Design of optimal polarimeters: maximization of signal-to-noise ratio and minimization of systematic error

J. Scott Tyo

The relationship between system condition and signal-to-noise ratio (SNR) in reconstructed Stokes parameter images is investigated for rotating compensator, variable retardance, and rotating analyzer Stokes vector (SV) polarimeters. A variety of optimal configurations are presented for each class of systems. The operation of polarimeters is discussed in terms of a four-dimensional conical vector space; and the concept of nonorthogonal bases, frames, and tight frames is introduced to describe the operation of SV polarimeters. Although SNR is an important consideration, performance of a polarimeter in the presence of errors in the calibration and alignment of the optical components is also important. The relationship between system condition and error performance is investigated, and it is shown that an optimum system from the point of view of SNR is not always an optimum system with respect to error performance. A detailed theory of error performance is presented, and the error of a SV polarimeter is shown to be related to the stability and condition number of the polarization processing matrices. The rms error is found to fall off as the inverse of the number of measurements taken. Finally, the concepts used to optimize SV polarimeters are extended to be useful for full Mueller matrix polarimeters.

OCIS codes: 280.0280, 260.5430, 230.5440.

1. Introduction

Imaging Stokes vector (SV) polarimeters have been designed and built for many remote sensing applications including medical, atmospheric, and underwater remote sensing.[1–3] Systems have been designed for use in all portions of the optical spectrum from visible to long-wave IR. SV sensors have been demonstrated to improve target contrast, reduce clutter, aid in target detection and identification, provide surface orientation information, give information about surface materials, and aid in the defeat of intervening scatterers for remote sensing in turbid media.

Although polarization is a potentially powerful tool, it can be quite difficult to measure the polarization properties of optical radiation across a scene. Polarimetric optics typically lower system throughput, thereby reducing the signal-to-noise ratio (SNR), and they are often difficult or costly to manufacture in a broadband sense. Although the SV measurement strategies are motivated from laboratory-based active ellipsometry,[4] the practical aspects of real-time, remotely sensed imagery limit frame integration time and necessitate simultaneous sensing at as many as 10^6 pixels. Factors such as these drive imaging polarimeters to make as few measurements as possible to reconstruct the desired polarization information.

Recently, a number of studies have explored optimization of polarimeters designed to measure Stokes parameters. Ambirajan and Look,[5,6] Sabatke *et al.*,[7,8] and Tyo[9,10] analyzed the system matrices that govern polarimeter operation and obtained optimum configurations by minimizing various condition numbers of the polarization processing matrices (L_1, L_2, L_∞, and Frobenius condition numbers). They determined that optimum configurations are obtained when the polarization states analyzed to determine the SV are equally spaced throughout the Poincaré sphere. Tyo[11] examined multichannel linear polarimeters and showed that optimal linear polarimeters have their measurements equally spaced about the equator of the Poincaré sphere by minimizing the correlation between successive measurements. It was asserted initially that the effect of these optimizations was to maximize SNR and minimize the effect of systematic errors, such as miscalibrations in the optical elements.[6] It has been shown recently with

J. S. Tyo (tyo@ieee.org) is with the Electrical and Computer Engineering Department, University of New Mexico, Albuquerque, New Mexico 87131.

Received 1 May 2001; revised manuscript received 24 September 2001.

0003-6935/02/040619-12$15.00/0

synthetic[7,9,10] and experimental data[8,12] that the benefit of increased SNR is obtained. However, to my knowledge, no studies have yet examined the general effect that element miscalibration and misalignment have on the error in the reconstructed SV.

In this paper I address two main points. In Section 2 a new geometric description of polarization information is derived and an underlying optimization strategy is presented that is fundamentally related to the previous optimization studies. An analysis in terms of the L_2 condition number of the processing matrices is developed and related to the previous optimizations that considered a variety of parameters. In Section 3 a detailed theory of systematic error is derived. Predictions and simulations of the mean-square error as a function of system errors are presented. The SNR optimization results are extended to Mueller matrix polarimeters in Section 4, and Section 5 concludes the paper.

2. Geometric Descriptions

The SV is a widely used means to fully characterize the time-averaged polarization properties of incoherent radiation.[4] The SV can be expressed in terms of canonical polarization states as

$$\mathbf{S} = \begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{bmatrix} \langle |E_x(t)|^2 \rangle + \langle |E_y(t)|^2 \rangle \\ \langle |E_x(t)|^2 \rangle - \langle |E_y(t)|^2 \rangle \\ \langle |E_{45}(t)|^2 \rangle - \langle |E_{-45}(t)|^2 \rangle \\ \langle |E_{\mathrm{lcp}}(t)|^2 \rangle - \langle |E_{\mathrm{rcp}}(t)|^2 \rangle \end{bmatrix}, \tag{1}$$

where x, y, 45, and -45 denote linear polarization directions and *lcp* and *rcp* are left- and right-circular polarization, respectively. The Stokes parameter s_0 is related to the total intensity, s_1 and s_2 describe the partial linear polarization information, and s_3 describes the partial circular polarization information. It is straightforward to show that[4]

$$s_1^2 + s_2^2 + s_3^2 \le s_0^2, \tag{2}$$

where the equality holds only for completely polarized radiation. When the equality does not hold, the radiation is said to be partially polarized. Because the intensities of incoherent beams of radiation add, the SV can be decomposed as

$$\begin{aligned} \mathbf{S} &= \mathbf{S}_u + \mathbf{S}_p \\ &= [s_{0,u} \quad 0 \quad 0 \quad 0]^T + [s_0 - s_{0,u} \quad s_1 \quad s_2 \quad s_3]^T, \end{aligned} \tag{3}$$

where $\mathbf{S}_u$ and $\mathbf{S}_p$ are the SVs of the unpolarized and completely polarized portions of the radiation. (Note that all vectors in this paper are shown in bold.)

A common graphical method to depict the polarization state is the Poincaré sphere.[4] We can obtain the position within the Poincaré sphere by plotting s_1/s_0, s_2/s_0, and s_3/s_0 along orthogonal axes in three-dimensional (3-D) Euclidean space. The circular polarization axis is typically taken as the polar axis, so the north pole represents completely left-circular polarization, the south pole represents completely right-circular polarization, and the equator gives the locus of all completely linearly polarized states. Completely elliptically polarized states lie on the surface of the Poincaré sphere, and partially polarized states are in the interior.

Although the concept of the Poincaré sphere is extremely powerful, there is one major counterintuitive property about it—orthogonal polarization states are not represented by orthogonal vectors on the Poincaré sphere. Instead, orthogonal polarization states occupy opposite ends of a single diameter of the Poincaré sphere. This problem can easily be remedied when we generalize the 3-D Poincaré sphere into a four-dimensional (4-D) Stokes cone, as is presented below.

A. *N*-Dimensional Cones and Spheres

For the purposes of this paper, an N-dimensional sphere is defined by

$$\sum_{i=1}^{N} \xi_i^2 \le r^2, \qquad N \ge 1, \tag{4}$$

where r is the radius and ξ_i are the coordinates. An $(N + 1)$-dimensional circular cone is defined by

$$\sum_{i=1}^{N} \xi_i^2 \le (\alpha \xi_0)^2, \qquad N \ge 1. \tag{5}$$

When $\alpha = 1$, the cone will be termed a unit cone. Only the upper half of the cone will be considered here, i.e., $\xi_0 \ge 0$. An important property of circular cones is that the cross section of an $(N + 1)$-dimensional circular cone perpendicular to the ξ_0 axis is an N-dimensional sphere. A cross section of an $(N + 1)$-dimensional circular cone that includes the ξ_0 axis is an N-dimensional cone. These relationships are highlighted in Fig. 1.

The Stokes parameters $\{s_0, s_1, s_2, s_3\}$ define a 4-D cone, as inequality (2) is equivalent to inequalities (5) with $\alpha = 1$. This cone represents all physically realizable polarization states and is called here the 4-D Stokes cone. The s_0 axis is the polar axis of this cone, and a cross section perpendicular to the s_0 axis yields the 3-D Poincaré sphere. Similarly, if a cross section of this cone is taken perpendicular to the s_3 axis, we have a 3-D cone that represents all partially linearly polarized states. If this 3-D cone is now projected on a plane perpendicular to the s_0 axis, a two-dimensional (2-D) sphere (circle) remains that represents partial linear polarization. This 2-D circle is equivalent to the equator of the Poincaré sphere in three dimensions. Other cross sections of the general 4-D cone might be useful for other applications.[15,16] Although it is difficult to visualize the full 4-D Stokes cone, the 3-D projection for linear polarization information is presented in Fig. 1.

