

Optimum selection of input polarization states in determining the sample Mueller matrix: a dual photoelastic polarimeter approach

D. Layden,^{1,*} M. F. G. Wood,² and I. A. Vitkin^{2,3,4}

¹Department of Physics and Astronomy, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada

²Division of Biophysics and Bioimaging, Ontario Cancer Institute, University Health Network, 610 University Avenue, Toronto, Ontario M5G 2M9, Canada

³Department of Medical Biophysics, University of Toronto, 610 University Avenue, Toronto, Ontario M5G 2M9, Canada

⁴Department of Radiation Oncology, University of Toronto, 610 University Avenue, Toronto, Ontario M5G 2M9, Canada

[*dlayden@uwaterloo.ca](mailto:dlayden@uwaterloo.ca)

Abstract: Dual photoelastic modulator polarimeter systems are widely used for the measurement of light beam polarization, most often described by Stokes vectors, that carry information about an interrogated sample. The sample polarization properties can be described more thoroughly through its Mueller matrix, which can be derived from judiciously chosen input polarization Stokes vectors and correspondingly measured output Stokes vectors. However, several sources of error complicate the construction of a Mueller matrix from the measured Stokes vectors. Here we present a general formalism to examine these sources of error and their effects on the derived Mueller matrix, and identify the optimal input polarization states to minimize their effects in a dual photoelastic modulator polarimeter configuration. The input Stokes vector states leading to the most robust Mueller matrix determination are shown to form Platonic solids in the Poincaré sphere space; we also identify the optimal 3D orientation of these solids for error minimization.

© 2012 Optical Society of America

OCIS codes: (120.5410) Polarimetry; (120.2130) Ellipsometry and polarimetry; (170.1470) Blood or tissue constituent monitoring; (170.4090) Modulation techniques; (260.5430) Polarization; (230.4110) Modulators.

References and links

1. G. G. Stokes, "On the composition and resolution of streams of polarized light from different sources," Trans. Cambridge Phil. Soc. **9**, 399–416 (1852).
2. E. Collett, *Field Guide to Polarization* (SPIE Press, 2005).
3. H. Poincaré, *Théorie mathématique de la lumière* (Gauthiers-Villars, 1892).
4. H. Mueller, "Memorandum on the polarization optics of the photoelastic shutter," Report No. 2 of the OSRD project OEMsr-576, (1943).
5. N. Ghosh, M. F. G. Wood, S. Li, R. D. Weisel, B. C. Wilson, R.-K. Li, and I. A. Vitkin, "Mueller matrix decomposition for polarized light assessment of biological tissues," J. Biophoton. **2**, 145–156 (2009).
6. D. Côté and I. A. Vitkin, "Balanced detection for low-noise precision polarimetric measurements of optically active, multiply scattering tissue phantoms," J. Biomed. Opt. **9**, 213–220 (2004).

7. X. Guo, M. F. G. Wood, and I. A. Vitkin, "Angular measurements of light scattered by turbid chiral media using linear Stokes polarimeter," *J. Biomed. Opt.* **11**, 041105 (2006).
8. S. H. Friedberg, A. J. Insel, and L. E. Spence, *Linear Algebra*, 2nd ed. (Prentice-Hall, 1989).
9. A. Ambirajan and D. C. Look, Jr., "Optimum angles for a polarimeter: part 1," *Opt. Eng.* **34**, 1651–1655 (1995).
10. E. Garcia-Caurel, A. D. Martino, and B. Dré villon, "Spectroscopic Mueller polarimeter based on liquid crystal devices," *Thin Solid Films* **455**, 120–123 (2004).
11. P. A. Letnes, I. S. Nerbø, L. M. S. Aas, P. G. Ellingsen, and M. Kildemo, "Fast and optimal broad-band Stokes/Mueller polarimeter design by the use of a genetic algorithm," *Opt. Express* **18**, 23095–23103 (2010).
12. A. D. Martino, E. Garcia-Caurel, B. Laude, and B. Dré villon, "General methods for optimized design and calibration of Mueller polarimeters," *Thin Solid Films* **455**, 112–119 (2004).
13. A. D. Martino, Y.-K. Kim, E. Garcia-Caurel, B. Laude, and B. Dré villon, "Optimized Mueller polarimeter with liquid crystals," *Opt. Lett.* **28**, 616–618 (2003).
14. D. S. Sabatke, M. R. Descour, E. L. 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).
15. S. N. Savenkov, "Optimization and structuring of the instrument matrix for polarimetric measurements," *Opt. Eng.* **41**, 965–972 (2002).
16. M. H. Smith, "Optimization of a dual-rotating-retarder Mueller matrix polarimeter," *Appl. Opt.* **41**, 2488–2493 (2002).
17. K. M. Twietmeyer and R. A. Chipman, "Optimization of Mueller matrix polarimeters in the presence of error sources," *Opt. Express* **16**, 11589–11603 (2008).
18. J. S. Tyo, "Design of optimal polarimeters: maximization of signal-to-noise ratio and minimization of systematic error," *Appl. Opt.* **41**, 619–630 (2002).
19. J. S. Tyo, "Noise equalization in Stokes parameter images obtained by use of variable-retardance polarimeters," *Opt. Lett.* **25**, 1198–1200 (2000).
20. I. J. Vaughn and B. G. Hoover, "Noise reduction in a laser polarimeter based on discrete waveplate rotations," *Opt. Express* **16**, 2091–2108 (2008).
21. J. Zallat, S. Ainouz, and M. P. Stoll, "Optimal configurations for imaging polarimeters: impact of image noise and systematic errors," *J. Opt. A, Pure Appl. Opt.* **8**, 807–814 (2006).
22. G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. (The Johns Hopkins University Press, 1996).
23. W. Guan, G. A. Jones, Y. Liu, and T. H. Shen, "The measurement of the Stokes parameters: a generalized methodology using a dual photoelastic modulator system," *J. Appl. Phys.* **103**, 043104 (2008).
24. G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM J. Numer. Anal.* **10**, 413–432 (1973).
25. J. Dattorro, *Convex Optimization & Euclidean Distance Geometry* (Meboo Publishing USA, 2005).
26. J. Stewart, *Calculus Early Transcendentals*, 6th ed. (Thompson Brooks/Cole, 2008).
27. A. Ambirajan and D. C. Look, Jr., "Optimum angles for a polarimeter: part 2," *Opt. Eng.* **34**, 1656–1658 (1995).
28. R. M. A. Azzam, I. M. Elminyaw, and A. M. El-Saba, "General analysis and optimization of the four-detector photopolarimeter," *J. Opt. Soc. Am. A* **5**, 681–689 (1988).
29. M. Atiyah and P. Sutcliffe, "Polyhedra in physics, chemistry and geometry," *Milan J. Math.* **71**, 33–58 (2003).
30. A. P. Arya, *Introduction to Classical Mechanics*, 2nd ed. (Prentice-Hall, 1998).

