Bounding the relative errors associated with a complete Stokes polarimeter

Emanuel Chironi* and Claudio Iemmi

Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, 1428 Buenos Aires, Argentina

**Corresponding author: manuchironi@hotmail.com*

Received July 19, 2013; revised November 21, 2013; accepted November 21, 2013; posted November 22, 2013 (Doc. ID 194210); published December 6, 2013

In this paper, we propose a criterion for bounding the relative errors associated with the determination of the Stokes vector that describes the state of polarization of a light beam. No assumptions about the magnitude, origin, or statistical behavior of the errors are made. It is shown that figures of merit such as the condition number and the equally weighted variance naturally arise as optimization parameters. Moreover, a third optimization parameter emerges, which takes into account errors associated with the matrix that represents the selected configuration of analyzers. Finally, a new and more general figure of merit is derived from this analysis and is applied in an optimization process of a very well known polarimeter.

OCIS codes: (120.5410) Polarimetry; (120.2130) Ellipsometry and polarimetry; (100.3190) Inverse problems; (230.5440) Polarization-selective devices.

http://dx.doi.org/10.1364/JOSAA.31.000075

1. INTRODUCTION

The knowledge about the polarimetric properties of a sample, or the state of polarization (SOP) of a light beam, allows us to obtain information related to different physical phenomena. In this sense, polarimetry is a topic of interest in many research fields, such as biological tissue analysis [1,2], glucose concentration sensing [3], determination of optical active molecules in turbid media [4], vegetation remote sensing [5], or cloud detection [6].

The SOP of a light beam is well determined if the corresponding four Stokes parameters are known. They can be obtained from radiometric measurements performed with a polarimeter. This device is said to be complete if it allows us to determine the four parameters.

A Stokes polarimeter is conformed by polarizing optical elements such as polarizers and waveplates, whose axis orientations and/or phase retardations can be modified. Each possible configuration of these elements constitutes a particular polarization state detector (PSD).

The measurement technique consists of projecting the SOP of the incoming light beam on different PSDs and registering the corresponding intensity values. The set of radiometric measurements yields a linear equation system. Therefore, at least four linearly independent equations, obtained from adequate configurations, are required in order to determine all parameters.

Mathematically, the equation system can be written in matrix form as

$$\mathbf{I} = \mathcal{Q}\mathbf{S}, \tag{1}$$

where $\mathbf{I}$ is the intensity measured vector, $\mathbf{S}$ is the Stokes vector, and $\mathcal{Q}$ is the matrix of analyzers of the system (the matrix that contains all the configurations over which the input vector is projected). In order to find the Stokes vector, the matrix $\mathcal{Q}$ should be inverted. When four measurements are performed, the $\mathcal{Q}$ matrix has the dimensions 4×4 and the system can be solved by multiplying by the inverse matrix on both sides of Eq. (1).

However, while four measurements performed in adequate configurations are enough to get a complete polarimeter, it has been reported in previous references [7–9] that precision in the results improves when redundant information is added, i.e., by conducting $N > 4$ projections. In that case, the matrix $\mathcal{Q}$ from Eq. (1) is rectangular (of $N \times 4$ dimensions) and therefore has no inverse.

Then the system represented by Eq. (1) may be solved by defining the *Moore–Penrose pseudoinverse* [10]

$$\tilde{\mathcal{Q}} \equiv (\mathcal{Q}^T \mathcal{Q})^{-1} \mathcal{Q}^T,$$

so that the Stokes vector we want to determine may be written as

$$\mathbf{S} = \tilde{\mathcal{Q}}\mathbf{I}.$$

Experimentally, both matrix $\mathcal{Q}$ and the set of measured intensities have associated uncertainties that affect the accuracy of the instrument in determining the Stokes vector. This accuracy can be improved by adequately choosing the number of measurements to be performed and the distribution of PSDs over the space of possible configurations of the polarimeter.

Thus, it is necessary to find a set of figures of merit that allow us to evaluate the performance of the instrument in relation to the selected configurations. Different parameters have been proposed in previous papers, such as the *condition number* (CN) [9,11], the *equally weighted variance* (EWV)

[7,9,11], the *absolute reciprocal determinant* [7], and the *determinant of the matrix of analyzers* [12]. Nevertheless, to our knowledge, the relation of these parameters to the accuracy of the instrument has not been explored, nor why they are appropriated to be used in an optimization process. Moreover, in previous articles, besides prioritizing a particular source of error, it is assumed that the inaccuracies are small. Some papers take into account only intensity errors [9], while others only consider uncertainties related to the matrix of analyzers, but they are not propagated to the pseudoinverse [8].

In this paper, we propose a criterion for bounding the relative errors associated with the determination of the Stokes vector **S**. No assumptions about the magnitude, origin, or statistical behavior of the errors are made. It is shown that figures of merit such as the CN and the EWV naturally arise as optimization parameters. Moreover, the influence of these quantities on the optimization process is explained through the proposed bounding relationship. Finally, a new and more general figure of merit is derived from this analysis. An example is performed in order to show how this new optimization parameter should be used and some of its benefits.

The general strategy that we propose has been applied in other references to analyze the errors associated with Mueller matrix polarimeters [13–15]. However, in this paper we introduce a particular approach based on the power series expansion of errors.