Next I derive a theory concerning orthogonal directions in N-dimensional cones. The space $\mathscr{C}$ represents an $(N + 1)$-dimensional cone. The space $\mathscr{C} \subset \mathbb{R}^{N+1}$, so not all vectors $\mathbf{x} \in \mathscr{C}$ are guaranteed to have an orthogonal vector $\mathbf{y} \in \mathscr{C}$.

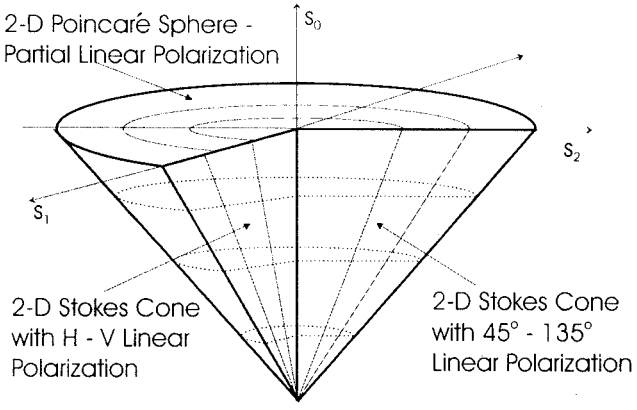

Fig. 1. Schematic of the 3-D Stokes cone that represents all possible partially linearly polarized states. The three Cartesian axes are given by s_0, s_1, and s_2, but only states that satisfy inequality (2) are physically realizable, resulting in the conical space. States on the surface of the cone are completely linearly polarized. This figure presents three cross sections. The first is parallel to the s_1 and s_2 axes and represents a 2-D Poincaré sphere (circle) of partially linearly polarized states. This plane is the projection of the traditional 3-D Poincaré sphere onto its equatorial plane. The second and third projections are in the s_0–s_1 and s_0–s_2 planes. These projections are 2-D Stokes cones that represent polarization differences between orthogonal linear polarization states.[11,13,14] The 3-D conical space is itself a projection of the 4-D Stokes cone onto the s_0–s_1–s_2 hyperplane.

Theorem 1: Given an $(N + 1)$-dimensional circular cone $\mathcal{C}$ with $\alpha < 1$, for any $\mathbf{x} \in \mathcal{C}$ there does not exist a nonzero vector $\mathbf{y} \in \mathcal{C}$ such that $\mathbf{x} \cdot \mathbf{y} = 0$.

Proof: Without loss of generality, we choose $\mathbf{x}$ to lie in the ξ_0–ξ_1 plane with $\xi_{1,x} \leq \alpha\xi_{0,x} < \xi_{0,x}$. Next, we choose an arbitrary vector $\mathbf{y} \in \mathcal{C}$, form the inner product, and assume that the vectors are orthogonal as

$$\mathbf{x} \cdot \mathbf{y} = \xi_{0,x}\xi_{0,y} + \xi_{1,x}\xi_{1,y} = 0. \tag{6}$$

Equation (6) is satisfied when

$$\xi_{1,y}{}^2 = (-\xi_{0,x}/\xi_{1,x})^2\xi_{0,y} > \xi_{0,y}{}^2. \tag{7}$$

The requirement in Eq. (7) violates inequalities (5). Q.E.D.

Corollary 1.1: Given an $(N + 1)$-dimensional circular unit cone $\mathcal{U}$, for $\mathbf{x} \in \mathcal{U}$ there exists a nonzero vector $\mathbf{y} \in \mathcal{U}$ such that $\mathbf{x} \cdot \mathbf{y} = 0$ if and only if $\mathbf{x}$ lies on the surface of $\mathcal{U}$, i.e., $\sum_{i=1}^{N} \xi_i{}^2 = \xi_0{}^2$. Furthermore, the vector $\mathbf{y} \in \mathcal{U}$ orthogonal to $\mathbf{x}$ is unique, lies on the surface of $\mathcal{U}$, and is in the same plane as $\mathbf{x}$ and the ξ_0 axis.

Proof: Without loss of generality, we choose $\mathbf{x}$ to lie in the first quadrant of the ξ_0–ξ_1 plane. For $\mathbf{x}$ not on the surface of $\mathcal{U}$, see Eqs. (6) and (7) above. For $\mathbf{x}$ on the surface of $\mathcal{U}$, $\xi_{1,x} = \xi_{0,x}$. A vector $\mathbf{y}$ is orthogonal to $\mathbf{x}$ when

$$\xi_{1,y} = -(\xi_{0,x}/\xi_{1,x})\xi_{0,y} = -\xi_{0,y}. \tag{8}$$

Equation (8) is satisfied by many vectors, but inequalities (5) dictate that for $\mathbf{y} \in \mathcal{U}$, $\xi_2, \ldots, \xi_N = 0$. Q.E.D.

As discussed above, the Stokes cone of physically realizable states, which is referred to hereafter as $\mathcal{S}$, is a 4-D subset of $\mathbb{R}^4$. The above theorem and corollary provide a mathematical verification (and explanation) for the well-known facts that a partially polarized state has no orthogonal polarization state and that a completely polarized state has exactly one orthogonal polarization state.[17] The completely polarized portion of Eq. (3) lies on the surface of $\mathcal{S}$ and has an orthogonal state. When the unpolarized portion of Eq. (3) is added, the total vector moves to the interior of $\mathcal{S}$.

B. Use of Polarimeters to Measure the Stokes Parameters

At optical wavelengths, detectors often respond to intensity only. They are usually only weakly sensitive to polarization information and are almost always insensitive to relative phase information. For this reason, it is usually impossible to measure the Stokes parameters directly; instead, an optical system must be designed that is polarization sensitive and tunable. By making several observations of intensity at various settings of the system parameters, we can build up a set of linear equations that can then be solved in a least-squares sense (assuming that $N \geq M$, where N is the number of measurements and M is the number of Stokes parameters to be reconstructed). Many different strategies have been developed to reconstruct the Stokes parameters, including rotating $\lambda/4$ wave-plate systems,[18] general rotating retarder (RR) systems,[7] variable retardance (VR) systems,[19] rotating analyzer systems with mechanical and electro-optical rotation,[13,14,20,21] photoelastic modulator-based systems,[15] four-photodetector systems with no additional polarization-sensitive optics,[22] and interlaced pixel arrays.[23]

Regardless of the strategy employed, the concept behind the polarimeter is the same. The optical system is composed of polarization-sensitive elements, each of which has a Mueller matrix that relates the incident SV to the output SV, and the entire system can be cascaded into a single Mueller matrix.[4] Common strategies for the design of an ideal optical system are composed of one or more ideal retarder elements that do not affect intensity and an ideal polarizer that extinguishes one polarization state (usually linear) and passes the orthogonal state without attenuation. With a combination of ideal retarders and a single ideal analyzer, the entire polarimeter can be thought of as an ideal elliptical diattenuator that passes one elliptically polarized state without modifying its intensity (the output state of the diattenuator is often not the same as the input state that passes with maximum intensity transmission) and completely extinguishes the orthogonal elliptically polarized state. For optical detectors that respond only to intensity, only the s_0 term of the output SV need be calculated. We can accomplish this by tak-

ing the inner product of the first row of the composite Mueller matrix with the input SV as

$$s_{0,\mathrm{out}}^{(i)} = M_{00}^{(i)} s_0 + M_{01}^{(i)} s_1 + M_{02}^{(i)} s_2 + M_{03}^{(i)} s_3 = [\underline{\mathbf{M}}_0^{(i)}]^T \cdot \mathbf{S}, \tag{9}$$

where $[\underline{\mathbf{M}}_0^{(i)}]$ is the first row of the composite Mueller matrix for the ith configuration of the polarimeter.[24] (Note that all matrices in this paper are shown in bold with an underline.) By taking a set of measurements, we can build up a system of linear equations allowing a solution for the unknown Stokes parameters in terms of the measured intensities. The linear system can be described as

$$\mathbf{I} = \underline{\mathbf{A}} \cdot \mathbf{S}, \qquad \mathbf{S} = \underline{\mathbf{B}} \cdot \mathbf{I}, \tag{10}$$

where $\mathbf{I}$ is the vector of the observed intensities and $\underline{\mathbf{B}} = \underline{\mathbf{A}}^{-1}$. The ith row of $\underline{\mathbf{A}}$ is $[\underline{\mathbf{M}}_0^{(i)}]^T$. Here the matrix $\underline{\mathbf{A}}$ is termed the analysis matrix, and the matrix $\underline{\mathbf{B}}$ is the synthesis matrix.