1. Introduction

The polarization of a beam of light can be characterized by the four so-called 'Stokes parameters': I , Q , U , and V , where I encodes the intensity, Q and U give the degree and orientation of linear polarization, and V gives the degree and direction of circular polarization [1, 2]. These parameters are often combined into a 'Stokes vector', denoted $S = [I, Q, U, V]^T$, which is said to be 'normalized' when a factor of I is divided out (thus making the first element equal to 1). While Stokes vectors are non-Euclidian (i.e. they have no magnitude or direction in a physical sense) it is still possible to interpret them geometrically: when normalized, the parameters $q = Q/I$, $u = U/I$, and $v = V/I$ (referred to as 'polarization parameters') all fall between -1 and 1 . In a Stokes vector representing fully polarized light, the normalized polarization parameters satisfy $q^2 + u^2 + v^2 = 1$, and so if we consider a 3-dimensional plot where $x = q$, $y = u$ and $z = v$, a vector $\vec{\xi} = [q, u, v]^T$ representing full polarization lies on a sphere of unit radius centred about the origin, which is known as the 'Poincaré sphere' [2, 3]. If $\vec{\xi}$ represents partial polarization, it will lie inside the Poincaré sphere, and if it represents random polarization, it will be equal to $\vec{0}$.

When a beam of polarized light, represented by S^{in} , interacts with a material, its polarization state usually changes, and the modified beam will have a corresponding Stokes vector S^{out} . The change in polarization can be represented by the matrix/vector equation

$$S^{out} = MS^{in}, \quad (1)$$

where M is called the ‘Mueller matrix’, and it fully describes the material’s effect on polarized light [2, 4]. A material’s Mueller matrix can be used to extract its physical characteristics, and the process of determining this matrix (usually by examining the change in polarization induced in several Stokes vectors), called ‘Mueller Polarimetry’, is a promising field with a wide range of potential applications. Notably, it has been used for biomedical purposes to measure the progress of stem-cell therapies in myocardial infarction animal models [5], and it may one day be used to measure blood glucose levels [6, 7]. However, while polarimetry may hold great potential for such applications, determining the Mueller matrix of a biological sample can be especially difficult due to the turbid nature of such samples, which often causes multiple scattering and severe polarization loss, as well as significant noise. Thus, to successfully determine the Mueller matrix of a biological material (or any material for that matter), it is necessary to use a very sensitive polarimeter and to make measurements in such a way that the determined matrix is minimally sensitive to measurement noise.

A common way of quantifying the error-sensitivity of the determined Mueller matrix to measurement errors is to use the so-called ‘condition number’ (see [8]), usually denoted κ , which quantifies the numerical robustness of the system [9–21]. With this approach, one usually describes the measurement system’s interaction with polarized light using a matrix A , and then calculates the corresponding condition number

$$\kappa(A) \equiv \|A\| \|A^{-1}\|, \quad (2)$$

where $\|\cdot\|$ is generally an induced norm (as shown in Eq. (3)). The condition number quantifies how ‘invertible’ A is, where a small value of κ means very invertible, and a large value means nearly singular. In practice, the determination of a sample’s Mueller matrix usually requires one to numerically invert a matrix which represents the system or the measured data; a popular strategy is to orient/adjust the system elements such that $\kappa(A)$ is minimized, so as to make this inversion as robust as possible, which in turn minimizes the resulting sample Mueller matrix’s sensitivity to measurement errors.

One potential problem with such an approach is the fact that Eq. (2) requires one to choose a specific norm (see [22]). Generally, vector p -norms are used to induce a matrix norm on A (as in Eq. (3)), although this approach lacks a clear physical interpretation: suppose the matrix A acts on a Stokes vector S (as is often the case), then it is difficult to assign any physical meaning to the p -norm $\|S\|_p = (I^p + Q^p + U^p + V^p)^{1/p}$, and thus even more difficult to interpret the induced matrix norm

$$\|A\|_p = \max_S \frac{\|AS\|_p}{\|S\|_p}. \quad (3)$$

There are several induced p -norms in common use (usually $p = 1, 2$ and ∞ , often chosen for mathematical convenience), and some authors have even applied entry-wise matrix norms directly to Eq. (2) [15, 20]. There has been some debate over the merits of these choices; Sabatke et al. [14], for example, concluded that the using $p = 2$ with Eq. (3) was a poor choice for overdetermined systems (i.e. systems in which there are more independent equations, or rather, pairs of measurements, than unknown parameters). The Frobenius norm, when applied directly to Eq. (2), is subject to the same problem: overdetermined systems usually involve non-square matrices of various shapes and sizes, which tend to skew it (e.g., this choice implies $\kappa(I_n) = n$,

leading to the peculiar conclusion that the 2×2 identity matrix is somehow more robust under inversion than the 3×3).

It is usually assumed that the various ways of defining the condition number produce roughly equivalent results, although Vaughn and Hoover [20] found that optimizing κ with different norms resulted in significantly different system configurations. The state of affairs was well summarized by Twietmeyer and Chipman [17] when they noted that while the condition number can be useful in providing general guidance, it is not sufficient to fully describe a polarimeter, and that further analysis is required in the presence of specific known error sources.

This paper will consider a dual photoelastic modulator (PEM) system, as described by Guan et al. [23], shown in Fig. 1. Such a polarimeter can determine a beam's Stokes vector very accurately within ~ 5 ms by measuring all four parameters simultaneously using polarization modulation and phase-sensitive synchronous detection. We will examine how such a system can be used to derive Mueller matrices using polarization modulations and phase-sensitive synchronous detection, and we will investigate the unique sources of error associated with it and how they translate from Stokes to Mueller polarimetry. We will examine the root mean square of the errors in the Mueller matrix (and in doing so, avoid condition numbers and their associated ambiguities) and show how this straightforward and physically interpretable approach can be used to thoroughly analyze the propagation of Mueller matrix errors in a dual PEM polarimeter.

2. Analysis

The dual PEM system described by Guan et al. [23] shown in Fig. 1 is a polarization analyzer: it can quickly and accurately measure the Stokes vector corresponding to a beam of light, but it is not designed to produce polarized light. In order to measure a material's Mueller matrix with such a system, it is necessary to produce beams with various known polarization states so as to quantify their interactions with the system. Throughout this paper, it is assumed that the dual PEM Stokes polarimeter is used in conjunction with a polarization state generator (PSG) capable of creating any fully polarized incident state (i.e. any Stokes vector on the Poincaré sphere). Then, one can use the PSG to create n different states, which should first be passed directly into the analyzer (the dual PEM system) so as to measure the Stokes vector of each, before repeating the process with the sample in the beam path. The first set of Stokes vectors (those coming directly from the PSG) will be referred to as 'input states', and the second set (the ones measured with the sample in place) as 'output states'. *Since the experimenter is free to choose the input Stokes vectors, it is desirable to find the set(s) of such vectors which will produce the most error-resistant Mueller matrix results.* Note that while the signal intensity I can be useful in polarimetry, we chose to work with normalized Stokes vectors, partly for mathematical simplicity and partly to account for its uncontrollable fluctuations, and so all input states considered in this paper are assumed to be normalized (i.e., $I = 1$).

