is leakage. The frequency leakage behavior can be included by convolving the ideal results with the sinc function that models the leakage, and will be ignored for now.

With the Fourier transform over f complete, now we need to take the discrete Fourier transform (DFT) of Eq. 4.11 over the index n. First, substituting the Fourier series of the sampled analyzer vector and $\underline{\underline{\mathbf{Z}}}$ yields

$$\hat{\mathbf{S}}[n] = w[n] * \sum_{p=0}^{P-1} \mathbf{d}_p \exp\left(\frac{2\pi i p n}{P}\right) I[n]. \tag{4.24}$$

Taking the DFT of $\hat{\mathbf{S}}[n]$ and using properties of delta functions[48] provides

$$\tilde{\hat{\mathbf{S}}}[f_k] = \sum_{p=0}^{P-1} \mathbf{d}_p \tilde{w}[f_k] \tilde{I}[f_k - f_p]. \tag{4.25}$$

Now that the total polarimeter operator is represented in the Fourier domain, the spectral density response can be calculated. The set of power spectra for the estimated Stokes parameters (or spectral density response) is defined as the Fourier transform of the auto correlation function. This can be written as

$$\underline{\underline{\hat{\mathbf{R}}}}[f_k] = \mathcal{F}_{n \to f_k}\left\{\hat{\mathbf{S}}[n] \star \hat{\mathbf{S}}[n]\right\} = \tilde{\hat{\mathbf{S}}}[f_k]\tilde{\hat{\mathbf{S}}}^{\dagger}[f_k], \tag{4.26}$$

where † represents the Hermitian adjoint and $\star$ is the correlation operator.[48] This quantity can be calculated in the Fourier domain as

$$\begin{aligned}\underline{\underline{\hat{\mathbf{R}}}}[f_k] = \sum_{p,p'=0}^{P-1} \sum_{b,b'=-\infty}^{\infty} & \tilde{h}(f - f_k + f_p)\tilde{h}^*(f - f_k + f_{p'}) \left\|\tilde{w}[f_k]\right\|^2 \\ & \times \mathbf{d}_p \mathbf{c}_b^{\mathrm{T}} \underline{\underline{\mathbf{R}}}(f - f_1, f - f_2)\mathbf{c}_{b'}^* \mathbf{d}_{p'}^{\dagger},\end{aligned} \tag{4.27}$$

where $f_1 = f_k - f_p + f_b$, $f_2 = f_k - f_{p'} + f_{b'}$, and the object SDM $\underline{\underline{\mathbf{R}}}(f - f_1, f - f_2) = \left\langle \tilde{\mathbf{S}}(f - f_1)\mathbf{S}(f - f_2)^{\dagger} \right\rangle$ describes the cross spectral density of the Stokes parameters and the cross spectral density of differently modulated Stokes parameters. Note that * denotes the non transposed complex conjugate of a vector. Equation 4.27 provides a solution to the problem that we set out to solve originally: given a known SDM for the object Stokes parameters, what is the spectral density response of the estimated Stokes parameters? Using these methods, and with a known estimate of the incident Stokes parameters, the spectral density response $\underline{\underline{\tilde{\mathbf{R}}}}[f_k]$ can be directly computed.

4.2.2 Illustration

To illustrate the concepts leading up to Eq. 4.27, consider the case of a rotating analyzer polarimeter, which can be described by

$$\mathbf{a}(t) = \frac{1}{2}\begin{bmatrix} 1 \\ \cos(2\pi \frac{f_s}{8} t) \\ \sin(2\pi \frac{f_s}{8} t) \\ 0 \end{bmatrix}, \tag{4.28}$$

where f_s is the sampling rate of the sensor. The Fourier transform of the rotating analyzer vector is shown in Fig. 4.1. We shall also assume that the system performs ideal periodic point sampling, such that $h(t) = \delta(t)$, or $\tilde{h}(f) = 1$.

Even though the aim of this paper is to describe the operation of a modulated polarimeter on stochastic signals, the steps from the previous subsection are more easily visualized using a deterministic signal. In this illustration, we assume that

$$\mathbf{S}(t) = \begin{bmatrix} \sqrt{2} \\ 1 \\ 1 \end{bmatrix} \mathrm{sinc}^2(t), \tag{4.29}$$

where[48]

$$\mathrm{sinc}(t) = \frac{\sin \pi t}{\pi t}. \tag{4.30}$$

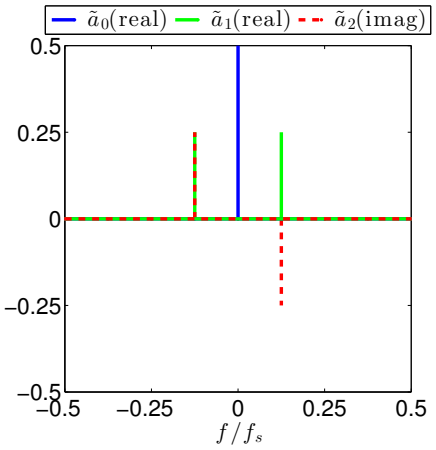


Figure 4.1: Analyzer vector for a rotating polarizer polarimeter in the Fourier domain. Since the analyzer is comprised of periodic functions, in the Fourier domain it is simply comprised of shifted delta functions.

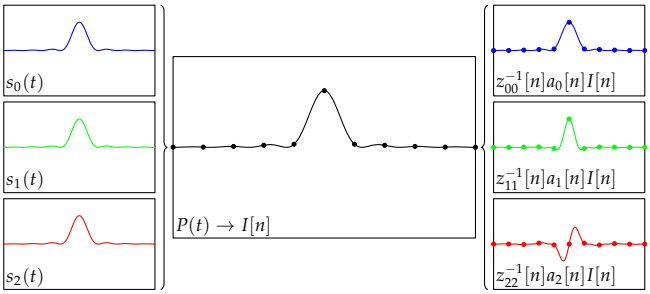


Figure 4.2: The components of the Stokes parameters s_0 – s_2 modulate the carriers a_0 – a_2 to produce the final sampled signal $I[n]$. The demodulation process is composed of remodulation using a_0 – a_2, then unmixing and equalizing the channels using $\underline{\underline{\mathbf{Z}}}^{-1}$. The colors allow us to track the contributions from s_0 – s_2 through the modulation and demodulation processes.

The steps in Eq. 4.5 – Eq. 4.8 in the time domain are shown in Fig. 4.2, and the demodulation steps are depicted in the frequency domain in Fig. 4.3 with an arbitrary low-pass filter. The figure shows how the information from each Stokes parameter channel maps into the components of the final signal.

The computation of the resulting spectral densities for s_0, s_1, and s_2 given in Eq. 4.27 is shown in Fig. 4.4.

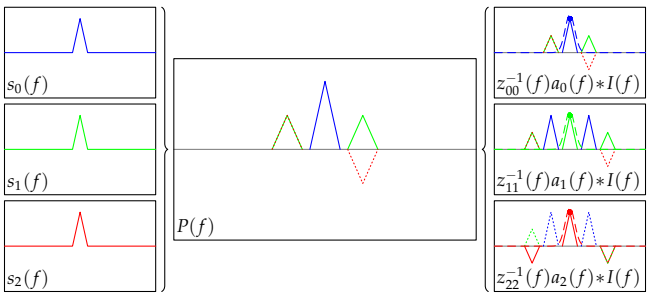


Figure 4.3: In the frequency domain, the three baseband signals corresponding to the Stokes parameters are placed into orthogonal channels by the carriers. In the demodulation process, each of the carriers is used to place the corresponding desired side band at baseband. A filter is used to isolate the desired signal to create the final estimate.

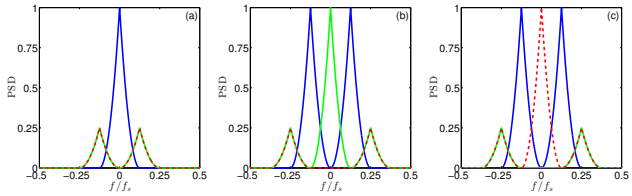


Figure 4.4: Graphical representation of the estimation of the spectral density response function before filtering, showing the contributions due to all three parameters. The total spectral density for each parameter is obtained by summing all contributions. Note that the signal for s_3 is shown with a dashed red line. (a) Reconstruction of s_0. (b) Reconstruction of s_1. (c) Reconstruction of s_2.

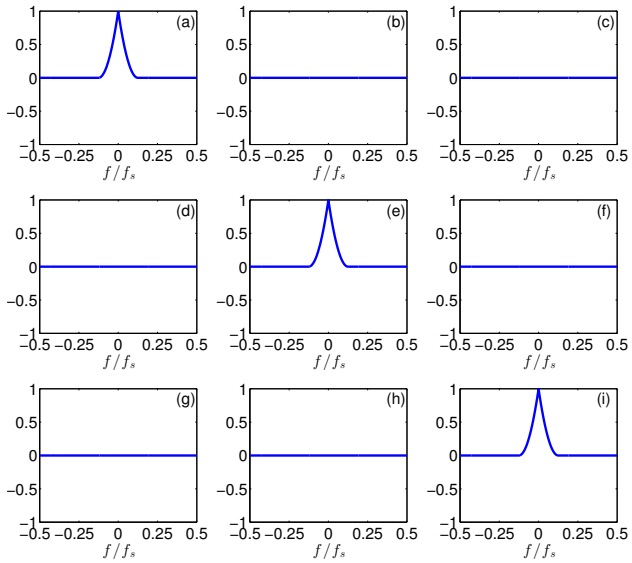


Figure 4.5: The assumed object SDM for this exercise. The diagonal elements correspond to the spectral density for (a) S_0, (e) S_1, and (i) S_2. The off diagonal elements represent the cross-spectral densities, for example (b) is the cross-spectral density between S_1 and S_0.

4.3 Example: Computation of the Wiener filter for modulated polarimeters

Finally, we assume that the SDM for the object to be measured is band-limited and described according to the equation

$$\underline{\underline{\mathbf{R}}}(f) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathrm{tri}^2\left(\frac{f}{5}\right). \tag{4.31}$$

The object SDM is assumed to be uncorrelated for differently modulated Stokes parameters $f_1 \neq f_2$, and the matrix is shown in Fig. 4.5. Note that because $a_3(t) = 0$, the S_3 information is lost, thereby allowing us to reduce $\underline{\underline{\mathbf{Z}}}^{-1}$ and $\underline{\underline{\mathbf{R}}}(f)$ to 3×3 matrices. It would also be possible to retain the dimensionality, but because $\underline{\underline{\mathbf{Z}}}$ would no longer be square, we would have to resort to using a pseudoinverse instead.

The classical DRM method[28] (referred to here as "the sliding pseudo inverse") and the ideal band-limited filters[9] have been presented in the literature, but for practical applications neither filter performs adequately for imaging polarimeters.[49] The sliding pseudo inverse produces strong artifacts at edges in the object, while the ideal band-limited filter only performs well on band-limited, noiseless data; estimation from noisy or non-band-limited data often produces strong ringing artifacts in the space-time domain. A more "visually pleasing" band-limited filter would be one that has a smooth transition to zero, rather than the sharp cutoff that the ideal band-limited filter has. One approach to building a smooth band-limited filter is the Wiener filter. The Wiener filter is well studied; one in-depth version of the Wiener filter and Wiener-Helstrom theory is in.[50] This filter is

derived for non correlated white noise using the metric of least square error. The Wiener filter is used here only as an example to illustrate the application of the current theory, since it is a well-defined, optimal linear filter. Other filters or nonlinear reconstruction methods may be more desirable for a particular application, and the spectral density response formalism developed above can be used for those as well.

The Wiener filter is designed to solve the estimation problem described by

$$\hat{y}(t) = w(t) * g(t), \tag{4.32}$$

where $g(t)$ is the measured signal

$$g(t) = s(t) + n(t). \tag{4.33}$$

In Eq. 4.33, $s(t)$ is the true signal and $n(t)$ is additive, uncorrelated noise. The Wiener filter is found by solving

$$w(t) \left(g(t) \star g(t)\right) = s(t) \star g(t), \tag{4.34}$$

and in the Fourier domain the Wiener filter is

$$\tilde{w}(f) = \frac{\|s(f)\|^2}{\|s(f)\|^2 + \|n(f)\|^2}. \tag{4.35}$$

The Wiener-Helstrom filter, which makes use of knowledge of the transfer function $\tilde{h}(f)$ of the system (also derived in[50]), can be calculated using

$$\tilde{w}(f) = \frac{\tilde{h}^*(f) \|s(f)\|^2}{\left\|\tilde{h}(f)\right\|^2 \|s(f)\|^2 + \|n(f)\|^2}. \tag{4.36}$$

The estimator in Eq. 4.11 unfortunately has two instances of the filter, the second of which is contained in the calculation of $\underline{\underline{\mathbf{Z}}}$. The analytic calculation of a filter that minimizes square error is difficult with the filter included in the calculation of the inverse. However, if we impose some loose constraints on $\tilde{w}(f)$, the problem simplifies, and we can ignore the fact that $\tilde{w}(f)$ is included in the calculation of $\underline{\underline{\mathbf{Z}}}$. Consider the portion of the operator inside of the inverse:

$$\underline{\underline{\mathbf{Z}}}^{-1}(t) = \left(w(t) * \mathbf{A}(t)\mathbf{A}^{\mathrm{T}}(t)\right)^{-1}. \tag{4.37}$$

To ignore the $w(t)$ in the definition of $\underline{\underline{\mathbf{Z}}}(t)$, the condition

$$\left(w(t) * \mathbf{A}(t)\mathbf{A}^{\mathrm{T}}(t)\right) = \int_{-\infty}^{\infty} \left(\mathbf{A}(t')\mathbf{A}^{\mathrm{T}}(t')\right) \mathrm{d}t' \tag{4.38}$$

must be satisfied. If we assume that $\mathbf{A}(t)$ is periodic, it can be rewritten in terms of its Fourier series

$$a_j(t) = \sum_{n=-\infty}^{\infty} z_{jm} \exp\left(2\pi i \frac{m}{T} t\right). \tag{4.39}$$

Using Eq. 4.39 in the right side of Eq. 4.38 yields

$$Z_{jj'} = \sum_{m'=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} z_{jm} z_{j'm'} \delta[m + m']. \tag{4.40}$$

Using Eq. 4.39 in the left side of Eq. 4.38 yields

$$Z_{jj'}(t) = \sum_{m'=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} z_{jm} z_{j'm'} w\left(t\right) * \exp\left(2\pi i \frac{m+m'}{T} t\right). \tag{4.41}$$

Taking the Fourier transform of Eq. 4.41 yields

$$Z_{jj'}(f) = \sum_{m'=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} z_{jm} z_{j'm'} \tilde{w}(f)\, \delta\left(f - \frac{m+m'}{T}\right). \tag{4.42}$$

For Eqs. 4.40 and 4.42 to be equivalent, the restrictions imposed on the filter are

$$\tilde{w}(f) = \begin{cases} 1, & \text{for } f = 0 \\ 0, & \text{for } f = \frac{m+m'}{T} \\ \text{arbitrary}, & \text{otherwise} \end{cases}. \tag{4.43}$$

These restrictions allow us to consider a large set of inverse operators for which the calculation of $\underline{\underline{\mathbf{Z}}}^{-1}$ yields identical results, and we can ignore $\tilde{w}(f)$ in its calculation.

For the moment, assume the impulse response functions associated with the system optics and detector are ideal. With the conditions imposed on $\tilde{w}(f)$ in Eq. 4.43, as well as ideal point impulse response functions, the unfiltered estimate of the Stokes parameters in the Fourier domain can be written as

$$\tilde{\mathbf{g}}[f_k] = \sum_{p=0}^{P-1} \sum_{b=-\infty}^{\infty} \mathbf{d}_p \mathbf{c}_b^{\mathrm{T}} \tilde{\mathbf{S}}(f - f_1), \tag{4.44}$$

with f_1 and f_2 defined in Section 4.2.1. Using this $\tilde{\mathbf{g}}[f_k]$ in the calculation of the Wiener filter without noise provides

$$\tilde{w}_i[f_k] = \frac{\sum_{p,b} z_{p,i} \mathbf{c}_b^{\mathrm{T}} \left\langle \tilde{\mathbf{S}}(f - f_1) \tilde{S}_i^*(f) \right\rangle}{\sum_{p,p',b,b'} z_{p,i} \mathbf{c}_b^{\mathrm{T}} \left\langle \tilde{\mathbf{S}}(f - f_1) \tilde{\mathbf{S}}^*(f - f_2) \right\rangle \mathbf{c}_{b'}^* z_{p',i}^{\dagger}}. \tag{4.45}$$

Rewriting this in terms of the cross-spectral density matrix defined in Eq. 4.27

$$\tilde{w}_i[f_k] = \frac{\sum_p \sum_b z_{p,i}^* \mathbf{c}_b^{\dagger} \underline{\underline{\mathbf{R}}}^*(f; f - f_1)}{\sum_{p,p',b,b'} z_{p,i} \mathbf{c}_b^{\mathrm{T}} \underline{\underline{\mathbf{R}}}(f - f_1; f - f_2) \mathbf{c}_{b'}^* z_{p',i}^{\dagger}}. \tag{4.46}$$

Under the assumption of completely uncorrelated Stokes parameters, the Wiener filter calculation simplifies to the form

$$\tilde{w}_i[f_k] = \frac{\left\| \tilde{S}_i(f) \right\|^2}{\sum_{p,p',b,b'} z_{p,i} \mathbf{c}_b^{\mathrm{T}} \underline{\underline{\mathbf{R}}}(f - f_1; f - f_2) \mathbf{c}_{b'}^* z_{p',i}^{\dagger}}. \tag{4.47}$$

Assuming that the Stokes parameters are completely uncorrelated with one another is a poor assumption from a physical perspective. Even if the polarization channels of the Stokes parameters are uncorrelated with each other, S_0 must have some degree of correlation with S_1, S_2, and S_3. This is because physically realizable signals always have DOLP less than 1. However, this assumption led us to a form of the Wiener filter that compared well analytically to the known result for conventional imaging, and is useful to discuss this result further.

Consider the system to be corrupted by additive white noise, changing the unfiltered estimation to be

$$\tilde{\mathbf{g}}_n[f_k] = \sum_{p=0}^{P-1} \sum_{b=-\infty}^{\infty} \mathbf{d}_p \left[\mathbf{c}_b^{\mathrm{T}} \tilde{\mathbf{S}}(f - f_1) + n(f - f_k + f_p) \right]. \tag{4.48}$$

The spectral density response found with only noise (zero contribution from Stokes parameters) will from now on be referred to as the noise response function (NRF),

$$\mathbf{NRF}(f) = \mathrm{diag}\left(\sum_{p,p'} \mathbf{d}_p n(f - f_k + f_p) n^*(f - f_k + f_{p'}) \mathbf{d}_{p'} \right), \tag{4.49}$$

where the quantity $n(f - f_k + f_p)n^*(f - f_k + f_{p'})$, describing the power spectrum of the noise, must be known or estimated for the application. This definition is similar to the required knowledge of the SDM for the Stokes parameters. With added noise, the Wiener filter calculation becomes

$$\tilde{w}_i[f_k] = \frac{\left\|\tilde{S}_i(f)\right\|^2}{\sum_{p,p',b,b'} z_{p,i}\mathbf{c}_b^{\mathrm{T}}\underline{\underline{\mathbf{R}}}(f - f_1; f - f_2)\mathbf{c}_{b'}^* z_{p',i}^\dagger + NRF_i(f)}. \tag{4.50}$$

Once more including the transfer functions of the detector, the unfiltered spectral density response is described by

$$\begin{aligned}\|\tilde{\mathbf{g}}[f_k]\|^2 = \sum_{p,p'}\sum_{b,b'} &\tilde{h}(f - f_k + f_p)\tilde{h}^*(f - f_k + f_{p'}) \\ &\times \mathbf{d}_p\mathbf{c}_b^{\mathrm{T}}\underline{\underline{\mathbf{R}}}(f - f_1; f - f_2)\mathbf{c}_{b'}^*\mathbf{d}_{p'}^*,\end{aligned} \tag{4.51}$$

and the Wiener-Helstrom filter is

$$\tilde{w}_i[f_k] = \frac{\tilde{h}^*(f)\left\|\tilde{S}_i(f)\right\|^2}{\|\tilde{g}_i[f_k]\|^2 + NRF_i(f)}. \tag{4.52}$$

The Wiener-Helstrom filter has the advantage that it can attempt to recover information that was deterministically attenuated due to the system transfer function.

4.4 Discussion

4.4.1 Infinitesimally small detector integration time

Spectral Density Response When the statistical object is measured by the rotating analyzer system and subsequently estimated as described above, the estimation results in a weighted superposition of shifted copies of the object SDM (Eq. 4.27). Figure 4.4 shows that the estimation of each Stokes parameter is correctly centered at the low frequency, but there are also higher frequency copies of all of the other Stokes parameters that interfere with the estimation; these copies must be filtered out with a low pass filter $\tilde{w}$, which will replace the filter with uniform frequency response. The spectral density response function using a uniform response filter allows us to find the Wiener filter that is optimal for estimating objects that have the same statistical properties as the object we have modeled. The Wiener filter calculation is shown in Eq. 4.47; graphically the Wiener filter calculation for the i^{th} Stokes parameter will be performed by dividing $R_{ii}(f)$ (the diagonal elements in Fig. 4.5) by the corresponding spectral density response obtained from Fig. 4.4. The resulting Wiener filters are shown in Fig. 4.6. These filters are band-limited filters. In fact, the Wiener filter for a system that has no noise and no overlapping signals converges to the ideal rectangular band-limited filter. However, the Wiener filter smoothly approaches zero in the presence of overlapping signals, which allows the filter to minimize undesirable ringing artifacts in the temporal domain in an objectively determined way.

We can also calculate a Wiener filter for functions that are assumed to be completly correlated. We will assume that the Stokes parameters are described by the same power spectra and that differently modulated Stokes parameters are uncorrelated, but that each of the Stokes parameters are completly correlated with each other. The graphical representation of the spectral density response function estimation for this case is shown in Fig. 4.7. The effect of a given cross-term is cancelled out by the effect of its complex conjugate, which is also added to the spectral density response estimation. This is represented by the magenta line at the bottom of the figure. Thus, as long as differently modulated parameters remain uncorrelated, the estimation of the spectral density response remains the same, whether the Stokes parameters are correlated or uncorrelated with each other.

Frequency-domain map of the SDM To further our understanding of the a system's spectral response, we note that the SDM in Eq. 4.27 can be represented as a frequency-domain map relating the spectral density response of the estimated Stokes parameters at different sampled frequencies to the estimated spectral density response of the continuous Stokes parameters. Figure 4.8 is an example of elements R_{00}, R_{11}, and R_{22} of the estimated unfiltered SDM for the object represented in Eq. 4.31. In this case, unfiltered means that the filter in

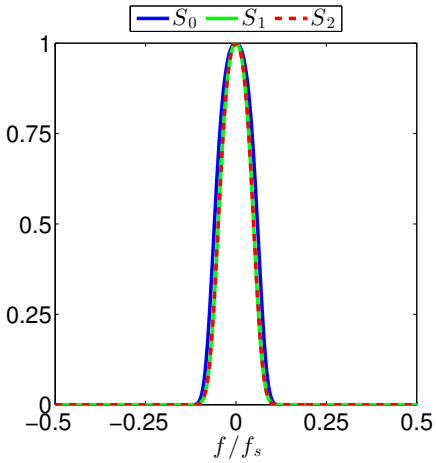


Figure 4.6: Wiener filters calculated for the assumed Stokes parameters power spectra being measured by the rotating analyzer polarimeter. The Wiener filters follow the signal to noise ratio of the estimation as a function of frequency. These filters are band-limited, but approach the cutoff smoothly to avoid undesirable temporal ringing artifacts.

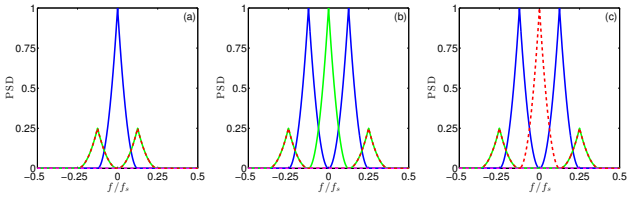


Figure 4.7: Graphical representation of the spectral density response function, assuming a filter with uniform frequency response, for (a) S_0, (b) S_1, and (c) S_2, and completely correlated Stokes parameters. Curves in blue, green, and red are components originally due to S_0, S_1, and S_2, respectively. The effect of the cross terms is represented by the magenta line at the bottom of the figure. The total spectral density for each parameter is obtained by summing all contributions.

the reconstruction is taken as $\tilde{w} = 1$ over the whole frequency domain. Since we assumed that different Stokes parameters are uncorrelated, elements R_{ij} are zero for $i \neq j$. In these figures, the vertical axis is the frequency of the continuous signal, whereas the horizontal axis is the sampled frequency. The total spectral density response obtained with the superposition of shifted elements from the object SDM represented in Fig. 4.4 is a cross section of Fig. 4.8 along the line $f_k = 0$. Therefore, the SDM frequency-domain map is a more general way of representing the system's response, since it contains the response at different sampling frequencies.

The secondary maxima in Fig. 4.8 constitute artifacts that limit our ability to reconstruct the SDM of the input signal. However, the Wiener filter presented in Section 4.3 allows us to filter out the higher frequency replicas in the unfiltered signal and can also be represented as a frequency-domain map. Figure 4.8 is the modulus square of the Wiener filter frequency map for each Stokes parameter. Just as for the SDM frequency maps, a cross section of this figure at $f_k = 0$ results in the Wiener filter results in Fig. 4.6. Applying this filter during the estimation of the Stokes parameters represented in Eq. 4.27 results in the filtered signal that yields the SDM in Fig. 4.8. The filtered SDM elements closely resemble a $\mathrm{tri}^2(f)$ along the diagonal, which would be the spectral density response function of an ideal system with no polarization artifacts.

4.4.2 Numerical comparison between filter types

In this section we quantitatively compare filter performance for the uncorrelated example using the mean squared error metric (under which the Wiener filter is derived). The filters we use for this section are the sliding pseudo inverse, the ideal band-limited filter, and the polarimetric Wiener filter. The frequency response functions of the filters are shown alongside the S_0 spectral density response in Fig. 4.9. To begin the numerical comparison, we generated an object with the statistical properties that were assumed in Eq. 4.31. To do this, many instantiations of white noise were added to the Stokes parameters of the object to build the desired SDM with a signal-to-noise

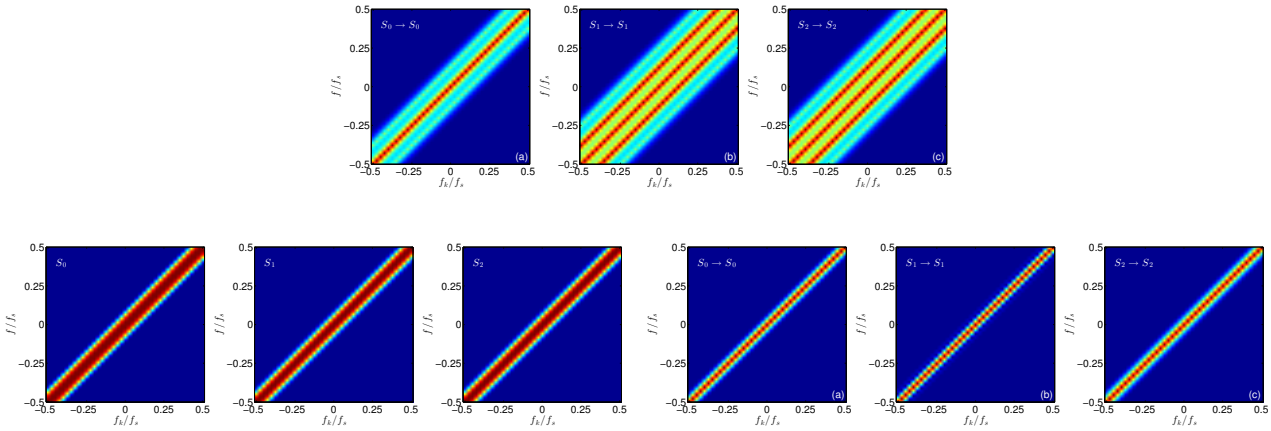


Figure 4.8: Frequency-domain map for the estimated (a) R_{00}, (b) R_{11}, and (c) R_{22} unfiltered elements of the SDM.

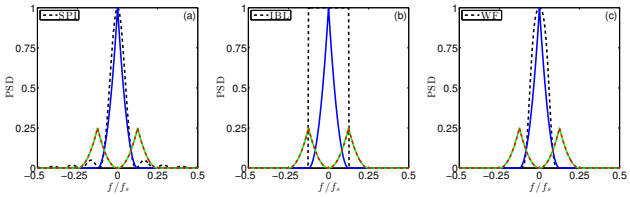


Figure 4.9: An illustration of the frequency repose of (a) the sliding pseudo inverse, (b) the ideal band-limited filter, and (c) a Wiener filter (smooth band-limited filter) compared to the graphical representation of the spectral density response of the S_0 channel.

ratio of 20 dB. A single instantiation of the white noise is shown in Fig. 4.10. Note that this simulation does not guarantee physical results (degree of polarization less than 1); physicality could be enforced by adding a constant offset to the S_0 object. However, the purpose of this discussion is to describe the method of application of these techniques.