Lu and Chipman[25] showed that the first row of the Mueller matrix for an ideal diattenuator is the SV of the state that passes the diattenuator. This is termed here the principal axis of the elliptical polarizer. Therefore the ith element of the vector $\mathbf{I}$ represents $\mathbf{S}_D^{(i)} \cdot \mathbf{S}$, where $\mathbf{S}_D^{(i)}$ is the principal axis of the ith configuration of the polarimeter and $\mathbf{S}$ is the input SV. As the parameters of the optical system are varied and the Mueller matrix changes, so does the principal axis of the composite Mueller matrix. The operation of a polarimeter is therefore equivalent to one taking projections of $\mathbf{S}$ onto several different unit vectors in $\mathbb{R}^4$. Because these vectors must also be elements of $\mathcal{S}$, they cannot form an orthogonal basis if $N > 2$ (see Theorem 1); and when $N > M$, the decomposition is in terms of a frame rather than a basis as discussed in Subsection 2.C.

C. Bases and Frames on the Stokes Cone

The operation of a polarimeter is suggestive of the input SV being decomposed over a basis in $\mathbb{R}^4$. It is clear from the above discussion that $\mathcal{S}$ is a 4-D subset of $\mathbb{R}^4$; however, it is not possible to construct an orthogonal basis for $\mathcal{S}$ out of elements of $\mathcal{S}$. Instead, a nonorthogonal set must be chosen. Furthermore, there is nothing restricting one to making only four measurements to reconstruct the SV. Many systems have been developed that introduce redundancy by making more measurements than are strictly necessary (i.e., making five or more measurements to reconstruct the four-element SV), in which case the decomposition vectors in Eq. (9) are linearly dependent and cannot be a basis. The mathematical tool of frames has been introduced to deal with such situations.

Frames and tight frames are extensions of bases and orthogonal bases. A frame is a set of vectors that spans a linear space, but is not necessarily linearly independent. There is a rich literature on frames in the signal processing literature,[26,27] and the theory is only briefly examined here.

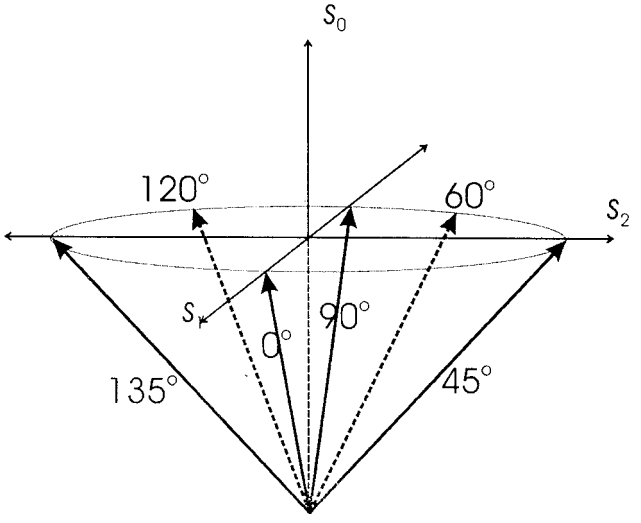

Fig. 2. Schematic of the 3-D Stokes cone with some important bases and frames. The basis formed by the vectors marked 0°, 60°, and 120° form the optimum three-measurement system.[11] The vectors marked 0°, 45°, and 90° form a second common basis that was shown to be suboptimal.[11] This basis, along with the fourth vector marked 135°, forms a four-element frame, the optimal system for $N = 4$, as proposed by Walraven.[20]

Consider the set of N basis vectors $\{\mathbf{x}_i\}_1^N$ in $\mathbb{R}^N$. An arbitrary vector $\mathbf{y} \in \mathbb{R}^N$ can be represented as

$$\mathbf{y} = \sum_{i=1}^{N} (\tilde{\mathbf{x}}_i \cdot \mathbf{y}) \mathbf{x}_i, \tag{11}$$

where $\{\tilde{\mathbf{x}}_i\}_1^N$ is the dual basis of $\{\mathbf{x}_i\}_1^N$, such that $\tilde{\mathbf{x}}_i \cdot \mathbf{x}_j = \delta_{ij}$, and δ_{ij} is the Kronecker delta. When $\{\mathbf{x}_i\}_1^N$ is orthonormal, $\tilde{\mathbf{x}}_i = \mathbf{x}_i$. The dual basis in Eq. (11) serves as the analysis tool; i.e., it gives the amount of each of the basis vectors that is needed to reconstruct $\mathbf{y}$. For a frame, the number of unit vectors might be larger than the dimensionality of the space. In some cases, a minimum-energy decomposition of the form of Eq. (11) can still be written. When such a decomposition exists, a Parseval-like relation exists[27]:

$$A(\mathbf{y} \cdot \mathbf{y}) \le \sum_{i=1}^{N} |\tilde{\mathbf{x}}_i \cdot \mathbf{y}|^2 \le B(\mathbf{y} \cdot \mathbf{y}), \tag{12}$$

where N is the number of unit vectors in the frame and $0 \le A \le B < \infty$. For tight frames $A = B$, and for orthogonal bases $A = B = 1$. Note that in the structure of Eq. (10) the dual basis contains the principal axes of the polarimeter.

As an example that can be visualized, consider the decomposition of the linear polarization information contained in s_0, s_1, and s_2 by a system composed of only linear polarization-sensitive elements. The 3-D cone in Fig. 2 represents all possible partially linearly polarized states. The set of three measurements that forms the optimum polarimeter (unique to a rotation of all three measurements by a constant angle) occurs when three linearly polarized measurements are made at $\phi_1 = 0°$, $\phi_2 = 60°$, and $\phi_3 = 120°$.[11]

This set of vectors is depicted in Fig. 2. A second common basis set where $\phi_1 = 0°$, $\phi_2 = 45°$, and $\phi_3 = 90°$ is also depicted in Fig. 2. The former case forms a frame with

$$A = 0.5625, \qquad B = 0.75. \tag{13}$$

The latter case makes a frame with

$$A = 0.25, \qquad B = 0.8536. \tag{14}$$

The first set is the optimum configuration for three measurements in 3-D polarization space.[11] An explanation of how this optimization is obtained is presented in Subsection 2.D. It should be noted that the reconstruction bases for both of these cases are formed of vectors that are not elements of $\mathcal{S}$. If more than three measurements are made, as done by Walraven[20] and others, then the optimal configuration is to space the linear polarization measurements out evenly between 0° and 180°.[11]

D. Optimization of Stokes Vector Polarimeters

Sabatke *et al.*[7,8] and Tyo[9] have shown that the SNR in reconstructed Stokes parameter images is maximized and equalized when the various condition numbers of the analysis and synthesis matrices defined in Eqs. (10) are minimized.[28] The condition number of a matrix is defined in terms of the matrix norms as

$$\kappa(\underline{\mathbf{A}}) = \|\underline{\mathbf{A}}\| \, \|\underline{\mathbf{A}}^{-1}\|. \tag{15}$$

The choice of matrix norm is left to the user,[29] but the one used here is the L_2 norm:

$$\|\underline{\mathbf{A}}\|_2 = \sup_{\mathbf{x}} \frac{\|\underline{\mathbf{A}} \cdot \mathbf{x}\|_2}{\|\mathbf{x}\|_2}, \tag{16}$$

where $\|\mathbf{x}\|_2$ is the Euclidean length of the vector $\mathbf{x}$. The reasons for this choice of norm are discussed in detail by Tyo,[9] and a detailed discussion of the relative merits of various norms is given by Sabatke *et al.*[8]

Optimum configurations for 4-D systems to make four measurements were obtained when the unknown input SV was analyzed with four elliptical diattenuators whose principal directions form a regular tetrahedron on the Poincaré sphere.[30] The effect of such an optimization is demonstrated by Tyo[9] with simulated data and by Sabatke *et al.*[8] and Tyo and Turner[19] with experimental data. Not all polarimeter configurations can produce a global optimum, and the optimum configuration for a particular polarimeter strategy might not be unique.[9] For rotating compensator systems,[7] the retardance of the compensator is fixed, and various fast-axis orientations are chosen to construct the system matrices $\underline{\mathbf{A}}$ and $\underline{\mathbf{B}}$. For a given retardance, the condition number can be minimized, and the minimum condition number is presented as a function of retardance in Fig. 3. There is a unique global optimum retardance at $\delta = 0.3661\lambda$,[7] and at this retardance there are two sets of angles that achieve the optimal condition. It is clear from Eqs. (15) and (16) that a matrix with a small L_2 norm will map the small vector due to observation noise into a small error vector upon reconstruction. Because both the analysis and the synthesis matrices influence $\kappa_2(\underline{\mathbf{A}})$, the matrix must be well conditioned with respect to inversion.