2.1. Determining Mueller matrices from Stokes vectors

The interaction of a polarized beam with a sample is described in Eq. (1), where M is the sought-after sample Mueller matrix. Of course, it is not possible to fully determine it using only one input and one output Stokes vector; rather, several authors have shown [11, 14, 15, 18] that it takes at least 16 measurements of output state parameters (as well as 16 corresponding input parameters), which can be cast in the form of 4 input and 4 output Stokes vectors. In order to develop a method of finding Mueller matrices from said Stokes vectors, we will denote the vectors representing the various input beams as $S_1^{in}, \dots, S_n^{in}$ and those representing the corresponding output states as $S_1^{out}, \dots, S_n^{out}$. Then, the relationship between the input and the output states, $S_i^{out} = MS_i^{in}$, can be fully described by the matrix equation analogous to Eq. (1):

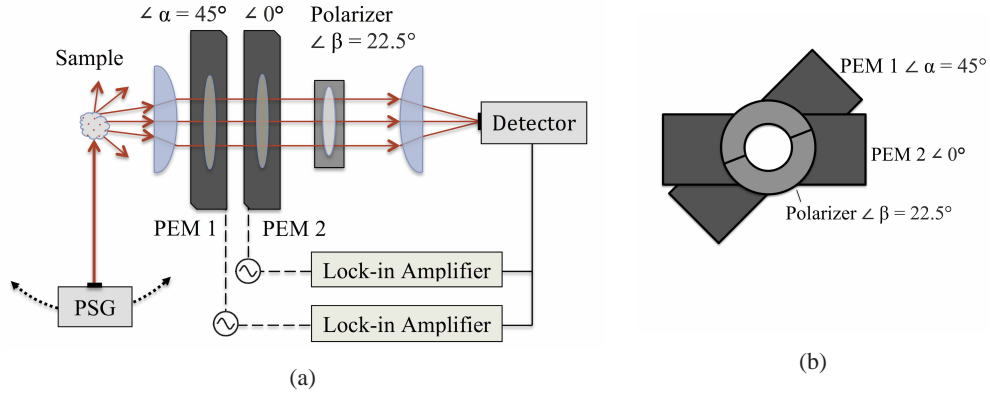


Fig. 1. The dual PEM Stokes polarimeter schematic. Panel (a) provides a top view of the system, which is comprised of a movable PSG to illuminate the sample and dual photoelastic modulators followed by a linear polarizer and a photodetector, which form the PSA. The two lock-in amplifiers recover the polarization parameters Q , U , and V using the PEM modulation frequencies as references. Panel (b) depicts the PSA as seen from the photodetector. The fast axis of PEM 1 is at an angle $\alpha = 45^\circ$ above the horizontal (laboratory frame), and the fast axis of PEM 2 is parallel to the optical table. The linear polarizer's transmission axis is at an angle $\beta = 22.5^\circ$ above the horizontal.

$$[S_1^{out} \quad \dots \quad S_n^{out}] \simeq M [S_1^{in} \quad \dots \quad S_n^{in}]. \quad (4)$$

For notational convenience, we will define the matrices $\mathbb{S}^{in} = [S_1^{in} \dots S_n^{in}]$ and $\mathbb{S}^{out} = [S_1^{out} \dots S_n^{out}]$. In the case of $n \leq 4$ this is a true (rather than approximate) equality, although for $1 \leq n \leq 3$ it is not possible to isolate M . When $n = 4$, the Mueller matrix can be determined as $M = \mathbb{S}^{out} (\mathbb{S}^{in})^{-1}$, provided that \mathbb{S}^{in} is invertible. Taking $n > 4$ can be useful, as such an approach can potentially reduce noise and generally improve accuracy [11, 16, 20], although in this more general case, the system in Eq. (4) will almost surely be overdetermined, and so it becomes an approximate equality. Thus, rather than seeking a matrix M which satisfies the equality, we must instead find M which provides the best fit, or rather, which minimizes the RMS size of the discrepancy $\mathbb{S}^{out} - M\mathbb{S}^{in}$, where the RMS of a matrix A will be denoted $\langle A \rangle$. Note that for the rest of this paper the notation $\|\cdot\|$ will be understood to represent the Frobenius norm, and $\langle \cdot, \cdot \rangle$ the matrix inner product from which it follows.

Defining a function $f: \mathbb{M}_{4 \times 4} \rightarrow \mathbb{R}$ (where $\mathbb{M}_{4 \times 4}$ denotes the set of all 4×4 matrices) as $f(X) = \langle \mathbb{S}^{out} - X\mathbb{S}^{in} \rangle$, we have that $f(M)$ is a minimum. In analogy to variational calculus, we then perturb this minimum with an arbitrary matrix η multiplied by a scalar ϵ , and set the derivative $\partial_\epsilon f(M + \epsilon\eta)^2$ (we consider the RMS squared for convenience) to zero:

$$\frac{\partial}{\partial \epsilon} f(M + \epsilon\eta)^2 \Big|_{\epsilon=0} = \frac{1}{4} \langle \eta, \mathbb{S}^{in} (\mathbb{S}^{in})^\top M^\top - \mathbb{S}^{in} (\mathbb{S}^{out})^\top \rangle = 0 \quad (5)$$

for any η . It follows immediately from the properties of inner products ([8] Section 6.1) that $\mathbb{S}^{in} (\mathbb{S}^{in})^\top M^\top - \mathbb{S}^{in} (\mathbb{S}^{out})^\top = 0$ if the equality is to hold for arbitrary $\eta \in \mathbb{M}_{4 \times 4}$. Isolating for M , we find that the Mueller matrix providing the best RMS fit is

$$M = \mathbb{S}^{out} (\mathbb{S}^{in})^\top [\mathbb{S}^{in} (\mathbb{S}^{in})^\top]^{-1} = \mathbb{S}^{out} (\mathbb{S}^{in})^+, \quad (6)$$

where $(\mathbb{S}^{in})^+ = (\mathbb{S}^{in})^\top [\mathbb{S}^{in} (\mathbb{S}^{in})^\top]^{-1}$ is the well-known Moore-Penrose pseudoinverse of \mathbb{S}^{in}

[11, 14–18, 20, 22]. Note that in the case where $n = 4$, M^+ reduces to M^{-1} , and so Eq. (6) becomes $M = \mathbb{S}^{out}(\mathbb{S}^{in})^{-1}$ when \mathbb{S}^{in} is square. Thus, the more general pseudoinverse encompasses the traditional matrix inverse.

2.2. Quantifying Mueller matrix errors

In practice, measurements are imperfect, and even with a well-calibrated system, there will always be a certain level of noise associated with the measurements, often originating from the sample itself. Suppose we measure a beam of light described by the Stokes vector S , but that our measurements are subject to some error, and rather than obtaining the ‘true’ vector S , we measure $S + \delta S$, where δS is composed of the random errors associated with each Stokes parameter of S . While it is straightforward to measure and quantify the random errors associated with measured Stokes vectors, it is somewhat more complicated to do so with the Mueller matrix calculated from these vectors.

Incorporating errors into the framework developed in Section 2.1, every input vector S_i^{in} becomes $S_i^{in} + \delta S_i^{in}$ and every output vector S_j^{out} becomes $S_j^{out} + \delta S_j^{out}$. Thus, in the presence of error, Eq. (4) becomes

$$[S_1^{out} + \delta S_1^{out} \quad \dots \quad S_n^{out} + \delta S_n^{out}] \simeq (M + \delta M) [S_1^{in} + \delta S_1^{in} \quad \dots \quad S_n^{in} + \delta S_n^{in}], \quad (7)$$

where δM is the deviation from the ‘true’ Mueller matrix of the sample that results from error-prone Stokes vector measurements. We will define the error matrices $\delta \mathbb{S}^{in} = [\delta S_1^{in} \dots \delta S_n^{in}]$ and $\delta \mathbb{S}^{out} = [\delta S_1^{out} \dots \delta S_n^{out}]$ so that Eq. (7) takes the concise form

$$\mathbb{S}^{out} + \delta \mathbb{S}^{out} \simeq (M + \delta M)(\mathbb{S}^{in} + \delta \mathbb{S}^{in}). \quad (8)$$