The matrix obtained after 1000 iterations is shown in Fig. 4.11, which is similar to the ideal SDM in Fig. 4.5 but with added noise. This SDM approximates the desired object, as the off diagonal elements are approximately within 2.5% of zero. The object is measured and estimated by the polarimeter using the various filters in Fig. 4.9. The mean square error for the estimation across all instantiations is shown in Fig. **??**. The two bottom rows are quite similar and show an overall better performance than the ideal band-limited filter. The Wiener filter distribution is slightly flatter than that of the sliding pseudo inverse, and continues decreasing until the edge of the frequency region covered by the original Stokes parameters power spectra, whereas the sliding pseudo inverse reaches a minimum and then rises at the edge of the considered frequency region. Hence, the Wiener filter outperforms either of the other two filters, which was expected since it was designed for optimal performance

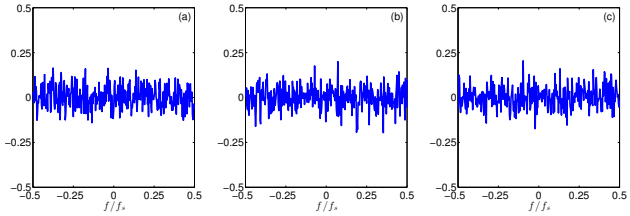


Figure 4.10: Single instantiation of the white noise for (a) S_0, (b) S_1, and (c) S_2. The noise is not guaranteed to produce a physically realizable signal since the object SDM employed in this problem does not enforce that.

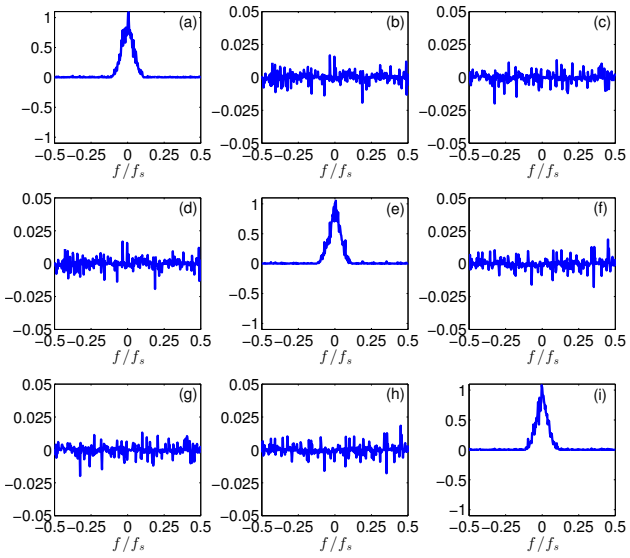


Figure 4.11: Calculated SDM for the simulated Stokes parameters with added white noise. Here the SDM (diagonal) elements for (a) S_0, (e) S_1, and (i) S_2 are approximately the desired triangle squared functions. The off diagonal cross spectral density elements ideally should be zero, and they are within 2.5% of the correct answer.

under the MSE metric. Note that this analysis is not meant to demonstrate the superiority of the Wiener filter, rather to demonstrate that the theory presented here results in a Wiener filter that is consistent with expectations.

4.4.3 Finite detector integration time

Now consider a case where there is a non-unity detector transfer function, such that the detector integration time has a duty cycle of 1 and is described in time as

$$h(t) = \text{rect}\left(\frac{t}{T}\right). \tag{4.53}$$

In the Fourier domain, the detector transfer function is written as

$$\tilde{h}(f) = \text{sinc}\left(\frac{f}{f_s}\right). \tag{4.54}$$

The first zero of the detector transfer function for this problem is then at f_s, while we are concerning ourselves with signals that oscillate at rates of less than $0.5f_s$. Figure 4.12 shows the graphical representation of the spectral density response R_{00} measured with a detector with finite integration time, together with the detector transfer function, compared to the ideally measured spectral density response and the difference between these two responses. We observe that the S_0 channel has been symmetrically affected by the detector transfer function since they have the same center of symmetry, but the high frequency shifted S_1 and S_2 channels have not been symmetrically affected. Many conventional processing techniques for modulated polarimetry implicitly ignore this non symmetrical effect, since conventional calibration is done at the zero frequency. The lower the duty cycle of the sampling detector, the less effect this asymmetry has since the detector is closer to ideal point sampling. Had the modulation scheme been at a higher frequency, the shifted copies would be further away from the symmetrical point and the estimation would be further degraded. The Wiener-Helstrom filter can be computed in this case, although for this particular polarimeter little advantage is gained. The effect of the detector transfer function being included in the simulation had a maximum effect of about 2.5% at certain frequencies, and this translates to having a maximum effect on the filter of less than 0.5% at those same frequencies. The Wiener-Helstrom filter is shown in Fig. 4.12, as well as the difference between the Wiener and the Wiener-Helstrom

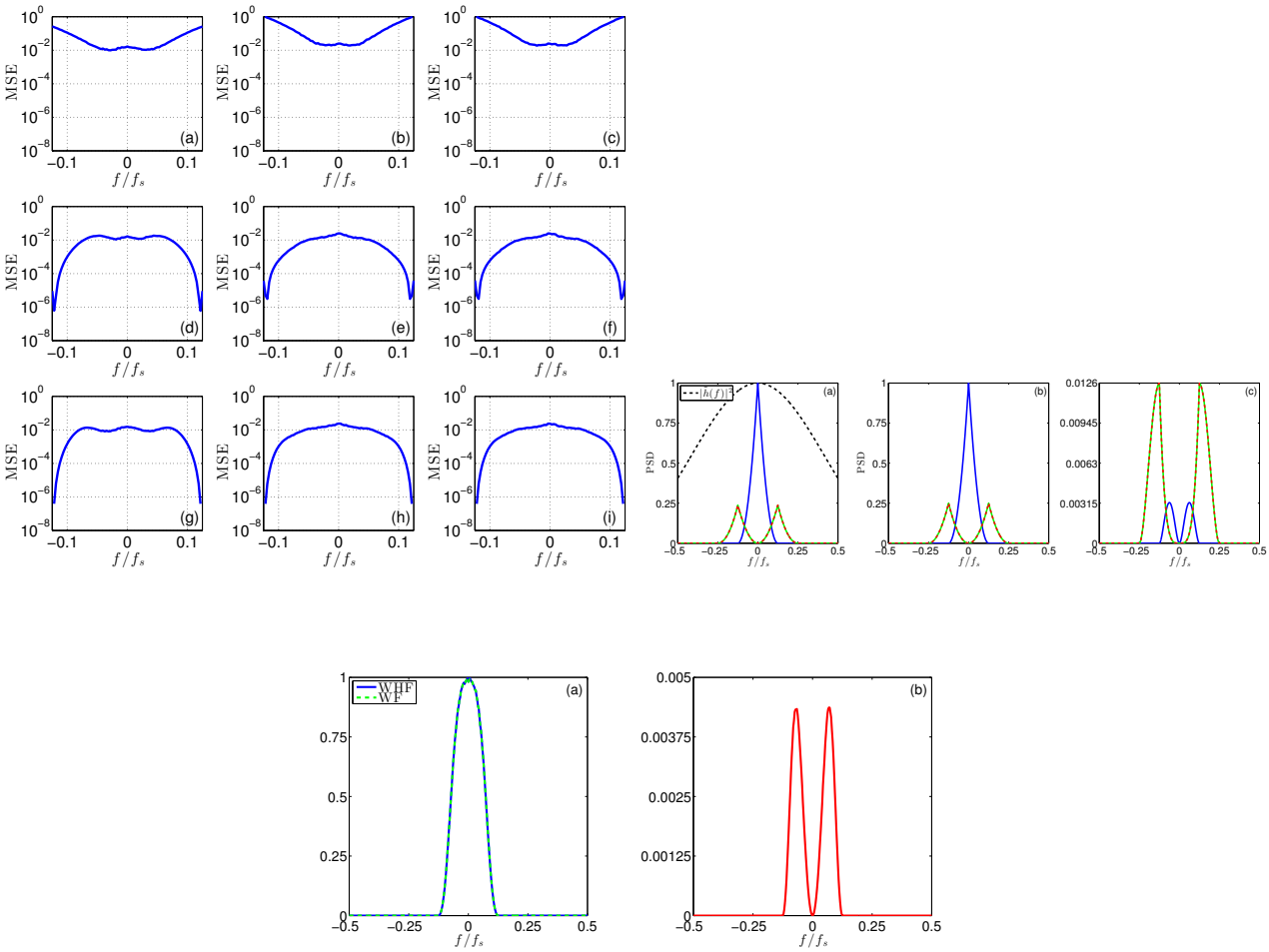


Figure 4.12: (a) Graphical representation of the spectral density response of a system with a detector with finite integration time of duty cycle 1. The modulus square of the detector transfer function is shown for reference. (b) Graphical representation of the spectral density response using an ideal detector with uniform frequency response. (c) The difference between the spectral density response with an ideal detector and the detector with finite integration time in (a), showing the non symmetrical effect on the high frequency copies.

filters. This difference is quite small, but for other applications there may be differing results. The importance of the detector transfer function will increase for applications where the modulation scheme introduces several side bands, such as the MSPI polarimeter described by Diner, *et al.*[4] Here, the advantage of calculating the Wiener-Helstrom filter may lie in helping to control channel cross introduced by the spatial bandwidth (converted to temporal bandwidth by the moving platform) of the scene being measured.

The techniques presented here provide a way to study the effect measurement has given models of the SDM of objects to be measured. The development of the optimal Wiener filter for a particular application is highly dependent on the accuracy of the object models. However, even without accurate object models, rough guesses of the object model may yield effective first-guess filters than can be iteratively improved upon as empirical data become available. Intuitively we know that a smooth band-limited filter should provide better results for periodic analyzer vectors than the ideal band-limited or sliding pseudo inverse: it is band-limited to prevent channel crosstalk and smooth to prevent ringing in the space time domain. The optimal Wiener filter will perform better than the other linear filters in terms of the mean squared error for a given well defined set of object statistics, because the Wiener filter is derived by minimizing this metric. The detection task drives the selection and optimization of the filter. Therefore, for any given application other filters than the Wiener filter may be more appropriate. The approach developed here allows for the comparison of the performance of different polarimeter modulation schemes and processing algorithms simultaneously given a defined application (object SDM). The spectral density response presented can be employed to perform objective comparisons between different polarimeter designs for suitability for an imaging task.

5. CHANNELED POLARIMETERS

In the class of channeled or modulated polarimeters, the polarization states that define the measurement are modulated either spatially, temporally, or spectrally.[9] Every one of those harmonic modulations will split the information in the corresponding Fourier domains, creating weighted copies of the Fourier Transform of the data at the modulating frequencies. These multiplexed copies are called channels. Oka and his coworkers[40,51,52] have popularized the design concepts that go into making a channeled system, which were then further developed by Hagen,[45] Kudenov[53,54] and their coworkers before going into wide use by many others. In this paper we will introduce a toolkit to describe, analyze and optimize such systems, and we will use it to investigate channeled polarimeters from the literature to show how they can be improved.

There has been a number of proposed channeled systems in the past,[40,45,51–54] whose designs and the corresponding reconstruction techniques were derived by hand. By limiting the number of parameters available during design, the likelihood of a suboptimal design is inadvertently increased. Lemaillet *et al.*[55] proposed a way to optimize a spectrally channeled system by introducing linear algebra methods to map the information. Their effort, however, focused on the particulars of one kind of system and stopped short of providing a complete solution to deal with any channeled polarimeter. This paper describes the generalized methods that can be used to model channeled information mapping and guide the reconstruction.

A great advantage posed by a channeled system is the possibility of constructing a snapshot system that is able to simultaneously record information pertaining to different polarization information channels. This removes the need for complex image registration that would be required in a temporally modulated system. In terms of object bandwidth, a snapshot channeled system favors temporal resolution at the cost of introducing stricter band limit constraints in other domains.

We consider a common 2D FPA detector, which enables access to up to two modulation types to be mapped onto the two orthogonal axes. In addition to having no modulation, we can also map spatial and spectral modulations into either x- or y-axes of the detector. Although the methods introduced here are general enough to be used with any channel structure on any orthogonal coordinates systems, this paper will focus on Cartesian coordinates, implying that the channels lie on a rectangular grid.

For the sake of completeness, we will also consider temporal modulation. Such a system will obviously lose its snapshot nature, but it is conceivable that some middle ground solution could be found, whereby a very limited number of temporal measurements are made with intent of balancing the resolution loss among all possible modulation dimensions.[7,9] Thus, using a conventional detector will allow us to split polarization

information into a three dimensional structure of channels that we can manipulate to reconstruct the polarization information. It is useful to recognize that when we have a small number of temporal measurements, the resultant temporal frequency channels may contain as few as one data point. In those cases it may be more prohibitive and unnecessary to work with the data in the Fourier domain. We can instead use measurements as information "channels" themselves with a clear benefit that they will contain modulation information more compactly.

5.1 Sinusoidal Channel Splitting

Every sinusoidal modulation in either a_i or g_j splits the element information in m_{ij} into two channels at certain frequencies within the Fourier domain of the modulation. For the available modulation dimensions of x, y, σ (wavenumber) and t, we will call the corresponding frequency dimensions ξ, η, τ (OPD) and ν. We will only show the relevant equations for the x–ξ pair, since all others can be obtained by taking any other sequential pairing of the above mentioned dimensions. The following transform pairs are well known:[56]

$$1(x) \longleftrightarrow \delta(\xi), \tag{5.1a}$$

$$\cos(2\pi\xi_i x) \longleftrightarrow \tfrac{1}{2}[\delta(\xi+\xi_i)+\delta(\xi-\xi_i)], \tag{5.1b}$$

$$\sin(2\pi\xi_i x) \longleftrightarrow \tfrac{j}{2}[\delta(\xi+\xi_i)-\delta(\xi-\xi_i)]. \tag{5.1c}$$

In the general case, the modulation functions are more complicated and involve more than just one modulating frequency. The following example is typical of the modulations we encounter and is directly applicable for many channeled polarimeters,

$$f_M(x) = \prod_{m=1}^{M} \substack{\cos\\\sin}(2\pi\xi_m x), \tag{5.2}$$

where $\substack{\cos\\\sin}$ denotes that the function could either be a cosine or a sine. Since we have M sinusoids multiplied together, we can create a $2^M \times 2^M$ matrix that will describe all the possible combinations of either $\pm\xi_m$ of the δ-function, as well as distinguish between a cosine and a sine. Each sub-function will have a phasor that, when multiplied together, will yield the net phase of the particular channel-weight. Employing matrix notation, we can create this "look-up-table" by means of an outer product of two matrices:

$$\underline{\underline{\mathbf{F}}}_M \equiv \left[\begin{array}{cccc} \underline{\mathbf{f}}_1 & \underline{\mathbf{f}}_2 & \cdots & \underline{\mathbf{f}}_M \end{array}\right], \tag{5.3a}$$

$$\underline{\underline{\mathbf{O}}}_M \equiv \left[\begin{array}{cccc} \underline{\mathbf{o}}_1 & \underline{\mathbf{o}}_2 & \cdots & \underline{\mathbf{o}}_M \end{array}\right], \tag{5.3b}$$

where $f_{m,k}$ is 0 for cosine and 1 for sine, while $o_{m,\ell}$ is -1 for $-\xi_i$ and $+1$ for $+\xi_i$. $\underline{\underline{\mathbf{F}}}_M$ and $\underline{\underline{\mathbf{O}}}_M$ are both $2^M \times M$ in size. The frequency phase matrix (FPM) is then,

$$\underline{\underline{\mathbf{P}}}_M \equiv \frac{1}{2^M} \exp\left[\frac{-j\pi}{2}\left(\underline{\underline{\mathbf{F}}}_M \underline{\underline{\mathbf{O}}}_M^{\mathrm{T}}\right)\right]. \tag{5.4}$$

Cases of $M = 1, \ldots, 4$ can be seen in Figure 5.1.

For some polarimeters the induced modulation may be more complicated than the one prescribed by Equation 5.2 and the one directly solved by creating a single FPM. In order to treat those modulation schemes, we can allow for addition of modulating functions, effectively treating individual FPMs as a basis set. The total FPM can be decomposed into L sub-FPMs,

$$\underline{\underline{\mathbf{P}}}_{\text{total}} = \underline{\underline{\mathbf{P}}}_{M_0} + \underline{\underline{\mathbf{P}}}_{M_1} + \cdots + \underline{\underline{\mathbf{P}}}_{M_{L-1}}, \tag{5.5}$$

which can be calculated separately and simply added together. The involved coefficients will be more complicated, but they should present no additional challenge within the prescribed methods.

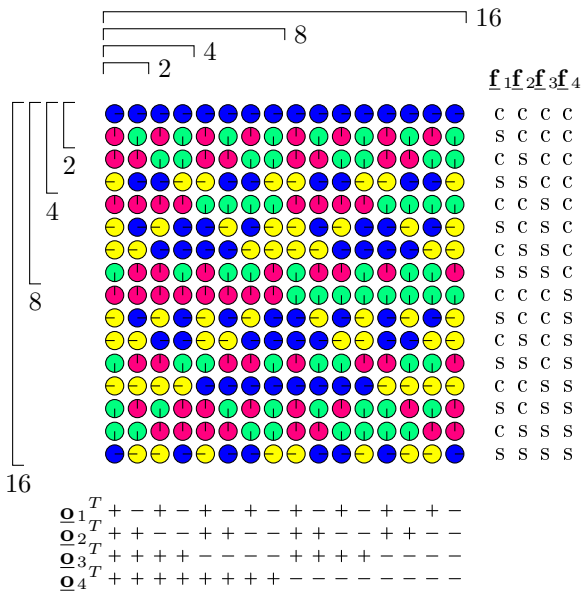


Figure 5.1: First four frequency phase matrices. The circles represent the polar form of the coefficients. Each element has an implied weight of 2^{-M}, which is omitted to avoid confusion. Also note that if $\xi_i = \xi_j$, then certain δ-functions will combine at $\delta(\xi \pm \cdots + \xi_i - \xi_j \pm \cdots)$ and $\delta(\xi \pm \cdots - \xi_i + \xi_j \pm \cdots)$. The magnitudes of those impulses will change. The side brackets denoted with 2/4/8/16 can be used as crop guidelines for obtaining FPMs for $M = 1/2/3/4$. The circles correspond in the following way: $\bullet = +1$, $\bullet = +j$, $\bullet = -1$, $\bullet = -j$.

We have several ways by which to combine multiple dimension modulations into a total structure of channels. In the case that we have determined each dimension's structure, we can combine them using a Kronecker product, namely,

$$\underline{\mathbf{q}}_{\{\tau\}} \otimes \underline{\mathbf{q}}_{\{\omega\}} \otimes \underline{\mathbf{q}}_{\{\xi\}} \otimes \underline{\mathbf{q}}_{\{\eta\}}. \tag{5.6}$$

On the other hand, if the order of modulation dimensions is alternating between elements, we could use convolution to create the N-dimensional cloud of channels that we would then unfold into a vector. As an example, we will consider four polarization modulation elements that operate over $x/y/x/y$ or modulate into $\xi/\eta/\xi/\eta$. The total vector is then

$$\operatorname{vec}\left(\underline{\mathbf{q}}_{\{\xi_{e_1}\}} *_n \underline{\mathbf{q}}_{\{\eta_{e_2}\}} *_n \underline{\mathbf{q}}_{\{\xi_{e_3}\}} *_n \underline{\mathbf{q}}_{\{\eta_{e_4}\}}\right), \tag{5.7}$$

where $*_n$ redundantly implies that the vectors are differently oriented or, more generally, can be described as degenerate N-dimensional structures. In this example, $\underline{\mathbf{q}}_{\{\xi_{e_1}\}}$ and $\underline{\mathbf{q}}_{\{\xi_{e_3}\}}$ are row vectors, while $\underline{\mathbf{q}}_{\{\eta_{e_2}\}}$ and $\underline{\mathbf{q}}_{\{\eta_{e_4}\}}$ are column vectors. The result of the convolution operation is a matrix and needs to be unfolded using vec operation that we defined above. The choice of row/column over column/row addressing is arbitrary at first, but once chosen must be maintained consistently throughout the analysis.

Another way of generating the sought-after total vector is by recognizing that the Mueller element modulation patterns can be alternately viewed as either a test dyad or a projection target. By treating it as a dyad, $\underline{\underline{\mathbf{D}}} = \underline{\mathbf{A}}\,\underline{\mathbf{G}}^{\mathrm{T}}$, we can look at its Fourier transform, $\mathcal{F}\{\underline{\underline{\mathbf{D}}}\} = \mathcal{F}\{\underline{\mathbf{A}}\} * \mathcal{F}\{\underline{\mathbf{G}}\}^{\mathrm{T}}$, with $*$ now being a matrix convolution (same as multiplication, but every product is replaced with convolution between the same elements and added as before). That allows us to combine PSA and PSG modulations as

$$\underline{\mathbf{q}}_{m_{ij}} = \operatorname{vec}\left(\underline{\mathbf{q}}_{g_i} * \underline{\mathbf{q}}_{a_j}\right). \tag{5.8}$$

Finally, using any of the prescribed methods to construct the Mueller element modulation vectors, we need to combine all of them into the corresponding $\underline{\underline{\mathbf{Q}}}$ matrix

$$\underline{\underline{\mathbf{Q}}} = \begin{bmatrix} \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{00}} \\ \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{01}} \\ \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{02}} \\ \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{03}} \\ \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{10}} \\ \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{11}} \\ \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{12}} \\ \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{13}} \\ \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{20}} \\ \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{21}} \\ \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{22}} \\ \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{23}} \\ \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{30}} \\ \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{31}} \\ \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{32}} \\ \underline{\mathbf{q}}^{\mathrm{T}}_{\{\tau,\omega,\xi,\eta\};m_{33}} \end{bmatrix}^{\mathrm{T}} \tag{5.9}$$

that maps an input Mueller vector into a channel vector,

$$\mathcal{F}\{\underline{\mathbf{C}}\} = \underline{\underline{\mathbf{Q}}}\mathcal{F}\{\underline{\mathbf{M}}'\}, \tag{5.10}$$

where $\underline{\mathbf{C}}$ describes the channels contents. However, we want to perform the opposite operation, since we measure channels directly. To do that, we can simply take the pseudo inverse of $\underline{\underline{\mathbf{Q}}}$, like we do for the DRM in regular polarimeters,

$$\underline{\underline{\mathbf{Q}}}^{+} = (\underline{\underline{\mathbf{Q}}}^{\dagger}\underline{\underline{\mathbf{Q}}})^{-1}\underline{\underline{\mathbf{Q}}}^{\dagger} = (\underline{\underline{\mathbf{Q}}}^{\dagger}\underline{\underline{\mathbf{Q}}})\backslash\underline{\underline{\mathbf{Q}}}^{\dagger}. \tag{5.11}$$

By correctly arranging Fourier transform operations around the multiplication, we can use that reverse mapping to get back to the Mueller elements' information,

$$\underline{\mathbf{M}}' = \mathcal{F}^{-1}\left\{\underline{\underline{\mathbf{Q}}}^{+}\mathcal{F}\{\underline{\mathbf{C}}\}\right\}. \tag{5.12}$$

An important piece of insight can be obtained if we recognize that $\underline{\underline{\mathbf{Q}}}$ is not that much different from $\underline{\mathbf{D}}'$; however, if before we relied on constructing multiple dyads against which to test the Mueller object, we can now have a very limited number of dyads. This is because the particular modulation choices create a multi-dimensional "pointer" that can be unfolded to the full $\underline{\underline{\mathbf{Q}}}$ representation.

5.2 Snapshot Channels

A snapshot measurement implies that there is no temporal modulation, which consequently means that no additional (other than the exposure time) temporal band-limit constraints are placed on the captured scene. If we consider that each dimension on the 2D detector can carry spatial, spectral or no modulation, we can create a verbose set of nine snapshot channeled systems that can be seen in Figure 5.2.

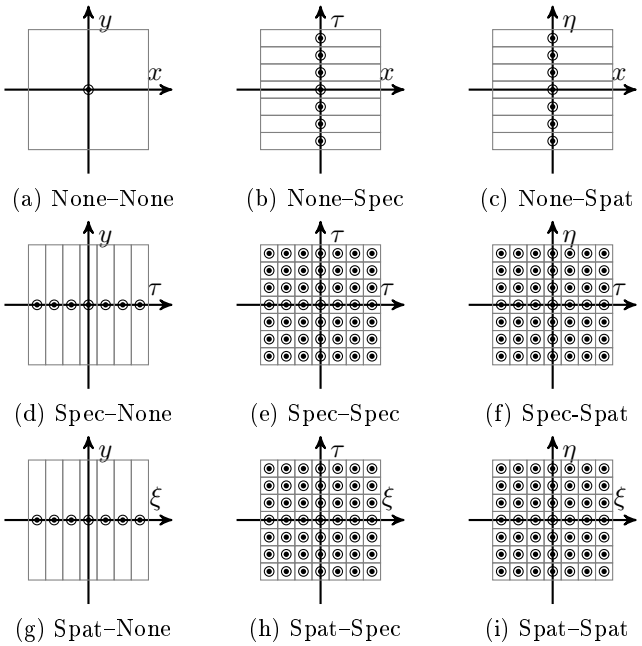


Figure 5.2: Snapshot systems. Case (a) provides no modulation. Case (e) is not straightforward to implement physically. Case pairs (b)/(d), (c)/(g) and (f)/(h) are essentially equivalent.

5.3 Multiple-Snapshot Channels

Further developing the consideration of physical realizability, we can take several snapshot measurements. This gives an easy access to a third modulation dimension — time. Provided that the temporal modulation is captured in even time steps, we can use the Fourier transformation to create our channels. However, since in most cases this will create more channels than the original data, with all channels being a single pixel, using the Fourier coefficients does not present any advantage. Instead, the captured snapshots themselves can be used as "direct channels", or simply, projection targets like in the $\underline{\underline{\mathbf{W}}}'$ formalism, namely,

$$\underline{\underline{\mathbf{Q}}}_{\text{total}} = \left[\begin{array}{cccc} \underline{\underline{\mathbf{Q}}}_{t_0}^{\mathrm{T}} & \underline{\underline{\mathbf{Q}}}_{t_1}^{\mathrm{T}} & \cdots & \underline{\underline{\mathbf{Q}}}_{t_{N-1}}^{\mathrm{T}} \end{array} \right]^{\mathrm{T}}. \tag{5.13}$$

This removes the need to have evenly spaced samples, yet maintains the compressed nature of $\underline{\underline{\mathbf{Q}}}$. Note that even though we selected direct channels only for temporal modulation, it is possible to treat other domain modulations similarly.

5.4 Channeled Reconstruction

We use an SVD method to calculate the pseudoinverse for reasons of its numerical stability. First, $\underline{\underline{\mathbf{Q}}}$ is decomposed as

$$\underline{\underline{\mathbf{Q}}} = \underline{\underline{\mathbf{U}}}_{\underline{\underline{\mathbf{Q}}}} \underline{\underline{\mathbf{\Sigma}}}_{\underline{\underline{\mathbf{Q}}}} \underline{\underline{\mathbf{V}}}_{\underline{\underline{\mathbf{Q}}}}^{\dagger}, \tag{5.14}$$

where the matrices $\underline{\underline{\mathbf{U}}}$ and $\underline{\underline{\mathbf{V}}}$ are $N \times N$ and 16×16 complex, orthogonal matrices, respectively, and $\underline{\underline{\mathbf{\Sigma}}}$ is a $N \times 16$ reduced diagonal matrix containing the N singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_N > 0$. The pseudoinverse can be written as

$$\underline{\underline{\mathbf{Q}}}^{+} = \underline{\underline{\mathbf{V}}}_{\underline{\underline{\mathbf{Q}}}} \underline{\underline{\mathbf{\Sigma}}}_{\underline{\underline{\mathbf{Q}}}}^{+} \underline{\underline{\mathbf{U}}}_{\underline{\underline{\mathbf{Q}}}}^{\dagger}, \tag{5.15}$$

where $\underline{\underline{\mathbf{\Sigma}}}^{+}$ is the $16 \times N$ reduced diagonal matrix containing the inverse of the singular values. The rank of the measurement can be calculated as $\mathrm{tr}(\underline{\underline{\mathbf{Q}}}^{+}\underline{\underline{\mathbf{Q}}})$. It is interesting to look at the diagonal elements of this matrix

to see how the information from the N measurements is distributed in the estimated Mueller matrix. We define the "reconstructables" matrix

$$\underline{\underline{\mathbf{B}}}' = \mathrm{vec}(\underline{\underline{\mathbf{B}}}) = \mathrm{diag}(\underline{\underline{\mathbf{Q}}}^{+}\underline{\underline{\mathbf{Q}}}). \tag{5.16}$$

For each Mueller k-th element, $\sqrt{b'_k}$ tells us the fraction of energy that is maintained after reconstruction. When $b'_k = 0$, sensor space and scene space are orthogonal, $\mathcal{R} \perp \mathcal{S}$. Equivalently, it can be said that the information lies in the null space of the measurement.[9] When $b'_k = 1$, then $\mathcal{R} \subseteq \mathcal{S}$ and the information can be reconstructed to within noise limitations.