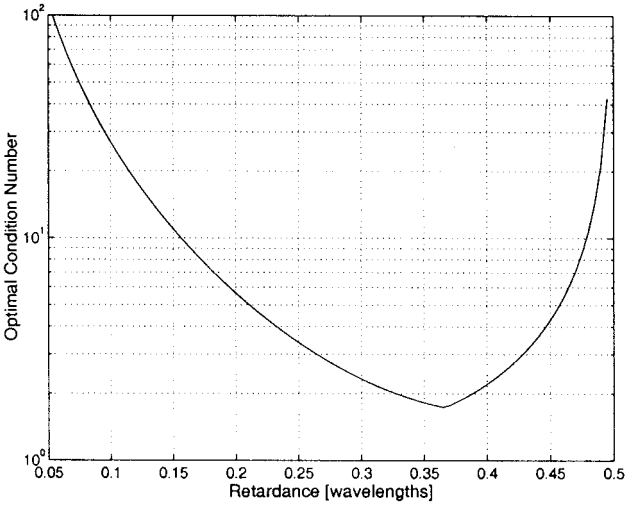

Fig. 3. Optimal condition number for four-measurement rotating compensator systems. There is a clear optimum at $\delta = 0.3661\lambda$. I obtained the optimization by minimizing the L_2 condition number of the system matrices, but achieved the same result as Sabatke *et al.*,[7] where the equal-weighted variance—equivalent to the Frobenius condition number of the system matrices—was minimized.

For VR systems, the orientations of two retarders are fixed, and sets of retardance values are chosen to construct $\underline{\mathbf{A}}$ and $\underline{\mathbf{B}}$.[19,31] For each pair of retarder orientation angles, the system can be optimized when we choose the best four pairs of retardance values. These results are presented in Fig. 4. As can be seen from Fig. 4, the VR system has many more possible optimum configurations, and this is related to the extra degree of freedom that is available in the design of a VR system as opposed to a RR system. A wider range of elliptical diattenuators can be synthesized with a VR system than a RR system, providing access to more of the Poincaré sphere. If the RR system were changed to a rotating variable retarder (adding a degree of freedom), similar coverage of the Poincaré sphere could be realized. A similar optimization can be performed for linear polarimeters. These devices make three (or more) measurements to reconstruct the first three Stokes parameters. The 2 three-measurement systems described in Eqs. (13) and (14) have condition numbers of $\sqrt{2}$ and $1 + \sqrt{2}$, respectively. It can be seen that the tighter frame is the better conditioned system.

Systems are often developed that make more than $N = 4$ measurements to introduce redundancy and make the system less sensitive to errors and noise. I show in Section 3 that making more measurements can decrease the susceptibility of the polarimeter to errors such as random angular positioning errors in a RR system. Because of the improvement in error performance realized by an increase in N, the optimum angles for $N = 4$, 6, and 8 measurements in a

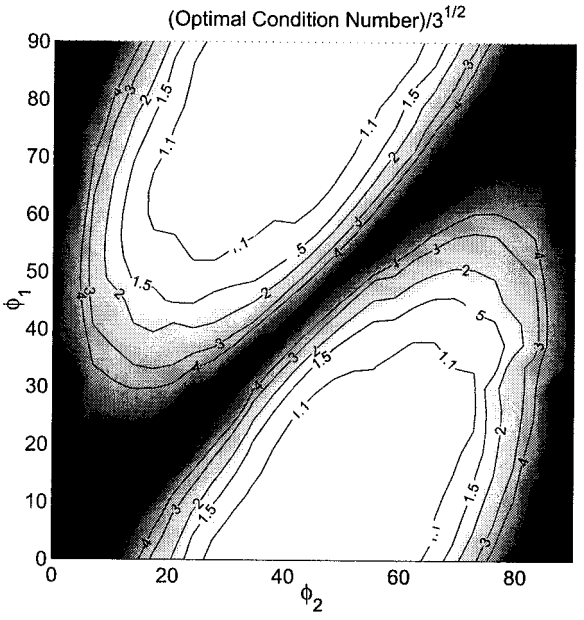

Fig. 4. Optimal condition number for VR systems showing the lowest possible condition number for a VR polarimeter with the given fast-axis orientation values of ϕ_1 and ϕ_2. The jaggedness of the contours is due to coarse sampling in ϕ_1 and ϕ_2, but that should not detract from the overall message. All ϕ_1–ϕ_2 pairs inside the lowest contour have a continuum of optimal configurations.[9] Although the condition number of all optimal configurations is the same, the best set of angles is $\phi_1 = 45°$, $\phi_2 = 0°$. This point is the geometric center of the optimal region and provides access to the entire Poincaré sphere. For a detailed description of this optimization, see Tyo.[10]

RR polarimeter were calculated and are presented in Table 1. These optimum configurations were all realized with $\delta = 0.3661\lambda$.

An interesting feature of Figs. 3 and 4 and Table 1 is the actual value of the optimum condition number. For both VR and RR systems, the absolute minimum condition number obtained numerically is approximately $\sqrt{3}$. Furthermore, the optimal condition number for the 3-D rotating analyzer systems was $\sqrt{2}$. It is shown in Appendix A that the minimum L_2 condition number of the system matrices is $(M - 1)^{1/2}$, where M is the number of reconstructed Stokes parameters. In Appendix B, relationships for the frame bounds in inequality (12) for optimized systems are derived.

3. Relationship between System Condition and Systematic Errors

To this point, the optimization has been in terms of SNR with respect to additive noise that is due to the detection process. Additional sources of errors in polarimeters are element misalignment and miscalibration.[4,32] In this section, expressions describing the sensitivity to such errors are derived, and the relationship between system condition and error minimization is discussed.

A. Theoretical Formulation

The theory developed here is applicable to a polarimeter with any number of parameters. The specific example of a one-parameter system, a RR polarimeter, is examined for simplicity. The process to analyze an unknown SV with an imperfect polarimeter to make N measurements and reconstruct an estimate of the SV is

$$\mathbf{S}' = \underline{\mathbf{B}}\underline{\mathbf{A}}' \cdot \mathbf{S}, \tag{17}$$

where $\underline{\mathbf{A}}'$ is the perturbed analysis matrix ($N \times 4$) when we consider errors in the parameters of the system (e.g., azimuthal settings of the retarder in a RR system), and $\underline{\mathbf{B}}$ is the ideal synthesis matrix ($4 \times N$). If $\underline{\mathbf{A}}' = \underline{\mathbf{A}} + \underline{\mathbf{\Delta}}$, then

$$\boldsymbol{\epsilon} = \mathbf{S}' - \mathbf{S} = (\underline{\mathbf{B}}\underline{\mathbf{A}}' - \mathbf{I})\mathbf{S} = \underline{\mathbf{B}}\underline{\Delta} \cdot \mathbf{S} \tag{18}$$

is the 4-D vector representing the error in the reconstructed SV, with $\mathbf{I}$ the 4×4 identity matrix. For a RR polarimeter, when we assume only errors in alignment,

$$\begin{aligned}\underline{\mathbf{\Delta}}_{ij} &= (\underline{\mathbf{A}}' - \underline{\mathbf{A}})_{ij} \\ &= \delta_i \left[\frac{\underline{\mathbf{M}}_{0j}(\phi_i + \delta_i) - \underline{\mathbf{M}}_{0j}(\phi_i)}{\delta_i} \right] \\ &\approx \delta_i \left. \frac{\partial \underline{\mathbf{M}}_{0j}(\phi)}{\partial \phi} \right|_{\phi_i},\end{aligned} \tag{19}$$

where ϕ_i is the nominal azimuthal setting of the retarder for the ith measurement, δ_i is the angular position error, and $\underline{\mathbf{M}}$ is the composite Mueller matrix of the system. The approximation in Eq. (19), which is the first term of a Taylor-series expansion of $\underline{\mathbf{A}}$ about the nominal settings, is good when the angular position error is small (compared with π radians). Note that $\underline{\mathbf{\Delta}}_{i0} = 0$, as $\underline{\mathbf{M}}_{01} = 1/2$ for all ideal elliptical

Table 1. Optimum Retarder Positioning Angles for Four-, Six-, and Eight-Measurement RR Systems[a]

N	$\{\phi\}_{i=1}^{N}$	CN
4	(±15.12°, ±51.69°) (±74.88°, ±38.31°)	1.7321
6	±(10.06°, ±36.76°, ±59.63°) (30.37°, ±53.24°, ±79.94°)	1.7321
8	(±12.40°, ±36.14°, ±49.66°, ±72.07°) (±17.98°, ±40.34°, ±53.86°, ±77.60°)	1.7322

[a]The two sets of angles are complements of each other. The third column gives the condition number (CN) of the optimal configuration for each N with $\delta = 0.3661\lambda$, each of which is approximately $\sqrt{3}$. For rotating analyzer systems, the optimum configuration is obtained when the N measurements are equally spaced between 0° and 180°. For RR systems, the optimum angles are not equally spaced, as the path traced out on the surface of the Poincaré sphere is more complicated than in rotating analyzer systems.

diattenuators. When errors exist in responsivity calibration, $\underline{\Delta}_{i0} \neq 0$. If this is the case, then Eq. (19) should be modified. To extend this theory to a polarimeter with more than one parameter (e.g., a VR system with two variable parameters, the retardance of the two retarders), the Taylor series in Eq. (19) should be generalized to a multivariable Taylor series.