It is useful to recall that in practice the Stokes errors ($\delta \mathbb{S}^{in}$ and $\delta \mathbb{S}^{out}$) and Mueller error (δM) are unknown; the experimenter will determine the Mueller matrix of the sample in question as $M + \delta M$, by applying Eq. (6) to get

$$(M + \delta M) = (\mathbb{S}^{out} + \delta \mathbb{S}^{out})(\mathbb{S}^{in} + \delta \mathbb{S}^{in})^+. \quad (9)$$

Here we have the complete relationship between the Stokes errors and the resulting error in the Mueller matrix. Unfortunately, it is very difficult to extract any useful dependence of δM on \mathbb{S}^{in} from this equation, and so it is necessary to invoke the matrix analogue of a Taylor series, which will be applied to the pseudoinverse in Eq. (9). For a reasonably well-behaved function (more precisely, a C^1 function) $g : \mathbb{R} \rightarrow \mathbb{R}$, the Taylor series truncated to first order provides a very good approximation for g on a small interval. (In more exact terms, $g(x_0 + dx) \approx g(x_0) + dg = g(x_0) + g'(x_0)dx$ is a good approximation for small dx .) In order to linearize the pseudoinverse, we will use the analogous relation

$$(A + B)^+ \approx A^+ + \partial(A^+) \big|_{\partial A=B}, \quad (10)$$

where $\|B\|$ is small. The differential of the pseudoinverse, as found by Golub and Pereyra [24], is given by

$$\partial(A^+) = -A^+(\partial A)A^+ + (I_n - A^+A)(\partial A^\top)(A^+)^\top A^+, \quad (11)$$

where I_n is the $n \times n$ identity matrix. Substituting Eq. (11) into Eq. (10) and evaluating for $A = \mathbb{S}^{in}$ and $\partial A = \delta \mathbb{S}^{in}$ we can linearize Eq. (9). When expanding said equation, it is possible to make some simplifications. By carrying out the multiplication, we find that we can use Eq. (6) to get additive M terms on both sides of the equation, which can be cancelled. Also, given the physical nature of the problem, we can assume that $\mathbb{S}^{out} \simeq M\mathbb{S}^{in}$ is a very good approximation, and thus we can treat it as an equality when performing simplifications. Finally, we find that

there are cross terms involving both $\delta\mathbb{S}^{in}$ and $\delta\mathbb{S}^{out}$. Terms formed with both of these matrices should be exceedingly small, and since we have already truncated our approximation to first order by linearizing, such terms can be neglected. Applying these simplifications leaves

$$\delta M \approx [\delta\mathbb{S}^{out} - M(\delta\mathbb{S}^{in})](\mathbb{S}^{in})^+. \quad (12)$$

This is an important result that serves as a starting point for further analysis. It quantifies, on an element-by-element basis, how the errors in each of the measured Stokes vectors affect the error on the experimentally determined Mueller matrix. Specifically, while the square bracketed terms are out of the experimenter's control, they are all multiplied by the pseudoinverse of \mathbb{S}^{in} , the matrix of input Stokes vectors, over which the experimenter has full control. This result indicates that the propagation of random errors into the Mueller matrix depends strongly on choice of input polarization states.

2.3. Maximizing the robustness of determined Mueller matrices

Having derived the dependence of Mueller matrix errors on Stokes vector measurement noise in a physically meaningful way, it remains to be decided which, and how many, optimal input polarization states should be used to most robustly determine the Mueller matrix of a sample.

2.3.1. Root mean square of Mueller errors

The main concern when determining Mueller matrices is that each of their calculated elements be as close to the 'true' values as possible, or more precisely, that the RMS of the Mueller error, $\langle\delta M\rangle$, be minimized. This global error reduction is essential since Mueller matrices are frequently used to measure physical properties of a sample, often by considering individual matrix elements [6] or by decomposing it into physically meaningful 'basis matrices' [5]. Thus, the RMS of the Mueller error matrix provides a way to quantify errors in the determined Mueller matrix that is physically interpretable and consistent with applications.

To determine the RMS error, begin by taking the norm of both sides of Eq. (12) and then converting each element to RMS. In order to separate the terms (which is necessary in interpreting the results), we must apply the consistency inequality $\|AB\| \leq \|A\| \|B\|$ (which follows quickly from the Cauchy-Schwarz inequality) and the triangle inequality $\|A+B\| \leq \|A\| + \|B\|$, which gives us the upper bound on the RMS Mueller error

$$\langle\delta M\rangle \lesssim \frac{n^{1/2}}{2} (\langle\delta\mathbb{S}^{out}\rangle + 4\langle M\rangle\langle\delta\mathbb{S}^{in}\rangle) \|(\mathbb{S}^{in})^+\|. \quad (13)$$

Now we have a clear relation between the accuracy of the derived matrix and the level of noise in each set of measured Stokes vectors. (We have considered the norm of pseudoinverse rather than the RMS since the former is more convenient for the analysis in Section 2.3.2.) The Mueller error depends on the size of the input and output errors in the Stokes vector measurements, as well as on the sample's Mueller matrix itself, all of which are out of the experimenter's control. However, it is also proportional to the norm of $(\mathbb{S}^{in})^+$, which is dependent on the polarization states chosen as inputs. The goal then becomes to find a set of input Stokes vectors such that $\|(\mathbb{S}^{in})^+\|$ is minimized, so as to minimize the effect of the Stokes errors on $\langle\delta M\rangle$.

We will see that in general, $\|(\mathbb{S}^{in})^+\|$ decreases as n grows, which can be interpreted physically as an 'averaging out' of errors. On the other hand, $\langle\delta M\rangle$ grows proportionally to $n^{1/2}$, which can be interpreted as resulting from accumulation of additional noisy measurements. To analyze the balance between these two effects, we will first consider $\|(\mathbb{S}^{in})^+\|$, before combining the two factors.

2.3.2. Optimal selections of n input Stokes vectors

The matrix \mathbb{S}^{in} contains the Stokes vectors representing the input beams used to determine the sample's Mueller matrix, and so the experimenter has full control over it. However, it is not immediately clear how the choice of input Stokes vectors relates to $\|(\mathbb{S}^{in})^+\|$, the norm of the pseudoinverse, and thus the resultant uncertainty in M via Eq. (13).

Let us examine the norm squared for convenience. We have that

$$\|(\mathbb{S}^{in})^+\|^2 = \text{Tr}\{[\mathbb{S}^{in}(\mathbb{S}^{in})^\top]^{-1}\}. \quad (14)$$

The matrix $\mathbb{S}^{in}(\mathbb{S}^{in})^\top$ is symmetric, which implies that there exist matrices Ω and D such that $\mathbb{S}^{in}(\mathbb{S}^{in})^\top = \Omega D \Omega^{-1}$ where D is a 4×4 diagonal matrix containing the eigenvalues of $\mathbb{S}^{in}(\mathbb{S}^{in})^\top$: $\lambda_1, \dots, \lambda_4$ ([8] Section 6.5). Taking the inverse of both sides, we find that $[\mathbb{S}^{in}(\mathbb{S}^{in})^\top]^{-1} = \Omega D^{-1} \Omega^{-1}$, where $D^{-1} = \text{diag}(1/\lambda_1, \dots, 1/\lambda_4)$. Since the trace is invariant under eigendecomposition ([8] Section 6.5), the norm of the pseudoinverse becomes

$$\|(\mathbb{S}^{in})^+\|^2 = \text{Tr}(D^{-1}) = \sum_{i=1}^4 \frac{1}{\lambda_i}. \quad (15)$$

Consider the eigenvalue-eigenvector relation $\mathbb{S}^{in}(\mathbb{S}^{in})^\top \vec{v} = \lambda \vec{v}$. Multiplying both sides to the left by \vec{v}^\top , this becomes $\|(\mathbb{S}^{in})^\top \vec{v}\|^2 = \lambda \|\vec{v}\|^2$, which implies that the eigenvalues cannot be negative (since $\|\vec{x}\| \geq 0 \ \forall \vec{x}$). Thus, to minimize the norm of $\|(\mathbb{S}^{in})^+\|$ we must maximize the eigenvalues of $\mathbb{S}^{in}(\mathbb{S}^{in})^\top$.

