Polarimetry from the Ground Up

Polarimetry from the Ground Up

Christoph Keller & Frans Snik

Utrecht University

Outline

1. Introduction
2. Statistical Errors in Polarimetry
3. Polarization Calibration
4. Polarimetry Error Budget

Introduction

Polarimetry Systems Engineering

goal: model and understand performance of polarimeter designs

Give answers to the following questions

- how to define polarimeter performance?
- how to compare different polarimeter designs?
- how to maximize performance?
- how to do all of this for different polarimeter designs?

Statistical Errors

- photon statistics
- detector read-out noise

Some Systematic Errors (instrumental errors)

- Atmospheric seeing and guiding errors
- Instrumental polarization due to
  - Telescope and instrument optics
  - Polarized scattered light in telescope and instrument
  - Spectrograph slit polarization
  - Angle, wavelength, temperature dependence of retarders
  - Crystal aberrations
  - Polarized fringes
- Ghost images
- Variable sky background
- Unpolarized scattered light in atmosphere and optics
- Limited calibration accuracy
- ...

Assumptions to Derive Expected Noise

- linear relation between Stokes parameters and measured signals
- noise in the various measurements is independent
- noise is independent of the signal, if
  - noise is dominated by signal-independent detector noise (e.g. read-out noise)
  - incoming vector is only slightly polarized
- noise statistic has a Gaussian distribution

Signal Matrix

- *m* intensity measurements combined into *signal vector* $\boldsymbol{S}$
- related to incoming Stokes vector, $\boldsymbol{I}$ by $4 \times m$ *signal matrix* X (also *modulation matrix*

$$\boldsymbol{S} = \mathrm{X}\boldsymbol{I}$$

- $\mathrm{X}(\boldsymbol{v})$ is function of free design parameters $\boldsymbol{v}$
- each row of X corresponds to first row of Mueller matrix describing the particular intensity measurement

Measuring Stokes Parameters

- estimate of incoming Stokes vector $\boldsymbol{I}$

$$\boldsymbol{I}' = \mathsf{Y}\boldsymbol{S} = \mathsf{X}^{-1}S$$

- Y is called *synthesis* or *demodulation matrix*
- error propagation provides standard deviations of Stokes parameters

$$\sigma_{I'_i} = \sqrt{\sum_{j=1}^{4} \mathsf{Y}_{ij}^2 \sigma_{S_j}^2}$$

- σ_{S_j} is standard deviation of intensity in measurement j

Polarimetric Efficiency

desired properties:

- comparable between different polarimeter designs and measurement approaches
- larger values should correspond to better designs
- independent of the intensity throughput
- consist of 4 quantities (“Stokes efficiency”)
- theoretical maximum efficiency shall be 1

Noise propagation

- if all measurements have same noise

$$\sigma_{I'_i} = \sigma_S \sqrt{\sum_{j=1}^{m} \mathrm{Y}_{ij}^2}\,.$$

- polarimetric efficiency independent of number of measurements (del Toro Iniesta and Collados 2000)

$$\epsilon_i = \left(m \sum_{j=1}^{m} Y_{ij}^2 \right)^{-\frac{1}{2}}$$

- in most cases $\mathrm{X}_{i1} = 1$, i.e. all measurements have the same throughput, normalize the measurements

$$\epsilon_1 \leq 1\,;\sum_{i=2}^{4} \epsilon_i^2 \leq 1$$

Analytic Optimization

- equations for properties of optimum polarimeter by e.g. del Toro Iniesta and Collados (2000) and Tyo (2002)
- *maximum performance*: best performance of all polarimeter designs
- *optimum performance*: best performance that can be achieved with given polarimeter design
- choose X($\boldsymbol{v}$), Y to minimize difference between $\boldsymbol{I}'$ and $\boldsymbol{I}$, given σ_{S_j}
- 3 steps:
    - derive equation for optimum Y given X($\boldsymbol{v}$)
    - using the equation for optimum Y, derive optimum X
    - choose optimum X and Y to obtain polarimeter with maximum performance

Optimum synthesis matrix

- if $m = 4$, four measurements represent linearly independent combinations of Stokes parameters, then Y is uniquely determined by
$$\mathsf{Y} = \mathsf{X}^{-1} .$$
- if $m > 4$, matrix inverse is not uniquely defined
- *generalized inverse* (e.g. Albert 1972) of X
$$\mathsf{Y} = \left(\mathsf{X}^T\mathsf{X}\right)^{-1}\mathsf{X}^T$$
- among all possible Y that fulfill $\mathsf{YX} = 1$, the generalized inverse minimizes the sum of squares of its rows
- optimum polarimetric efficiencies given by
$$\epsilon_{\mathrm{opt},i} = \sqrt{\frac{1}{m\left(\mathsf{X}^T\mathsf{X}\right)^{-1}_{ii}}}$$