Although $\underline{\underline{\mathbf{B}}}$ can be useful, it is a summary metric and does not tell the full story about the polarimeter. If we investigate the $\underline{\underline{\mathbf{Q}}}^{+}\underline{\underline{\mathbf{Q}}}$ multiplication, we can gain more insight, since

$$\underline{\underline{\mathbf{Q}}}^{+}\underline{\underline{\mathbf{Q}}} = \underline{\underline{\mathbf{V}}}_{\underline{\underline{\mathbf{Q}}}}\,\underline{\underline{\mathbf{\Sigma}}}^{+}_{\underline{\underline{\mathbf{Q}}}}\,\underline{\underline{\mathbf{U}}}^{\dagger}_{\underline{\underline{\mathbf{Q}}}}\,\underline{\underline{\mathbf{U}}}_{\underline{\underline{\mathbf{Q}}}}\,\underline{\underline{\mathbf{\Sigma}}}_{\underline{\underline{\mathbf{Q}}}}\,\underline{\underline{\mathbf{V}}}^{\dagger}_{\underline{\underline{\mathbf{Q}}}} \tag{5.17}$$

$$= \underline{\underline{\mathbf{V}}}_{\underline{\underline{\mathbf{Q}}}}\,\underline{\underline{\mathbf{\Sigma}}}^{+}_{\underline{\underline{\mathbf{Q}}}}\,\underline{\underline{\mathbf{\Sigma}}}_{\underline{\underline{\mathbf{Q}}}}\,\underline{\underline{\mathbf{V}}}^{\dagger}_{\underline{\underline{\mathbf{Q}}}}. \tag{5.18}$$

$\underline{\underline{\mathbf{\Sigma}}}^{+}_{\underline{\underline{\mathbf{Q}}}}\,\underline{\underline{\mathbf{\Sigma}}}_{\underline{\underline{\mathbf{Q}}}}$ will reveal a sub-identity matrix that will have first R (rank of $\underline{\underline{\mathbf{Q}}}$) elements of the diagonal equal to unity and the remaining $16-R$ elements equal to zero. This has the effect of cropping $\underline{\underline{\mathbf{V}}}$ into two — the column space that the polarimeter supports and the null space that the polarimeter rejects. By keeping only the column space vectors, we can truncate $\underline{\underline{\mathbf{V}}}$, namely into $\underline{\underline{\mathbf{V}}}'$, that can be used to rewrite Equation 5.16 as

$$\underline{\underline{\mathbf{B}}}' = \mathrm{diag}\left(\underline{\underline{\mathbf{V}}}'_{\underline{\underline{\mathbf{Q}}}}\,\underline{\underline{\mathbf{V}}}'^{\dagger}_{\underline{\underline{\mathbf{Q}}}}\right). \tag{5.19}$$

A full system will have a reconstructables vector of all ones, while partial systems could design b'_k based on prior knowledge of the scene in question.[57] This concept is developed more fully for general pMMPs elsewhere.[18]

5.5 Examples

By simplifying the generation of a polarimeter design, we can calculate the idealized SNR from $\underline{\underline{\mathbf{Q}}}$ directly, without performing full simulations of the system. Simply changing the way the problem is written allows us to introduce more optimization parameters that can help us find an optimal polarimeter. In this section we will look at systems and see how the introduced concepts could help increase performance.

Sabatke introduced Equally Weighted Variance (EWV) as an appropriate metric to evaluate Stokes polarimeters,[32] and Twietmeyer later adopted a similar metric for use with Mueller polarimeters.[58]

$$\mathrm{EWV} = \mathrm{tr}\left[\underline{\underline{\mathbf{K}}}_{\underline{\underline{\mathbf{W}}}'^{+}}\right] = \textstyle\sum_{k=0}^{15} 1/\sigma^2_{\underline{\underline{\mathbf{W}}}'^{+},k}, \tag{5.20}$$

with $\underline{\underline{\mathbf{K}}}_{\underline{\underline{\mathbf{W}}}'^{+}}$, the covariance matrix, $\underline{\underline{\mathbf{W}}}'^{+}\underline{\underline{\mathbf{W}}}'^{+\mathrm{T}}$. For the purposes of designing pMMPs, it is necessary to generalize the combination of Mueller element variances with arbitrary preference. Instead of implying a uniform weight for all elements, we can introduce a weighing vector, $\underline{\mathbf{u}}$, to calculate a Weighted Variance (WV),

$$\mathrm{WV} = \underline{\mathbf{u}}^{\mathrm{T}}\,\mathrm{diag}\left[\underline{\underline{\mathbf{K}}}_{\underline{\underline{\mathbf{W}}}'^{+}}\right] = \textstyle\sum_{k=0}^{15} u_k/\sigma^2_{\underline{\underline{\mathbf{W}}}'^{+},k}. \tag{5.21}$$

For channeled systems, we can calculate the WV similarly as

$$\mathrm{WV} = \underline{\mathbf{u}}^{\mathrm{T}}\,\mathrm{diag}\left[\underline{\underline{\mathbf{K}}}_{\underline{\underline{\mathbf{Q}}}^{+}}\right] = \textstyle\sum_{k=0}^{15} u_k/\sigma^2_{\underline{\underline{\mathbf{Q}}}^{+},k}, \tag{5.22}$$

with $\underline{\underline{\mathbf{K}}}_{\underline{\underline{\mathbf{Q}}}^{+}} = \underline{\underline{\mathbf{Q}}}^{+}\underline{\underline{\mathbf{Q}}}^{+\dagger}$, the product which contains the Mueller element reconstruction variances within its diagonal.

5.5.1 Spectrally Channeled Polarimeter

Consider the spectral-none channeled polarimeter proposed by Hagen,[45] an example of a system in Figure 5.2d. The measured intensity can be written as

$$I(\vec{\vartheta}) = \sum_{i=0}^{3}\sum_{j=0}^{3} f_{a_i}(\vec{\vartheta}) m_{ij}(\vec{\vartheta}) f_{g_j}(\vec{\vartheta}), \tag{5.23}$$

where $\vec{\vartheta}$ is used to denote a set of domains where the information is modulated. For this system, there is only modulation in wavenumber, meaning $\vec{\vartheta} = \{\sigma\}$ and

$$\underline{\mathbf{f}}_{\underline{\mathbf{A}}}(\sigma) = \begin{bmatrix} 1 \\ \cos(c_4\sigma) \\ \sin(c_3\sigma)\sin(c_4\sigma) \\ \cos(c_3\sigma)\sin(c_4\sigma) \end{bmatrix}, \tag{5.24a}$$

$$\underline{\mathbf{f}}_{\underline{\mathbf{G}}}(\sigma) = \begin{bmatrix} 1 \\ \cos(c_1\sigma) \\ \sin(c_1\sigma)\sin(c_2\sigma) \\ \sin(c_1\sigma)\cos(c_2\sigma) \end{bmatrix}, \tag{5.24b}$$

The argument

$$c_i\sigma = 2\pi\tau_i\sigma = 2\pi d_o d_i \lambda_o B\sigma,\ \tau_i = d_o d_i \lambda_o B \tag{5.25}$$

contains the global thickness factor, d_o, individual retarder thickness factor, d_i, center wavelength, λ_o, and birefringence, B. The vector $\underline{\mathbf{d}}$ contains all the modulation information and will be our optimization target. Hagen chooses $\underline{\mathbf{d}} = (\ 1\ \ 2\ \ 5\ \ 10\ )$, with the resulting channels and the proposed reconstruction scheme.

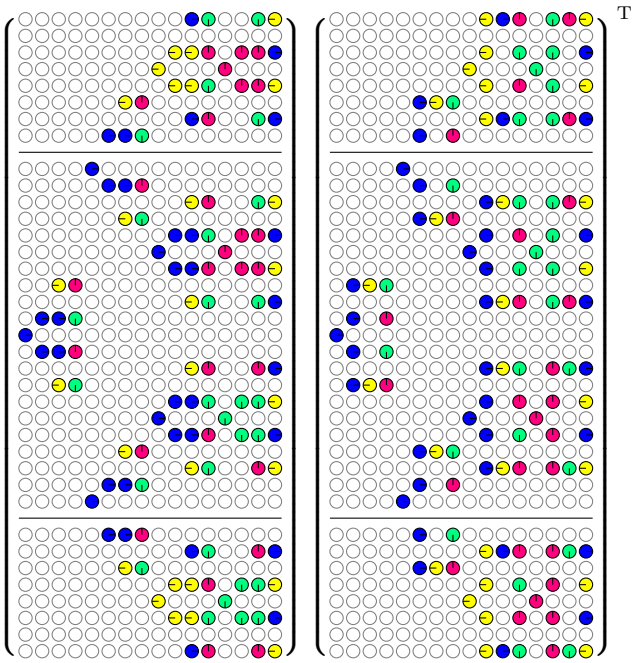
From the proposed reconstruction, we see that some measurements are ignored for the sake of algebraic simplicity — only channels $c_0 - c_{10}$ are referenced, with real and imaginary operators constituting the use of conjugates. Thus, instead of using all 37 channels, only 21 are used. An alternative method would be to recognize the modulation induced by retarders, construct an appropriate FPM, look up the coefficients and construct $\underline{\underline{\mathbf{Q}}}$ by placing them at the contributing frequencies. The resultant $\underline{\underline{\mathbf{Q}}}$ and its inverse can be seen in Figure 5.3.

c_n	Channel Content $\times(64/S_{in,0})$
0	$16m_{00}$
±1	$8m_{01} + 4m_{02} \pm 4im_{03}$
±2	$-m_{22} \pm im_{23} \mp im_{32} - m_{33}$
±3	$-4m_{02} \mp 4im_{03}$
±4	$2m_{21} + m_{22} \mp im_{23} \pm 2im_{31} \pm im_{32} + m_{33}$
±5	$4m_{20} \pm 4im_{30}$
±6	$2m_{21} + m_{22} \pm 1m_{23} \pm 2im_{31} \pm im_{32} - m_{33}$
±7	$-2m_{12} \pm 2im_{13}$
±8	$-m_{22} \mp im_{23} \mp im_{32} + m_{33}$
±9	$4m_{11} + 2m_{12} \mp 2im_{13}$
±10	$8m_{10}$
±11	$4m_{11} + 2m_{12} \pm 2im_{13}$
±12	$m_{22} \mp im_{23} \mp im_{32} - m_{33}$
±13	$-2m_{12} \mp 2im_{13}$
±14	$-2m_{21} - m_{22} \pm im_{23} \pm 2im_{31} + im_{32} + m_{33}$
±15	$-4m_{20} \pm 4im_{30}$
±16	$-2m_{21} - m_{22} \mp im_{2}3 \pm 2im_{31} \pm im_{32} - m_{33}$
±18	$m_{22} \pm im_{23} \mp im_{32} + m_{33}$

Table 1: Hagen polarimeter's channels.

$m_{ij}(\sigma)$	
$m_{00}(\sigma) =$	$4c_0$
$m_{01}(\sigma) =$	$8(c_1 + c_3)$
$m_{02}(\sigma) =$	$-16\Re[c_3]$
$m_{03}(\sigma) =$	$16\Re[c_1]$
$m_{10}(\sigma) =$	$8c_{10}$
$m_{11}(\sigma) =$	$16(c_7 + c_9)$
$m_{12}(\sigma) =$	$-32\Re[c_7]$
$m_{13}(\sigma) =$	$32\Im[c_7]$
$m_{20}(\sigma) =$	$16\Re[c_5]$
$m_{21}(\sigma) =$	$32\Re[c_2 + c_4]$
$m_{22}(\sigma) =$	$-32\Re[c_2 + c_8]$
$m_{23}(\sigma) =$	$32\Im[c_2 - c_8]$
$m_{30}(\sigma) =$	$16\Re[c_5]$
$m_{31}(\sigma) =$	$16\Im[c_2 + c_4 + c_6 + c_8]$
$m_{32}(\sigma) =$	$-32\Im[c_2 + c_8]$
$m_{33}(\sigma) =$	$32\Re[c_8 - c_2]$

Table 2: Hagen's proposed reconstruction.



(a) $\underline{\underline{\mathbf{Q}}}$ (b) $\underline{\underline{\mathbf{Q}}}^{+}$

Figure 5.3: Hagen's spectropolarimeter, $\underline{\mathbf{d}} = (\ 1\ \ 2\ \ 5\ \ 10\ )$. The matrix containing 21 cropped channels can be seen between the two horizontal lines and has an EWV of 355; including the other 16 channels lowers the EWV to 187. Note that these extra channels must be measured in order to prevent aliasing. The distinction is whether the data contained within these channels is used in reconstruction, after the Fourier transform of the measured intensity was determined.

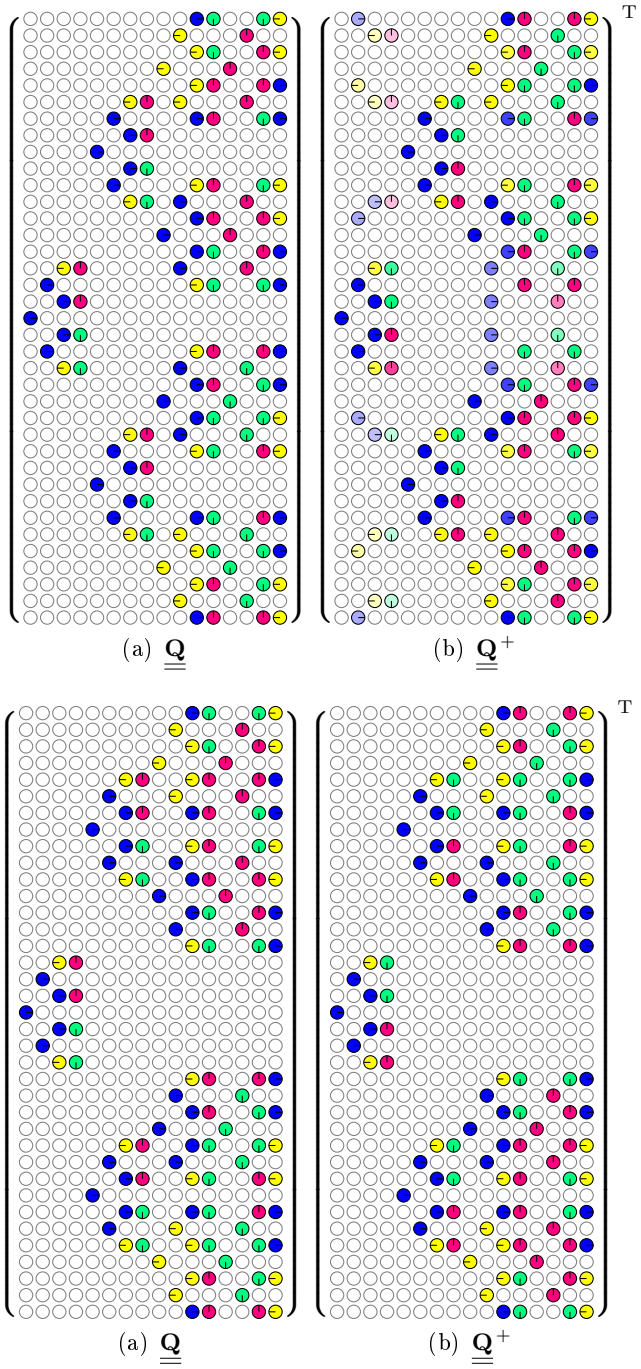


Figure 5.4: By swapping the first two elements, i.e., $\underline{\mathbf{d}} = (\ 2\ \ 1\ \ 5\ \ 10\ )$, we can lower the EWV down to $130\frac{4}{7}$.

The system in Figure 5.4 was found by trial-and-error to see if any other arrangement of exactly the same elements will produce better results. The system in Figure 5.4 was found by optimizing EWV with the preference towards fewer channels. Changing to $\underline{\mathbf{d}} = (\ 2\ \ 1\ \ 4\ \ 11\ )$ has the effect of "orthogonalizing" the channels in a way that the PSA-channels are available independently from PSG-channels, a characteristic that we have empirically observed to be indicative of optimality. Considering all the channels in Figure 5.3 lowers EWV by 47.3%, while the systematic approach to measurement selection brings another 35.3% reduction to EWV. In total, EWV was

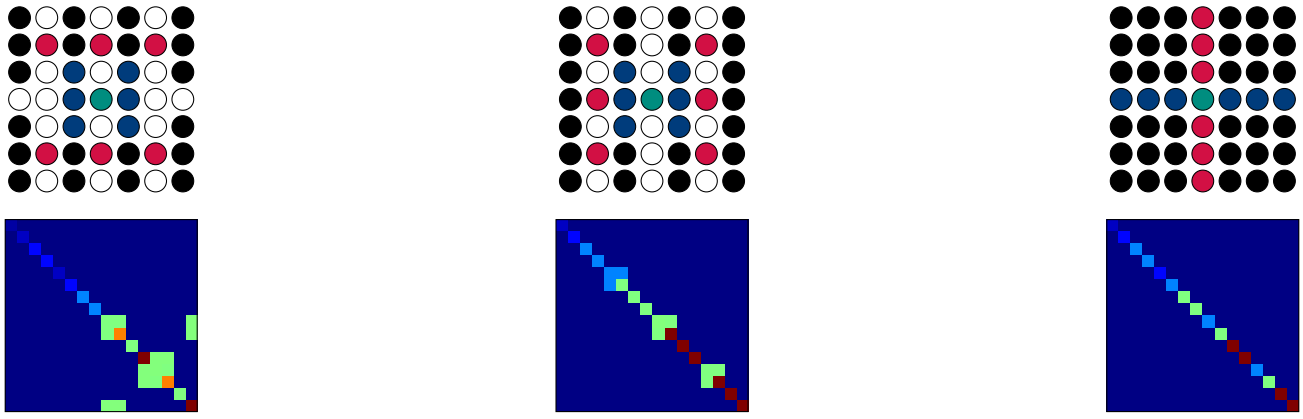
reduced to 34.1% of its original value, suggesting that the polarimeter's SNR is almost three times higher. Since the number of channels did not increase, no spectral resolution was lost.

5.5.2 Spatially Channeled Polarimeter

Next consider the spatial-spatial channeled polarimeter described by Kudenov,[54] an example of a system in Figure 5.2i. The modulation is achieved via polarization gratings that separate the different Mueller matrix elements onto patterns of frequencies that are determined by the spacing of the elements. The intensity can be similarly represented as in Equation 5.23 with $\vec{\vartheta} = \{x, y\}$. Kudenov used the following modulations:

$$\underline{\mathbf{f}}_{\underline{\mathbf{A}}}(x,y) = \begin{bmatrix} 1 \\ \cos(4\pi x) \\ \sin(4\pi x)\cos(4\pi y) \\ \sin(4\pi x)\sin(4\pi y) \end{bmatrix}, \tag{5.26a}$$

$$\underline{\mathbf{f}}_{\underline{\mathbf{G}}}(x,y) = \begin{bmatrix} 1 \\ \cos(2\pi y) \\ \sin(2\pi y)\cos(2\pi x) \\ \sin(2\pi y)\sin(2\pi x) \end{bmatrix}. \tag{5.26b}$$



(c) $a_x = a_y = 2g_x = 2g_y = 2$, $x/y/y/x$, 33/49 channels used, EWV = 209

(d) $a_x = a_y = 2g_x = 2g_y = 2$, $x/y/x/y$, 35/49 channels used, EWV = 151

(e) $a_1 = 2a_2 = g_1 = 2g_2 = 2$, $x/x/y/y$, 49/49 channels used, EWV = 121

Figure 5.5: Top row shows the ξ/η plane of channels of each configuration. Bottom row shows $\underline{\underline{\mathbf{K}}}_{\underline{\underline{\mathbf{Q}}}^+}$.

Figure 5.5 shows a comparison between three systems with the only difference being the order of modulation. The merit of introducing $\underline{\underline{\mathbf{Q}}}$ is clear; better performance is achieved virtually for free, using the same polarization elements in a different order. Systems in Figure 5.5 were found by optimizing using genetic algorithm, while continuously relaxing the design restrictions.

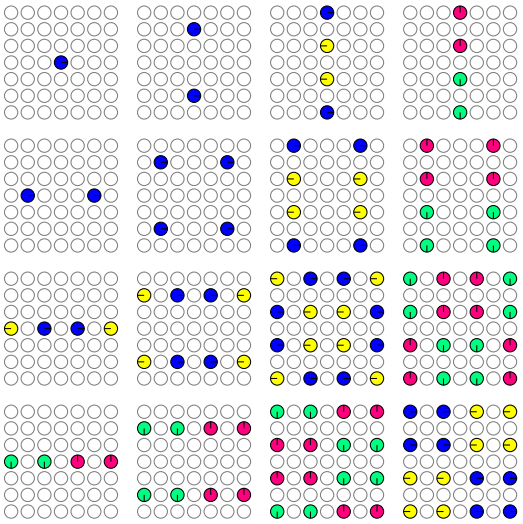


Figure 5.6: Frequency grid of the Mueller modulation. ξ and η are the x- and y-axes, respectively.

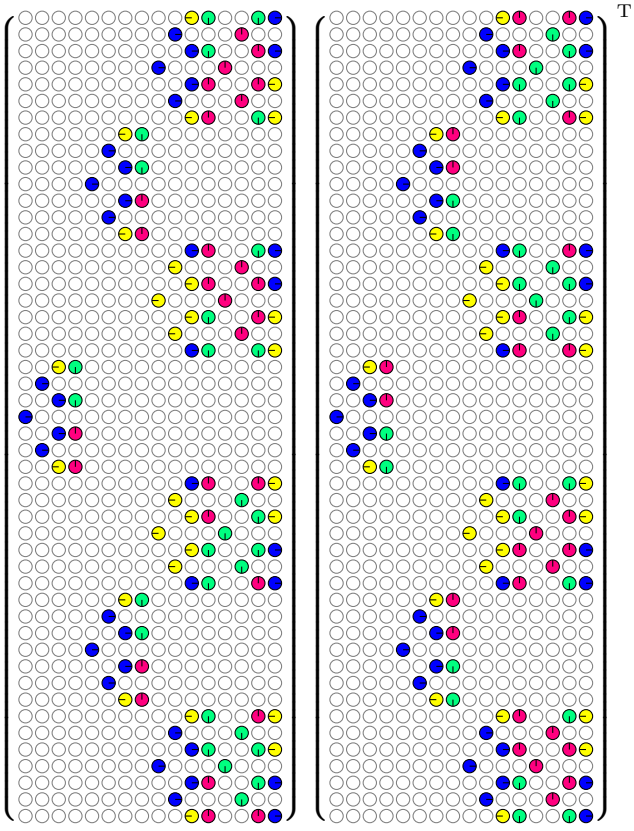


Figure 5.7: Optimal Spatial–Spatial Polarimeter.

Although a symmetrical (x/y/y/x) modulation design may seem intuitively logical, it is possible to improve

the design as evidenced by the polarimeter in Figure 5.5e, which is shown in greater detail in Figure 5.6 and Figure 5.7. First, in Figure 5.5d, an asymmetrical order of modulations improves EWV by 27.8%. Then, in Figure 5.5e, it is improved by another 19.9% by splitting the modulation into one dimensional structures for the PSG and the PSA. In total, EWV was reduced to 57.9% of its original value. Although not as large of an improvement as in the previous example, it is, nonetheless, significant.

The reason for the EWV improvements lies in how the channels interfere. From the comparison of Figure 5.5 we can note that the better systems "focus" the reconstruction into the diagonal of $\underline{\underline{\mathbf{K}}}_{\underline{\underline{\mathbf{Q}}}}$. Upon further investigation, it is possible to make a general statement that we want $\underline{\underline{\mathbf{Q}}}$ to be a matrix that is unitary through a scalar. When this condition holds, the channel structures created for each of the Mueller elements forms an orthogonal basis set, which will force all the variances to lie on the diagonal, thereby ensuring the minimum achievable EWV for the number of modulations introduced. In fact, we can calculate the minimum EWV simply as

$$\mathrm{EWV}_{\min} = \sum \underline{\mathbf{n}}_G \otimes \underline{\mathbf{n}}_A, \tag{5.27}$$

where $\underline{\mathbf{n}}$ is a vector that contains the number of channels that the PSG's and the PSA's Stokes parameters are split into. If we are to assume a spherical modulation, i.e.,

$$\underline{\mathbf{f}}_{\underline{\mathbf{A}}} = \begin{bmatrix} 1 \\ \cos(2\pi x\xi_1) \\ \sin(2\pi x\xi_1)\cos(2\pi x\xi_2) \\ \sin(2\pi x\xi_1)\sin(2\pi x\xi_2) \end{bmatrix}, \tag{5.28a}$$

$$\underline{\mathbf{f}}_{\underline{\mathbf{G}}} = \begin{bmatrix} 1 \\ \cos(2\pi y\eta_1) \\ \sin(2\pi y\eta_1)\cos(2\pi y\eta_2) \\ \sin(2\pi y\eta_1)\sin(2\pi y\eta_2) \end{bmatrix}, \tag{5.28b}$$

then $\underline{\mathbf{n}}_A = \underline{\mathbf{n}}_G = \begin{bmatrix} 1 & 2 & 4 & 4 \end{bmatrix}^{\mathrm{T}}$ and the minimum EWV is 121, like in Figure 5.5e. A better EWV is mathematically possible if for example, $\underline{\mathbf{n}}_A = \underline{\mathbf{n}}_G = \begin{bmatrix} 1 & 2 & 2 & 4 \end{bmatrix}^{\mathrm{T}}$, in which case $\mathrm{EWV}_{\min} = 81$. However, there appears to exist no such physically realizable modulation scheme as it would require analyzing and generating vectors to reach a DoP of $\sqrt{2}$.

5.5.3 Multiple Snapshot Polarimeter

As mentioned before, our matrix $\underline{\underline{\mathbf{Q}}}$ is nothing more than a dyad that contains modulations within each Mueller channel projection target. Thus, once we determine the system's spectral and/or spatial modulations, we can rotate the channel structures via a unitary transformation that will remix the channel structures in PSG and PSA

$$\underline{\underline{\mathbf{Q}}}_\theta = \left[\underline{\underline{\mathbf{U}}}_A \mathcal{F}\{\underline{\mathbf{A}}\}\right] * \left[\mathcal{F}\{\underline{\mathbf{G}}\}^{\mathrm{T}}\, \underline{\underline{\mathbf{U}}}_G\right]. \tag{5.29}$$

Applying $\underline{\underline{\mathbf{U}}}_G$ and $\underline{\underline{\mathbf{U}}}_A$ will keep the channels' relative orientation, thereby maintaining the EWV. Looking at the Mueller matrix of a linear retarder[28] we can show that $|\det(\underline{\underline{\mathbf{LR}}}(\delta,\theta))| = 1$, which means a unitary transformation. Thus, we can construct a multiple snapshot system simply by enclosing the system between two retarders that have different orientations for each measurement. Using Equation 5.13 gives us the total $\underline{\underline{\mathbf{Q}}}$.

		1	2	1	2
		1	1	2	2
1	1	441.0000	171.6667	214.9412	151.0000
2	1	171.6667	121.0000	147.6667	147.6667
1	2	214.9412	147.6667	133.0000	214.9412
2	2	151.0000	147.6667	214.9412	441.0000

(a) One snapshot — 49 channels

		1	2	1	2
		1	1	2	2
1	1	59.9992	53.5044	53.5987	53.5987
2	1	53.5044	53.5667	54.2778	53.5044
1	2	53.5987	54.2778	56.2201	59.0824
2	2	53.5987	53.5044	59.0824	59.9992

(b) Two snapshots — 98 channels

		1	2	1	2
		1	1	2	2
1	1	36.2603	35.5047	35.3180	34.9990
2	1	35.5047	34.0003	34.4980	34.7655
1	2	35.3180	34.4980	35.2025	35.9454
2	2	34.9990	34.7655	35.9454	36.4899

(c) Three snapshots — 147 channels

		1	2	1	2
		1	1	2	2
1	1	25.3344	25.4774	25.5234	25.6483
2	1	25.4774	25.2196	25.6139	25.6146
1	2	25.5234	25.6139	25.7783	26.1636
2	2	25.6483	25.6146	26.1636	25.9362

(d) Four snapshots — 196 channels

Table 3: Optimization results for different number of temporal snapshots

We ran optimizations of 64 differently configured spatial–spatial channeled polarimeters. The distinction is in the frequencies used for a_1, a_2, g_1, g_2 and the corresponding dimension into which the data was mapped. Retardances and orientations of the retarders were the other variables. A genetic algorithm was used to find the lowest EWV.

From these results we gather that as the number of temporal measurements grows, the importance of the spatial frequencies and order of modulations diminishes. This bodes well if we are to understand this phenomenon as a continuously growing temporal bandwidth constraint allowing us to simplify the spatial multiplexing.

5.6 Conclusion

Introducing $\underline{\underline{\mathbf{Q}}}$ and methods for generating it automatically allows us to describe a wide range of similar systems with a handful of parameters and removes the need to handle reconstruction by hand. Furthermore, analysis of $\underline{\underline{\mathbf{Q}}}$ reveals certain design metrics immediately instead of having to run an elaborate simulation. The end result is that a more optimal system can be found often without requiring the use of any extra elements, while injecting the optimization procedure before element selection will allow us to choose better elements.

6. PARTIAL POLARIMETERS

Numerous authors have studied the structure of the Mueller matrix, and much is known about how the various Mueller matrix elements relate to the physical properties of diattenuation, retardance, and depolarization.[31,59] It should be clear that not all 4×4 real matrices are physically realizable. A physical Mueller matrix must map real sets of Stokes parameters into real sets of Stokes parameters, but there are other conditions that must also be met as recently discussed by Gil.[59]

Much of the literature on Mueller matrices is concerned with methods to decompose the Mueller matrix in order to understand its structure and relate it to scattering properties. In the class of series decompositions, the Mueller matrix is broken up into discrete diattenuating, retarding, and depolarizing layers, and the result is a product of Mueller matrices that describe the effects of the whole. Lu and Chipman[60] developed a series decomposition that writes the Mueller matrix as a non-unique cascade of pure diattenuation, retardance, and depolarization Mueller matrices. Ossikovski and colleagues developed a different decomposition that eliminated the order-dependence of the Lu-Chipman decomposition by creating a decomposition that is symmetric through the Minkowski metric tensor $\underline{\underline{\mathbf{G}}} = \mathrm{diag}(1,-1,-1,-1)$.[61] It's clear that while one can use either of these decompositions (or any other), they may not actually represent the physics of any particular process.