Observe that $\det[\mathbb{S}^{in}(\mathbb{S}^{in})^\top] = \prod_{i=1}^4 \lambda_i$ ([8] Section 6.5), and so a large determinant is indicative of collectively large eigenvalues. To motivate our search for a large determinant, recall that geometrically, the determinant of a matrix represents the area swept out by the vectors which form its columns ([8] Section 4.1). In 2 and 3 dimensions, this occurs when the vectors are orthogonal, so as to form a rectangle or a rectangular prism in \mathbb{R}^2 and \mathbb{R}^3 respectively. To extend this idea to 4 dimensions, we would like to find a matrix \mathbb{S}^{in} such that $\mathbb{S}^{in}(\mathbb{S}^{in})^\top$ is diagonal. (Technically, we only need the columns of $\mathbb{S}^{in}(\mathbb{S}^{in})^\top$ to be orthogonal, although since the determinant is invariant under eigendecomposition, we can consider the diagonal case without loss of generality.)

For a $4 \times n$ matrix \mathbb{S}^{in} whose rows are formed by the n -dimensional vectors $\vec{r}_1^\top, \dots, \vec{r}_4^\top$, we have that

$$[\mathbb{S}^{in}(\mathbb{S}^{in})^\top]_{\mu\nu} = \vec{r}_\mu \cdot \vec{r}_\nu \quad (16)$$

and so to create a diagonal matrix, we need the rows of \mathbb{S}^{in} to be orthogonal. To confirm that such a diagonal matrix does in fact produce the largest possible determinant, it is useful to examine the matrix derivative [25]

$$\frac{\partial \det(X)}{\partial X} = \det(X)(X^{-1})^\top. \quad (17)$$

Substituting $X = \mathbb{S}^{in}(\mathbb{S}^{in})^\top = \text{diag}(\lambda_1, \dots, \lambda_4)$, we see that the derivative is also a diagonal matrix. This implies that perturbing any of the off-diagonal elements of $\mathbb{S}^{in}(\mathbb{S}^{in})^\top$ will decrease $\det[\mathbb{S}^{in}(\mathbb{S}^{in})^\top]$ since their current values (of zero) already represent critical points. The elements lying along the diagonal, however, are non-zero since they are not critical points (i.e. $\det(X)$ increases along with any of the diagonal elements), although we shall see that they are constrained by the physical nature of the problem.

Here it becomes necessary to reconsider the original interpretation of the matrix \mathbb{S}^{in} : its columns are formed by the input Stokes vectors (assumed to be normalized), which represent

fully polarized light. This physical interpretation limits the magnitude of the rows of \mathbb{S}^{in} , and so for a given number of input states (columns of \mathbb{S}^{in}), n , this puts an upper bound on the determinant in question.

The matrix \mathbb{S}^{in} is composed as

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ q_1 & q_2 & \cdots & q_n \\ u_1 & u_2 & \cdots & u_n \\ v_1 & v_2 & \cdots & v_n \end{bmatrix}, \quad (18)$$

and so for all $1 \leq i \leq n$ we have that $q_i^2 + u_i^2 + v_i^2 = 1$. Thus

$$\|\vec{r}_2\|^2 + \|\vec{r}_3\|^2 + \|\vec{r}_4\|^2 = \sum_{i=1}^n q_i^2 + \sum_{i=1}^n u_i^2 + \sum_{i=1}^n v_i^2 = \sum_{i=1}^n (q_i^2 + u_i^2 + v_i^2) = n, \quad (19)$$

and so the norms of the bottom three rows are collectively bounded by n , the number of Stokes vectors used as inputs. (The norm squared of the first row is necessarily equal to n .)

As we have seen, it is necessary for the rows of \mathbb{S}^{in} to be orthogonal in order to minimize the Mueller error. Since the first row is composed entirely of ones, the dot product $\vec{r}_1 \cdot \vec{r}_i = \sum_j (\vec{r}_i)_j$ must be equal to zero for $2 \leq i \leq 4$, and so the sum of each particular polarization parameter (e.g. $\sum_{i=1}^n q_i$) must equal 0. Physically, this indicates that the input Stokes vectors must be as ‘spread out’ as possible over the Poincaré sphere so as to increase the overall stability of the result.

There remains one degree of freedom which we must examine: while Eq. (19) limits the norm of the rows collectively, it does not limit them individually; that is, under the imposed condition, it would be permissible, for example, to have $\|\vec{r}_2\| = n$ and the other rows be zero. Of course, such a solution would result in a determinant of zero, in which case $(\mathbb{S}^{in})^+$ would not exist, and there would be no way to find the Mueller matrix. Clearly it is necessary to examine the rows individually, and to do so we will make use of Lagrange multipliers ([26] Section 14.8).

The original goal of maximizing the determinant was to minimize the norm of $(\mathbb{S}^{in})^+$, which was shown to be equivalent to minimizing $\sum_{i=1}^4 1/\lambda_i$, where λ_i are the eigenvalues of $\mathbb{S}^{in}(\mathbb{S}^{in})^\top$, for $1 \leq i \leq 4$. Having imposed the condition of orthogonality, we know that $\lambda_1 = n$, and $\lambda_j = \|\vec{r}_j\|^2$, for $2 \leq j \leq 4$. Thus, we wish to minimize $\|(\mathbb{S}^{in})^+\|^2 = n^{-1} + \sum_{j=2}^4 \|\vec{r}_j\|^{-2}$, subject to the constraint from Eq. (19). The corresponding Lagrange function, Γ , is defined by

$$\Gamma(\lambda_1, \dots, \lambda_4, \gamma) = \sum_{i=1}^4 \frac{1}{\lambda_i} + \gamma \left[\left(\sum_{i=2}^4 \lambda_i \right) - \lambda_1 \right]. \quad (20)$$

In order to optimize this constrained system, we must impose the condition $\vec{\nabla} \Gamma = \vec{0}$. Solving the resulting vector equation, we find that $\|\vec{r}_j\|^2 = n/3$ for all $2 \leq j \leq 4$.

Thus, we arrive at the result that for a set of input Stokes vectors $\{S_i^{in}\}_{i=1}^n$, where $S_i^{in} = [1, q_i, u_i, v_i]^\top$, the conditions

$$\sum_{i=1}^n q_i = \sum_{i=1}^n u_i = \sum_{i=1}^n v_i = 0, \quad (21)$$

$$\sum_{i=1}^n q_i u_i = \sum_{i=1}^n u_i v_i = \sum_{i=1}^n v_i q_i = 0, \quad (22)$$

$$\sum_{i=1}^n q_i^2 = \sum_{i=1}^n u_i^2 = \sum_{i=1}^n v_i^2 = \frac{n}{3}, \quad (23)$$

guarantee a minimized norm $\|(\mathbb{S}^{in})^+\|$ for a given n (for fully polarized input states). Unfortunately, finding various sets of Stokes vectors which satisfy these conditions is not trivial. However, in the process of designing optimal polarimeters, several authors have concluded that Stokes vectors forming a tetrahedron on the Poincaré sphere are an optimal solution to this related problem [14, 15, 17, 18, 27, 28] when $n = 4$. It is easy to show that when \mathbb{S}^{in} is composed of 4 Stokes vectors which form the vertices of such a tetrahedron, Eqs. (21)–(23) are satisfied.