Signal matrix for maximum performance

- follow scheme of del Toro Iniesta and Collados (2000)
- minimize sum of squares of rows of Y with $XY = 1$
- squares of maximum possible polarimetric efficiencies

$$\epsilon^2_{\max,i} = \frac{1}{m}\sum_{j=1}^{m} X^2_{ji} = \frac{1}{m}\left(X^T X\right)_{ii} \; .$$

- signal matrix X of maximum performance polarimeter

$$X^T X = m \begin{pmatrix} \epsilon^2_{\max,1} & 0 & 0 & 0 \\ 0 & \epsilon^2_{\max,2} & 0 & 0 \\ 0 & 0 & \epsilon^2_{\max,3} & 0 \\ 0 & 0 & 0 & \epsilon^2_{\max,4} \end{pmatrix}$$

- optimum efficiency for measuring polarized Stokes components with equal efficiencies given by $\frac{1}{\sqrt{n}}$, where n is the number of polarized Stokes parameters

Generation of maximum performance signal matrices

- elements of first column of signal matrix X are 1, $X^T X = \mathrm{diagonal}$
- each row of the signal matrix: point on the Poincaré sphere
- maximize average distance squared between points on Poincaré sphere for signal matrix with maximum polarimetric efficiency
- distance $\Delta = \sqrt{2 + \frac{2}{m-1}}$
  - $m = 2$: $\Delta = 2$ (opposite sides of Poincaré sphere)
  - $m = 3$: $\Delta = \sqrt{3}$ (side length of triangle whose plane includes the origin of Poincaré sphere)
  - $m = 4$: $\Delta = \sqrt{\frac{8}{3}}$ (side length of tetrahedron inside Poincaré sphere)
- average distance is indepenent of the number of polarized Stokes components
- four measurements for linear polarization only, spaced 90 deg apart, square root of average distance squared is also $\Delta = \sqrt{\frac{8}{3}}$

Caveat

- maximizing average square distance of points is a necessary but not a sufficient criteria for generating a signal matrix corresponding to a polarimeter with maximum performance
- signal matrix

$$X = \begin{pmatrix} 1 & x & x & x \\ 1 & x & x & x \\ 1 & -x & -x & -x \\ 1 & -x & -x & -x \end{pmatrix} \tag{1}$$

  also has an average distance of $\sqrt{\frac{8}{3}}$, but $X^T X$ is obviously not diagonal

Optimum Calibration

Basic Approach

- optimum way to calibrate polarimeter, i.e. measure X

$$\boldsymbol{S} = \mathrm{X}\boldsymbol{I}$$

- at least 16 measurements of signal vector elements $\boldsymbol{S}_i, i = 1..m$
- at least 4 different input Stokes vectors $\boldsymbol{I}_i^c, i = 1..m$
- group calibration input Stokes vectors, signal vectors into 4 by 4 matrices (Azzam et al. 1988)

$$\begin{aligned} \mathrm{S} &= \mathrm{X}\mathrm{I}^c \\ \mathrm{S} &= (\boldsymbol{S}_1 \boldsymbol{S}_2 \ldots \boldsymbol{S}_m) \\ \mathrm{I}^c &= (\boldsymbol{I}_1^c \boldsymbol{I}_2^c \ldots \boldsymbol{I}_m^c) \end{aligned}$$

- signal matrix $\mathrm{X} = \mathrm{S}\mathrm{J}$ with $\mathrm{I}^c\mathrm{J} = 1$

Equivalence with Maximum Polarimetric Performance

- choose calibration Stokes vectors $\boldsymbol{I}_i^c$ to minimize error in X, given errors in measurements S
- equivalent to optimizing polarimetric efficiency
- calibration input Stokes vectors $\boldsymbol{I}_i^c$ equivalent to rows of the signal matrix X of maximum performance polarimeter
- vector-polarimeter: corners of a tetrahedron as suggested by Azzam et al. (1988)
- only true if there are no systematic errors
- better to use many more measurements and also model non-ideal calibration optics

Polarization Error Budget

Optical Error Budgets

- classical tool in systems engineering
- sets requirements of parts such that system as a whole meets requirements
- much faster than end-to-end simulations
- provides more insight than end-to-end simulations

Example: 80% Encircled Energy in arcsec

level 1 item	level 2 item	level 3 item	level 1	level 2	level 3
atmosphere			0.50		
telescope			0.25		
	primary mirror			0.17	
		mirror polishing			0.10
		mirror support			0.10
		thermal distortion			0.10
	secondary mirror			0.17	
		mirror polishing			0.10
		mirror support			0.10
		thermal distortion			0.10
instrument			0.25		
total			0.61		