The limit of series decompositions is the class of differential decompositions.[62] These split the Mueller matrix into differential slices in an attempt to identify its fundamental characteristics. Noble and Chipman[31,63] use the method of matrix roots to uncover a fundamental differential Mueller matrix that can be written in terms of 15 Mueller matrix generators - three for retardance, three for diattenuation, and nine for depolarization. Ossikovski developed a logarithmic decomposition of the Mueller matrix[64] that operates using a different formalism, but produces an equivalent outcome to that of the matrix roots decomposition.[62]

A third class of decomposition is the class of additive decompositions that consider the Mueller matrix as an ensemble average of parallel scattering processes that are added incoherently. Gil provides a recent review that covers the general cases of the trivial, spectral, and arbitrary decompositions.[59] The most famous parallel decomposition is that of Cloude, who demonstrated that an arbitrary Mueller matrix could be written as a superposition of not more than four pure Mueller-Jones matrices.[65] Ossikovski has demonstrated rigorously that in the limit of weakly depolarizing Mueller matrices, all decompositions return identical polarization properties to first order.[62] However, for more general depolarizing matrices, the various methods return different results for the "fundamental" properties or retardance and diattenuation of a Mueller matrix under test.

All of these classes of decompositions are important for understanding the fundamental properties of the Mueller matrix. However, measurement of the Mueller matrix requires us to consider a different basis set altogether. A Mueller matrix polarimeter operates by using a polarization state generator (PSG) to illuminate the sample with a controlled state of polarization. The polarimeter then measures the irradiance passed through a polarization state analyzer (PSA) set to a second polarization state. Through a suitably diverse set of illumination and analysis states, the elements of the Mueller matrix can be determined.[28] Much as is the case in Stokes polarimetry,[37] the measurement corresponding to each pair of PSG/PSA states can be thought of as a projection onto a basis vector, and then the unknown Mueller matrix can be estimated through a least-squares inversion process that produces an additive decomposition. Once the problem is cast in this manner, the design of a measurement system then becomes an optimization problem where a particular measurement basis is chosen in order to highlight specific aspects of the Mueller matrix. At least 16 measurements are needed in order to reconstruct the full Mueller matrix in general,[28] while the choice of specific illumination states can help balance the signal-to-noise ratio (SNR) and/or error in particular Mueller matrix elements.[58,66] Going one step further, we can even design a partial Mueller matrix polarimeter (pMMP) that allows certain elements or combinations of elements of the Mueller matrix to be recovered with fewer than 16 measurements while ignoring other elements that might not be necessary for a particular sensing problem.[20] Hoover and his coworkers[19,57] have demonstrated that reduced dimensionality subspaces of Mueller matrix space can be used to perform invariant target detection through nonlinear model fitting. Goudail and his coworkers[67–69] have demonstrated that a single-measurement pMMP is optimal for maximizing polarization contrast in a two-class detection problem with known class Mueller matrices.

In this paper we consider the design of pMMPs that seek to measure certain aspects of the Mueller matrix that might be dictated by a particular sensing task. The pMMP could be an imaging or non-imaging device, but

the design of the instrument proceeds from knowledge of linear combinations of Mueller matrix elements that allow a particular task to be performed.[19,20] It is well known that it is not possible to measure a single Mueller matrix element or a single arbitrary combination of Mueller matrix elements in a single measurement due to the restrictions on the structure of the Stokes parameters of the PSG and PSA states. Previous authors[20,70] have considered specific optimization strategies designed to maximize performance on a particular task. In this paper we approach the more general two-part problem of a) identifying the proper subspace in which to make a detection decision and b) designing a pMMP to get as close as possible to a specified subspace of Mueller matrix space through careful selection of measurement states. In order to accomplish this, we discuss some of the details of the structure of the Mueller matrix and how it interacts with the PSG and PSA before developing a numerical optimization method that produces the desired pMMP design.

The remainder of this manuscript is organized as follows. Section 6.1 reviews the mathematics of Mueller matrix polarimetry and discusses the modifications necessary to consider pMMPs. Section 6.2 considers the structure of a few pMMPs in a way that allows us to understand how the PSG and PSA interact with the Mueller matrix to build up a pMMP basis. In Section 6.3 we generalize patterns seen in Section 6.2 to a general class of $4ij$ pMMP systems, as well as develop various metrics by which to evaluate the noise resilience and the proximity of a K-dimensional subspace of Mueller matrix space to an N-measurement pMMP. In Section 6.4 we apply the developed concepts to an object discrimination task from the literature[19] and discuss the results. Section 6.5 concludes the paper.

6.1 Notation Overview

Before discussing the topics of pMMP design, we must first review the concepts of polarimetric sensing and expand them in a way so as to expose its structure and thereby give us tools to handle the information mapping successfully.

Mueller Matrix Polarimetry
Consider an unknown Mueller matrix $\underline{\underline{\mathbf{M}}}$ that modifies the Stokes parameters. The system PSG generates an incident beam with Stokes parameters $\underline{\mathbf{G}}$ and the PSA analyzes the scattered light by taking a projection of the scattered Stokes parameters onto the state described by the Stokes parameter set $\underline{\mathbf{A}}$. The n^{th} measurement of irradiance in the polarimeter is

$$I_n = \underline{\mathbf{A}}_n^{\mathrm{T}} \underline{\underline{\mathbf{M}}}\, \underline{\mathbf{G}}_n = \underline{\mathbf{D}}'^{\mathrm{T}}_n \underline{\mathbf{M}}'. \tag{6.1}$$

In Eq. 6.1 we define the 16×1 vectors:

$$\underline{\mathbf{D}}'_n = \mathrm{vec}(\underline{\mathbf{A}}_n \underline{\mathbf{G}}_n^{\mathrm{T}}) = \underline{\mathbf{A}}_n \otimes \underline{\mathbf{G}}_n, \tag{6.2}$$

and

$$\underline{\mathbf{M}}' = \mathrm{vec}(\underline{\underline{\mathbf{M}}}), \tag{6.3}$$

where $\otimes$ is the Kronecker (direct) product, and $\mathrm{vec}(\underline{\underline{\mathbf{M}}})$ creates a column vector by reordering the matrix $\underline{\underline{\mathbf{M}}}$ into a vector in a row-by-row fashion. Eq. 6.1 shows that a single measurement of the pMMP is a projection of the unknown Mueller matrix onto a known basis vector in $\mathbb{R}^{16}$. By taking a collection of such projections, the unknown matrix – or portions of it in the case of a pMMP – can be determined in a least-squares sense. The series of N measurements in a polarimeter is

$$\underline{\mathbf{I}} = \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_N \end{bmatrix} = \begin{bmatrix} \underline{\mathbf{D}}'^{\mathrm{T}}_1 \underline{\mathbf{M}}' \\ \underline{\mathbf{D}}'^{\mathrm{T}}_2 \underline{\mathbf{M}}' \\ \vdots \\ \underline{\mathbf{D}}'^{\mathrm{T}}_N \underline{\mathbf{M}}' \end{bmatrix} + \vec{\mathbf{n}} = \underline{\underline{\mathbf{W}}}\, \underline{\mathbf{M}}' + \vec{\mathbf{n}}. \tag{6.4}$$

In Eq. 6.4, the $N \times 16$ measurement matrix $\underline{\underline{\mathbf{W}}}$ is

$$\underline{\underline{\mathbf{W}}} = \begin{pmatrix} \underline{\mathbf{D}}'_1 & \underline{\mathbf{D}}'_2 & \cdots & \underline{\mathbf{D}}'_N \end{pmatrix}^{\mathrm{T}}, \tag{6.5}$$

and $\vec{\underline{\mathbf{n}}}$ is considered to be additive noise with variance ϖ^2. The Mueller matrix estimation is typically accomplished by using the pseudoinverse:

$$\hat{\underline{\mathbf{M}}}' = \underline{\underline{\mathbf{W}}}^{+}\underline{\mathbf{I}} = \underline{\underline{\mathbf{W}}}^{+}\underline{\underline{\mathbf{W}}}\,\underline{\mathbf{M}}' + \underline{\underline{\mathbf{W}}}^{+}\vec{\underline{\mathbf{n}}}. \tag{6.6}$$

The performance of $\hat{\underline{\mathbf{M}}}'$ in Eq. 6.6 is dictated by how close $\underline{\underline{\mathbf{W}}}^{+}\underline{\underline{\mathbf{W}}}$ is to the identity matrix and how the pseudoinverse operates on the noise.

Structure of Partial Mueller Matrices
In the case of partial polarimeters, $N < 16$, and the maximum rank that the polarimeter can achieve is N. It is easy to demonstrate

$$\mathrm{tr}\left(\underline{\underline{\mathbf{W}}}^{+}\underline{\underline{\mathbf{W}}}\right) = \mathrm{rank}\left(\underline{\underline{\mathbf{W}}}\right) = R. \tag{6.7}$$

In this paper we will use the singular value decomposition (SVD)[71] to compute the pseudoinverse. The SVD of $\underline{\underline{\mathbf{W}}}$ yields

$$\underline{\underline{\mathbf{W}}} = \underline{\underline{\mathbf{U}}}\,\underline{\underline{\mathbf{\Sigma}}}\,\underline{\underline{\mathbf{V}}}^{\mathrm{T}}. \tag{6.8}$$

The matrices $\underline{\underline{\mathbf{U}}}$ and $\underline{\underline{\mathbf{V}}}$ are $R \times R$ and 16×16 real, orthogonal matrices, respectively, and $\underline{\underline{\mathbf{\Sigma}}}$ is a $R \times 16$ reduced diagonal matrix containing the R singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_R > 0$. The columns of $\underline{\underline{\mathbf{U}}}$ span the range of the pseudoinverse, and the columns of $\underline{\underline{\mathbf{V}}}$ span Mueller matrix space. The first R columns correspond to the non-zero singular values and span the portion of Mueller matrix space that the pMMP can reconstruct. The pseudoinverse can be written as

$$\underline{\underline{\mathbf{W}}}^{+} = \underline{\underline{\mathbf{V}}}\,\underline{\underline{\mathbf{\Sigma}}}^{+}\underline{\underline{\mathbf{U}}}^{\mathrm{T}}, \tag{6.9}$$

where $\underline{\underline{\mathbf{\Sigma}}}^{+}$ is the $16 \times R$ reduced diagonal matrix containing the inverse of the singular values. The SVD pseudoinverse creates a "maximally orthogonal" inverse.

Examining the diagonal elements of $\underline{\underline{\mathbf{W}}}^{+}\underline{\underline{\mathbf{W}}}$ matrix tells how the information from the N measurements contributes to the rank and how that information is distributed in the estimated Mueller matrix. Define the reconstructables matrix as:

$$\underline{\mathbf{B}}' = \mathrm{vec}(\underline{\mathbf{B}}) = \mathrm{diag}\left(\underline{\underline{\mathbf{W}}}^{+}\underline{\underline{\mathbf{W}}}\right). \tag{6.10}$$

We will consider examples of this matrix in subsequent subsections, but at this point we can say that $\underline{\mathbf{B}}$ relates the percentage of each Mueller matrix element that is reconstructed in the pMMP. In the limit of $N = 16$, the pMMP becomes a full polarimeter, $\underline{\underline{\mathbf{W}}}^{+}\underline{\underline{\mathbf{W}}} = \mathbb{I}_{16\times16}$, and $\underline{\mathbf{B}}$ is a 4×4 matrix of all ones; all elements of the Mueller matrix can be reconstructed.

To understand the function of the pMMP, consider the multiplication of the matrix and its pseudoinverse

$$\underline{\underline{\mathbf{W}}}^{+}\underline{\underline{\mathbf{W}}} = \underline{\underline{\mathbf{V}}}\,\underline{\underline{\mathbf{\Sigma}}}^{+}\underline{\underline{\mathbf{U}}}^{+}\underline{\underline{\mathbf{U}}}\,\underline{\underline{\mathbf{\Sigma}}}\,\underline{\underline{\mathbf{V}}}^{+} = \underline{\underline{\mathbf{V}}}\,\underline{\underline{\mathbf{\Sigma}}}^{+}\underline{\underline{\mathbf{\Sigma}}}\,\underline{\underline{\mathbf{V}}}^{+}. \tag{6.11}$$

The matrix $\underline{\underline{\mathbf{\Sigma}}}^{+}\underline{\underline{\mathbf{\Sigma}}}$ is diagonal with the first R elements equaling unity and the last $16 - R$ equaling zero. This permits the claim made in Equation 6.7. Thus,

$$\underline{\underline{\mathbf{W}}}^{+}\underline{\underline{\mathbf{W}}} = \underline{\underline{\mathbf{V}}}'\underline{\underline{\mathbf{V}}}'^{\mathrm{T}}, \tag{6.12}$$

where $\underline{\underline{\mathbf{V}}}'$ is the $16 \times R$ matrix composed of the first R columns of $\underline{\underline{\mathbf{V}}}$. Another way of interpreting the SVD of $\underline{\underline{\mathbf{W}}}$ is that $\underline{\underline{\mathbf{V}}}'$ forms an orthogonal basis that spans the subspace of Mueller matrix space that forms the domain of the particular pMMP represented by $\underline{\underline{\mathbf{W}}}$. Likewise, the columns of $\underline{\underline{\mathbf{V}}}$ discarded by the SVD (corresponding to singular values of zero) span the null space of the particular pMMP. However, as shown later, knowledge of the domain alone is not sufficient to predict performance, as the conditioning of the matrix $\underline{\underline{\mathbf{W}}}$ is important in the presence of noise and error.

6.2 Examples of Partial Mueller Matrix Polarimeters

This analysis is restricted to pMMPs that use fully polarized PSG and PSA states. Goudail and Tyo[69] demonstrated that partially polarized PSG and PSA states never improve contrast. Below we will consider cases where one or more PSG or PSA state is unpolarized, allowing reconstruction of particular elements of the Mueller matrix with fewer measurements than would be necessary if all PSG and PSA states were fully polarized.

Canonical 4-Measurement pMMP

Consider the simple $N = 4$ measurement pMMP that measures the co-polarized and cross-polarized return for both vertically and horizontally polarized illumination. For compactness, we introduce the following notation for the analyzer and generator matrices

$$\underline{\underline{\mathbf{A}}} \Rightarrow \tfrac{1}{2}\left[\ \rightarrow\ \ \rightarrow\ \ \uparrow\ \ \uparrow\ \right], \tag{6.13a}$$

$$\underline{\underline{\mathbf{G}}} \Rightarrow \ \left[\ \rightarrow\ \ \uparrow\ \ \rightarrow\ \ \uparrow\ \right], \tag{6.13b}$$

where $\rightarrow = \left[\begin{array}{cccc} 1 & 1 & 0 & 0 \end{array}\right]^{\mathrm{T}}$ is the set of Stokes parameters for ideally horizontally polarized light and $\uparrow = \left[\begin{array}{cccc} 1 & -1 & 0 & 0 \end{array}\right]^{\mathrm{T}}$ is the set of Stokes parameters for ideally vertically polarized light. The presence of $\frac{1}{2}$ in the definitions of the analyzing vector is needed for rigor — the polarization sensing systems in consideration dismiss half of the light if the input is unpolarized. The set of four PSG/PSA pairs in Eqs. 6.13a and 6.13b results in the instrument matrix

$$\underline{\underline{\mathbf{W}}} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{6.14}$$

and the reconstructables matrix

$$\underline{\underline{\mathbf{B}}} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{6.15}$$

This is the well known result that four measurements are needed to reconstruct four Mueller matrix elements, and that these four elements must come in a "block" pattern within the Mueller matrix.[20] A similar polarimeter could be obtained with a 4-measurment combination of any two of the six canonical states $\rightarrow$, $\uparrow$, $\nearrow$, $\nwarrow$, $\circlearrowright$, $\circlearrowleft$, where $\nearrow$ and $\nwarrow$ represent 45° and −45°, and $\circlearrowright$ and $\circlearrowleft$ represent right- and left circular polarization, respectively.

While this well-known result tells how to design a pMMP to reconstruct one of these groupings of four elements, it is not obvious how to add additional measurements to reconstruct additional elements or how to design a pMMP to reconstruct linear combinations of elements rather than isolated elements within the pMMP.

Diagonal Depolarization Elements

Depolarization is a rich physical process that contains significant information about the random scattering properties.[72] Noble and Chipman[31,63] recently described the nine degrees of freedom for depolarization. Three of these correspond to randomness in the diattenuation properties of the Mueller matrix, three to randomness in the retardance properties of the Mueller matrix, and three to "diagonal depolarization," which is related to randomness in geometric transformations as would happen in multiple scattering or rough surface scattering processes. Often, the diagonal depolarization elements are important for discrimination in both optical and radar tasks.[73]

Subsection 6.2 shows that each canonical four-measurement polarimeter provides one diagonal element (in addition to m_{00}, which is involved in all pMMPs). One obvious way to reconstruct the diagonal elements then would be to use a 12-measurement pMMP defined by the analyzer and generator matrices

$$\underline{\underline{\mathbf{A}}} \Rightarrow \tfrac{1}{2}\left[\ \rightarrow\ \ \rightarrow\ \ \uparrow\ \ \uparrow\ \ \nearrow\ \ \nearrow\ \ \nwarrow\ \ \nwarrow\ \ \circlearrowleft\ \ \circlearrowleft\ \ \circlearrowright\ \ \circlearrowright\ \right], \tag{6.16a}$$

$$\underline{\underline{\mathbf{G}}} \Rightarrow \ \left[\ \rightarrow\ \ \uparrow\ \ \rightarrow\ \ \uparrow\ \ \nearrow\ \ \nwarrow\ \ \nearrow\ \ \nwarrow\ \ \circlearrowleft\ \ \circlearrowright\ \ \circlearrowleft\ \ \circlearrowright\ \right], \tag{6.16b}$$

which produces the following reconstructables matrix

$$\underline{\underline{\mathbf{B}}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}. \tag{6.17}$$

In addition to the desired diagonal elements, the diattenuation and polarizance vectors[60] are also measured. This 12-measurement pMMP only reconstructs 10 Mueller matrix elements, since each of the three canonical pMMPs redundantly reconstructs m_{00}.

We can address this redundancy by eliminating one of the cross-polarized measurements in two of the canonical pMMPs so that

$$\underline{\underline{\mathbf{A}}} = \tfrac{1}{2}\big[\ \rightarrow\ \rightarrow\ \uparrow\ \uparrow\ \nearrow\ \nwarrow\ \nwarrow\ \circlearrowright\ \circlearrowright\ \circlearrowright\ \big], \tag{6.18a}$$

$$\underline{\underline{\mathbf{G}}} = \big[\ \rightarrow\ \uparrow\ \rightarrow\ \uparrow\ \nearrow\ \nearrow\ \nwarrow\ \circlearrowright\ \circlearrowright\ \circlearrowright\ \big], \tag{6.18b}$$

which produces the same reconstructables matrix as Eq. 6.17. In this case the elimination of redundancy allows 10 Mueller matrix elements to be reconstructed from 10 measurements.

The matrix of Eq. 6.17 still unnecessarily reconstructs the first column and the first row of the Mueller matrix. The number of measurements can be lowered by further reducing two of the three canonical 4-measurement pMMPs to two-measurement pMMPs that make co-polarized measurements only, e. g.

$$\underline{\underline{\mathbf{A}}} \Rightarrow \tfrac{1}{2}\big[\ \rightarrow\ \rightarrow\ \uparrow\ \uparrow\ \nearrow\ \nwarrow\ \circlearrowright\ \circlearrowright\ \big], \tag{6.19a}$$

$$\underline{\underline{\mathbf{G}}} \Rightarrow \big[\ \rightarrow\ \uparrow\ \rightarrow\ \uparrow\ \nearrow\ \nwarrow\ \circlearrowright\ \circlearrowright\ \big], \tag{6.19b}$$

which produces the reconstructables matrix

$$\underline{\underline{\mathbf{B}}} = \begin{bmatrix} 1 & 1 & \frac{1}{2} & \frac{1}{2} \\ 1 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 0 \\ \frac{1}{2} & 0 & 0 & 1 \end{bmatrix}. \tag{6.20}$$

Examination of $\underline{\underline{\mathbf{V}}}$ can help to determine how the elements of $\underline{\underline{\mathbf{B}}}$ correspond to reconstructed Mueller matrix channels as discussed in section 6.3. This pMMP can reconstruct the diagonal elements $m_{00}, m_{11}, m_{22}, m_{33}$ as well as the elements m_{10} and m_{01}. In addition to these individual elements, the polarimeter can also reconstruct the linear combination channels $(m_{20} + m_{02})/\sqrt{2}$ and $(m_{30} + m_{03})/\sqrt{2}$. Note that existence of reconstruction channels does not guarantee that these channels will have acceptable SNR. These items are discussed in greater detail below.

The polarimeter described by Eq. 6.19a and Eq. 6.19b is the lowest dimensionality that we have been able to find that reconstructs all three of the diagonal elements with fully polarized analyzer and generator states. However, use of unpolarized measurements adds another degree of freedom and provides capacity for fewer measurements. Consider a system that makes six canonical, co-polarized measurements and one unpolarized measurement

$$\underline{\underline{\mathbf{A}}} \Rightarrow \tfrac{1}{2}\big[\ \bigcirc\ \rightarrow\ \uparrow\ \nearrow\ \nwarrow\ \circlearrowright\ \circlearrowright\ \big], \tag{6.21a}$$

$$\underline{\underline{\mathbf{G}}} \Rightarrow \big[\ \bigcirc\ \rightarrow\ \uparrow\ \nearrow\ \nwarrow\ \circlearrowright\ \circlearrowright\ \big], \tag{6.21b}$$

where $\bigcirc = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}}$. The reconstructables matrix for this polarimeter is

$$\underline{\underline{\mathbf{B}}} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 0 \\ \frac{1}{2} & 0 & 0 & 1 \end{bmatrix}. \tag{6.22}$$

The addition of the one unpolarized measurement allows the m_{00} term to be reconstructed directly, obviating the need for the cross-polarized measurements indicated in Eq. 6.19a and Eq. 6.19b.

6.3 Partial Mueller Matrix Polarimeter Design

Based on the understanding gained in Section 6.2, we can remove some prior constraints, and generalize pMMP systems to have arbitrary fully polarized PSG and PSA states within the requirement that $N = R$.

Interpreting the Reconstructables Matrix

It is unclear from $\underline{\underline{\mathbf{B}}}$ alone, which Mueller elements are grouped together into combinatorial channels. When a particular element of $\underline{\underline{\mathbf{B}}}$ is unity, then we can reconstruct the corresponding Mueller matrix element. But when it is other than unity, the element must appear in combination with other Mueller matrix elements. However, the fact that $\underline{\underline{\mathbf{B}}}$ is derived from the columns of $\underline{\underline{\mathbf{V}}}'$ provides insight into the overall subspace spanned by the pMMP. If two or more columns of $\underline{\underline{\mathbf{V}}}'$, say $\underline{\mathbf{v}}_n$ and $\underline{\mathbf{v}}_m$ correspond to equal singular values $\sigma_n = \sigma_m$, then they span a hyperplane with identical geometrical characteristics in the context of $\underline{\underline{\mathbf{W}}}$. In this case, any set of orthogonal basis vectors in that hyperplane can be used, allowing for a more intuitive grouping of Mueller matrix elements if desired.

All of the previous examples featured four-block measurements. Prior work was already familiar with $m_{00}/m_{i0}/m_{0j}/m_{ij}$ measurement, but a similar reconstruction exists for off-Mueller grid measurement. Consider the arbitrary analyzing and generating vectors:

$$\underline{\mathbf{A}}_{\pm} = \begin{bmatrix} 1 & \pm a_1 & \pm a_2 & \pm a_3 \end{bmatrix}, \tag{6.23a}$$

$$\underline{\mathbf{G}}_{\pm} = \begin{bmatrix} 1 & \pm g_1 & \pm g_2 & \pm g_3 \end{bmatrix}. \tag{6.23b}$$

A four-block polarimeter would go through the following four combinations of measurements:

$$\underline{\underline{\mathbf{A}}}_4 \Rightarrow \tfrac{1}{2}\begin{bmatrix} \underline{\mathbf{A}}_+ & \underline{\mathbf{A}}_+ & \underline{\mathbf{A}}_- & \underline{\mathbf{A}}_- \end{bmatrix}^{\mathrm{T}}, \tag{6.24a}$$

$$\underline{\underline{\mathbf{G}}}_4 \Rightarrow \begin{bmatrix} \underline{\mathbf{G}}_+ & \underline{\mathbf{G}}_- & \underline{\mathbf{G}}_+ & \underline{\mathbf{G}}_- \end{bmatrix}^{\mathrm{T}}. \tag{6.24b}$$

The resulting measurement matrix is

$$\underline{\underline{\mathbf{W}}}_4 = \begin{bmatrix} (\underline{\mathbf{A}}_+ \otimes \underline{\mathbf{G}}_+)^{\mathrm{T}} \\ (\underline{\mathbf{A}}_+ \otimes \underline{\mathbf{G}}_-)^{\mathrm{T}} \\ (\underline{\mathbf{A}}_- \otimes \underline{\mathbf{G}}_+)^{\mathrm{T}} \\ (\underline{\mathbf{A}}_- \otimes \underline{\mathbf{G}}_-)^{\mathrm{T}} \end{bmatrix}, \tag{6.25}$$

The SVD of $\underline{\underline{\mathbf{W}}}_4$ in Eq. 6.25 will have four column-space vectors with four identical singular values. If a particular representation of $\underline{\underline{\mathbf{V}}}$ is chosen, then there is only one $\underline{\underline{\mathbf{U}}}$ to go along with it. A linear combination of these vectors corresponds to rotation of the underlying vectors. This operation does not alter the space, but merely rotates the axes by which that space is described. We will show that we can write down a particular decomposition that is relatively easy to treat analytically. When faced with more complex $\underline{\underline{\mathbf{V}}}$ matrices that have non-equal singular values, adding and subtracting the underlying vectors is also possible, but special care needs to be taken.

For the four-block polarimeter, the column space can be described with the following set of vectors

$$\underline{\underline{\mathbf{V}}}'_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & g_1 & 0 & 0 \\ 0 & g_2 & 0 & 0 \\ 0 & g_3 & 0 & 0 \\ 0 & 0 & a_1 & 0 \\ 0 & 0 & 0 & a_1 g_1 \\ 0 & 0 & 0 & a_1 g_2 \\ 0 & 0 & 0 & a_1 g_3 \\ 0 & 0 & a_2 & 0 \\ 0 & 0 & 0 & a_2 g_1 \\ 0 & 0 & 0 & a_2 g_2 \\ 0 & 0 & 0 & a_2 g_3 \\ 0 & 0 & a_3 & 0 \\ 0 & 0 & 0 & a_3 g_1 \\ 0 & 0 & 0 & a_3 g_2 \\ 0 & 0 & 0 & a_3 g_3 \end{bmatrix}, \tag{6.26}$$

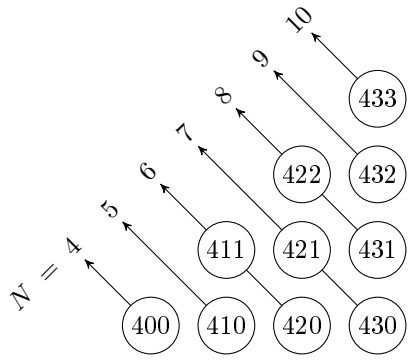


Figure 6.1: Possible sets of measurements that maintain the optimal $N = R$.

that correspond to four unity singular values and

$$\underline{\underline{\mathbf{U}}} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}. \tag{6.27}$$

The corresponding reconstructables matrix is

$$\underline{\underline{\mathbf{B}}}_4 = \begin{bmatrix} 1 & {g_1}^2 & {g_2}^2 & {g_3}^2 \\ {a_1}^2 & {a_1}^2{g_1}^2 & {a_1}^2{g_2}^2 & {a_1}^2{g_3}^2 \\ {a_2}^2 & {a_2}^2{g_1}^2 & {a_2}^2{g_2}^2 & {a_2}^2{g_3}^2 \\ {a_3}^2 & {a_3}^2{g_1}^2 & {a_3}^2{g_2}^2 & {a_3}^2{g_3}^2 \end{bmatrix}. \tag{6.28}$$

Additional Measurements

In order to expand the space coverage, we need to add more measurements. To do so while keeping system rank equal to the number of measurements means that each additional column in $\underline{\underline{\mathbf{V}}}'$ needs to be orthogonal to every pre-existing one. If we are limited by the assumption of fully-polarized measurements, then the new analyzing vector pair $\underline{\mathbf{A}}_{2,\pm}$ needs to be orthogonal to the pre-existing analyzing vector pair $\underline{\mathbf{A}}_{1,\pm}$ in the Poincaré sphere space. Thus, once the first pair is selected, the new pair is bound to the space of the orthogonal circle. Once $\underline{\mathbf{A}}_{2,\pm}$ is chosen, there is only one more orthogonal set of vectors $\underline{\mathbf{A}}_{3,\pm}$ that can be added.

It is important to make the connection between this general case and the one discussed in Subsection 4.B. If we make four measurements from each combination in $\underline{\mathbf{A}}_{1,\pm}$, $\underline{\mathbf{A}}_{2,\pm}$ and $\underline{\mathbf{A}}_{3,\pm}$, then we will end up with $N = 12$ and $R = 10$. This is because each block is capable of reconstructing m_{00} on its own and measuring it three times will have the effect of averaging, and thereby lowering the noise in its reconstruction. In order to keep $N = R$ we must make only one four-measurement set specified by $\underline{\mathbf{A}}_{1,\pm}$ and up to two additional fewer-than-four measurements specified by $\underline{\mathbf{A}}_{2,\pm}$ and $\underline{\mathbf{A}}_{3,\pm}$. This produces 16 possible measurement schemes. By denoting the set as $4ij$ and requiring that $4 > i \geq j$, we can ignore the six redundant schemes as can be seen in Fig. 6.1.