The error metric computed here is

$$\|\boldsymbol{\epsilon}\|_2^2 = \|\underline{\mathbf{B}}\underline{\boldsymbol{\Delta}} \cdot \mathbf{S}\|_2^2 = \sum_{i=0}^{3} (\underline{\mathbf{B}}\underline{\boldsymbol{\Delta}} \cdot \mathbf{S})_i^2, \qquad (20)$$

where the 2 norm was chosen arbitrarily, and Eq. (20) corresponds to the mean-square Euclidean length of the reconstruction error vector. The azimuthal positioning error of the retarder is assumed to be independent and identically distributed, i.e., $E[\delta_i \delta_j] = \sigma^2 \delta_{ij}$, where δ_{ij} is the Kronecker delta, σ is the rms error in azimuthal position (in radians), and $E[\cdot]$ is an expectation. Forming $(\underline{\mathbf{B}}\underline{\boldsymbol{\Delta}} \cdot \mathbf{S})_i^2$ and taking the expected value over the error, we obtain

$$\begin{aligned} E_\delta[(\underline{\mathbf{B}}\underline{\boldsymbol{\Delta}}\mathbf{S})_i^2] &= \left[\sum_j \underline{\mathbf{B}}_{ij}\delta_j(\underline{\Delta}_{j1}s_1 + \underline{\Delta}_{j2}s_2 + \underline{\Delta}_{j3}s_3)\right] \\ &\times \left[\sum_k \underline{\mathbf{B}}_{ik}\delta_k(\underline{\Delta}_{k1}s_1 + \underline{\Delta}_{k2}s_2 + \underline{\Delta}_{k3}s_3)\right] \\ &= \sum_j \sum_k \underline{\mathbf{B}}_{ij}\underline{\mathbf{B}}_{ik}(\underline{\Delta}_{j1}s_1 + \underline{\Delta}_{j2}s_2 + \underline{\Delta}_{j3}s_3) \\ &\times (\underline{\Delta}_{k1}s_1 + \underline{\Delta}_{k2}s_2 + \underline{\Delta}_{k3}s_3)E(\delta_j\delta_k) \\ &= \sigma^2 \sum_{j=0}^{3} (\underline{\mathbf{B}}_{ij})^2[(\underline{\Delta}_{j1})^2 s_1^2 + (\underline{\Delta}_{j2})^2 s_2^2 \\ &+ (\underline{\Delta}_{j3})^2 s_3^2 + 2\underline{\Delta}_{j1}\underline{\Delta}_{j2}s_1s_2 + 2\underline{\Delta}_{j1}\underline{\Delta}_{j3}s_1s_3 \\ &+ \underline{\Delta}_{j2}\underline{\Delta}_{j3}s_2s_3]. \end{aligned} \qquad (21)$$

The relationship $\underline{\Delta}_{j0} = 0$ is used to simplify Eq. (21). When the error is not independent and identically distributed, Eq. (21) will be more complicated.

Equation (21) can be further simplified when we assume that the input polarization states are uniformly distributed over the Poincaré sphere. In that case, $E(s_i s_j) = (1/3)\delta_{ij}$ with $i, j \neq 0$.[33] Taking the expectation of $(\underline{\mathbf{B}}\underline{\boldsymbol{\Delta}} \cdot \mathbf{S})_i^2$ over the Poincaré sphere, we obtain

$$E[(\underline{\mathbf{B}}\underline{\boldsymbol{\Delta}} \cdot \mathbf{S})_i^2] = \frac{\sigma^2}{3} \sum_{j=0}^{3} (\underline{\mathbf{B}}_{ij})^2[(\underline{\Delta}_{j1})^2 + (\underline{\Delta}_{j2})^2 + (\underline{\Delta}_{j3})^2]. \qquad (22)$$

Summing over i yields

$$E(\|\boldsymbol{\epsilon}\|_2^2) = E(\|\underline{\mathbf{B}}\underline{\boldsymbol{\Delta}} \cdot \mathbf{S}\|_2^2) = \frac{\sigma^2}{3} \sum_{i=0}^{3} \sum_{j=0}^{3} (\underline{\mathbf{B}}_{ij})^2 \sum_{k=0}^{3} (\underline{\Delta}_{jk})^2. \qquad (23)$$

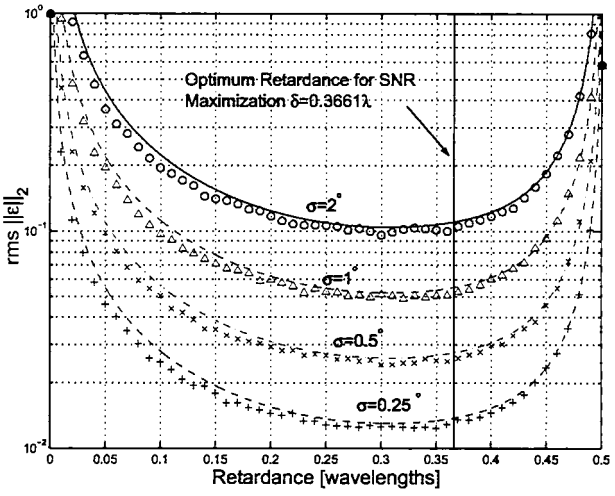

Fig. 5. Numerical simulation and analytic prediction of rms error in the reconstructed SV for RR systems. I calculated the rms error by generating 10^6 random polarization states and simulating the polarimeter operation with 50 different realizations of error in the angular setting of the compensator. The error was generated by use of normal statistics with a standard deviation of σ. Analytic curves were generated with Eq. (23).

A straightforward calculation can be used to show that

$$\frac{\sigma^2}{3} \sum_{i=0}^{3} \sum_{j=0}^{3} (\underline{\mathbf{B}}_{ij})^2 \sum_{k=0}^{3} (\underline{\Delta}_{jk})^2 = E(\|\underline{\mathbf{B}}\underline{\boldsymbol{\Delta}}\|_F^2), \qquad (24)$$

where

$$\|\underline{\mathbf{A}}\|_F^2 = \sum_{ij} |\underline{\mathbf{A}}_{ij}|^2 \qquad (25)$$

is the Frobenius norm of the matrix $\underline{\mathbf{A}}$.[29] The implications of Eqs. (23) and (24) are discussed in Subsection 3.B.

B. Numerical Simulation

A Monte Carlo simulation was prepared to test the above results for rotating compensator systems. The retardance of the rotating compensator was taken to be between 0 and $\lambda/2$. For each retardance, the system was designed with the optimum angles given in Table 1 for $N = 4$, and 50 realizations of the system were generated with rms azimuthal error σ. Each system operated on 10^6 randomly generated SVs that were uniformly distributed over the surface of the Poincaré sphere, and the Euclidean length of the reconstructed error was computed. The rms error was them computed for several values of σ, and the results are presented in Fig. 5 along with the predictions from Eq. (23).

The analytic predictions closely match the computed values of error from the simulation. There is a broad optimum that occurs from a retardance value of 0.22λ to 0.38λ where the rms error is within 20% of the optimum value. This is in contrast to the SNR optimization presented in Fig. 3 where the condition number increases rapidly away from the optimum.

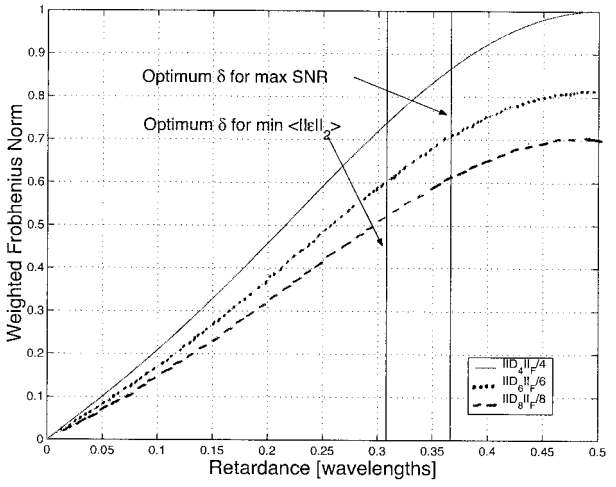

Fig. 6. Frobenius norm of the perturbation matrix $\underline{\mathbf{\Delta}}$ as a function of retardance of the rotating compensator. $\underline{\mathbf{\Delta}}$ was computed for $N = 4$, 6, and 8 at the optimum angles given in Table 1 and divided by the number of measurements. Note that at $\delta = 0.3661\lambda$ where the condition number of $\underline{\mathbf{A}}$ and $\underline{\mathbf{B}}$ is minimized, $\|\underline{\mathbf{\Delta}}\|_F$ is relatively large.