Of course, we are not limited to the case $n = 4$: the tetrahedron is one of five ‘Platonic Solids’ [29], and since the former provides an optimal solution, it seems natural to consider the other solids. It is similarly easy to show that for a given number of Stokes vectors, each of the Platonic solids represent an optimal configuration of Stokes vectors on the Poincaré sphere, when these vectors form the vertices of said solids. For $n = 4$ this corresponds to a tetrahedron, for $n = 6$ it is an octahedron, for $n = 8$ it is a cube, for $n = 12$ an icosahedron, and finally for $n = 20$, a dodecahedron. These solutions are represented graphically in Fig. 2.

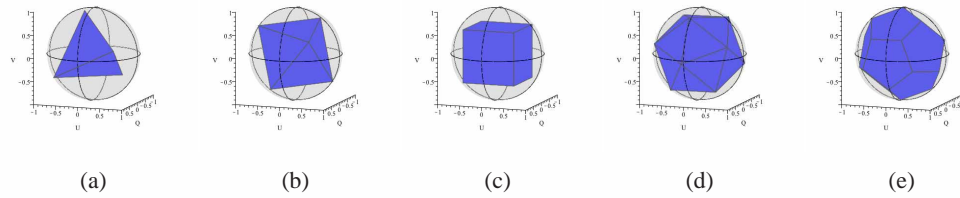


Fig. 2. The five Platonic solids inscribed in the Poincaré sphere, where the equator represents linear polarization and the poles represent circular polarization. The three great circles on each sphere represent the areas where Q , U , or V are zero. When the input Stokes vectors form the vertices of any of these shapes, Eqs. (21)–(23) are satisfied, and so the error-sensitivity of the determined Mueller matrix will be minimized. In (a) the $n = 4$ input Stokes vectors form a tetrahedron in polarization space ([Media 1](#)), in (b) where $n = 6$, they form an octahedron ([Media 2](#)), in (c) where $n = 8$, they form a cube ([Media 3](#)), in (d) where $n = 12$, they form an icosahedron ([Media 4](#)), and finally, in (e) where $n = 20$, they form a dodecahedron ([Media 5](#)).

Combining Eq. (15) with Eq. (23) we find that the minimum value of $\|(\mathbb{S}^{in})^+\|$ for a given $n \geq 4$ is

$$\|(\mathbb{S}^{in})^+\|_{min}(n) = \left(\frac{10}{n}\right)^{1/2}. \quad (24)$$

While we have only presented the configurations for certain n here (e.g. the ones forming the vertices of Platonic solids on the Poincaré sphere), there are almost surely other solutions to Eqs. (21)–(23) for values of n which we have not considered. However, Eq. (24) allows us to examine the effects of varying n analytically, without having to find other solutions to the derived conditions, thus allowing us to compare the merits of all optimal configurations without having to know the exact (geometrical) form of each.

2.3.3. Optimal number of measurements

Assuming that for a given $n \geq 4$ one is able to find an optimal configuration of Stokes vectors (such as the ones presented in Fig. 2), the effect of n , the number of input/output Stokes vectors, on the Mueller error has yet to be determined.

Substituting Eq. (24) into Eq. (13) we have that for an optimal configuration of n Stokes vectors

$$\langle \delta M \rangle \lesssim \left(\frac{5}{2} \right)^{1/2} (\langle \delta S^{out} \rangle + 4 \langle M \rangle \langle \delta S^{in} \rangle). \quad (25)$$

Thus, to first order approximation, the upper bound RMS Mueller error is independent of the number of measurements when the input Stokes vectors form an optimal set for a given n . And so, the question of *which* configurations (in Fig. 2 or otherwise) to use remains unaddressed using this method. As discussed in Section 1, the answer must come from considering the unique types of errors which can occur in dual PEM Stokes polarimeters.

2.3.4. Errors in relative modulation phase

Dual photoelastic modulator Stokes polarimeters use polarization modulation and synchronous phase-sensitive detection (e.g. lock-in amplifiers) to measure polarization parameters from an alternating photocurrent by monitoring pre-determined frequencies and harmonics. The Stokes parameters Q , U , and V (i.e. the polarization parameters) are measured by multiplying the magnitude of the frequency-specific oscillations in the signal by the sign of the relative phase between the signal oscillations and the PEM reference frequency. Here we consider a system like the one described by Guan et al. [23] (with $\alpha = 45^\circ$ and $\beta = 22.5^\circ$, as shown in Fig. 1(b)), and so to measure U , for example, we use $U = |U| \text{sgn}(\theta)$, where the absolute value of U is the amplitude measured by the lock-in amplifier and θ is the relative phase. Thus far we have been considering noise in the final Stokes parameters rather than in their constituent quantities: the measured amplitudes and relative phases. While this approach was useful and remains valid, there is another type of error specific to this kind of system. If the relative phase θ is close to 0, then fluctuations (noise in θ) which cause it to jump above and below 0 can result in enormous errors if the amplitude is nonzero.

In practice, the relative phase of each polarization parameter approaches zero as the measured amplitude becomes small. Thus, phase errors become more probable as the polarization parameters approach zero, as shown in Fig. 3 (Media 6).

In order for all three polarization parameters to be equally free of phase errors (i.e. equidistant from the red zones in Fig. 3 Media 6), we should have $|Q| = |U| = |V|$ for all measured Stokes vectors. Of course, the experimenter has no control over the output vectors without *a priori* knowledge of the Mueller matrix being examined, and so in order to develop a general framework, we must settle for a reduction of input vector phase errors.

Notice that there are 8 Stokes vectors on the Poincaré sphere where $|Q| = |U| = |V|$ (or rather, where $|q| = |u| = |v|$, although the two conditions are equivalent for any given Stokes vector since the dividing I term is the same for all of the parameters), and that these form the vertices of a cube—one of the optimal configurations found in Section 2.3.2 (shown in Fig. 2(c) Media 3). However, before concluding that said cubic configuration provides a global optimum, it is important to note that there are infinitely many ways to rotate a cube inside the Poincaré sphere (this holds true for all solutions in Fig. 2), and it is not immediately clear what effect such a rotation has.