Purely for purposes of simplifying the notation, we define the analyzing and generating vectors of the one-, two- and three-measurement cases as:

$$\underline{\underline{\mathbf{A}}}_3 \Rightarrow \tfrac{1}{2}\begin{bmatrix} \underline{\mathbf{A}}_+ & \underline{\mathbf{A}}_- & \underline{\mathbf{A}}_- \end{bmatrix}^{\mathrm{T}}, \tag{6.29a}$$

$$\underline{\underline{\mathbf{G}}}_3 \Rightarrow \begin{bmatrix} \underline{\mathbf{G}}_+ & \underline{\mathbf{G}}_+ & \underline{\mathbf{G}}_- \end{bmatrix}^{\mathrm{T}}, \tag{6.29b}$$

$$\underline{\underline{\mathbf{A}}}_2 \Rightarrow \tfrac{1}{2}\begin{bmatrix} \underline{\mathbf{A}}_+ & \underline{\mathbf{A}}_- \end{bmatrix}^{\mathrm{T}}, \tag{6.29c}$$

$$\underline{\underline{\mathbf{G}}}_2 \Rightarrow \begin{bmatrix} \underline{\mathbf{G}}_+ & \underline{\mathbf{G}}_- \end{bmatrix}^{\mathrm{T}}, \tag{6.29d}$$

$$\underline{\underline{\mathbf{A}}}_1 \Rightarrow \tfrac{1}{2}\begin{bmatrix} \underline{\mathbf{A}}_+ \end{bmatrix}^{\mathrm{T}}, \tag{6.29e}$$

$$\underline{\underline{\mathbf{G}}}_1 \Rightarrow \begin{bmatrix} \underline{\mathbf{G}}_+ \end{bmatrix}^{\mathrm{T}}. \tag{6.29f}$$

It can be shown that this selection considers all possible combinations. The $\pm$ only denotes the operation on the a_1/g_1, a_2/g_2, and a_3/g_3, but a selection of a different vector can effectively construct all other combinations within the syntax implied by $\pm$.

The denoted characteristics of these measurements are only correct if m_{00} is known. Thus, they can only be used as additional measurements and their measurement sub-matrices are:

$$\underline{\underline{\mathbf{W}}}_3 = \begin{bmatrix} (\underline{\mathbf{A}}_+ \otimes \underline{\mathbf{G}}_+)^{\mathrm{T}} \\ (\underline{\mathbf{A}}_- \otimes \underline{\mathbf{G}}_+)^{\mathrm{T}} \\ (\underline{\mathbf{A}}_- \otimes \underline{\mathbf{G}}_-)^{\mathrm{T}} \end{bmatrix}, \tag{6.30a}$$

$$\underline{\underline{\mathbf{W}}}_2 = \begin{bmatrix} (\underline{\mathbf{A}}_+ \otimes \underline{\mathbf{G}}_+)^{\mathrm{T}} \\ (\underline{\mathbf{A}}_- \otimes \underline{\mathbf{G}}_-)^{\mathrm{T}} \end{bmatrix}, \tag{6.30b}$$

$$\underline{\underline{\mathbf{W}}}_1 = \begin{bmatrix} (\underline{\mathbf{A}}_+ \otimes \underline{\mathbf{G}}_+)^{\mathrm{T}} \end{bmatrix}, \tag{6.30c}$$

and the sub-reconstructables matrices are:

$$\underline{\underline{\mathbf{B}}}_3 = \begin{bmatrix} 0 & {g_1}^2 & {g_2}^2 & {g_3}^2 \\ {a_1}^2 & {a_1}^2{g_1}^2 & {a_1}^2{g_2}^2 & {a_1}^2{g_3}^2 \\ {a_2}^2 & {a_2}^2{g_1}^2 & {a_2}^2{g_2}^2 & {a_2}^2{g_3}^2 \\ {a_3}^2 & {a_3}^2{g_1}^2 & {a_3}^2{g_2}^2 & {a_3}^2{g_3}^2 \end{bmatrix}, \tag{6.31a}$$

$$\underline{\underline{\mathbf{B}}}_2 = \frac{1}{2}\begin{bmatrix} 0 & {g_1}^2 & {g_2}^2 & {g_3}^2 \\ {a_1}^2 & 2{a_1}^2{g_1}^2 & 2{a_1}^2{g_2}^2 & 2{a_1}^2{g_3}^2 \\ {a_2}^2 & 2{a_2}^2{g_1}^2 & 2{a_2}^2{g_2}^2 & 2{a_2}^2{g_3}^2 \\ {a_3}^2 & 2{a_3}^2{g_1}^2 & 2{a_3}^2{g_2}^2 & 2{a_3}^2{g_3}^2 \end{bmatrix}, \tag{6.31b}$$

$$\underline{\underline{\mathbf{B}}}_1 = \frac{1}{3}\begin{bmatrix} 0 & {g_1}^2 & {g_2}^2 & {g_3}^2 \\ {a_1}^2 & {a_1}^2{g_1}^2 & {a_1}^2{g_2}^2 & {a_1}^2{g_3}^2 \\ {a_2}^2 & {a_2}^2{g_1}^2 & {a_2}^2{g_2}^2 & {a_2}^2{g_3}^2 \\ {a_3}^2 & {a_3}^2{g_1}^2 & {a_3}^2{g_2}^2 & {a_3}^2{g_3}^2 \end{bmatrix}. \tag{6.31c}$$

The total measurement matrix of a $4ij$ system is

$$\underline{\underline{\mathbf{W}}}_{4ij} = \begin{bmatrix} \underline{\underline{\mathbf{W}}}_4^{\mathrm{T}} & \underline{\underline{\mathbf{W}}}_i^{\mathrm{T}} & \underline{\underline{\mathbf{W}}}_j^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}. \tag{6.32}$$

The constraints placed on $\underline{\mathbf{A}}_{2,\pm}$ and $\underline{\mathbf{A}}_{3,\pm}$ mean that the reconstructables matrix is the sum of the sub-matrices

$$\underline{\underline{\mathbf{B}}}_{4ij} = \underline{\underline{\mathbf{B}}}_4 + \underline{\underline{\mathbf{B}}}_i + \underline{\underline{\mathbf{B}}}_j. \tag{6.33}$$

Structured Decomposition

As before, we can perform SVD on the matrix to find the space coverage and noise resilience of any given polarimeter. However, in the case of being limited to the defined class of $4ij$ pMMP systems, we can introduce a structured decomposition

$$\underline{\underline{\mathbf{W}}}_{4ij} = \underline{\underline{\mathbf{U}}}_{s,4ij}\underline{\underline{\mathbf{\Sigma}}}_{s,4ij}\underline{\underline{\mathbf{V}}}_{s,4ij}^{\mathrm{T}}, \tag{6.34}$$

where s differentiates this decomposition from the typical SVD. The goal of this decomposition is to be easily parsable by a human and provide an intuitive view of pMMP properties. The following are the structured matrices for any $4ij$ system:

$$\underline{\underline{\mathbf{U}}}_{s,4ij} = \begin{bmatrix} \underbrace{\begin{bmatrix} \sqrt{\frac{\mathbf{1}}{\mathbf{N}}} \end{bmatrix}}_{N\times 1} & \underline{\underline{\mathbf{U}}}'_4 & \underline{\underline{\mathbf{0}}}_{4\times i} & \underline{\underline{\mathbf{0}}}_{4\times j} \\ & \underline{\underline{\mathbf{0}}}_{i\times 3} & \underline{\underline{\mathbf{U}}}'_i & \underline{\underline{\mathbf{0}}}_{i\times j} \\ & \underline{\underline{\mathbf{0}}}_{j\times 3} & \underline{\underline{\mathbf{0}}}_{j\times i} & \underline{\underline{\mathbf{U}}}'_j \end{bmatrix}, \tag{6.35}$$

$$\underline{\underline{\mathbf{\Sigma}}}_{s,4ij} = \sqrt{\frac{1}{4}\,\mathrm{diag}(\begin{array}{cccc} N & \vec{\varsigma}_4 & \vec{\varsigma}_i & \vec{\varsigma}_j \end{array})}, \tag{6.36}$$

$$\underline{\underline{\mathbf{V}}}_{s,4ij} = \begin{bmatrix} \underline{\underline{\mathbf{V}}}'_4 & \underline{\underline{\mathbf{V}}}'_i & \underline{\underline{\mathbf{V}}}'_j \end{bmatrix}, \tag{6.37}$$

where the left structured sub-matrices are:

$$\underline{\underline{\mathbf{U}}}'_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \\ -1 & -1 & 1 \end{bmatrix}, \tag{6.38a}$$

$$\underline{\underline{\mathbf{U}}}'_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ -1 & -1 & 1 \end{bmatrix}, \tag{6.38b}$$

$$\underline{\underline{\mathbf{U}}}'_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \tag{6.38c}$$

$$\underline{\underline{\mathbf{U}}}'_1 = \frac{1}{\sqrt{1}} \begin{bmatrix} 1 \end{bmatrix}, \tag{6.38d}$$

the effectively rotated singular values are:

$$\vec{\varsigma}_4 = \{4, 4, 4\}, \tag{6.39a}$$

$$\vec{\varsigma}_3 = \{3, 3, 3\}, \tag{6.39b}$$

$$\vec{\varsigma}_2 = \{4, 2\}, \tag{6.39c}$$

$$\vec{\varsigma}_1 = \{3\}, \tag{6.39d}$$

while the right structured sub-matrices can be defined in terms of Eq. 6.26. If we write $\underline{\underline{\mathbf{V}}}'_4 = \begin{bmatrix} \underline{\mathbf{v}}_1 & \underline{\mathbf{v}}_2 & \underline{\mathbf{v}}_3 & \underline{\mathbf{v}}_4 \end{bmatrix}$, then the corresponding structured sub-matrices are:

$$\underline{\underline{\mathbf{V}}}'_4 = \begin{bmatrix} \underline{\mathbf{v}}_1 & \underline{\mathbf{v}}_2 & \underline{\mathbf{v}}_3 & \underline{\mathbf{v}}_4 \end{bmatrix}, \tag{6.40a}$$

$$\underline{\underline{\mathbf{V}}}'_3 = \begin{bmatrix} \underline{\mathbf{v}}_2 & \underline{\mathbf{v}}_3 & \underline{\mathbf{v}}_4 \end{bmatrix}, \tag{6.40b}$$

$$\underline{\underline{\mathbf{V}}}'_2 = \begin{bmatrix} \frac{1}{\sqrt{2}}(\underline{\mathbf{v}}_2 + \underline{\mathbf{v}}_3) & \underline{\mathbf{v}}_4 \end{bmatrix}, \tag{6.40c}$$

$$\underline{\underline{\mathbf{V}}}'_1 = \begin{bmatrix} \frac{1}{\sqrt{3}}(\underline{\mathbf{v}}_2 + \underline{\mathbf{v}}_3 + \underline{\mathbf{v}}_4) \end{bmatrix}. \tag{6.40d}$$

Procedures defined above create matrices that are orthogonal in both dimensions, but normalizable only in one. To quantify noise resilience, we need to know the product $\underline{\underline{\mathbf{\Sigma}}}^+_{s,4ij} \underline{\underline{\mathbf{U}}}^+_{s,4ij}$. Calculating $\underline{\underline{\mathbf{\Sigma}}}^+_{s,4ij}$ is trivial because $\underline{\underline{\mathbf{\Sigma}}}_{s,4ij}$ is diagonal, while calculating $\underline{\underline{\mathbf{U}}}^+_{s,4ij}$ is more challenging. However, for $4ij$ systems, we can show that

$$\underline{\underline{\mathbf{U}}}^+_{s,4ij} = \begin{bmatrix} \underline{\underline{\left[\sqrt{\frac{\mathbf{N}}{\mathbf{16}}}\right]}}^{1\times 4} & \underline{\underline{\mathbf{0}}}_{1\times i} & \underline{\underline{\mathbf{0}}}_{1\times j} \\ \underline{\underline{\mathbf{U}}}'^{+}_4 & \underline{\underline{\mathbf{0}}}_{3\times i} & \underline{\underline{\mathbf{0}}}_{3\times j} \\ \underline{\underline{\mathbf{R}}}_i & \underline{\underline{\mathbf{U}}}'^{+}_i & \underline{\underline{\mathbf{0}}}_{i\times j} \\ \underline{\underline{\mathbf{R}}}_j & \underline{\underline{\mathbf{0}}}_{j\times i} & \underline{\underline{\mathbf{U}}}'^{+}_j \end{bmatrix} \tag{6.41}$$

is a correct inverse, where

$$\underline{\underline{\mathbf{U}}}'^{+}_4 = \underline{\underline{\mathbf{U}}}'^{\mathrm{T}}_4, \tag{6.42a}$$

$$\underline{\underline{\mathbf{U}}}'^{+}_3 = \frac{\sqrt{3}}{2} \begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & -1 \\ 1 & 0 & 1 \end{bmatrix}, \tag{6.42b}$$

$$\underline{\underline{\mathbf{U}}}'^{+}_2 = \underline{\underline{\mathbf{U}}}'^{\mathrm{T}}_2, \tag{6.42c}$$

$$\underline{\underline{\mathbf{U}}}'^{+}_1 = \underline{\underline{\mathbf{U}}}'^{\mathrm{T}}_1, \tag{6.42d}$$

and

$$\underline{\underline{\mathbf{R}}}_3 = \sqrt{\frac{3}{16}} \begin{bmatrix} -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \end{bmatrix}, \tag{6.43a}$$

$$\underline{\underline{\mathbf{R}}}_2 = \sqrt{\frac{2}{16}} \begin{bmatrix} 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & -1 \end{bmatrix}, \tag{6.43b}$$

$$\underline{\underline{\mathbf{R}}}_1 = \sqrt{\frac{1}{16}} \begin{bmatrix} -1 & -1 & -1 & -1 \end{bmatrix}. \tag{6.43c}$$

Noise Resilience

Following the same additive noise model as prescribed by Eq. 6.4, we can project the noise into the respective directions of the *sensor space*,[74] $\mathcal{V}$:

$$\underline{\vec{\mathbf{n}}}' = \underline{\underline{\mathbf{W}}}^{+}_{4ij} \underline{\vec{\mathbf{n}}}, \tag{6.44}$$

where the pseudoinverse can be written within the context of $4ij$ systems as

$$\underline{\underline{\mathbf{W}}}^{+}_{4ij} = \underline{\underline{\mathbf{V}}}_{s,4ij} \underline{\underline{\mathbf{\Sigma}}}^{+}_{s,4ij} \underline{\underline{\mathbf{U}}}^{+}_{s,4ij} = \underline{\underline{\mathbf{V}}}_{s,4ij} \underline{\underline{\mathbf{L}}}_{\underline{\underline{\mathbf{V}}}_{s,4ij}}. \tag{6.45}$$

$\underline{\underline{\mathbf{L}}}_{\underline{\underline{\mathbf{V}}}_{s,4ij}}$ contains the mapping weights of information for each of the vectors of $\underline{\underline{\mathbf{V}}}_{s,4ij}$:

$$\underline{\underline{\mathbf{L}}}_{\underline{\underline{\mathbf{V}}}_{s,4ij}} = \begin{bmatrix} \underline{\ell}^{\mathrm{T}}_{\underline{\mathbf{v}}_1} & \underline{\ell}^{\mathrm{T}}_{\underline{\mathbf{v}}_2} & \cdots & \underline{\ell}^{\mathrm{T}}_{\underline{\mathbf{v}}_N} \end{bmatrix}^{\mathrm{T}}. \tag{6.46}$$

Since the pMMP's sensor space contains N vectors, $\underline{\underline{\mathbf{L}}}_{\underline{\underline{\mathbf{V}}}_{s,4ij}}$ also contains N channels. The Euclidean length of each of those vectors represents the noise magnitude in each of the vectors of $\underline{\underline{\mathbf{V}}}_{s,4ij}$,

$$p_{\underline{\mathbf{v}}} = ||\underline{\ell}_{\underline{\mathbf{v}}}||_2. \tag{6.47}$$

For each measurement set, the matrix multiplications reveal that each of the vectors making up $\underline{\underline{\mathbf{V}}}_{s,4ij}$ will have easily identifiable noise magnitudes:

$$\underline{\mathbf{P}}_{\underline{\underline{\mathbf{V}}}_4} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^{\mathrm{T}}, \tag{6.48a}$$

$$\underline{\mathbf{P}}_{\underline{\underline{\mathbf{V}}}_3} = \begin{bmatrix} \sqrt{3} & \sqrt{3} & \sqrt{3} \end{bmatrix}^{\mathrm{T}}, \tag{6.48b}$$

$$\underline{\mathbf{P}}_{\underline{\underline{\mathbf{V}}}_2} = \begin{bmatrix} 1 & \sqrt{3} \end{bmatrix}^{\mathrm{T}}, \tag{6.48c}$$

$$\underline{\mathbf{P}}_{\underline{\underline{\mathbf{V}}}_1} = \begin{bmatrix} \sqrt{\frac{5}{3}} \end{bmatrix}^{\mathrm{T}}. \tag{6.48d}$$

Finally, the total noise magnitude vector is the concatenation of the ones defined above,

$$\underline{\mathbf{P}}_{\underline{\underline{\mathbf{V}}}_{s,4ij}} = \begin{bmatrix} \underline{\mathbf{P}}^{\mathrm{T}}_{\underline{\underline{\mathbf{V}}}_4} & \underline{\mathbf{P}}^{\mathrm{T}}_{\underline{\underline{\mathbf{V}}}_i} & \underline{\mathbf{P}}^{\mathrm{T}}_{\underline{\underline{\mathbf{V}}}_j} \end{bmatrix}^{\mathrm{T}}. \tag{6.49}$$

However, since the intent of this exercise is to build systems that perform the best for a given task, it also follows that it would be desired to evaluate system performance not for the entire sensor space, but for the *scene space* instead.[74] We will denote that space $\mathcal{Y}$ to match its computational representation, $\underline{\underline{\mathbf{Y}}}$. We can then define a transformation,

$$\underline{\underline{\mathbf{T}}} = \underline{\underline{\mathbf{Y}}} \backslash \underline{\underline{\mathbf{V}}}_{s,4ij}, \tag{6.50}$$

which can be used to combine N measurements into the vectors approaching $\mathcal{Y}$ or estimating $\underline{\underline{\mathbf{Y}}}$,

$$\hat{\underline{\underline{\mathbf{Y}}}} = \underline{\underline{\mathbf{T}}} \underline{\underline{\mathbf{V}}}^{\mathrm{T}}_{s,4ij}. \tag{6.51}$$

Note that while $\underline{\underline{\mathbf{P}}}_{\underline{\underline{\mathbf{V}}}_{s,4ij}}$ represents the noise magnitude for the reconstructable vectors represented by $\underline{\underline{\mathbf{V}}}_{s,4ij}$, it would be incorrect to use these absolute magnitudes to map noise from reconstructables to the desired channels. Instead, the noise characteristics contained within $\underline{\underline{\mathbf{L}}}_{\underline{\underline{\mathbf{V}}}_{s,4ij}}$ need to be similarly mapped:

$$\underline{\underline{\mathbf{L}}}_{\underline{\underline{\mathbf{Y}}}} = \underline{\underline{\mathbf{T}}}\,\underline{\underline{\mathbf{L}}}_{\underline{\underline{\mathbf{V}}}_{s,4ij}} = \begin{bmatrix} \underline{\ell}^{\mathrm{T}}_{\underline{\mathbf{y}}_1} & \underline{\ell}^{\mathrm{T}}_{\underline{\mathbf{y}}_2} & \cdots & \underline{\ell}^{\mathrm{T}}_{\underline{\mathbf{y}}_K} \end{bmatrix}^{\mathrm{T}}, \tag{6.52}$$

where K is the total number of vectors in $\underline{\underline{\mathbf{Y}}}$. The resulting noise magnitudes within those vectors can be evaluated in a philosophically equivalent way,

$$p_{\underline{\mathbf{y}}} = ||\underline{\ell}_{\underline{\mathbf{y}}}||_2, \tag{6.53}$$

and then can be combined into a total magnitude vector

$$\underline{\underline{\mathbf{P}}}_{\underline{\underline{\mathbf{Y}}}} = \begin{bmatrix} p_{\underline{\mathbf{y}}_1} & p_{\underline{\mathbf{y}}_2} & \cdots & p_{\underline{\mathbf{y}}_K} \end{bmatrix}^{\mathrm{T}}. \tag{6.54}$$

Space Coverage

To properly evaluate a given partial system, it is important to know not only the system's noise resilience, but also the closeness of the sensor space to the scene space, which can be described by K ordered canonical angles $\zeta_1 \leq \zeta_2 \leq \ldots \leq \zeta_K$.[75] The first canonical angle ζ_1 is

$$\zeta_1 = \cos^{-1}\left(\min_{\hat{\underline{\mathbf{v}}}_1 \in \mathcal{V}, \hat{\underline{\mathbf{y}}}_1 \in \mathcal{Y}} \left(\hat{\underline{\mathbf{v}}}_1 \cdot \hat{\underline{\mathbf{y}}}_1\right)\right). \tag{6.55}$$

Subsequent canonical angles are computed by evaluating Eq. 6.55 with the portions of subspace $\mathcal{V}$ remaining after the elimination of $\hat{\mathbf{v}}_1$. The best case scenario is when $\zeta_K = 0$, which means $\mathcal{Y} \subset \mathcal{V}$, and the pMMP spans the desired channels.

While Eq. 6.55 provides an intuitive interpretation of the canonical angles, there are more efficient ways of computing the angles. We form the auxiliary matrix

$$\underline{\underline{\mathbf{X}}} = \underline{\underline{\mathbf{Y}}} - \underline{\underline{\mathbf{V}}}(\underline{\underline{\mathbf{V}}}^{\mathrm{T}}\underline{\underline{\mathbf{Y}}}) = \begin{bmatrix} \underline{\mathbf{x}}_1 & \underline{\mathbf{x}}_2 & \cdots & \underline{\mathbf{x}}_K \end{bmatrix}. \tag{6.56}$$

and compute the canonical angles from the singular values of this auxiliary matrix as

$$\zeta_k = \arcsin(\sigma_{\underline{\mathbf{x}}_k}). \tag{6.57}$$

6.4 Example of pMMP Optimization

To find the best pMMP design for a given task, we need to optimize for both noise resilience and space coverage. Because those properties are not inherently guaranteed to have overlapping minimums, we are invariably bound for the solution to be a point on the Pareto surface of a multi-objective optimization problem. We have found that the following metric

$$\arg\min_{\vec{\xi}} \left[\underbrace{\sum_{k=1}^{K} (\alpha_k p_{\underline{\mathbf{y}}_k})^2}_{\varepsilon_1} + w \underbrace{\sum_{k=1}^{K} (\beta_k \sigma_{\underline{\mathbf{x}}_k})^2}_{\varepsilon_2} \right] \tag{6.58}$$

successfully finds appropriate pMMP designs. The choice of w, $\{\alpha_k\}$ and $\{\beta_k\}$ provides handles to adjust the importance of all the various parameters, while the optimization variable vector $\vec{\xi}$ contains six values to construct three generating and three analyzing vector pairs. The first four variables define $\phi_{\underline{\mathbf{G}}_1}$, $\theta_{\underline{\mathbf{G}}_1}$, $\phi_{\underline{\mathbf{A}}_1}$ and $\theta_{\underline{\mathbf{A}}_1}$ to produce vectors $\underline{\mathbf{G}}_{1,\pm}$ and $\underline{\mathbf{A}}_{1,\pm}$, while the second two variables define ψ_G and ψ_A to prescribe where $\underline{\mathbf{G}}_{2,\pm}$, $\underline{\mathbf{A}}_{2,\pm}$, $\underline{\mathbf{G}}_{3,\pm}$ and $\underline{\mathbf{A}}_{3,\pm}$ reside on the orthogonal circles with respect to $\underline{\mathbf{G}}_{1,\pm}$ and $\underline{\mathbf{A}}_{1,\pm}$.

To illustrate the design of pMMPs, we consider the example presented in Hoover and Tyo.[19] Four different coupons of an ABS plastic material were exposed to different fluences of high energy laser energy, and the

resulting damaged samples had their monostatic Mueller matrices measured at a range of angles from $-20°$ to $20°$. Performing SVD of the data reveals that the most fundamental three measurement channels are

$$\underline{\underline{\mathbf{Y}}} = \begin{bmatrix} -0.9204 & 0.3097 & 0.2378 \\ -0.0347 & 0.0410 & -0.2480 \\ -0.0010 & 0.0034 & -0.0136 \\ -0.0003 & 0.0007 & 0.0088 \\ -0.0318 & 0.0524 & -0.2356 \\ -0.2757 & -0.4730 & -0.4418 \\ -0.0010 & -0.0043 & -0.0050 \\ -0.0004 & 0.0033 & -0.0085 \\ 0.0035 & -0.0039 & 0.0207 \\ 0.0013 & -0.0043 & 0.0138 \\ 0.2703 & 0.4860 & 0.3996 \\ -0.0019 & -0.0033 & -0.0220 \\ -0.0004 & 0.0008 & -0.0008 \\ 0.0001 & 0.0023 & -0.0037 \\ 0.0028 & 0.0017 & 0.0292 \\ 0.0398 & 0.6630 & -0.6850 \end{bmatrix}. \tag{6.59}$$

Note that the original paper used covariance matrix principal component analysis that resulted in a different set of channels, which did not include m_{00} in any of the measurements. If we add $m00$ back, then the maximum canonical angle between the two spaces is $3.1011°$. The difference is small enough to be accounted for by the extra idealization step taken in Hoover and Tyo.[19]

We used MATLAB's built-in genetic algorithm routine together with Eq. 6.58 to optimize each of the $4ij$ pMMP designs with $\alpha_k = \beta_k = 1$ and $w = 25$. Note that there is nothing fundamental about our choice of w — we tried a number of different weights between 1 and 100, and found that for this data set the value of 25 provided a good solution where the space coverage penalty was just significant enough for the reconstruction of relevant information to be prioritized over the noise resilience.

Table 4 shows the system performances that we were able to find for each of the defined classes of polarimeters. We can point out that the space coverage seems to be marginally better for the 422 system than it is for the 432 or the 433, despite the latter two making more measurements and having a capacity only to expand the space coverage if the 422 design is used as the base. That, however, is purely an artifact of our choice of w, which leads to the optimizer finding a solution with slightly better noise resilience by sacrificing some space coverage. Practically, the designs should be evaluated on whether or not they can separate the different objects classes. To determine which of these pMMPs accomplish that, it is necessary to look at the object projections onto $\hat{\mathcal{Y}}$. We can capture this by looking at how the proximity of each of the 25 objects from each of the four types of objects to the nearby classes changes. Instead of comparing data points directly, we will instead piece-wise interpolate the comparison class and determine the separation for each object/class both in $\mathcal{Y}$ and $\hat{\mathcal{Y}}$:

$$d_{\alpha,\beta,\gamma,\delta} = \frac{|(\vec{r}_{\alpha,\gamma} - \vec{r}_{\beta,\delta}) \times (\vec{r}_{\alpha,\gamma} - \vec{r}_{\beta,\delta+1})|}{|\vec{r}_{\beta,\delta+1} - \vec{r}_{\beta,\delta}|}, \tag{6.60a}$$

$$\hat{d}_{\alpha,\beta,\gamma,\delta} = \frac{|(\hat{\vec{r}}_{\alpha,\gamma} - \hat{\vec{r}}_{\beta,\delta}) \times (\hat{\vec{r}}_{\alpha,\gamma} - \hat{\vec{r}}_{\beta,\delta+1})|}{|\hat{\vec{r}}_{\beta,\delta+1} - \hat{\vec{r}}_{\beta,\delta}|}, \tag{6.60b}$$

where α and β represent the object classes, γ represents one of the 25 points within class α, and δ represents one of the 24 line segments created for class β. We evaluate the geometric mean of the ratios of least separation,

$$h_{\alpha,\beta} = \left[\prod_{\gamma=1}^{25} \frac{\hat{d}_{\alpha,\beta,\gamma,\min}}{d_{\alpha,\beta,\gamma,\min}} \right]^{\frac{1}{25}}. \tag{6.61}$$

When $h_{\alpha,\beta} = 0$, classes α and β have collapsed to lie on top of each other, while when $h_{\alpha,\beta} = 1$, the separation between classes α and β has remained unchanged. In the case that $h_{\alpha,\beta} > 1$, the separation within the reconstruction is greater than the original separation. Although this presents a seemingly interesting scenario, this

Design	N	ε_1	ε_2	$\varepsilon_1 + w\varepsilon_2$	ζ_K	$h_{1,2}$	$h_{2,3}$	$h_{3,4}$	$h_{1,3}$	$h_{2,4}$	$h_{1,4}$
400	4	1.911	1.089	29.142	89.65°	0.0003	0.0006	0.0006	0.0001	0.0001	0.0000
410	5	2.546	0.717	20.475	56.90°	0.1747	0.2207	0.2698	0.3144	0.1922	0.0355
411	6	2.669	0.556	16.569	47.53°	0.1770	0.2417	0.4435	0.2232	0.4506	0.1438
420	6	3.520	0.469	15.247	43.18°	0.5305	0.4743	0.5594	0.2750	0.5191	0.5163
421	7	3.528	0.248	9.727	29.78°	0.5184	0.2960	0.8056	0.2087	0.8787	0.5441
422	8	3.967	0.002	4.011	2.32°	1.0852	0.9811	0.9652	0.9893	0.9922	0.9851
430	7	3.318	0.469	15.042	43.17°	0.5133	0.4793	0.5625	0.2705	0.5435	0.5151
431	8	3.493	0.249	9.713	29.84°	0.5155	0.3029	0.8132	0.1998	0.8583	0.5302
432	9	3.932	0.002	3.987	2.60°	1.0940	0.9921	0.9632	0.9893	0.9914	0.9869
433	10	3.897	0.002	3.949	2.57°	1.1676	0.9966	0.9637	0.9925	1.0005	0.9863

Table 4: Optimization results for the 10 pMMP system classes. The optimization targets, ε_1 and ε_2, are defined in Eq. 6.58, while ζ_K represents the largest canonical angle. The values for $h_{\alpha,\beta}$ are calculated via Eq. 6.61. Because of the way that the four classes are distributed in $\mathcal{Y}$, knowledge of $h_{1,2}$, $h_{2,3}$ and $h_{3,4}$ may suffice.

result is attributable to non-linearities introduced by the averaging of different space projections and would be compensated by another $h_{\alpha,\beta}$ elsewhere.