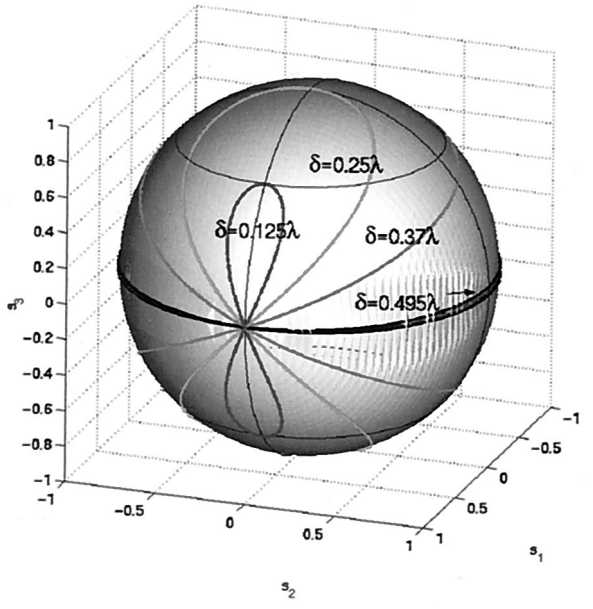

Fig. 7. Trajectories traced out on the surface of the Poincaré sphere for RR polarimeters as the retarder angle is altered. After Sabatke *et al.*,[7] Fig. 1. For $\delta = 0.125\lambda$, the condition number is high, but the total trajectory length is small as the retarder rotates. For $\delta = 0.37\lambda$, the condition number is minimized, but the trajectory on the Poincaré sphere is longer, hence $\|\underline{\mathbf{\Delta}}\|_F$ is large.

Furthermore, the retardance that produces the absolute minimum rms error is at $\delta = 0.3073\lambda$, which is noticeably different from the optimum of $\delta = 0.3661\lambda$ used to maximize SNR.

It is somewhat counterintuitive that the system with the minimum condition number would not produce the minimum rms error. The reason for this can be discovered when we reexamine Eqs. (17)–(25). The expected value of the error vector is controlled by both the synthesis matrix $\underline{\mathbf{B}}$ and the perturbation matrix $\underline{\mathbf{\Delta}}$. Equations (15) and (16) indicate that matrices with low L_2 norms will map small offset vectors into small error vectors. However, a well-conditioned matrix is not necessarily stationary with respect to the system parameters. The expected length of the error in the $\mathbf{I}$ measurements in Eqs. (10) is determined solely by the variance of the noise. In contrast, the error in $\mathbf{I}$ in Eq. (18) is determined solely by the matrix $\underline{\mathbf{\Delta}}$. Even when $\underline{\mathbf{B}}$ is well conditioned, the size of the reconstruction error $\|\boldsymbol{\epsilon}\|_2$ can be large if $\|\underline{\mathbf{\Delta}}\|_F$ is large.

The perturbation matrix is a measure of how sensitive the principal direction of the Mueller matrix (and hence the analysis matrix $\underline{\mathbf{A}}$) is to changes in the system parameters. As a measure of this sensitivity, $\|\underline{\mathbf{\Delta}}\|_F$ is plotted as a function of retardance in Fig. 6. At small values of retardance, $\|\underline{\mathbf{\Delta}}\|_F$ is small; as the retarder has little or no retardance, it does not significantly modify the polarization state of the radiation. As the retardance increases, so does $\|\underline{\mathbf{\Delta}}\|_F$, implying that the principal direction of the polarimeter is changing more rapidly. The trajectories traced on the Poincaré sphere as a retarder is rotated in front of an analyzer are presented in Fig. 7. The smaller the retardance, the shorter the total path length, and hence the principal direction is less sensitive to azimuthal error. The sensitivity effect is balanced by the fact that, at small (or large) retardances, the system is highly ill-conditioned, hence the optimum retardance for minimization of rms error is somewhat less than for maximization of SNR, representing a compromise between the two effects.

C. Overdetermined Systems

Many polarimeter systems have been designed to overspecify the SV by making more than four measurements to reconstruct the full SV. This was done to help minimize the effects of noise and error on any one measurement. Sabatke *et al.*[7] demonstrated that SNR could be reduced without increasing complexity by taking the extra time that would be necessary to make more than M measurements to simply increase the integration time at each of the M settings of the polarimeter. If SNR were the only consideration, there would be no need to make more than M measurements. However, Eq. (23) can be used to predict the benefit in reduced rms error that can be realized when we add redundancy to the polarimeter. Figure 8 shows analytic and Monte Carlo results for the rms error vector length when the SV is reconstructed from optimal four-, six-, and eight-measurement RR systems. The overall shape of the contours is similar, but the addition of redundancy enhances the rms error of the polarimeter. Figure 9 shows the predicted rms error for an optimum RR polarimeter ($\delta = 0.3661\lambda$) as a function of the number of measurements. The data fall almost exactly on a

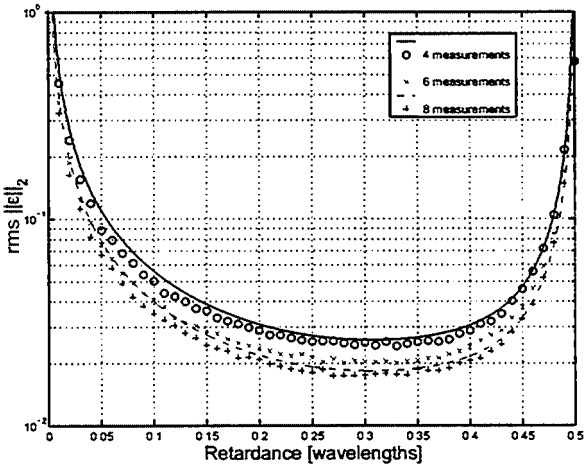

Fig. 8. Predicted and simulated rms error for the optimum RR system with $N = 4$, 6, and 8. Data were generated with the same procedures as in Fig. 5 with system parameters as given in Table 1.

line in log–log space, and a least-squares fit to the data reveals that, for optimized systems,

$$\|\epsilon\|_2 = \frac{b_0}{N}, \tag{26}$$

where N is the number of measurements made in the RR polarimeter, and those measurements are made at the optimal angles for that particular N. The parameter b_0 will depend on the rms error in azimuthal position. As expected, when we add additional, redundant measurements, the rms error performance improves for the same amount of azimuthal error in the parameter settings.

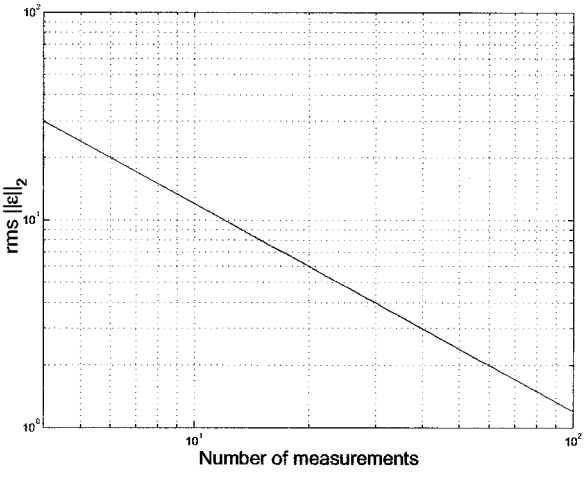

Fig. 9. Predicted rms error for optimum RR system ($\delta = 0.3661\lambda$) as a function of the number of measurements made. At large N, the optimization is slow, and it is not always possible to find an optimum where the condition number of $\underline{\mathbf{A}}$ is $\sqrt{3}$. However, in all cases the condition number used to generate this figure was within 0.1% of $\sqrt{3}$.