Consider a three-dimensional rotation matrix $R(\phi, \theta, \psi)$ which rotates vectors in \mathbb{R}^3 by the Euler angles ([30] Section 13.8). The analogous matrix for Stokes vectors which rotates the polarization components around the Poincaré sphere by the same angles but leaves the intensity unchanged is given by

$$R' = \begin{bmatrix} 1 & \vec{0}^\top \\ \vec{0} & R \end{bmatrix}, \quad (26)$$

where $\vec{0} = [0, 0, 0]^\top$. Thus, for a matrix S^{in} of input Stokes vectors, $R' S^{in}$ describes all possible rotations of the polarization components inside the Poincaré sphere. Since R is a rotation matrix,

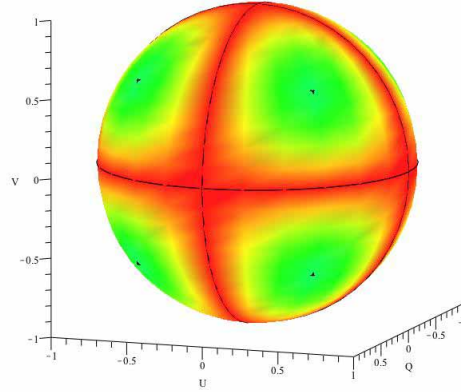


Fig. 3. (Media 6). The Poincaré sphere coloured according to the likelihood of phase errors in each region, where red represents a high probability and green represents a low probability. As in Fig. 2, the great circles in black represent the areas where the polarization parameters are 0. As Stokes vectors approach these areas (i.e. they enter the red/yellow zones) they become increasingly prone to phase errors. Here, the cubic configuration of Fig. 2(c) (Media 3) has been inscribed in the sphere to show that the Stokes vectors forming its vertices are all minimally prone to phase errors.

it is orthogonal ([8] Section 6.10), and it is trivial to show that R' must be orthogonal as well (i.e. that $(R')^\top = (R')^{-1}$). Consider a matrix of input states \mathbb{S}^{in} , whose column Stokes vectors are rotated arbitrarily in polarization space (the Poincaré sphere). The pseudoinverse of the resulting matrix is

$$(R'\mathbb{S}^{in})^+ = (\mathbb{S}^{in})^+ (R')^\top. \quad (27)$$

Taking the norm of this rotated pseudoinverse, we find that $\|(R'\mathbb{S}^{in})^+\| = \|(\mathbb{S}^{in})^+\|$, and so it is invariant under an arbitrary rotation. Thus, any set of input Stokes vectors, including the optimal configurations found in Section 2.3.2 can be freely rotated in polarization space. *Consequently, not only does a cubic configuration of Stokes vectors (Fig. 2(c)) yield maximum Mueller matrix robustness to noise, but it also minimizes the risk of phase errors, since it is the only solid that can be naturally oriented such that every vertex falls precisely in the areas which are least prone to phase errors.*

3. Simulated results

In order to verify the results of Section 2.3 we have simulated a dual photoelastic modulator Stokes polarimeter and used the model to examine the dependency of RMS Mueller errors on Stokes parameter errors. Due to the random nature of the errors considered in this paper, a computer model is a useful tool since it can quickly simulate a large number of noisy Mueller

matrices using different sets of input vectors, whereas experimental measurements take more time, which can make it difficult to gather enough data to accurately infer a distribution.

In order to simulate the dual PEM polarimeter, we began by considering a known Mueller matrix and then chose $n \geq 4$ Stokes vectors at random, all of which were normalized and represented fully polarized states. (The results presented here are based on the experimentally-derived matrix considered by Ghosh et al. [5] in their decomposition, since it is known to exhibit depolarization, retardance, and diattenuation – three important elements in biomedical applications, although in practice the choice of Mueller matrix is largely irrelevant to the results.) The idealized output Stokes vectors were obtained by multiplying the input states by the pre-determined Mueller matrix, as in Eq. (4), after which Gaussian errors and phase errors were added to both the input and the output Stokes vectors. Phase errors were simulated by changing the sign of each polarization parameter (Q , U , and V) with a probability shown in Fig. 4, such that as each corresponding normalized parameter approached ± 1 , the probability of a phase error approached 0, and as each approached 0 the probability of a phase error approached $1/2$.

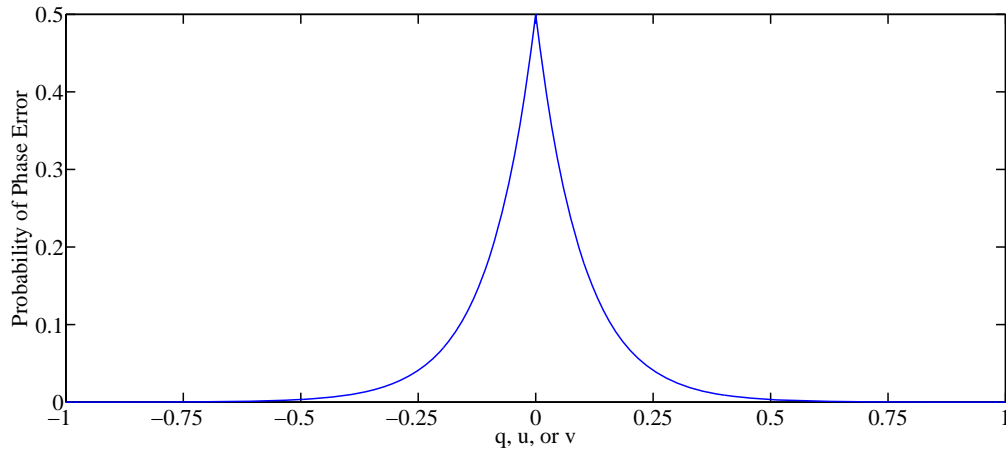


Fig. 4. The probability of a simulated phase inversions as a function of q , u , or v . The likelihood of a phase error approaches 0.5 when the polarization parameters are near 0, and it drops quickly to ~ 0 as said parameters move towards ± 1 .

Then, the error-infused Stokes vectors were used to calculate $M + \delta M$, as in Eq. (9), from which the original Mueller matrix was subtracted to isolate δM . This procedure was repeated for $\sim 10^4$ sets of randomly chosen Stokes vectors, uniformly distributed over $4 \leq n \leq 20$. The results of each set are shown as blue dots in Fig. 5 (Media 7). To test the validity of the results from Section 2.3, we also performed the same procedure 50 times using the derived cubic configuration, and plotted the maximum observed Mueller error, $\langle \delta M \rangle_{\max}$ in red on Fig. 5 (Media 7) for comparison.

The most striking result from Fig. 5 (Media 7) is that the variance in $\langle \delta M \rangle$ and $\|(\mathbb{S}^{in})^+\|$ is very large for 4 randomly chosen input Stokes vectors, but as n increases, the points become more tightly clustered and the Mueller error decreases. Fig. 6 shows the various 2-dimensional projections of this plot.

In Fig. 6(a) (Media 8) we see that both the upper and lower bounds of $\|(\mathbb{S}^{in})^+\|$ decrease as n grows. The theoretically predicted minimum in Eq. (25) has been overlaid on the plot and provides an excellent fit to the lower bound, which suggests that for every n , some of the randomly chosen sets of input vectors correspond approximately to optimal configurations, like the ones shown in Fig. 2. These optimal sets are the ones for which $\langle \delta M \rangle$ is minimized,

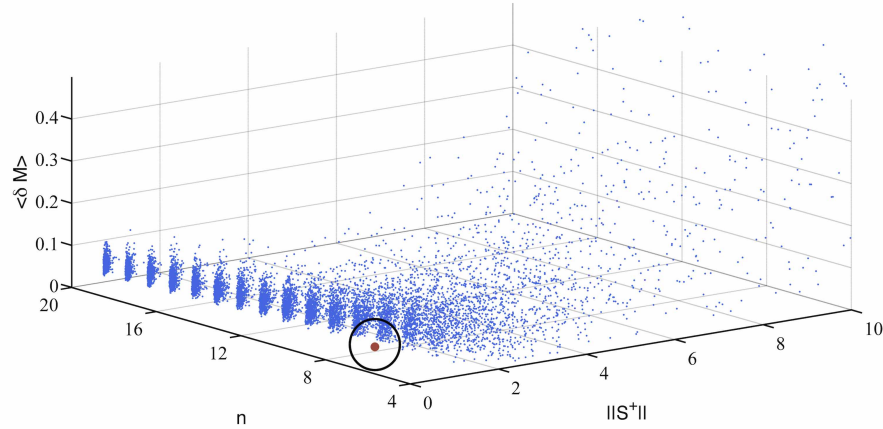


Fig. 5. (Media 7). The Mueller errors resulting from $\sim 10^4$ randomly chosen sets of n Stokes vectors, evenly distributed over $4 \leq n \leq 20$. Each blue dot represents one simulated Mueller matrix determination, and $\langle \delta M \rangle$ on the vertical axis shows the RMS error in said matrix. On the horizontal plane, n and $\|(\mathbb{S}^{in})^+\|$ serve to describe and quantify each configuration. The red dot surrounded by the black circle at the bottom of the plot shows the simulation result corresponding to the optimal cubic configuration in Fig. 3 (Media 6).