Examining the performance of each of the ten pMMPs optimized designs in Table 4 and Fig. 6.2, it becomes clear that the 422 system is the first design of the defined range of systems that accomplishes the task of matching the space coverage and thereby separating the object projections adequately for object detection.

There is room to make the optimization routine more elaborate. For example, instead of matching the scene and sensor spaces of any given pMMP class, object identification can be done in the measurement space itself. This would require constructing a manifold as a model for the object distribution in the N-dimensional space, applying the proper noise model and looking at the separability of the classes within the measurement space. Performing all of this in each of the optimization instantiations is computationally intensive, as well as outside the scope of this development. A separate discussion is warranted to address that level of optimization properly.

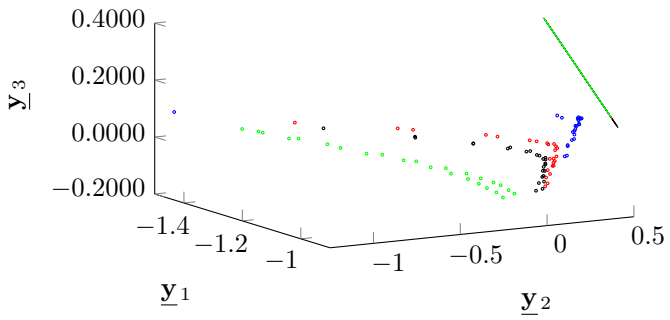
6.5 Conclusions

Mueller matrix polarimeters have demonstrated utility recently to assist in target identification, and the use of partial Mueller matrix polarimeters provides a way to develop a sensor that measures the polarization featured needed for a particular detection or classification task without having to measure the full Mueller matrix. Previous designs of pMMPs have been ad hoc, in that the polarimeters were developed by hand. In some instances, there was no real attention paid to whether or not the pMMP was even physically realizable.

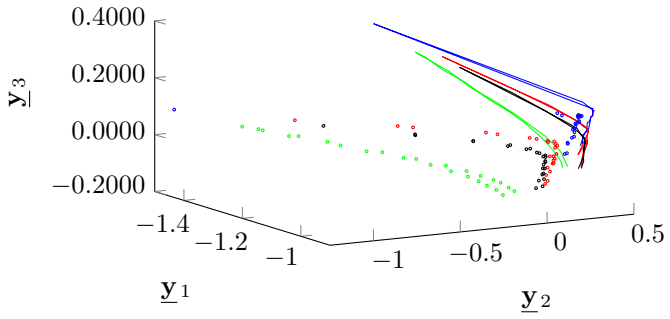
We have developed a theory of pMMPs that enables the structure of a pMMP to be determined from the actual generator/analyzer pairs used to form its instrument matrix $\underline{\underline{\mathbf{W}}}$. By proper analysis of $\underline{\underline{\mathbf{W}}}$, it is possible to determine the portion of Mueller matrix space that a particular pMMP measures. We developed metrics of optimality for pMMPs that are based on balancing their SNR performance with their closeness to the particular scene space at hand. The performance of this optimization method was demonstrated for a case previously presented in the literature.[19]

7. HYBRID DOMAIN MODULATED POLARIMETERS

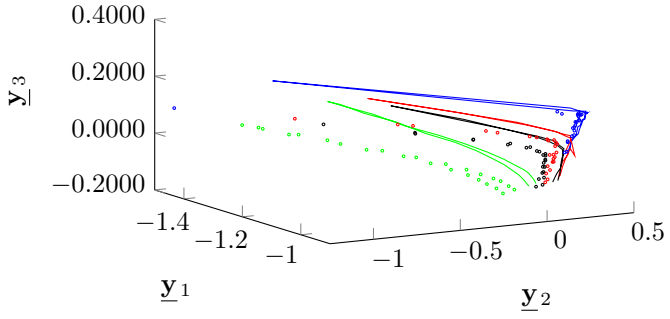
Channeled Mueller matrix polarimeters and the concept of using these channels was first introduced by Azzam.[76] Azzam published a very specific case, 1) a specific temporal framework was analyzed, 2) an implicit assumption about the object was made, *the object had no temporal bandwidth*, i.e., the object was stationary in time. Oka, Sabatke, Derniak, Kudenov, and Hagen then demonstrated both spectrally channeled and spatially (over spectrum) channeled systems,[40,41,45,77–80] mostly Stokes polarimeters. Dubreuil *et al*[81] then presented a spectrally channeled Mueller matrix polarimeter, which of course is non-imaging since the focal plane array is used to resolve the spectrum. LaCasse, Chipman, Tyo, and LeMaster and Hirakawa[9,10,82] then described micropolarizer array partial Stokes polarimeters as channeled systems, and LaCasse *et al* presented a spatio-temporally modulated



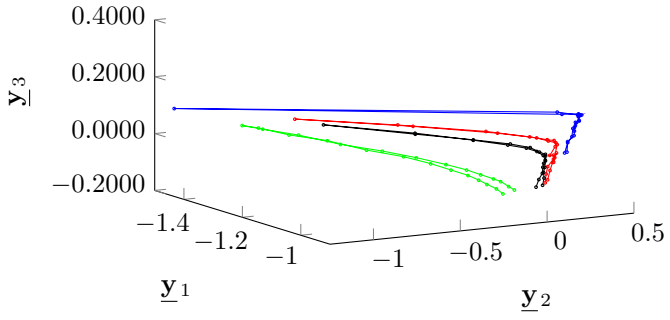
(a) 400 pMMP, $N = 4$



(b) 411 pMMP, $N = 6$



(c) 421 pMMP, $N = 7$



(d) 422 pMMP, $N = 8$

Figure 6.2: Space coverage of various optimized pMMP designs. Circles represent perfect Mueller matrix object projection, while lines are the approximation that each pMMP achieves. Note that the 400 pMMP collapses all measurements onto a single line, while the space coverage for 432 and 433 are virtually identical to that of 422, coupled with slightly improved noise resilience.

hybrid channeled Stokes system,[7] and subsequently both Myhre *et al*[83] and Zhao *et al*[84] presented spatially modulated full Stokes polarimeters. Finally Alenin and Tyo[17] formalized a general framework which describes channeled polarimeters almost completely, both Mueller and Stokes.

Prior to the work by LaCasse *et al*,[7,9,10] bandwidth in channeled polarimetric systems had not been addressed, or only addressed as a consequence of instrumental "error." Additionally, prior to Alenin and Tyo[17] channeled systems were designed in an *ad-hoc* manner. In this communication we address bandwidth using the *systematic* design tools introduced by Alenin and Tyo[17] for a hybrid spatio-temporally modulated channeled active polarimetric system.

7.1 Formalism and channels

Note that this section originally appeared in our other publication in this conference proceeding[24] and is derived/adapted from that section to address the topic of this communication for ease of reference. Portions may be reproduced verbatim, however quotes will not be used.

We use the Mueller-Stokes mathematical formalism here, as it is most commonly used in instrumental polarization and polarimeter design. This analysis is, however, agnostic to the formalism used, a coherence formalism[85] with periodic modulators could also be used and would have similar results. In the next sections, it should be kept in mind that modulations are done in some physical domain, they are periodic, i.e., a superposition of sinusoidal functions, and the "channels" are the resultant δ-functions which ensue from the Fourier transform of the sinusoidal modulations.

7.1.1 Modulated Mueller formalism

The Stokes parameters are described by

$$\mathbf{s} = \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} \propto \begin{bmatrix} \langle |E_x|^2 \rangle + \langle |E_y|^2 \rangle \\ \langle |E_x|^2 \rangle - \langle |E_y|^2 \rangle \\ 2\Re\langle E_x E_y^* \rangle \\ 2\Im\langle E_x E_y^* \rangle \end{bmatrix} = \begin{bmatrix} \bullet \\ \updownarrow \\ \nwarrow\!\!\searrow \\ \circlearrowright \end{bmatrix}, \text{ where } s_0 > 0, \;\; s_0^2 \geq s_1^2 + s_2^2 + s_3^2 \tag{7.1}$$

where $\langle \cdot \rangle$ denotes the time average, s_0 is proportional to the total irradiance, s_1 is proportional to the prevalence of horizontal ($0°$) over vertical ($90°$) polarization, s_2 is proportional to the prevalence of $+45°$ over $-45°$ polarization, and s_3 is proportional to the prevalence of right circular over left circular polarization.[86] Because optical sensors measure a quantity proportional to the time averaged Poynting vector, the phase information is lost, and only the incoherent time averaged polarization information can be obtained.[85,86]

For materials which can be described via linear optical interactions, we can use the Mueller-Stokes formalism. A Mueller matrix, $\underline{\underline{\mathbf{M}}}$, is a matrix which linear transforms one set of Stokes parameters, $\mathbf{s}_{\text{in}}$, into another set of Stokes parameters, $\mathbf{s}_{\text{out}}$:

$$\mathbf{s}_{\text{out}} = \underline{\underline{\mathbf{M}}} \cdot \mathbf{s}_{\text{in}} \tag{7.2}$$

Notice that $\underline{\underline{\mathbf{M}}} \in \mathbb{R}^{4\times4}$ but not every 4×4 real valued matrix is a Mueller matrix due to the constraints in Eqn.7.1, see Gil[87] for details.

With an active, or Mueller matrix, polarimeteric instrument, we must modulate in irradiance to infer the Mueller matrix of an object, $\underline{\underline{\mathbf{M}}}(\mathbf{x})$, where $\mathbf{x} = \begin{bmatrix} x & y & z & t & \sigma \end{bmatrix}^T$. We can then rewrite Eqn. 7.2 to have Mueller matrices and Stokes parameters be functions of space, time, and wavelength or wavenumber. Eqn.7.2 then becomes

$$\begin{bmatrix} s_{0,\text{out}}(\mathbf{x}) \\ s_{1,\text{out}}(\mathbf{x}) \\ s_{2,\text{out}}(\mathbf{x}) \\ s_{3,\text{out}}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} m_{00}(\mathbf{x}) & m_{01}(\mathbf{x}) & m_{02}(\mathbf{x}) & m_{03}(\mathbf{x}) \\ m_{10}(\mathbf{x}) & m_{11}(\mathbf{x}) & m_{12}(\mathbf{x}) & m_{13}(\mathbf{x}) \\ m_{20}(\mathbf{x}) & m_{21}(\mathbf{x}) & m_{22}(\mathbf{x}) & m_{23}(\mathbf{x}) \\ m_{30}(\mathbf{x}) & m_{31}(\mathbf{x}) & m_{32}(\mathbf{x}) & m_{33}(\mathbf{x}) \end{bmatrix} \cdot \begin{bmatrix} s_{0,\text{in}} \\ s_{1,\text{in}} \\ s_{2,\text{in}} \\ s_{3,\text{in}} \end{bmatrix} \tag{7.3}$$

where for simplicity we fix $\mathbf{s}_{\text{in}}$. Our detector then measures a quantity proportional to $s_{0,\text{out}}(\mathbf{x})$. For a Mueller matrix measuring instrument, we have an unknown object Mueller matrix, $\underline{\underline{\mathbf{M}}}_{\text{obj}}(\mathbf{x})$, and we write down the instrument equation which modulates Stokes parameters:[31]

$$\mathbf{s}_{\text{out}}(\mathbf{x}) = \underline{\underline{\mathbf{A}}}(\mathbf{x}) \cdot \underline{\underline{\mathbf{M}}}_{\text{obj}}(\mathbf{x}) \cdot \underline{\mathbf{G}}(\mathbf{x}) \cdot \mathbf{s}_{\text{in}} \tag{7.4}$$

$$= \underline{\underline{\mathbf{A}}}(\mathbf{x}) \cdot \underline{\underline{\mathbf{M}}}_{\text{obj}}(\mathbf{x}) \cdot \mathbf{s}_{\underline{\mathbf{G}}}(\mathbf{x}) \tag{7.5}$$

where $\underline{\mathbf{G}}(\mathbf{x}), \underline{\underline{\mathbf{A}}}(\mathbf{x})$ are the generator and analyzer Mueller matrices respectively, known and modulated via the physical instrument. The generator modulation can then be thought of as only a Stokes parameter modulation, $\mathbf{s}_{\underline{\mathbf{G}}}(\mathbf{x})$.

7.1.2 Channels

Eqn.7.5 can be expanded to obtain a linear equation[86] for $s_{0,\text{out}}(\mathbf{x})$.

$$\boxed{s_{0,\text{out}}(\mathbf{x}) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{0i}(\mathbf{x}) s_j(\mathbf{x}) m_{ij}(\mathbf{x})} \tag{7.6}$$

where $a_{0i}(\mathbf{x})$ are the elements of the first row of $\underline{\underline{\mathbf{A}}}(\mathbf{x})$, $s_j(\mathbf{x})$ are elements of $\mathbf{s}_{\underline{\mathbf{G}}}(\mathbf{x})$, and $m_{ij}(\mathbf{x})$ are elements of $\underline{\underline{\mathbf{M}}}_{\text{obj}}(\mathbf{x})$. We can then take the Fourier transform of $s_{0,\text{out}}(\mathbf{x})$ to obtain

$$\boxed{S_{0,\text{out}}(\rho) = \sum_{i=0}^{3}\sum_{j=0}^{3} A_{0i}(\rho) * S_j(\rho) * M_{ij}(\rho)} \tag{7.7}$$

Where $\mathbf{x} \rightarrow \rho$ in the Fourier transform, $*$ denotes convolution, and the shift to capital letters indicates a function has been Fourier transformed. If $a_{0i}(\mathbf{x})$ and $s_j(\mathbf{x})$ are superpositions of sinusoidal functions, then $A_{0i}(\rho) * S_j(\rho)$ is a set of δ-functions, and each $M_{ij}(\rho)$ is then convolved with each δ-function in the set. The complete set of δ-functions for the system

$$\sum_{i=0}^{3}\sum_{j=0}^{3} A_{0i}(\rho) * S_j(\rho) \tag{7.8}$$

are defined as the *channels* of the system, or the system's *channel structure*.[17]

7.1.3 System equation

All examples in this communication will assume a quad-retarder + micropolarizer array Mueller matrix polarimeter system. This results in a spatio-temporally modulated system channel structure. The details of the design and an actual instrument implementation are in Vaughn *et al*[24] and the system equation is reproduced below:

$$\mathbf{s}_{\text{out}} = \underline{\underline{\mathbf{P}}}(x,y) \cdot \underline{\underline{\mathbf{R}}}(\nu_4,\epsilon_4,\delta_4) \cdot \underline{\underline{\mathbf{R}}}(\nu_3,\epsilon_3,\delta_3) \cdot \underline{\underline{\mathbf{M}}}_{\text{obj}}(x,y,t) \cdot \underline{\underline{\mathbf{R}}}(\nu_2,\epsilon_2,\delta_2) \cdot \underline{\underline{\mathbf{R}}}(\nu_1,\epsilon_1,\delta_1)\mathbf{s}_{\text{in}} \tag{7.9}$$

where

$$\underline{\underline{\mathbf{P}}}(x,y) = \text{micropolarizer array Mueller matrix} \tag{7.10}$$

$$\underline{\underline{\mathbf{R}}}(\nu_j,\epsilon_j,\delta_j) = \text{retarder Mueller matrix} \tag{7.11}$$

$$\nu_j = \text{retarder frequency in } 2\pi\frac{radians}{s} \tag{7.12}$$

$$\epsilon_j = \text{retarder start position in } 2\pi\, radians \tag{7.13}$$

$$\delta_j = \text{retarder retardance in } radians \tag{7.14}$$

$$\underline{\underline{\mathbf{M}}}_{\text{obj}}(x,y,t) = \text{Mueller matrix of the object} \tag{7.15}$$

Where of course only the final irradiance value, proportional to the $s_{0,\text{out}}(\mathbf{x})$ element of $\mathbf{s}_{\text{out}}(\mathbf{x})$ is sampled. The channels are encoded in $s_{0,\text{out}}(\mathbf{x})$.

7.2 Channel design

Designing polarimetric instruments using a channeled framework is somewhat new in the field, especially for instruments which are not spectrally modulated. Spectral instrument designers have, however, utilized channel design in an *iterative* way, tweaking the system designs and then observing the channels which result, then again tweaking the system design. Here, we specify constraints and needs, and then optimize the system *directly in the channel space* for some cost function dependent on those specifications. This design paradigm allows for faster and (sometimes) conceptually simpler system configuration.

We will only address spatio-temporally modulated systems here, and we assume modulations are of the type

$$f(x, y, t) = h(x, y) \cdot g(t), \tag{7.16}$$

i.e., that any modulation is mathematically separable between time and space. Non-separable modulations can be constructed (envision a rotating focal plane array with a micropolarizer, or a spatial light modulator modulating in time and spatially), but we will address these in a future publication. A separable system only allows for bandwidth improvements from channel cancellations or combinations, while a non-separable system may allow for improvements from rotations of the channel structure, however the latter remains an open question. We emphasize once again that the channels are δ-functions in the Fourier domain which result from *sinusoidal modulations of irradiance in the physical domain*, i.e. space, time, wavelength, etc.

7.2.1 Notation

Visualization of channel structures can be accomplished by graphing the sets of channels, or δ-functions over the Fourier domain dual to the physical modulation domain. Another systematic way of graphing the modulations is the *frequency phase matrix* (FPM), introduced by Alenin and Tyo,[17] which is a book keeping method for the channel splitting behavior, the signs, and real and imaginary components of the δ-functions. In this communication we will graph the channels as they actually appear in the channel space to reinforce intuition and understanding.

	positive	negative
real	△	▽
imag.	▷	◁

Table 5: Notation for channels.

A δ-function can be characterized by its position, and its complex magnitude. Table 5 outlines the graphical notation that we will use, 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 size of each triangle represents the absolute value of the real part or the imaginary part. Note that we will only show the channels for a single Mueller matrix element for each visualization, with the channels for all other Mueller matrices represented by light gray-blue circles. Keep in mind that many channels, for each Mueller matrix element, end up being added together at the locations shown. Again, the relationship between the channels of different Mueller matrices is *additive*. An example of the channel structure for a spatio-temporally modulated system for m_{23} is shown in Fig.7.1.

Some of the examples presented here will be normalized to a temporal frequency range of $[-1, 1]$, this is because for instrument design only the relative frequencies are important, we are always limited by some maximum sampling rate in practice, so relative bandwidth with respect to a maximum absolute frequency of 1 is what must be optimized for. 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.

7.2.2 Optimization

Once some spatio-temporal modulation scheme is selected, and the free parameters of that scheme are known, an optimization over the parameters for some specific cost function can be carried out. Alenin and Tyo[17,88] have used cost functions which optimize the spectral channel structure for noise performance, but other cost functions may be used. In this communication, we jointly optimize for bandwidth and noise using the following cost function:

$$\mathcal{O}(\mathbf{p}) = \frac{[CN(\mathbf{p})]^n}{\text{dist}(\mathbf{p})}. \tag{7.17}$$

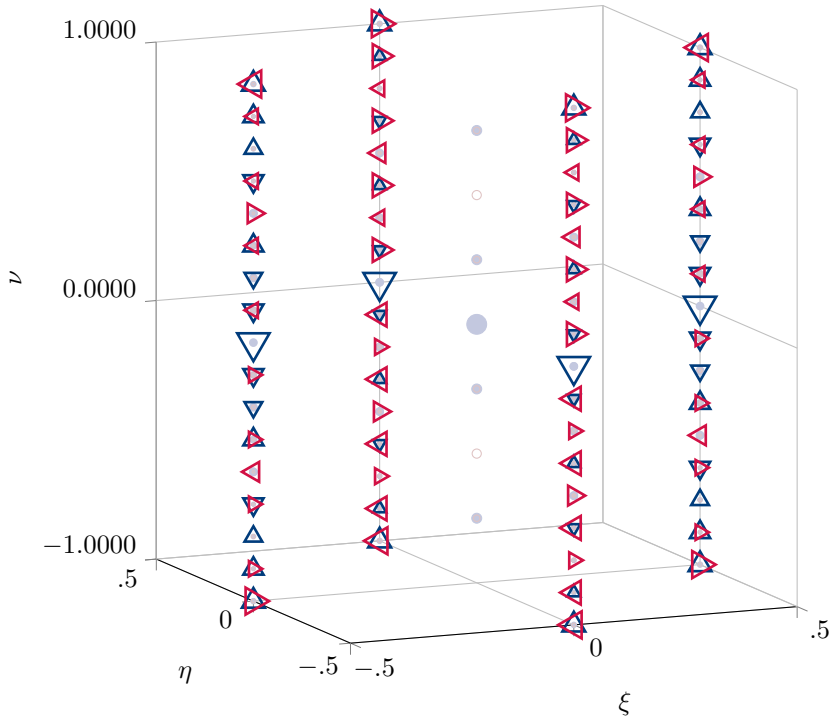


Figure 7.1: Example of a spatio-temporal channel structure with δ-functions specific to m_{23}. The maximum bandwidth corresponds to the minimum distance between two adjacent channels, taken over all possible adjacent channel pairs.

where $\mathbf{p}$ is a vector of free parameters which control the channel structure, the dist($\mathbf{p}$) function quantifies the distance between channels, i.e., bandwidth, $CN(\mathbf{p})$ is the condition number of the "mixing matrix" $\underline{\underline{\mathbf{Q}}}$ (details about $\underline{\underline{\mathbf{Q}}}$ are in Alenin and Tyo[17]), and $n \in \mathbb{R}^+$ is a weighting parameter for condition number corresponding to noise optimization. This is valid since $CN \geq 1$ by definition. The optimization then minimizes $\mathcal{O}(\mathbf{p})$ over $\mathbf{p}$. Note that as we increase n, the optimization tends to favor noise performance over bandwidth. A system conditioning metric must be included in the cost function to ensure reconstructablility of the full Mueller matrix, i.e., higher bandwidths can be found which result in some partial Mueller matrix polarimeter (pMMP) reconstruction, but the system conditioning would be infinite for the full Mueller matrix reconstruction. This fact has utility for pMMP designs, but will not be addressed here.

7.3 Bandwidth

The treatment of bandwidth for channeled systems is mature and well known in information and communications theory.[89–91] 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. 3) has mostly been addressed by Alenin and Tyo.[17] 2) is a complicated subject and we will not delve into details here, but we emphasize again that these constraints must be enforced when optimizing for some cost function.

We will discuss a spatio-temporal system, with modulation in the domain

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ t \end{bmatrix} \tag{7.18}$$

and examples will be from a quad-retarder + micropolarizer array system.[24] First we will briefly review convolution. Given some unknown quantity, $m(t)$, this quantity can be modulated with a sinusoidal function. Without

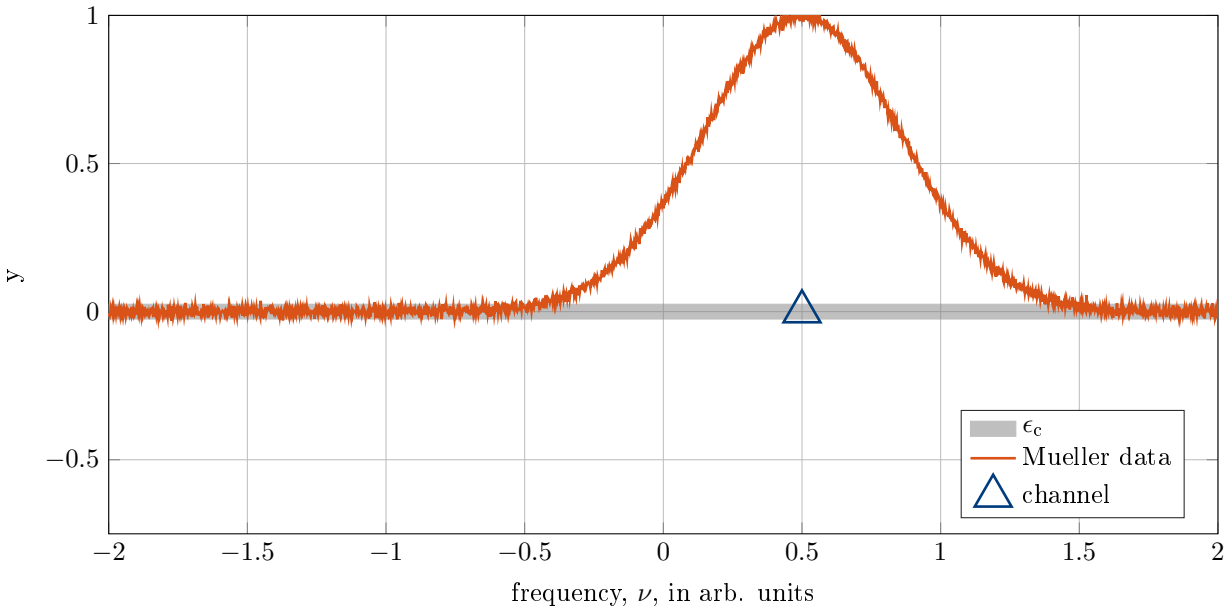


Figure 7.2: An example of convolution of data with a channel. The gray band represents the range of ϵ_c, resulting in data being outside of the ϵ_c range for $-0.5 < \nu < 1.5$. This implies a bandwidth of 2 arb. units for the Mueller data.

loss of generality, we choose cosine here:

$$f(t) = \cos 2\pi\nu_0 t \cdot m(t) \tag{7.19}$$

$$\implies \mathcal{F}\{f(t)\}_{t\to\nu} = \mathcal{F}\{\cos 2\pi\nu_0 t \cdot m(t)\}_{t\to\nu} \tag{7.20}$$

$$\implies F(\nu) = \mathcal{F}\{\cos 2\pi\nu_0 t\}_{t\to\nu} * \mathcal{F}\{m(t)\}_{t\to\nu} \tag{7.21}$$

$$\implies F(\nu) = \frac{1}{2}\Big[\delta(\nu-\nu_0) + \delta(\nu+\nu_0)\Big] * M(\nu) \tag{7.22}$$

$$\implies F(\nu) = \frac{M(\nu-\nu_0)}{2} + \frac{M(\nu+\nu_0)}{2} \tag{7.23}$$

where $\mathcal{F}\{\cdot\}$ is the Fourier transform, $*$ is convolution, and $\delta(\nu)$ is the Dirac delta function. The last line is due to property of convolution with delta functions. This gives us some tools for conceptual descriptions for the forward problem and hence the bandwidth. There will exist a set of channels (δ-functions) for each Mueller matrix element $M_{ij}(\rho)$ in the Fourier domain. For each channel in that set, $M_{ij}(\rho)$ will be copied at that channel's location with $M_{ij}(\mathbf{0})$ being located precisely where the delta function is located. We can then define the bandwidth of M_{ij} for some threshold $\epsilon_c \geq 0$ as the values of ρ where $|M_{ij}(\rho)| > \epsilon_c$. Fig.7.2 clarifies these concepts.

The polarimetric system channel structure contains constraints on bandwidth because there is a finite distance between channels as shown in Fig.7.1. 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.

7.3.1 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 Fig.7.3.

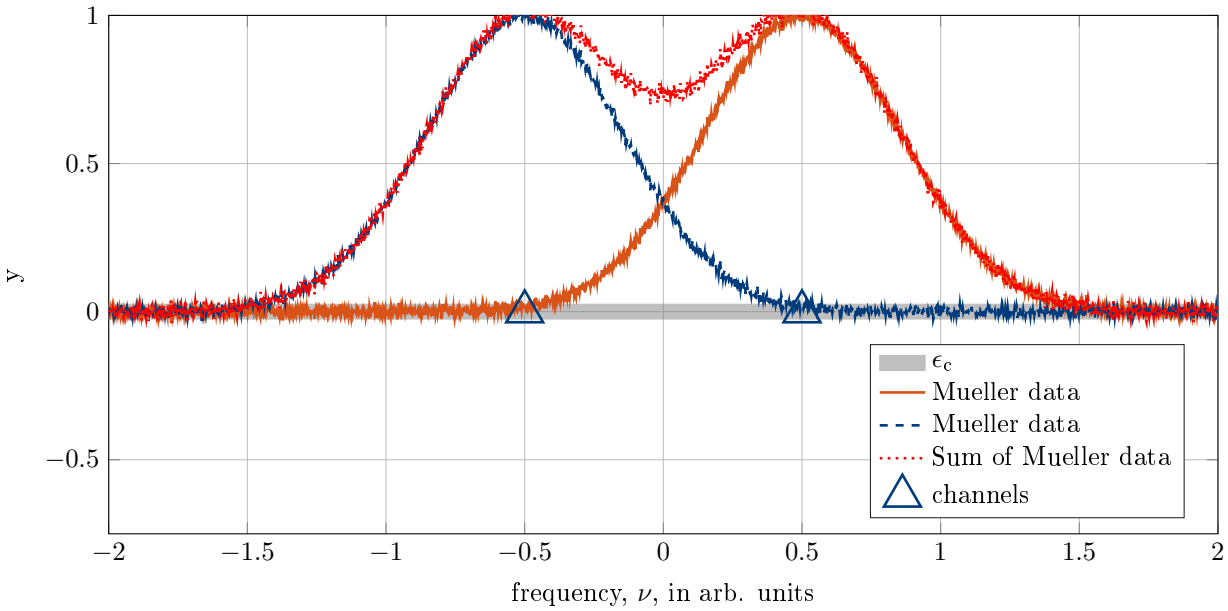


Figure 7.3: An example of channel crosstalk. Mueller data is placed at two channels, with the distance (bandwidth) between them less than the bandwidth of the Mueller data. When added, the Mueller data from different channels adds together, leaving no remedy to differentiate data between channels in the region of bandwidth crossover when given arbitrary Mueller data.