4. Optimization of Mueller Matrix Polarimeters

Design of a Mueller matrix imaging polarimeter is more complicated than the design of a SV polarimeter. The Mueller matrix at each pixel in an image reveals how the scattering process at that pixel changes the polarization state of light. The Mueller matrix is 4×4, all 16 elements of which are typically independent when no assumptions can be made about the nature of the target.[4] Laboratory-based nonimaging[32] and imaging[34] Mueller matrix polarimeters have been developed that use Fourier methods to reconstruct the Mueller matrix from the harmonics generated in the intensity signal as the polarization generator and analyzer sections are altered periodically. For imaging devices, as many as 80 intensity images have been stored to reconstruct the underlying Mueller matrix images.[34] For real-time systems with typical image sizes (of the order of 640×480), the time and storage required to capture a Mueller matrix image can be prohibitive. Even if parallel methods are used, the spatiotemporal resolution is beyond what is currently available for near-real-time imagery. Because of the huge additional storage and processing needed to form a Mueller matrix image, there is an even greater push to make as few measurements as possible to reconstruct the Mueller matrix on a pixel-by-pixel basis.

Here it is assumed that reconstruction of all 16 elements of the Mueller matrix is desired and that the reconstruction will be made from only 16 measurements. If either the number of desired elements or the number of measurements change, then the method is still applicable, but matrix inverses may need to be replaced with pseudoinverses, as solutions are obtained in a least-squares sense. Because $\underline{\mathbf{M}}$ is the unknown, at least four different pairs of $\mathbf{S}_i$ and $\mathbf{S}_o$ must be used. This results in the expression

$$\underline{\mathbf{M}} \cdot \underline{\mathbf{S}}_i = \underline{\mathbf{S}}_o, \quad \text{with} \tag{27}$$

$$\underline{\mathbf{S}}_i = [\mathbf{S}_i^{(1)} \quad \mathbf{S}_i^{(2)} \quad \mathbf{S}_i^{(3)} \quad \mathbf{S}_i^{(4)}], \tag{28}$$

$$\underline{\mathbf{S}}_o = [\mathbf{S}_o^{(1)} \quad \mathbf{S}_o^{(2)} \quad \mathbf{S}_o^{(3)} \quad \mathbf{S}_o^{(4)}], \tag{29}$$

where $\mathbf{S}_i^{(j)}$ is the jth input SV. The unknown Mueller matrix can be determined from Eq. (27) by

$$\underline{\mathbf{M}} = \underline{\mathbf{S}}_o \cdot \underline{\mathbf{S}}_i^{-1}, \tag{30}$$

where the inverse operation is replaced by the appropriate pseudoinverse when $\underline{\mathbf{S}}_i$ is not full rank. It is clear from the discussion above that the SNR in the reconstructed Mueller matrix images is maximized and equalized when the condition number of $\mathbf{S}_i$ is minimized. This implies that SNR is maximized when four input states are generated that form a regular tetrahedron on the Poincaré sphere, and each of the output states is analyzed through four elliptical diattenuators that form a regular tetrahedron on the Poincaré sphere. Available *a priori* information about $\underline{\mathbf{M}}$ can be used to help select the appropriate input and output states to use. It should be noted

that SNR is only one consideration. There might be operational restrictions that make a choice of suboptimal input and output polarization states more desirable, such as ease of implementation; systematic error is not addressed here.

5. Conclusion

In this paper several considerations in polarimeter design have been presented. The relationship between condition number and optimization of SNR in a SV polarimeter were reviewed[7,9]; and optimal configurations for RR, VR, and rotating analyzer systems were presented in Figs. 3 and 4 for a wide range of system parameters. A detailed theory was developed linking system design to error performance, and it was found that two items affect error propagation in a system: (1) condition number of the system matrices $\underline{\mathbf{A}}$ and $\underline{\mathbf{B}}$ and (2) sensitivity of the analysis matrix $\underline{\mathbf{A}}$ to changes in the system parameters. Finally, the extension of the ideas of optimization from SV polarimeters to full Mueller matrix polarimeters was discussed.

Optimizing polarimeters by use of SVs that inscribe a regular tetrahedron in the Poincaré sphere is not a new concept. Azzam *et al.*[22] introduced the concept for choosing the optimal set of four SVs to calibrate their four-detector photopolarimeter. Ambirajan and Look[6] used the concept to propose a best possible arrangement of principal directions for a SV polarimeter, although they were not able to show that such a configuration was necessarily optimal. Sabatke *et al.*[7] demonstrated that there are exactly two possible ways to inscribe a regular tetrahedron in the Poincaré sphere using a RR polarimeter, and Tyo[9] showed that a VR system can have many possible optimal configurations, a property that might make VR systems more desirable in some applications. Previously, Tyo[11] demonstrated that linear polarimeters with the measurements made at equally spaced intervals between 0° and 180° were optimal. Although that analysis was done in terms of measurement correlation, not system condition, it is straightforward to show that such a system has the lowest condition number for its analysis matrix. Walraven[20] may have been the first to employ such a strategy for a linear polarimeter, as he used four linear polarizers oriented at 0°, 45°, 90°, and 135°, although this was done primarily for ease of computation.

All the above analysis methods have come to fundamentally identical results for all classes of SV polarimeters so far investigated. The reason for this is as follows. Describing a polarimeter as a device that makes a linear decomposition of an unknown 4-D input vector onto the optimal N-element frame encompases all the above techniques. Such a system is by definition the best conditioned,[27] the projections have the least possible correlation between them, and the resulting reconstructed images can be expected to have the optimal input SNR. Only the issue of error performance as discussed here cannot be easily understood solely by the concept of frames on the Stokes cone. Ambirajan and Look[5] proposed that the primary benefit to minimizing the system condition is to use an optimally conditioned system to minimize the effect of such errors; however, the results presented here and by Sabatke *et al.*[7] and Tyo[9] make it clear that the primary benefit of to minimizing the system condition is the reduction of SNR that is due to noise introduced during the detection process, and that sensitivity of the system to changes in the parameters is equally important in rms error performance.

Appendix A

The L_2 condition number of a matrix is equal to the ratio of the largest and smallest singular values of the matrix.[29] The singular value decomposition of a $N \times M$ matrix $\underline{\mathbf{A}}$ is

$$\underline{\mathbf{A}} = \underline{\mathbf{U}}\underline{\mathbf{S}}\underline{\mathbf{V}}^T, \tag{A1}$$

where $\underline{\mathbf{U}}$ is an $N \times N$ unitary matrix, $\underline{\mathbf{V}}$ is a $M \times M$ unitary matrix, and $\underline{\mathbf{S}}$ is an $N \times M$ diagonal matrix of singular values. Each of the vectors (columns) of $\underline{\mathbf{V}}$ represents a direction in $\mathbb{R}^M$ that maps into the corresponding column of $\underline{\mathbf{U}}$ (set of intensity measurements) in $\mathbb{R}^N$ with lengths related by the corresponding singular value. Columns of $\underline{\mathbf{U}}$ greater than M in index are not in the range of $\underline{\mathbf{A}}$.

Recall that the rows of $\underline{\mathbf{A}}$ are the principal axes of the successive Mueller matrices used to decompose the unknown input. When $\underline{\mathbf{A}}$ is optimized, these vectors are uniformly spread out on the surface of the Stokes cone. Regardless of the configuration of the polarimeter, the first entry in the principal axis vector $[\underline{\mathbf{M}}_{00}^{(i)}]$ is always 1/2 when the diattenuator is ideal. Operation on the unpolarized SV $\mathbf{S}_u = [1 \quad 0 \quad 0 \quad 0]^T$ produces

$$\underline{\mathbf{A}} \cdot \mathbf{S}_u = [1/2 \ldots 1/2]^T = \mathbf{u}_1. \tag{A2}$$

When the system is optimized, the 2nd–Mth columns of each row of $\underline{\mathbf{A}}$ form a vector in $\mathbb{R}^{M-1}$ that is on the surface of the $(M-1)$-dimensional space normal to $\mathbf{S}_u$. It was shown that this space is a $(M-1)$-dimensional sphere, corresponding to the Poincaré sphere when $M = 4$. The 2nd–Mth columns of each row also correspond to the principal axes of the elliptical diattenuators, which are equally spaced in $\mathbb{R}^{M-1}$. Because there are more axes (N) than dimensions (M–1) and the system is optimal, these vectors form a tight frame in $\mathbb{R}^{M-1}$,[27] and the frame bounds in inequality (12) are equal to the redundancy ratio times the square length of the vectors, e.g., N vectors in $(M-1)$-dimensional space[27] multiplied by $(M_{00})^2$ [see inequality (12)]. Because all vectors in $\mathbb{R}^{M-1}$ have the same transformed length, and this space is orthogonal $\mathbf{S}_u$, they form the principal spaces spanned by the matrix $\underline{\mathbf{A}}$ with one singular value,

$$\sigma_1 = \frac{\|\underline{\mathbf{A}} \cdot \mathbf{S}_u\|_2}{\|\mathbf{S}_u\|_2} = \frac{\sqrt{N}}{2}, \tag{A3}$$

and three remaining singular values,

$$\sigma_j = \frac{\sqrt{N}}{2(M-1)^{1/2}} \, (j \neq 1), \tag{A4}$$

which yields for the condition number

$$\kappa_2(\underline{\mathbf{A}}) = \sigma_1/\sigma_M = (M-1)^{1/2}. \tag{A5}$$

Equation (A5) indicates that the minimum condition number for the processing matrices of a polarimeter is equal to $(M-1)^{1/2}$, where M is the dimensionality of the SV to be reconstructed. This result is independent of the number of measurements N that are made to reconstruct the M Stokes parameters and is in agreement with the results presented above for RR, VR, and rotating analyzer systems with various values of N.