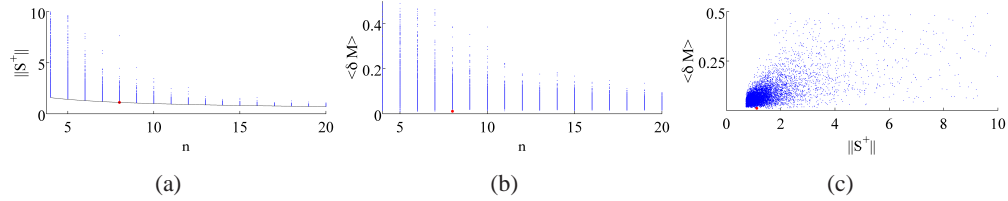


Fig. 6. The 2-dimensional projections of Fig. 5 (Media 7). (a) shows the norm of $(\mathbb{S}^{in})^+$ for randomly chosen sets of input states versus n (Media 8), (b) shows the RMS Mueller error associated with these sets of input states, again versus n (Media 9), and (c) plots the RMS Mueller error against the norm of $(\mathbb{S}^{in})^+$ for each simulated set (Media 10).

for a given n . Examining Fig. 6(b) (Media 9) we see that the lower bound of the Mueller error is approximately constant for all n . Since the states with the lowest Mueller error should correspond to the optimal configuration of Stokes vectors, the observed result that $\langle \delta M \rangle_{min}$ is largely independent of n is consistent with Eq. (25). Finally, examining Fig. 6(c) (Media 10) we see that the cubic configuration (the red dot) does in fact provide the lowest Mueller error of all of the simulated configurations, and therefore it allows for the most robust possible Mueller matrix measurements with a dual PEM Stokes polarimeter.

4. Discussion

The main problem considered in this paper is that of using a polarization state generator and analyzer to determine the Mueller matrix of a sample. More specifically, we have examined the propagation of errors from Stokes vector measurements to the determined Mueller matrices, and we have derived sufficient conditions (Eqs. (21)–(23)) for sets of Stokes vectors to be generated by the PSG which minimize the effect of these errors on the final Mueller matrix. While

several authors have considered similar problems, they have often done so by introducing additional mathematical artifacts to the experimentally-oriented Stokes/Mueller formalism, such as condition numbers and induced norms, which lack clear physical interpretations. To avoid this problem, we have based our analysis on the root mean square error in the determined Mueller matrix, which provides a physically meaningful quantification of accuracy with minimal additional mathematical formalism.

We found that, for a given number of input/output measurements, sets of input Stokes vectors forming the vertices of Platonic solids on the Poincaré sphere provide the most error-resistant results for their respective number of measurements, and that, to a first order approximation, the robustness of the results is independent of the number of different polarization states considered, so long as these states satisfy Eqs. (21)–(23). Thus, the question of how many different input states should be used must be answered based on the specific polarimeter in question and the intended applications. For example, if one wishes to determine a sample's Mueller matrix as quickly as possible with a PSA that is only prone to generic noise (rather than any instrument-specific errors, such as relative phase errors) then four input Stokes vectors forming a tetrahedron in polarization space (Fig. 2(a) [Media 1](#)) would suffice. This tetrahedral optimum is well known [14, 15, 17, 18, 27, 28], and an example of such a configuration (i.e. one of the infinitely many possible orientation on the Poincaré sphere) is the set of input states

$$\left\{ \begin{bmatrix} 1 \\ 0.68 \\ 0.70 \\ 0.21 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.43 \\ -0.75 \\ -0.50 \end{bmatrix}, \begin{bmatrix} 1 \\ -0.73 \\ 0.40 \\ -0.56 \end{bmatrix}, \begin{bmatrix} 1 \\ -0.39 \\ -0.35 \\ 0.85 \end{bmatrix} \right\}. \quad (28)$$

For a general polarimeter, the orientation of the optimal solid formed by the Stokes vectors in polarization space is unimportant, although specific instrumental considerations may strongly favour one particular configuration (e.g. non-tetrahedral) and orientation. For example, in the case of a dual photoelastic modulator polarization state analyzer, which is prone to phase errors, the set of input Stokes vectors

$$\left\{ \begin{bmatrix} 1 \\ 0.58 \\ 0.58 \\ 0.58 \end{bmatrix}, \begin{bmatrix} 1 \\ -0.58 \\ 0.58 \\ 0.58 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.58 \\ -0.58 \\ 0.58 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.58 \\ 0.58 \\ -0.58 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.58 \\ -0.58 \\ -0.58 \end{bmatrix}, \begin{bmatrix} 1 \\ -0.58 \\ 0.58 \\ -0.58 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.58 \\ -0.58 \\ -0.58 \end{bmatrix}, \begin{bmatrix} 1 \\ -0.58 \\ -0.58 \\ -0.58 \end{bmatrix} \right\}, \quad (29)$$

which form the vertices of a cube on the Poincaré sphere (as shown in Fig. 2(c) [Media 3](#)) is a global optimum in terms of robustness to measurement errors. Our simulation showed that out of over 10 000 sets of input states, this particular set did in fact provide the best results. The simulation also showed that as the number of different input states grows, the Mueller matrix robustness generally increases, and for certain sets of input Stokes vectors, the robustness approaches that of the cubic configuration. However, this configuration (Fig. 2(c) [Media 3](#)) remains advantageous as it allows one to determine the Mueller matrix using only 8 input Stokes vectors with a level of accuracy that slightly exceeds measurements made with as many as 20 other different input states.

5. Conclusions

We have discussed methods of using Stokes polarimeters for Mueller matrix measurements, and we used the root mean square of the Mueller matrix error elements to quantify the impact of noise in the Stokes parameter on the derived Mueller matrix. We then derived the conditions for *optimal configurations* of input Stokes vectors, and we found that vectors forming Platonic

solids on the Poincaré sphere satisfy these conditions. We found that, to first order, the number of Stokes vectors in an *optimal set* does not affect the robustness of the Mueller matrix measurements, and simulations confirmed this result. Finally, we deduced that in the case of a dual PEM Stokes polarimeter which is prone to phase errors, input Stokes vectors forming a cube on the Poincaré sphere whose vertices are equidistant from the areas prone to such errors represents a global optimum—a result which was also confirmed by simulation. We plan to report on experimental tests of these results in a future publication. Overall, this work presents a general framework for polarimetric optimization strategy, as well as furnishing practical ‘recipes’ for an optics researcher viz. experimental methodology for robust Mueller matrix determination with a dual-PEM polarimeter.