Example: 80% Encircled Energy in arcsec

level 1 item	level 2 item	level 3 item	level 1	level 2	level 3
atmosphere			0.50		
telescope			0.25		
	primary mirror			0.17	
		mirror polishing			0.10
		mirror support			0.10
		thermal distortion			0.10
	secondary mirror			0.17	
		mirror polishing			0.10
		mirror support			0.10
		thermal distortion			0.10
instrument			0.25		
total			0.61		

Structure of an Error Budget

- parts can consist of smaller parts $\Rightarrow$ several levels
- overall error estimated from combination of many error sources
- mistakes in estimates tend to average out
- errors can be allocated in different ways
- allocation minimizes complexity and cost
- main contributors to system error easily identified

Adding Wavefront Errors

- wavefront errors of individual elements are not correlated
- square of final error given by squares of errors of individual elements (*root sum of squares (RSS)*)

Polarization Error Budget Issues

- polarimeter: final errors often non-linear combinations of errors of very different nature (e.g. instrumental polarization and non-linearity of detector)
- some (but not all) errors can be drastically reduced by calibration
- *polarization* error budgets not (yet) used to design polarimeters

Schematic Polarization Error Budget

level 1 item	level 2 item	level 3 item
source variation		
atmosphere		
telescope		
polarimeter		
	polarizer	
	detector system	
		undetected bias variation
		nonlinearity
calibration		
	polarizer	
	retarder	
		positioning repeatability
		temperature change of retarder
data reduction		
total		

VLT SPHERE Polarimetric Error Budget

	Level 1	Level 2	Level 3	Value	Stokes Vector				Mueller Matrix			
					I	Q	U	V	M11	M12	M13	M14
									M21	M22	M23	M24
									M31	M32	M33	M34
									M41	M42	M43	M44
3	Telescope											
									1	0	0	0
									0	1	0	0
									0	0	1	0
									0	0	0	1
3.1		M1										
									2	0	0	0
									0	2	0	0
									0	0	2	0.01
									0	0	0.1	2
3.1.1			static						1	1.00E-05	0	0
									1.00E-05	1	0	0
									0	0	1	1.00E-04
									0	0	0.1	1

Errors in Mueller Matrices

- need standard deviation of Mueller matrix elements as a function of errors in parameters
- Mueller matrix with two parameters α and β

$$\mathsf{M}(\alpha, \beta)$$

- variance over uniform distribution of errors in α over $\pm\Delta$

$$\mathsf{M}_\alpha(\Delta) = \int_{-\Delta}^{+\Delta} (\mathsf{M}(\alpha+\epsilon, \beta) - \mathsf{M}(\alpha, \beta))^2 \, \delta\epsilon$$

- linearized standard deviation in Δ

$$\mathsf{m}_\alpha = \lim_{\Delta \Rightarrow 0} \frac{\partial}{\partial \Delta} \mathsf{M}_\alpha(\Delta)$$

- errors in Mueller matrix elements can be written as

$$\mathsf{M}(\alpha, \beta) \pm \mathsf{m}_\alpha \cdot \delta\alpha \pm \mathsf{m}_\beta \cdot \delta\beta$$

Example: Linear Retarder, fast axis angle θ, retardance ϕ

Mueller matrix:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2(2\theta) + \cos(\phi)\sin^2(2\theta) & \cos(2\theta)\sin(2\theta) - \cos(2\theta)\cos(\phi)\sin(2\theta) & \sin(2\theta)\sin(\phi) \\ 0 & \cos(2\theta)\sin(2\theta) - \cos(2\theta)\cos(\phi)\sin(2\theta) & \cos(\phi)\cos^2(2\theta) + \sin^2(2\theta) & -\cos(2\theta)\sin(\phi) \\ 0 & -\sin(2\theta)\sin(\phi) & \cos(2\theta)\sin(\phi) & \cos(\phi) \end{pmatrix}$$

uniform error distribution of retardance in interval $\pm\Delta$ leads to variance of Mueller matrix elements

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2}(4(\Delta - \sin(\Delta)) + \cos(2\phi)(2\Delta - 4\sin(\Delta) + \sin(2\Delta)))\sin^4(2\theta) & \frac{1}{8}(4(\Delta - \sin(\Delta)) + \cos(2\phi)(2\Delta - 4\sin(\Delta) + \\ 0 & \frac{1}{8}(4(\Delta - \sin(\Delta)) + \cos(2\phi)(2\Delta - 4\sin(\Delta) + \sin(2\Delta)))\sin^2(4\theta) & \frac{1}{2}\cos^4(2\theta)(4(\Delta - \sin(\Delta)) + \cos(2\phi)(2\Delta - 4 \\ 0 & \sin^2(2\theta)\left(2(\Delta - 2\sin(\Delta))\sin^2(\phi) + \Delta - \cos(\Delta)\cos(2\phi)\sin(\Delta)\right) & \cos^2(2\theta)\left(2(\Delta - 2\sin(\Delta))\sin^2(\phi) + \Delta - \cos(\end{pmatrix}$$