Appendix B

A second relationship can also be developed that links the frame bounds A and B in inequality (12) to the singular values of the processing matrices for optimized systems. It has already been shown that an optimal SV polarimeter has its decomposition basis spaced out equally along the surface of $\mathcal{S}$. The vector $\mathbf{S}_u$ representing unpolarized radiation can be shown to be the vector that satisfies the upper frame bound.[35] In that case,

$$\frac{\|\underline{\mathbf{A}} \cdot \mathbf{S}_u\|_2^{\,2}}{\|\mathbf{S}_u\|_2^{\,2}} = \sigma_1^{\,2} = N/4. \tag{B1}$$

A completely polarized vector can be shown to satisfy the lower frame bound (from among allowed SVs), and

$$\frac{\|\underline{\mathbf{A}} \cdot \mathbf{S}_p\|_2^{\,2}}{\|\mathbf{S}_p\|_2^{\,2}} = \frac{1}{2}\,(\sigma_1^{\,2} + \sigma_M^{\,2}) = \frac{1}{8}\left(N + \frac{N}{M-1}\right). \tag{B2}$$

Substitution verifies Eqs. (B1) and (B2) for the cases considered above.

References and Notes

1. D. H. Goldstein and R. A. Chipman, eds., *Polarization: Measurement, Analysis, and Remote Sensing*, Proc. SPIE **3121,** (1997).
2. D. H. Goldstein and D. B. Chenault, eds., *Polarization: Measurement, Analysis, and Remote Sensing II*, Proc. SPIE **3754,** (1999).
3. D. B. Chenault, M. J. Duggin, W. G. Egan, and D. H. Goldstein, eds., *Polarization Analysis, Measurement, and Remote Sensing III*, Proc. SPIE **4133** (2000).
4. R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light* (North-Holland, New York, 1977).
5. A. Ambirajan and D. C. Look, "Optimum angles for a polarimeter: part I," Opt. Eng. **34,** 1651–1655 (1995).
6. A. Ambirajan and D. C. Look, "Optimum angles for a polarimeter: part II," Opt. Eng. **34,** 1656–1659 (1995).
7. 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,** 802–804 (2000).
8. D. S. Sabatke, A. M. Locke, M. R. Descour, W. C. Sweatt, J. P. Garcia, E. Dereniak, S. A. Kemme, and G. S. Phipps, "Figures of merit for complete Stokes polarimeters," in *Polarization Analysis, Measurement, and Remote Sensing III*, D. B. Chenault, M. J. Duggin, W. G. Egan, and D. H. Goldstein, eds., Proc. SPIE **4133,** 75–81 (2000).
9. J. S. Tyo, "Noise equalization in Stokes parameter images obtained by use of variable-retardance polarimeters," Opt. Lett. **25,** 1198–2000 (2000).
10. J. S. Tyo, "Considerations in polarimeter design," in *Polarization Analysis, Measurement, and Remote Sensing III*, D. B. Chenault, M. J. Duggin, W. G. Egan, and D. H. Goldstein, eds., Proc. SPIE **4133,** 65–74 (2000).
11. J. S. Tyo, "Optimum linear combination strategy for an N-channel polarization-sensitive vision or imaging system," J. Opt. Soc. Am. A **15,** 359–366 (1998).
12. J. S. Tyo and T. S. Turner, "Imaging spectropolarimeters for use in visible and infrared remote sensing," in *Imaging Spectrometry V*, M. R. Descour and S. S. Shen, eds., Proc. SPIE **3753,** 214–225 (1999).
13. J. S. Tyo, E. N. Pugh, and N. Engheta, "Colorimetric representations for use with polarization-difference imaging of objects in scattering media," J. Opt. Soc. Am. A **15,** 367–374 (1998).
14. J. S. Tyo, "Polarization difference imaging," Ph.D. dissertation (University of Pennsylvania, Philadelphia, Pa., 1997).
15. M. P. Silverman and W. Strange, "Object delineation within turbid media by backscattering of phase modulated light," Opt. Commun. **144,** 7–11 (1997).
16. E. A. West, J. G. Porter, J. M. Davis, A. Gary, and M. Adams, "Development of a polarimeter for magnetic field measurements in the ultraviolet," in *Polarization Analysis, Measurement, and Remote Sensing*, D. H. Goldstein, D. B. Chenault, W. G. Egan, and M. J. Duggin, eds., Proc. SPIE **4481** (to be published).
17. In fact, there is only one orthogonal polarization direction. Two states are orthogonally polarized regardless of their relative intensities; only the direction within the Stokes cone matters.
18. M. H. Smith, J. D. Howe, J. B. Woodruff, M. A. Miller, G. R. Ax, T. E. Petty, and E. A. Sornsin, "Multispectral infrared Stokes imaging polarimeter," in *Polarization: Measurement, Analysis, and Remote Sensing II*, D. H. Goldstein and D. B. Chenault, eds., Proc. SPIE **3754,** 137–143 (1999).
19. J. S. Tyo and T. S. Turner, "Variable-retardance, Fourier-transform imaging spectropolarimeters for visible spectrum remote sensing," Appl. Opt. **40,** 1450–1458 (2001).
20. R. Walraven, "Polarization imagery," Opt. Eng. **20,** 14–18 (1981).
21. L. B. Wolff, "Polarization camera for computer vision with a beam splitter," J. Opt. Soc. Am. A **11,** 2935–2945 (1994).
22. R. M. A. Azzam, I. M. Elminyawi, and A. M. El-Saba, "General analysis and optimization of the four-detector photopolarimeter," J. Opt. Soc. Am. A **5,** 681–689 (1988).
23. G. P. Nordin, J. T. Meier, P. C. Deguzman, and M. Jones, "Diffractive optical element for Stokes vector measurement with a focal plane array," in *Polarization: Measurement, Analysis, and Remote Sensing II*, D. H. Goldstein and D. B. Chenault, eds., Proc. SPIE **3754,** 169–177 (1999).
24. Note that the rows and columns of the Mueller matrix are numbered 0, . . . , 3, in agreement with conventions.[4]
25. S.-Y. Lu and R. A. Chipman, "Interpretation of Mueller matrices based on the polar decomposition," J. Opt. Soc. Am. A **13,** 1106–1113 (1996).
26. R. M. Young, *An Introduction to Nonharmonic Fourier Series* (Academic, New York, 1980), pp. 184–196.
27. I. Daubecheis, *Ten Lectures on Wavelets* (Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1992), pp. 53–106.
28. Sabatke *et al.*[7,8] introduced the concept of equal-weighted variance for polarimeter optimization. The equal-weighted vari-

ance is equivalent to the Frobenius norm of the synthesis matrix $\underline{\mathbf{B}}$, defined in Eq. (25). Because the Frobenius norm of all analysis matrices is $N\sqrt{2}/2$, the Frobenius norm of the synthesis matrix is equivalent to the Frobenius condition number $\mathbf{A}$.
29. G. H. Golub and C. F. van Loan, *Matrix Computations* (Johns Hopkins U. Press, Baltimore, Md., 1983), Chap. 1, pp. 11–29.
30. Azzam *et al*.[22] first proposed the tetrahedron on the Poincaré sphere for choosing the optimal set of calibration states for a SV polarimeter.
31. T. S. Turner, K. W. Peters, and J. S. Tyo, "Portable, visible, imaging spectropolarimeters for remote sensing applications," in *Sensors, Systems, and Next-Generation Satellites II*, H. Fujisada, ed., Proc. SPIE **3498,** 223–230 (1998).
32. D. H. Goldstein and R. A. Chipman, "Error analysis of a Mueller matrix polarimeter," J. Opt. Soc. Am. A **7,** 693–700 (1990).
33. If the input states are not uniformly distributed, the expectation can simply be evaluated with the appropriate statistical distribution of polarization state.
34. J. L. Pezzaniti and R. A. Chipman, "Mueller matrix imaging polarimetry," Opt. Eng. **34,** 1558–1568 (1995).
35. Note that the full set of N vectors forms a frame in $\mathbb{R}^M$. The last three columns form a tight frame in $\mathbb{R}^{M-1}$.