7.3.2 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 essentially reduces the 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 depth here; the literature on filtering is vast and mature in control theory and electrical engineering. Keep in mind, however, that if the statistics of the objects being measured are known, then optimal filters can be designed.[49]

7.3.3 Maximizing bandwidth

Our focus will be on increasing relative bandwidth to reduce crosstalk, and subsequently increasing the system resolution or speed for spatio-temporally modulated 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. We also don't have instruments with arbitrary measurement bandwidth. In order to constrain the bandwidth maximization to relative frequencies, we can normalize all of the channels to be contained in a cube (or rectangular prism in certain cases) where the maximum frequency is normalized to be some fixed value. We can choose different norms to accomplish this as long as we are 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 channels are also normalized by this same length, or 2) normalize in an ∞-norm way, that is normalize each frequency domain coordinate by the respective maximum frequency channel domain coordinate. This is what we do for our system examples here. To clarify with an example,

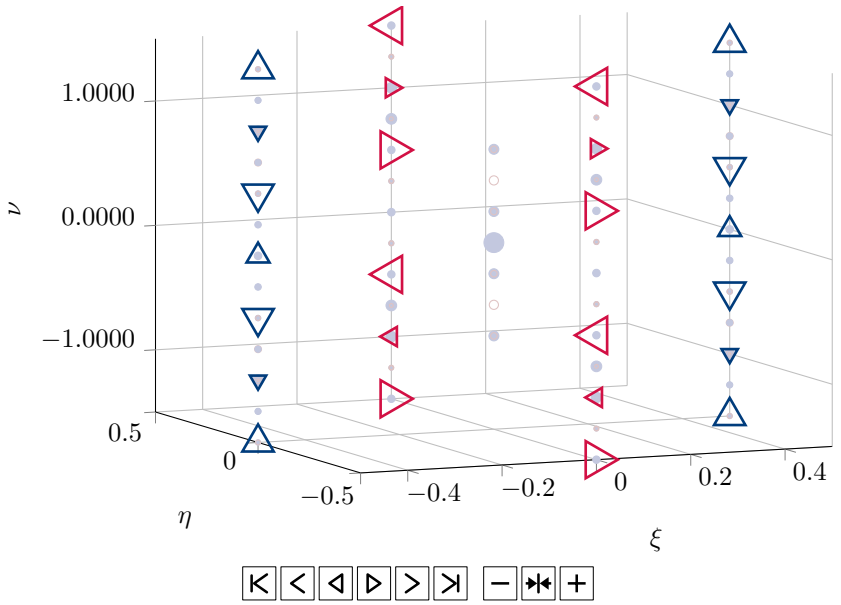


Figure 7.4: An example of varying frequency parameters to move channels around in the channel space. Note the cancellation/addition at certain parameter values which opens up more bandwidth between the channels. Animated in the electronic version, use the controls displayed to view.

suppose that our maximum frequency channel is located at $[0.5, 0.5, 60]^T$, then the normalization factors would be

$$n_{2\text{-norm}} = \sqrt{0.5^2 + 0.5^2 + 60^2} \tag{7.24}$$

$$n_{\infty\text{-norm},0} = 0.5 \tag{7.25}$$
$$n_{\infty\text{-norm},1} = 0.5 \tag{7.26}$$
$$n_{\infty\text{-norm},2} = 60 \tag{7.27}$$
$$\tag{7.28}$$

and for an arbitrary channel located at $\mathbf{c}_{\text{arb}} = \begin{bmatrix}\xi_{\text{arb}} & \eta_{\text{arb}} & \nu_{\text{arb}}\end{bmatrix}^T$ the two normalizations would be

$$\mathbf{c}_{\text{arb,2-norm}} = \frac{\mathbf{c}_{\text{arb}}}{\sqrt{0.5^2 + 0.5^2 + 60^2}} \quad \text{and} \quad \mathbf{c}_{\text{arb},\infty\text{-norm}} = \begin{bmatrix} \frac{\xi_{\text{arb}}}{0.5} \\ \\ \frac{\eta_{\text{arb}}}{0.5} \\ \\ \frac{\nu_{\text{arb}}}{60} \end{bmatrix} \tag{7.29}$$

In the examples here, we only normalize the locations of frequency corresponding to the temporal domain, ν, because the examples assume a fixed micropolarizer array which cannot be changed.[24] This results in normalization of the spatial frequency coordinates having no effect on the analysis, 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 we are making an *oranges to oranges* comparison as channel location changes. Note that the use of different normalization types will lead to different optimization outcomes.

7.3.4 Channel cancellation

The next step is to attempt to maximize the relative channel bandwidth, now that the channels are all normalized to a maximum frequency. This maximization can typically be accomplished by optimizing over the system channel

structure's free parameters $\mathbf{p}$ as described in Eqn. 7.17. Typical parameters for a spatio-temporally modulated polarimeter (which is not spectrally modulated) include

- retardance and retarders, spatially or temporally modulated.
- spatial or temporal analyzer/diattenuator modulation.
- rotation or rotators; these elements are typically combined with one of the types listed above to achieve a modulation.

For a separable channel structure, only channel cancellation/combination or reduction of overall channels may be used to increase the relative bandwidth. Fig. 7.4 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 can then find an optimal channel structure, given your free parameters $\mathbf{p}$. An example of an optimal (to the best of our knowledge) channel structure for the quad retarder + micropolarizer array system is shown in Fig.7.5.

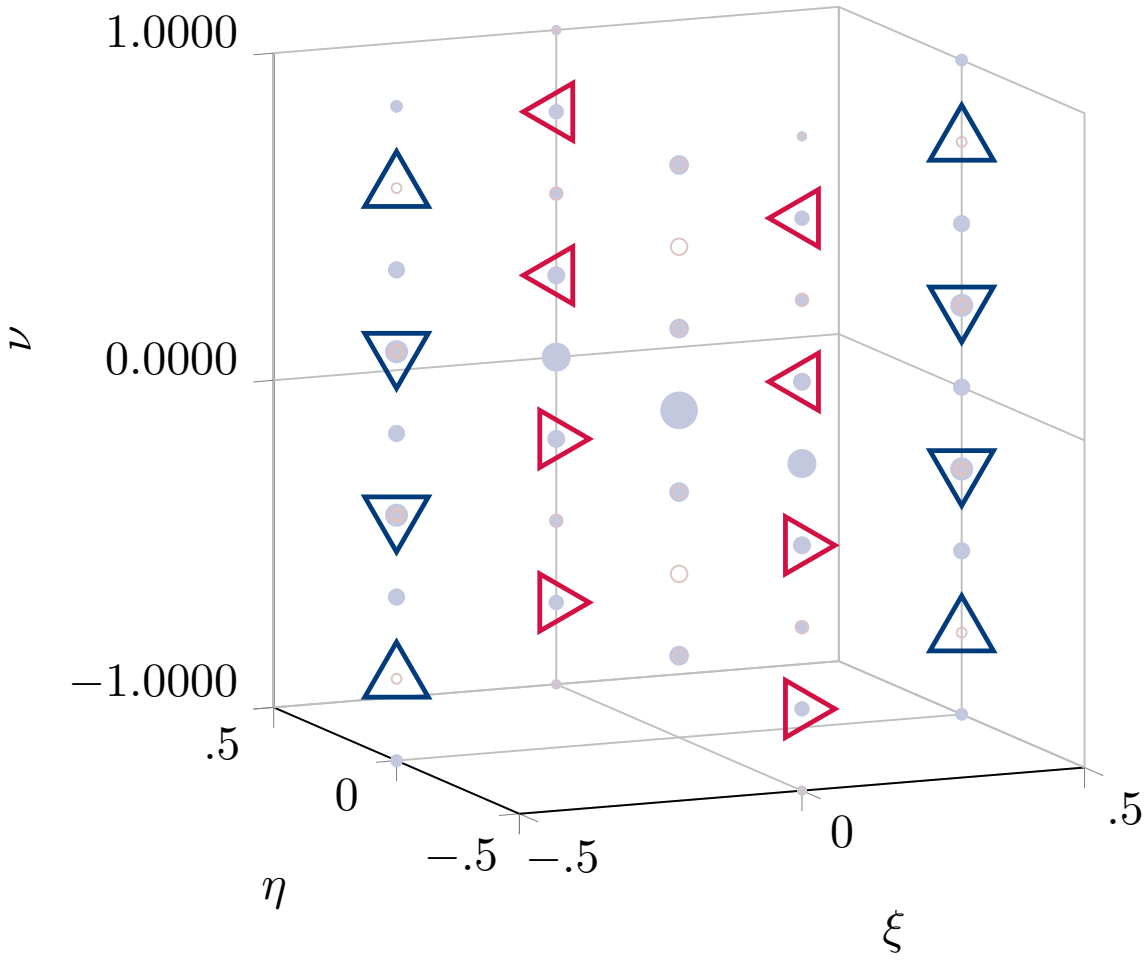


Figure 7.5: Optimal channel structure for a specific quad retarder + micropolarizer channel structure for m_{23}. Reproduced from a figure in Vaughn *et al.*[24]

7.3.5 Discussion

Once the channeled system framework is understood, and the free parameters, $\mathbf{p}$ of a spatio-temporally modulated 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, defining a cost function and 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. We have designed a model for the specific case of a quad-retarder + micropolarizer array system, but we hope to adapt our current model to generate generic spatio-temporally modulated systems in the near future.

Additionally, if the statistics of an object or set of objects are known, then non-uniform bandwidth can be maximized. All of the examples shown in this communication 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.

7.4 Noise and sytematic error

The channel bandwidth optimization discussed in the previous section must be paired with the sensitivity of the system to deviations from the ideal optimized parameters, i.e., an actual system will have non-ideal modulations and modulator elements. We present some preliminary results on systematic deviations here, but we have not fully worked out how to measure the sensitivity of a system's channel structure to systematic deviations (often called systematic error in the literature). The sensitivity to random noise sources is also discussed here, but again we have not worked out a general systematic way to compute the sensitivity of channel structures to random noise when bandwidth is taken into account (including filtering effects). Noise effects from the inversion of the $\underline{\underline{\mathbf{Q}}}$ matrix, however, have been addressed.[17]

7.4.1 Systematic error

Systematic error will occur when the actual system deviates from the designed system, or when there is some consistent bias due to physical instrument details. Here systematic error for a channeled system refers to the differences between the real system's channels as compared with the ideal channels from some designed channel structure. Issues arise for a separable channeled system when channel cancellation/combination has been used as a tool to increase system bandwidth.

For separable channeled systems a potentially serious problem arises: at the locations in the frequency space where the channel cancellation(s) occurred, channels will again be present due to deviations of real components. These *spurious channels* will be convolved with Mueller data, and introduce channel crosstalk. We really only have one option available, to minimize the magnitude of these spurious channels so that the crosstalk into the adjacent channels is low. An example of is shown in Fig. 7.6, the small triangles represent the spurious channels.

We have found a way to reduce the effects of the spurious channels for the quad retarder + micropolarizer system, which was to re-optimize over our channels using the remaining free parameters available while fixing the ones constrained by the physical instrument components. In our quad-retarder + micropolarizer system example, once our actual retardances were fixed, we re-optimized using the starting position of each retarder and added a parameter to our 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.

In fact, an iterative approach would likely yield the best results with a single component at a time being sourced, measured and characterized, then the system can be re-optimized to reset the other components specifications to minimize the crosstalk, and then the process is repeated.[8]

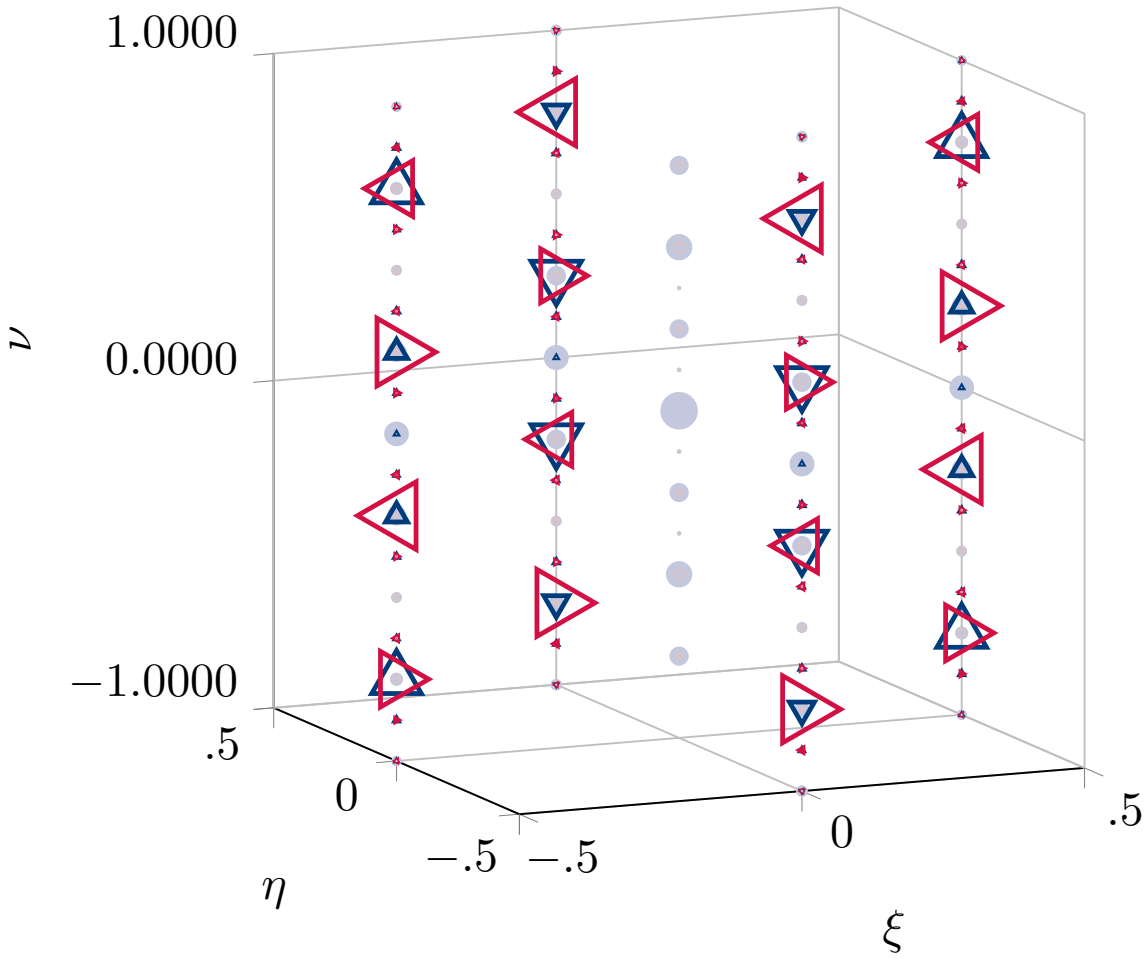


Figure 7.6: An example of systematic error in a real instrument, the small triangles represent channels which are present due to retardance deviation from the specifications.

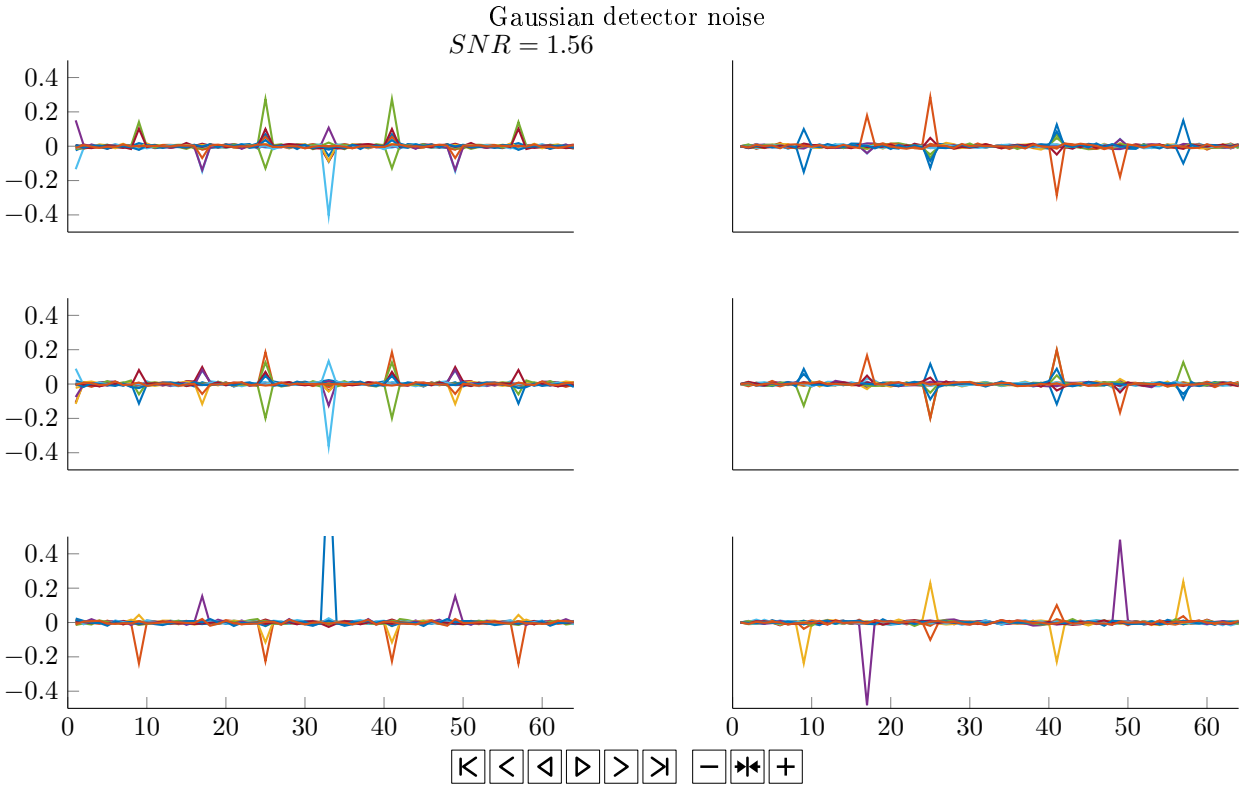


Figure 7.7: Gaussian detector noise effects on channel structure. Note that these are the same channels as shown in Fig. 7.6, but plotted along the ν lines. The left column is the real part, the right column is the imaginary part, there are 16 colors in each graph representing each Mueller matrix element. The first row is ν at $\xi = 0, \eta = -0.5$, the second row is ν at $\xi = -0.5, \eta = 0$ and the third row is ν at $\xi = 0, \eta = 0$. Animated in the electronic version.

7.4.2 Noise

The other impacts to real channeled systems are random noise, for instance detector noise. The results presented here are preliminary, and we only modeled detector Johnson (Gaussian like) noise for the detector. The results are simulated, but show that the channel structure itself is quite robust to detector noise. Here we switch to a different view of the channels to better view the noise effects. Each line of channels in the ν direction is plotted in Fig. 7.7 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. This result will need to be validated on the real instrument, and we plan on testing this in the near future.

We have not modeled the effects of noise on the reconstruction for our specific system yet, but noise effects have been addressed for channeled systems by others.[17, 32, 88]

7.5 Conclusion

Channeled polarimeter design has changed the way that instrument designers approach the design process, and allows engineers to systematically design both general and task specific polarimeters. We have presented examples and some general insight into using channeled system design for spatio-temporally modulated Mueller matrix polarimeters. We also address some of the intricacies of channeled design, and give some preliminary examples of possible systematic errors and noise effects on channel structures. Furthermore, we give an example of how to mitigate some of the systematic errors, and in the future will use a different approach when building a real instrument by re-optimizing the channel structure after each instrument component is sourced. In the future we hope to build a fully generic channeled system model for spatio-temporally modulated instruments, so that the

community can design their own instruments in a systematic way. We also need to validate our systematic error mitigation on our real instrument, and validate the effects of noise on the channel structure of a real instrument.

8. COHERENCE MANIPULATION

8.1 Polarization and its Intrinsic Relation to Coherence

Current design philosophy treats conventional imaging, polarimetry, spectrometry, depth imaging, etc., as separate sensing architectures. However, they are all fundamentally linked through the statistical description of the optical field embodied in the mutual coherence function at the sensor aperture. Once the optical field passes the system collection aperture, the information available to the observer has been fully determined; were it possible to know the full spatiotemporal distribution of the field on the aperture, then the output of any conventional sensor could be computed. Unfortunately, for a wide class of sensing systems that rely on incoherent or partially coherent radiation, this complete knowledge is impossible to obtain at optical wavelengths for any meaningful aperture size.[92] Since current optical detectors measure second moments of the optical field, the best the designer can hope to accomplish is to measure some meaningful projection of the field's mutual coherence function across the aperture.

High resolution imagers, spectrometers, light field imagers, polarimeters, and other current systems use field transducers (lenses, gratings, polarizers, etc.) to choose the projection of mutual coherence to measure. Lens-based imaging systems convert spatial Fourier modes on the aperture into irradiance measurements in the focal plane; spectrometer systems convert temporal Fourier modes into irradiance measurements on the detector plane; extended depth-of-field imagers use aperture coding to control the projections of the light field; and various field modulation strategies allow spectral, polarization, and depth information to be demodulated from channeled intensity data. Recently several advanced computational and compressive sensing strategies that are more closely tied to the coherence properties have emerged.

A key aspect of our research program is study of the theoretical fabric that unifies these apparently diverse sensing strategies through the theory of mutual coherence. Little is known about how to efficiently measure specific projections of the field coherence, and even less is known about how to control the mutual coherence properties of illumination systems. While classical statistical optics theory describes how the coherence properties propagate through space, the tools do not currently exist to describe the transformation of coherence upon interactions after scattering.

Variable coherence tomography (VCT) was developed as a method for directly measuring the second-order statistical properties of quasi-homogeneous media.[93,94] VCT uses an incident wave field with a specific mutual coherence structure that allows the average coherent two-point scattering from the medium to be probed. VCT develops a speckle field that is uncorrelated *except* at specific lateral offsets. The spatial correlation structure of the incident beam can be shaped by changing the parameters of the source, allowing a tomographic reconstruction of the spatial correlation of the scattering potential. As originally developed, VCT is based on scalar scattering by weak, quasi-homogeneous, isotropic scattering media. However, the underlying theory is equally valid for isotropic and anisotropic scatterers.[95] Variable coherence polarimetry (VCPol) was introduced by our group as an extension of the theory behind VCT to include polarimetric information.[16] VCPol makes it possible to probe the second-order statistics of the scattering object from monostatic measurements, which provides greater capabilities for the identification and classification of scene abnormalities in remote sensing applications.

8.1.1 Mutual Coherence Matrix

The second-order correlation properties of a stochastic, statistically stationary electromagnetic beam are represented by the 2×2 mutual coherence matrix. This matrix is the electromagnetic equivalent to the mutual coherence function of the scalar theory,[96] which is important in the study of imaging systems with partially coherent illumination. The generalized van Cittert–Zernike theorem is a fundamental result in statistical optics that relates the spatial coherence properties and irradiance distribution of a beam with the irradiance distribution and spatial coherence properties of the quasi-homogeneous source, respectively, via a Fourier transform. Equivalent results in the context of the joint theory of coherence and polarization have been proposed recently.[97–99] The generalized van Cittert–Zernike theorem for the cross-spectral density matrix developed in our group states that

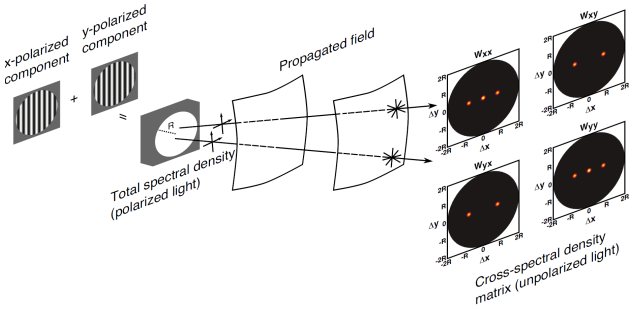


Figure 8.1: Diagram of a source designed using the generalized van Cittert-Zernike theorem for the cross-spectral density matrix and the resulting propagated field.

the spatial coherence and polarization properties in the far-field of a quasi-homogeneous electromagnetic source are given by the Fourier transform of the functions that describe its polarization and spatial coherence properties. The importance of this theorem in both the scalar and vectorial theories is that it allows the design of sources with prescribed spatial coherence distributions. In the case of the vectorial theory, this capability is further extended to the polarization state of the beam.[98] Therefore, knowledge of the propagation relations followed by the mutual coherence matrix is fundamental in the design of sources with prescribed spatial coherence and polarization properties that can be used in the implementation of the VCPol technique. Fig. 8.1 provides an example of a source with such characteristics. The cross-spectral density matrix depicted in the figure represents a beam that is unpolarized in the usual one-point sense but polarized in the two-point sense.[98]

8.1.2 Statistical description of scattering

The study of the interaction between an incident field and a random object is often carried out in a deterministic way for a pair of single realizations of the field and object and then repeated for a large number of different realizations. In this Monte Carlo approach, the final result is obtained as the ensemble average of the results obtained for the different realizations. An alternative strategy is to describe both the field and the random object as ensemble averages from the beginning and find the scattering relation that they satisfy. A first approximation to the description of this interaction between a stochastic, statistically stationary field and a random object, based on the first Born approximation, can be found in the literature for the scalar case.[92–94,100] The corresponding result for the electromagnetic case has been addressed in earlier work by our group.[16] This approach to the calculation of the scattered field in statistical optics is more computationally convenient than the Monte Carlo method because it uses the statistical properties of the random object and stochastic incident field to predict the scattered field directly. The main inconvenience of this approach is that the scattering potential and cross-spectral density matrix of the incident field must have a relatively simple mathematical representation in order to compute the scattered field in closed form. Examples of such relatively simple representations are the quasi-homogeneous fields and media, which constitute a good approximation in many applications.[95,101] To account for situations in which the first Born approximation is not applicable (e.g., multiple scattering and scattering in the near-field) the scattering theory must be extended. The extension of the statistical scattering theory to cover these situations is part of the work that will be done in future research.

We have developed a design for a novel active sensing system able to measure the elements of the mutual coherence matrix, allowing us to obtain a proof of concept of VCPol. Fig. 8.2 is an example of a sensing systems able to control the second order statistical properties of the illumination and measure the elements of the mutual coherence matrix resulting from its interaction with a specimen. The system combines a subsystem to control the second order statistical properties of the illumination with a Stokes polarimeter formed by variable retarders VR1 and VR2, and linear polarizer P4. Measurements with different second-order correlation properties of the source may be used to compute the elements of the mutual coherence matrix. The system in Fig. 8.2 is not the only configuration able to measure these quantities and we must examine other possibilities to determine an accurate and efficient way to measure the elements of the mutual coherence matrix. Furthermore, an analysis of the amount of data required and the development of the methods necessary to reconstruct the VCPol results from measurements of the mutual coherence matrix are two of our ongoing and future research objectives.

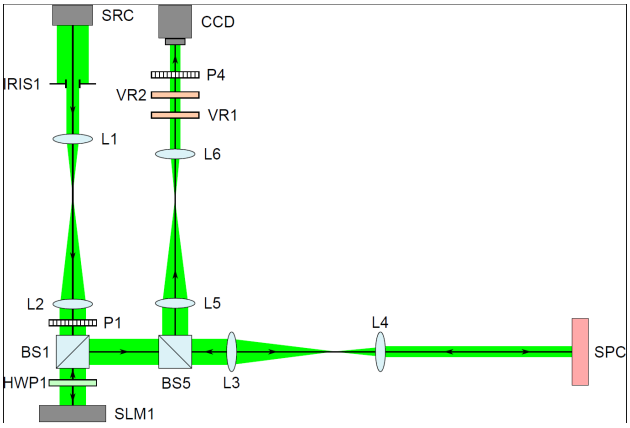


Figure 8.2: Diagram of a novel sensing system to indirectly measure the elements of the mutual coherence matrix.