derivative with respect to Δ

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\sqrt{\sin^4(2\theta)\sin^2(\phi)}}{\sqrt{3}} & \frac{\sqrt{\sin^2(4\theta)\sin^2(\phi)}}{2\sqrt{3}} & \frac{\sqrt{\cos^2(\phi)\sin^2(2\theta)}}{\sqrt{3}} \\ 0 & \frac{\sqrt{\sin^2(4\theta)\sin^2(\phi)}}{2\sqrt{3}} & \frac{\sqrt{\cos^4(2\theta)\sin^2(\phi)}}{\sqrt{3}} & \frac{\sqrt{\cos^2(2\theta)\cos^2(\phi)}}{\sqrt{3}} \\ 0 & \frac{\sqrt{\cos^2(\phi)\sin^2(2\theta)}}{\sqrt{3}} & \frac{\sqrt{\cos^2(2\theta)\cos^2(\phi)}}{\sqrt{3}} & \frac{\sqrt{\sin^2(\phi)}}{\sqrt{3}} \end{pmatrix}$$

Stenflo 1994

13.4. Weakly Polarizing Optical Train

The Mueller matrices of different optical components generally do not commute. Thus for the optical train described by Eq. (13.2) the order in which the components occur is essential. If the beam direction is reversed, the resulting polarization will usually be quite different. For instance, a linear polarizer followed by a quarter-wave retarder ($\lambda/4$ plate) with its fast optical axis at 45° to the polarizing direction produces a circularly polarized output beam. If the beam direction is reversed, linearly polarized output is obtained.

In the case when the optical components are only weakly polarizing, however, symmetry is restored, and the Mueller matrices commute. This greatly simplifies the analysis and calibration of almost polarization-free telescopes. Such telescopes will be increasingly used for precision solar polarimetry in the future. Although "polarization-free" in a first approximation, their small residual polarization (e.g. due to stresses or imperfections) needs to be accounted for.

Let us thus assume that all components of the optical train are weakly polarizing, which means that their Mueller matrices $\boldsymbol{M}_i$ may be written in the form

$$\boldsymbol{M}_i = \boldsymbol{E} + \boldsymbol{m}_i \,, \qquad (13.26)$$

where the components $m_{i,jk} \ll 1$ for all i, j, k. Let us next consider the two components with $i = 1$ and 2. With Eq. (13.26) we get

$$\boldsymbol{M}_2\boldsymbol{M}_1 = \boldsymbol{E} + \boldsymbol{m}_1 + \boldsymbol{m}_2 + \boldsymbol{m}_2\boldsymbol{m}_1 \,. \qquad (13.27)$$

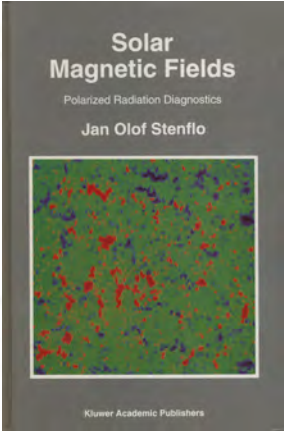
To first order in $\boldsymbol{m}_i$ we have

$$\boldsymbol{M}_2\boldsymbol{M}_1 \approx \boldsymbol{M}_1\boldsymbol{M}_2 \approx \boldsymbol{E} + \boldsymbol{m}_1 + \boldsymbol{m}_2 \,. \qquad (13.28)$$

Commutation has been achieved, because in the weakly polarizing limit matrix multiplication can be converted to matrix summation.

Repeating the above procedure we obtain for the full optical train of Eq. (13.2):

$$\boldsymbol{M} \approx \boldsymbol{E} + \sum_{i=1}^{n} \boldsymbol{m}_i \,. \qquad (13.29)$$



Adding Mueller matrix errors

- for weakly polarizing and retarding elements, RSS can be applied
- for strongly polarizing and retarding elements, we obtain

$$\prod_{i=1}^{n} (\mathsf{M}_i + \mathsf{m}_i) \approx \prod_{i=1}^{n} \mathsf{M}_i + \sum_{i=1}^{n} \left(\prod_{j=1}^{i-1} \mathsf{M}_j \right) \mathsf{m}_i \left(\prod_{j=i+1}^{n} \mathsf{M}_j \right)$$

- Mueller matrix errors need to be transformed using ideal Mueller matrices of elements before and after current element

Outlook

- first attempts at polarimetry systems engineering
- add calibration and data reduction influences
- determine Mueller matrix errors for various parameters
- test against end-to-end simulations and real instruments