REFERENCES

1. J. S. Tyo, "Hybrid division of aperture/division of focal plane polarimeter for real-time polarization imagery without IFOV error," *Opt. Lett.* **31**, pp. 2984 – 2986, 2006.
2. J. S. Tyo, J. K. Boger, B. M. Ratliff, W. T. Black, D. L. Bowers, and M. P. Fetrow, "The effects of thermal equilibrium and thermal contrast in polarimetric images in the lwir," *Opt. Express* , 2007.
3. J. S. Tyo, C. F. LaCasse, and B. M. Ratliff, "Total elimination of sampling errors in polarization imagery obtained with integrated microgrid polarimeters," *Opt. Lett.* **34**, pp. 3187 – 3189, 2009.
4. D. J. Diner, A. Davis, B. Hancock, G. Gutt, R. A. Chipman, and B. Cairns, "Dual-photoeleastic-modulator-based polarimetric imaging concept for aerosol remote sensing," *Appl. Opt.* **46**, pp. 8428 – 8445, 2007.
5. M. W. Kudenov, L. Pezzaniti, E. L. Dereniak, and G. R. Gerhart, "Prismatic imaging polarimeter calibration for the infrared spectral region," *Opt. Express* **16**, pp. 13720 – 13737, 2008.
6. J. Craven-Jones, M. W. Kudenov, and M. G. S. E. L. Dereniak, "Infrared hyperspectral imaging polarimeter using birefringent prisms," *Appl. Opt.* **50**, pp. 1170 – 1185, 2011.
7. C. F. LaCasse, J. S. Tyo, and R. A. Chipman, "Spatio-temporally modulated polarimetry," in *Proc. SPIE vol. 8160: Polarization Science and Remote Sensing V*, J. A. Shaw and J. S. Tyo, eds., p. 816020, SPIE, (Bellingham, WA), 2011.
8. F. Snik, G. can Harten, J. S. Tyo, A. S. Alenin, and I. Vaughn, "A multi-domain full-stokes polarization modulator that is efficient for 300-2500 nm spectropolarimetry," in *SPIE vol. 9613: Polarization Science and Remote Sensing VII*, J. A. Shaw and D. LeMaster, eds., p. 961315, SPIE, (Bellingham, WA), 2015.
9. C. F. LaCasse, , R. A. Chipman, and J. S. Tyo, "Band-limited reconstruction in modulated polarimeters," *Opt. Express* **19**, pp. 14976 –âĂŞ 14989, 2011.
10. C. F. LaCasse, J. S. Tyo, and R. A. Chipman, "The role of the null space in modulated polarimeters," *Opt. Lett.* **37**, pp. 1097 – 1099, 2012.
11. C. F. LaCasse, O. G. Rodrìguez-Herrera, R. A. Chipman, and J. S. Tyo, "Spectral density response functions for modulated polarimeters," *Appl. Opt.* **54**, pp. 9490 – 9499, 2015.
12. W. T. Black and J. S. Tyo, "Frequency-independent scene cancellation nonuniformity correction technique," *J. Electronic Imaging* **23**, p. 023005, 2014.
13. W. T. Black and J. S. Tyo, "Improving feedback-integrated scene cancellation nonuniformity correction through optimal selection of available camera motion," *J. Electronic Imaging* **23**, p. 053014, 2014.
14. W. T. Black and J. S. Tyo, "Frequency-domain scene-based non-uniformity correction and application to microgrid polarimeters," in *Proc. SPIE vol 8160: Polarization Science and Remote Sensing V*, p. 816003, SPIE, (Bellingham, WA), 2011.
15. O. G. R. Herrera and J. S. Tyo, "Generalized van cittertâĂŞzernike theorem for the cross-spectral density matrix of quasi-homogeneous planar electromagnetic sources," *J. Opt. Soc. Am. A* **29**, pp. 1939 – 1947, 2012.

16. J. S. Tyo and T. S. Turner, "Sensing polarization with variable coherence tomography," *J. Opt. Soc. Am. A* **25**, pp. 2383 – 2389, 2008.
17. A. S. Alenin and J. S. Tyo, "Generalized channeled polarimetry," *J. Opt. Soc. Am. A* **31**, pp. 1013–1022, May 2014.
18. A. S. Alenin and J. S. Tyo, "Structured decomposition design of partial mueller matrix polarimeters," *J. Opt. Soc. Am. A* **32**, pp. 1302–1312, Jul 2015.
19. B. G. Hoover and J. S. Tyo, "Polarization components analysis for invariant discrimination," *Appl. Opt.* **46**, pp. 8364 – 8373, 2007.
20. J. S. Tyo, Z. Wang, S. J. Johnson, and B. G. Hoover, "Design, analysis, and optimization of partial Mueller matrix polarimeters," *Appl. Opt.* **49**, pp. 2326 – 2333, 2010.
21. A. S. Alenin and J. S. Tyo, "Channeled partial mueller matrix polarimetry," in *Proc. SPIE vol. 9613: Polarization Science and Remote Sensing VII*, J. A. Shaw and D. LeMaster, eds., p. 96130M, SPIE, (Bellingham, WA), 2015.
22. D. A. LeMaster, A. Mahamat, B. M. Ratliff, A. Alenin, and J. S. Tyo, "SWIR active polarization imaging," in *Proc. SPIE vol. 8873: Polarization Science and Remote Sensing VI*, D. Lemaster and J. A. Shaw, eds., p. 88730O, SPIE, (Bellingham, WA), 2013.
23. I. J. Vaughn, O. G. Rodriguez-Herrera, M. Xu, and J. S. Tyo, "A portable imaging mueller matrix polarimeter based on a spatio-temporal modulation approach: theory and implementation," in *Proc. SPIE vol. 9613: Polarization Science and Remote Sensing VII*, J. A. Shaw and D. LeMaster, eds., p. 961304, SPIE, (Bellingham, WA), 2015.
24. I. J. Vaughn, M. Xu, and J. S. Tyo, "A portable imaging mueller matrix polarimeter based on a spatio-temporal modulation approach: theory and implementation," in *Proc. SPIE vol. 9613: Polarization Science and Remote Sensing VII*, J. A. Shaw and D. LeMaster, eds., p. 961337, SPIE, (Bellingham, WA), 2015.
25. T. Wakayama, O. G. Rodríguez-Herrera, J. S. Tyo, Y. Otani, M. Yonemura, and T. Yoshizawa, "Generation of achromatic, uniform-phase, radially polarized beams," *Opt. Express* **22**, pp. 3306–3315, Feb 2014.
26. T. Wakayama, T. Higashiguchi, H. Oikawa, K. Sakaue, M. Washio, M. Yonemura, T. Yoshizawa, J. S. Tyo, and Y. Otani, "Determination of the polarization states of an arbitrary polarized terahertz beam: Vectorial vortex analysis," *Scientific Reports* **5**, p. 9416, 2015.
27. H. H. Barrett and K. J. Myerrs, *Foundations of Image Science*, Wiley-Interscience, New York, 2003.
28. R. A. Chipman, "Polarimetry," in *Handbook of Optics*, M. Bass, ed., **2**, ch. 22, McGraw-Hill, 1995.
29. J. S. Tyo, D. H. Goldstein, D. B. Chenault, and J. A. Shaw, "Review of passive imaging polarimetry for remote sensing applications," *Appl. Opt.* **45**, pp. 5453 – 5469, August 2006.
30. J. L. Pezzaniti and D. B. Chenault, "A division of aperture MWIR imaging polarimeter," in *Proceedigns of SPIE vol. 5888: Polarizaiton Science and Remote Sensing II*, J. A. Shaw and J. S. Tyo, eds., p. 5888OV, SPIE, Bellingham, WA, 2005.
31. R. A. Chipman, "Mueller matrices," in *Handbook of Optics*, M. Bass, ed., **1**, ch. 15, pp. 15.1–15.41, McGraw Hill, third ed., 2009.
32. D. S. Sabatke, M. R. Descour, E. Dereniak, W. C. Sweatt, S. A. Kemme, and G. S. Phipps, "Optimization of retardance for a complete Stokes polarimeter," *Opt. Lett.* **25**, pp. 802–804, 2000.
33. J. S. Tyo and T. S. Turner, "Variable retardance, Fourier transform imaging spectropolarimeters for visible spectrum remote sensing," *Appl. Opt.* **40**, pp. 1450–1458, 2001.
34. A. Ambirajan and D. C. Look, "Optimum angles for a polarimeter: part I," *Opt. Eng.* **34**, pp. 1651–1655, 1995.
35. R. M. E. Illing, "High-speed fieldable imaging stokes vector polarimeter," in *Proceedigns of SPIE vol. 5888: Polarizaiton Science and Remote Sensing II*, J. A. Shaw and J. S. Tyo, eds., p. 58880X, SPIE, (Bellingham, WA), 2005.
36. M. P. Silverman and W. Strange, "Object delineation within turbid media by backscattering of phase modulated light," *Opt. Commun.* **144**, pp. 7–11, 1997.
37. J. S. Tyo, "Design of optimal polarimeters: maximization of SNR and minimization of systematic errors," *Appl. Opt.* **41**, pp. 619–630, 2002.

38. F. Goudail and A. Beniere, "Estimation precision of the linear degree of polarization and of the angle of polarization in the presence of different types of noises," *Appl. Opt.* **49**, pp. 683 – 693, 2010.
39. R. A. Chipman, "Polarization analysis of optical systems," *Opt. Eng.* **28**, pp. 90–99, 1989.
40. K. Oka and T. Kaneko, "Compact complete imaging polarimeter using birefringent wedge prisms," *Opt. Express* **11**, pp. 1510 – 1519, 2003.
41. K. Oka and T. Kato, "Spectroscopic polarimetry with a channeled spectrum," *Opt. Lett.* **24**, pp. 1475 – 1477, 1999.
42. J. P. McGuire and R. A. Chipman, "Diffraction image formation i noptical systems with polarization aberrations. I: Formulation and example," *J. Opt. Soc. Am. A* **7**, pp. 1614 – 1626, September 1990.
43. C. F. LaCasse and J. S. Tyo, "The effect of calibration error on polarimetric reconstruction in microgrid polarimetric imagery," in *Frontiers in Optics 2010: OSA Annual Meeting*, OSA, (Rochester, NY), 2010.
44. L. Gendre, A. Foulonneau, and L. Bigué, "High-speed imaging acquisition of stokes linearly polarized components using a single ferroelectric liquid crystal modulator," in *Proc. SPIE vol. 7461: Polarization Science and Remote Sensing IV*, J. A. Shaw and J. S. Tyo, eds., p. 74610G, SPIE, (Bellingham, WA), 2009.
45. N. Hagen, K. oka, and E. L. Dereniak, "Snapshot Mueller matrix spectropolarimeter," *Opt. Lett.* **32**, pp. 2100 – 2102, 2007.
46. D. J. Diner, A. Davis, B. Hancock, S. Geier, B. Rheingans, V. Jovanovic, M. Bull, D. M. R. R. A. Chipman, A.-B. Mahler, and S. C. McClain, "First results from a dual photoelastic-modulator-based polarimetric camera," *Appl. Opt.* **49**, pp. 2929 – 2946, 2010.
47. F. Snik, G. van Harten, R. Navarro, P. Groot, L. Kaper, and A. de Wijn, "Design of a full-Stokes polarimeter for VLT/X-shooter," in *Proc. SPIE vol. 8446*, 2012. Need to complete entry.
48. J. S. Tyo and A. S. Alenin, *Field Guide to Linear Systems in Optics*, SPIE, Bellingham, WA, 2015.
49. C. F. LaCasse and J. S. Tyo, "Modulated polarimeter operators in the presence of stochastic signals," in *Proc. SPIE vol. 8364: Polarization Measurement, Analysis, and Remote Sensing X*, D. B. Chenault and D. H. Goldstein, eds., p. 836402, SPIE, (Bellingham, WA), 2012.
50. R. L. J. Easton, *Fourier Methods in Imaging*, Wiley, New York, 2010.
51. K. Oka and T. Kato, "Spectroscopic polarimeter based on optical frequency-domain interferometry," in *Proceedings of SPIE vol. 3749*, A. J. Glass, J. W. Goodman, M. Chang, A. H. Guenther, and T. Asakura, eds., pp. 681–682, SPIE, (Bellingham, WA), 1999.
52. H. Okabe, M. Hayakawa, H. Naito, A. Taniguchi, and K. Oka, "Spectroscopic polarimetry using channeled spectroscopic polarization state generator (cspsg)," *Opt. Express* **15**, pp. 3093–3109, Mar 2007.
53. M. W. Kudenov, N. A. Hagen, E. L. Dereniak, and G. R. Gerhart, "Fourier transform channeled spectropolarimetry in the mwir," *Opt. Express* **15**, pp. 12792–12805, Oct 2007.
54. M. W. Kudenov, M. J. E. N. Hagena, E. L. Dereniak, and K. Oka, "Snapshot imaging mueller matrix polarimeter using polarization gratings," *Opt. Lett.* **37**, pp. 1367 – 1369, 2012.
55. P. Lemaillet, S. Rivet, and B. L. Jeune, "Optimization of a snapshot mueller matrix polarimeter," *Opt. Lett.* **33**, pp. 144–146, Jan 2008.
56. J. D. Gaskill, *Linear Systems, Fourier Transforms, and Optics*, John wiley & Sons, New York, 1978.
57. I. Vaughn, B. G. Hoover, and J. S. Tyo, "Classification using active polarimetry," in *Proc. SPIE vol. 8364: Polarization Measurement, Analysis, and Remote Sensing X*, D. B. Chenault and D. H. Goldstein, eds., p. 83610S, SPIE, (Bellingham, WA), 2012.
58. K. Twietmeyer and R. A. Chipman, "Optimization of Mueller polarimeters in the presence of error sources," *Opt. Express* **16**, pp. 11589 – 11603, 2008.
59. J. J. Gil, "Polarimetric characterization of light and media," *European J. of Physics* **40**, pp. 1–47, 2007.
60. S.-Y. Lu and R. A. Chipman, "Interpretation of Mueller matrices based on the polar decomposition," *J. Opt. Soc. Am. A* **13**, pp. 1106–1113, 1996.
61. R. Ossikovski, "Analysis of depolarizing Mueller matrices through a symmetric decomposition," *J. Opt. Soc. Am. A* **26**, pp. 1109 – 1118, 2009.
62. R. Ossikovski, "Differential and product Mueller matrix decompositions: a formal comparison," *Opt. Lett.* **37**, pp. 220 – 222, 2012.

63. H. D. Noble and R. A. Chipman, "Mueller matrix roots algorithm and computational considerations," *OX* **20**, pp. 17 – 31, 2012.
64. R. Ossikovski, "Differential matrix formalism for depolarizing anisotropic media," *Opt. Lett.* **36**, pp. 2330 – 2332, 2011.
65. S. R. Cloude, "Group theory and polarization algebra," *Optik (Stuttgart)* **75**, pp. 26–36, 1986.
66. I. J. Vaughn and B. G. Hoover, "Noise reduction in a laser polarimeter based on discrete waveplate rotations," *Opt. Express* **16**, pp. 2091 – 2108, 2008.
67. F. Goudail and A. Beniere, "Optimization of the contrast in polarimetric scalar images," *Opt. Lett.* **34**, pp. 1471 – 1473, 2009.
68. F. Goudail, "Comparative study of the best achievable contrast in scalar, stokes and mueller images," in *First NanoCharm workshop on advanced polarimetric instrumentation*, NanoCharm: Multifunctional Nanomaterials Characterization Exploiting EllipsoMetry and Polarimetry, 2009.
69. F. Goudail and J. S. Tyo, "When is polarimetric imaging preferable to intensity imaging for target detection?," *J. Opt. Soc. Am. A* **28**, pp. 46 – 53, 2011.
70. Anna and Goudail, "Optimal mueller matrix estimation in the presence of poisson shot noise," *Opt. Express* **20**, pp. 21331 – 21340, 2012.
71. G. H. Golub and C. F. van Loan, *Matrix Computations*, ch. 2, pp. 11–29. Johns Hopkins University Press, Baltimore, MD, 1983.
72. R. A. Chipman, "Metrics for depolarization," in *Proceedigns of SPIE vol. 5888: Polarizaiton Science and Remote Sensing II*, J. A. Shaw and J. S. Tyo, eds., p. 58880L, SPIE, (Bellingham, WA), 2005.
73. S. R. Cloude and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *Geoscience and Remote Sensing, IEEE Transactions on* **34**(2), pp. 498–518, 1996.
74. Z. Wang, J. S. Tyo, and M. M. Hayat, "Data interpretation for spectral sensors with correlated bands," *J. Opt. Soc. Am. A* **24**, pp. 2864 – 2870, 2007.
75. H. Hotelling, "Relations between two sets of variates," *Biometrika* **28**, pp. 312 – 377, 1936.
76. R. M. A. Azzam, "Photopolarimetric measurement of the Mueller matrix by Fourier analysis of a single detected signal," *Opt. Lett.* **2**, pp. 48–50, 1977.
77. K. Oka, "Singleshot spectroscopic polarimetry using channeled spectrum," in *Photonics Asia 2002*, pp. 167–175, International Society for Optics and Photonics, 2002.
78. D. S. Sabatke, A. Locke, E. L. Dereniak, M. R. Descour, J. Garcia, T. Hamilton, and R. W. McMillan, "Snapshot imaging spectropolarimeter," *Opt. Eng.* **41**, pp. 1048 – 1054, 2002.
79. M. W. Kudenov, M. J. Escuti, E. L. Dereniak, and K. Oka, "White-light channeled imaging polarimeter using broadband polarization gratings," *Applied optics* **50**(15), pp. 2283–2293, 2011.
80. N. Hagen and M. W. Kudenov, "Review of snapshot spectral imaging technologies," *Optical Engineering* **52**(9), pp. 090901–090901, 2013.
81. M. Dubreil, S. Rivet, B. L. Jeune, and J. Cariou, "Snapshot Mueller matrix polarimeter by wavelength polarization coding," *Opt. Express* **15**, pp. 13660 – 13668, 2007.
82. D. A. LeMaster and K. Hirakawa, "Improved microgrid arrangement for integrated imaging polarimeters," *Opt. Lett.* **39**, pp. 1811–1814, Apr 2014.
83. 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**, pp. 27393–27409, Dec 2012.
84. 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**(17), pp. 17776–17787, 2010.
85. E. Wolf, *Introduction to the Theory of Coherence and the Polarization of Light*, Cambridge University Press, New York, United States, 2007.
86. E. Collett, *Polarized Light, Fundamentals and Applications*, Marcel Dekker, New York, 1992.
87. J. J. Gil, "Characteristic properties of Mueller matrices," *J. Opt. Soc. Am. A* **17**, pp. 328–334, 2000.
88. A. Alenin and J. S. Tyo, "Task-specific snapshot mueller matrix channeled spectropolarimeter optimization," in *Proc. SPIE vol. 8364: Polarization Measurement, Analysis, and Remote Sensing X*, D. B. Chenault and D. H. Goldstein, eds., p. 836403, SPIE, (Bellingham, WA), 2012.

89. H. Nyquist, "Certain topics in telegraph transmission theory," *American Institute of Electrical Engineers, Transactions of the* **47**(2), pp. 617–644, 1928.
90. 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**(26), pp. 429–441, 1946.
91. C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review* **5**(1), pp. 3–55, 2001.
92. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics*, Cambridge, 1995.
93. E. Baleine and A. Dogariu, "Variable-coherence tomography," *Opt. Lett.* **29**, pp. 1233 – 1235, 2004.
94. E. Baleine and A. Dogariu, "Variable-coherence tomography for inverse scattering problems," *J. Opt. Soc. Am. A* **21**, pp. 1917 – 1923, 2004.
95. W. H. Carter and E. Wolf, "Scattering from quasi-homogeneous media," *Opt. Commun.* **67**, pp. 85 – 90, 1988.
96. E. Wolf, *Introduction to the Theory of Coherence and Polarization of Light*, Cambridge, 2007.
97. F. Gori, M. Santarsiero, R. Borghi, and G. Piquero, "Use of the van cittert–zernike theorem for partially polarized sources," *Opt. Lett.* **25**, pp. 1291–1293, 2000.
98. O. G. Rodríguez-Herrera and J. S. Tyo, "Generalized van cittert-zernike theorem for the cross-spectral density matrix of quasi-homogeneous planar electromagnetic sources," *J. Opt. Soc. Am. A* **29**, pp. 1939 – 1947, 2012.
99. O. Korotkova, B. G. Hoover, V. L. Gamiz, and E. Wolf, "Coherence and polarization properties of far fields generated by quasi-homogeneous planar electromagnetic sources," *J. Opt. Soc. Am. A* **22**, pp. 2547–2556, 2005.
100. A. M. Zarubin, "Three-dimensional generalization of the van cittert–zernike theorem to wave and particle scattering," *Opt. Commun.* **100**, pp. 491–507, 1993.
101. W. H. Carter and E. Wolf, "Coherence and radiometry with quasi-homogeneous planar sources," *J. Opt. Soc. Am.* **67**, pp. 785 – 796, 1977.

1.

1. Report Type

Final Report

Primary Contact E-mail
Contact email if there is a problem with the report.

tyo@optics.arizona.edu

Primary Contact Phone Number
Contact phone number if there is a problem with the report

+61262689534

Organization / Institution name

University of Arizona

Grant/Contract Title
The full title of the funded effort.

Imaging and Non-Imaging Polarimetric Methods for Remote Sensing

Grant/Contract Number
AFOSR assigned control number. It must begin with "FA9550" or "F49620" or "FA2386".

FA9550-10-1-0114

Principal Investigator Name
The full name of the principal investigator on the grant or contract.

J Scott Tyo

Program Manager
The AFOSR Program Manager currently assigned to the award

Dr. Julie Moses

Reporting Period Start Date

04/15/2010

Reporting Period End Date

12/15/2015

Abstract

During the lifetime of this project, the research team signi˝cantly advanced the state of understanding of modulated polarimeter systems for active and passive, imaging- and non-imaging polarimetry. Modulated polarimeters infer the polarization state of light by creating a set of polarization-dependent carriers that are modulated by the intensity signal. These carriers can be created in any independent domain, such as time, space, wavenumber, angle of incidence, and they can be created in combinations of domains simultaneously. The work supported on this project has solidified the theory behind such instruments, allowing for new design philosophies that improve state-of-the-art instruments. When the project began five years ago, only a cursory understanding of modulated instruments existed, and the data reduction matrix was the primary means of processing polarization data. In this project, the DRM has been expanded to include a full functional formalism, allowing for a range of new tools in polarimeter design to be brought to bear. During the course of this project, five PhD students, seven MS students, and two undergraduates were trained. Four PhD dissertations were primarily supported by funding from this project.

Distribution Statement
This is block 12 on the SF298 form.

Distribution A - Approved for Public Release

Explanation for Distribution Statement

If this is not approved for public release, please provide a short explanation. E.g., contains proprietary information.

SF298 Form

Please attach your SF298 form. A blank SF298 can be found here. Please do not password protect or secure the PDF The maximum file size for an SF298 is 50MB.

SF298.pdf

Upload the Report Document. File must be a PDF. Please do not password protect or secure the PDF . The maximum file size for the Report Document is 50MB.

afosr_2010_finalreport.pdf

Upload a Report Document, if any. The maximum file size for the Report Document is 50MB.

Archival Publications (published) during reporting period:

1. J. S. Tyo, Z. Wang, S. J. Johnson, and B. G. Hoover, Design, analysis, and optimization of partial Mueller matrix polarimeters, Appl. Opt. 49, pp. 2326 2333, 2010.

2. C. F. LaCasse, J. S. Tyo, and R. A. Chipman, Spatio-temporally modulated polarimetry, in Proc. SPIE vol. 8160: Polarization Science and Remote Sensing V, J. A. Shaw and J. S. Tyo, eds., p. 816020, SPIE, (Bellingham, WA), 2011.

3. C. F. LaCasse, , R. A. Chipman, and J. S. Tyo, Band-limited reconstruction in modulated polarimeters, Opt. Express 19, pp. 14976 14989, 2011.

4. A. Alenin, L. Morrison, C. Curiel, and J. S. Tyo, Hyperspectral measurement of the scattering of polarized light by skin, Proc. SPIE vol 8160: Polarization Science and Remote Sensing V, J. A. Shaw and J. S. Tyo, Eds., pp. 816014 (SPIE, Bellingham, 2011)

5. W. T. Black and J. S. Tyo, Frequency-domain scene-based non-uniformity correction and application to microgrid polarimeters, in Proc. SPIE vol 8160: Polarization Science and Remote Sensing V, p. 816003, SPIE, (Bellingham, WA), 2011.

6. O. G. R. Herrera and J. S. Tyo, Generalized van Cittert-Zernike theorem for the cross-spectral density matrix of quasi-homogeneous planar electromagnetic sources, 29, pp. 1939 1947, 2012.

7. C. F. LaCasse, J. S. Tyo, and R. A. Chipman, The role of the null space in modulated polarimeters, Opt. Lett. 37, pp. 1097 1099, 2012.

8. A. S. Alenin and J. S. Tyo, Task-Specic Snapshot Mueller Matrix Channeled Spectropolarimeter Optimization,
in Proc. SPIE vol. 8364: Polarization Measurement, Analysis, and Remote Sensing IX, D. B. Chenault and D. L. Goldstein, Eds., pp. 836402 (SPIE, Bellingham, WA, 2012)

9. C. F. LaCasse and J. S. Tyo, Modulated polarimeter operators in the presence of stochastic signals, Proc.
SPIE vol. 8364: Polarization Measurement, Analysis, and Remote Sensing IX, D. B. Chenault and D. L. Goldstein, Eds., pp. 836402 (SPIE, Bellingham, WA, 2012)

10. I. Vaughn, B. G. Hoover, and J. S. Tyo, Classication using active polarimetry, SPIE Polarization Measurement, Analysis, and Remote Sensing IX, D. B. Chenault and D. L Goldstein, Eds., p. 83640S (SPIE, Bellingham, WA, 2012)

11. D. A. LeMaster, A. Mahamat, B. M. Ratli, A. Alenin, and J. S. Tyo, SWIR active polarization imaging, in Proc. SPIE vol. 8873: Polarization Science and Remote Sensing VI, D. Lemaster and J. A. Shaw, eds.,

p. 88730O, SPIE, (Bellingham, WA), 2013.

12. T. Wakayama, K. Komaki, I. Vaughn, J. S. Tyo, Y. Otani, T. Yoshizawa, Evaluation of Mueller matrix of achromatic axially symmetric waveplate, SPIE Polarization Science and Remote Sensing VI D. A. LeMaster and J. A. Shaw, Eds., pp. 88730P (SPIE, Bellingham, WA, 2013)

13. W. T. Black and J. S. Tyo, Frequency-independent scene cancellation nonuniformity correction technique,
J. Electronic Imaging 23, p. 023005, 2014.

14. W. T. Black and J. S. Tyo, Improving feedback-integrated scene cancellation nonuniformity correction through optimal selection of available camera motion, J. Electronic Imaging 23, p. 053014, 2014.

15. A. S. Alenin and J. S. Tyo, Generalized channeled polarimetry, J. Opt. Soc. Am. A 31, pp. 10131022, May 2014.

16. A. S. Alenin and J. S. Tyo, Structured decomposition design of partial mueller matrix polarimeters, J. Opt. Soc. Am. A 32, pp. 13021312, Jul 2015.

17. F. Snik, G. can Harten, J. S. Tyo, A. S. Alenin, and I. Vaughn, A multi-domain full-stokes polarization modulator that is ecient for 300-2500 nm spectropolarimetry, in SPIE vol. 9613: Polarization Science and Remote Sensing VII, J. A. Shaw and D. LeMaster, eds., p. 961315, SPIE, (Bellingham, WA), 2015.

18. C. F. LaCasse, O. G. Rodrìguez-Herrera, R. A. Chipman, and J. S. Tyo, Spectral density response functions
for modulated polarimeters, Appl. Opt. 54, pp. 9490 9499, 2015.

19. A. S. Alenin and J. S. Tyo, Channeled partial mueller matrix polarimetry, in Proc. SPIE vol. 9613: Polarization Science and Remote Sensing VII, J. A. Shaw and D. LeMaster, eds., p. 96130M, SPIE, (Bellingham, WA), 2015.

20. I. J. Vaughn, O. G. Rodriguez-Herrera, M. Xu, and J. S. Tyo, A portable imaging mueller matrix polarimeter
based on a spatio-temporal modulation approach: theory and implementation, in Proc. SPIE
vol. 9613: Polarization Science and Remote Sensing VII, J. A. Shaw and D. LeMaster, eds., p. 961304, SPIE, (Bellingham, WA), 2015.

21. I. J. Vaughn, M. Xu, and J. S. Tyo, A portable imaging mueller matrix polarimeter based on a spatiotemporal
modulation approach: theory and implementation, in Proc. SPIE vol. 9613: Polarization
Science and Remote Sensing VII, J. A. Shaw and D. LeMaster, eds., p. 961337, SPIE, (Bellingham, WA), 2015.

22. T. Wakayama, T. Higashiguchi, H. Oikawa, K. Sakaue, M. Washio, M. Yonemura, T. Yoshizawa, J. S. Tyo,
and Y. Otani, Determination of the polarization states of an arbitrary polarized terahertz beam: Vectorial vortex analysis, Scientic Reports 5, p. 9416, 2015.

Changes in research objectives (if any):

None

Change in AFOSR Program Manager, if any:

Dr Kent Miller (2010 - 2013)
Dr Julie Moses (2013 - 2015)

Extensions granted or milestones slipped, if any:

Extension of end date from 14 April 2015 to 15 Dec 2015

AFOSR LRIR Number

LRIR Title

Reporting Period

Laboratory Task Manager

Program Officer

Research Objectives

Technical Summary

Funding Summary by Cost Category (by FY, $K)

	Starting FY	FY+1	FY+2
Salary			
Equipment/Facilities			
Supplies			
Total			

Report Document

Report Document - Text Analysis

Report Document - Text Analysis

Appendix Documents

2. Thank You

E-mail user

Feb 03, 2016 20:44:06 Success: Email Sent to: tyo@optics.arizona.edu