2. ERROR ANALYSIS FOR A COMPLETE STOKES POLARIMETER

Let us suppose that $\delta\mathbf{I}$ and $\delta\tilde{\mathcal{Q}}$ are uncertainties associated with the intensity measurements and the matrix of analyzers, respectively. The last one includes calibration errors, positioning errors, misalignments, etc. Then $\tilde{\mathcal{Q}}$ and $\mathbf{I}$ values will be contained in the intervals $[\tilde{\mathcal{Q}} - \delta\tilde{\mathcal{Q}}, \tilde{\mathcal{Q}} + \delta\tilde{\mathcal{Q}}]$ and $[\mathbf{I} - \delta\mathbf{I}, \mathbf{I} + \delta\mathbf{I}]$, respectively. Thus, if variables $\tilde{q} \subset [-\delta\tilde{\mathcal{Q}}, \delta\tilde{\mathcal{Q}}]$ and $\mathbf{i} \subset [-\delta\mathbf{I}, \delta\mathbf{I}]$ are defined, the equation corresponding to $\mathbf{S}$ could be written as

$$\mathbf{S}(\tilde{q}, \mathbf{i}) = (\tilde{\mathcal{Q}} + \tilde{q})(\mathbf{I} + \mathbf{i}). \tag{2}$$

This equation implicitly contains uncertainties in **S**. Let us define the mean value of **S** as $\bar{\mathbf{S}} = \tilde{\mathcal{Q}}\mathbf{I}$. Then, if this expression is subtracted from Eq. (2), fluctuations $\delta\mathbf{S}$ around the mean value can be expressed as

$$\delta\mathbf{S}(\tilde{q}, \mathbf{i}) = \mathbf{S}(\tilde{q}, \mathbf{i}) - \bar{\mathbf{S}} = (\tilde{\mathcal{Q}} + \tilde{q})(\mathbf{I} + \mathbf{i}) - \tilde{\mathcal{Q}}\mathbf{I}. \tag{3}$$

It should be taken into account that in error analysis what matters is the overall magnitude of the fluctuation, and that this quantity can be obtained from the norm of matrices and vectors, respectively. Thus by distributing the brackets and applying the appropriate norm on both sides of the previous equation, it results in

$$\|\delta\mathbf{S}(\tilde{q}, \mathbf{i})\| = \|\tilde{\mathcal{Q}}\mathbf{i} + \tilde{q}\mathbf{I} + \tilde{q}\mathbf{i}\|. \tag{4}$$

Let us write Eq. (3) in indicial form as

$$\delta S_j(\tilde{q}, \mathbf{i}) = \sum_k \tilde{\mathcal{Q}}_{jk} i_k + \sum_k \tilde{q}_{jk} I_k + \sum_k \tilde{q}_{jk} i_k,$$

where j and k represent the row and column indexes, respectively.

If variables i_k and $\tilde{q}_{jk}$ vary in a continuous way and they are bounded for all j and k, then by the continuity of the sum and product operations, it is found that δS_j is a bounded and continuous function of i_k and $\tilde{q}_{jk}$ for all j and k. Moreover, from the continuity of the norm, it follows that $\|\delta\mathbf{S}\|$ is a bounded and continuous function of $\tilde{q}$ and $\mathbf{i}$. Then the Weierstrass generalized theorem states that there exist $\tilde{q}_+$ ($\tilde{q}_-$) and $\mathbf{i}_+$ ($\mathbf{i}_-$) such that the function reaches a maximum (minimum) in the considered region [16]. It is supposed in this argument that uncertainty intervals are of a finite size. If any of them (or both) were infinite, the maximum value of the function would tend to infinity, because

$$\lim_{\delta\mathbf{I}, \delta\tilde{\mathcal{Q}} \to \infty} \|\delta\mathbf{S}\| = \infty.$$

Thus, if we define $\Delta\mathbf{S} \equiv \delta\mathbf{S}(\tilde{q}_+, \mathbf{i}_+)$, then $\|\Delta\mathbf{S}\| = \max\{\|\delta\mathbf{S}\|\}$ and $\Delta\mathbf{S}$ is the maximum available error in computing **S**. This implies that the real value of **S** belongs to the interval $[\mathbf{S} - \Delta\mathbf{S}, \mathbf{S} + \Delta\mathbf{S}]$.

As $\Delta\mathbf{S}$ depends on the values of $\tilde{q}_\pm$ and $\mathbf{i}_\pm$, which in turn depend on the values of **I** and $\tilde{\mathcal{Q}}$, its magnitude depends on the error associated with each particular measurement. However, if an optimization process were carried out in order to improve the performance of the instrument, it would be desirable to find a bounding criterion related to the extreme values of the error interval instead of one related to the errors associated with particular measurements.

In order to do so, it is enough to remember that

$$\|\tilde{q}_+\| \leq \|\delta\tilde{\mathcal{Q}}\|, \quad \|\mathbf{i}_+\| \leq \|\delta\mathbf{I}\|,$$

and by the matrix norm properties (Appendix A), Eq. (4) may be bounded as

$$\begin{aligned} \|\delta\mathbf{S}(q, \mathbf{i})\| &\leq \|\tilde{\mathcal{Q}}\mathbf{i}_+\| + \|q_+\mathbf{I}\| + \|q_+\mathbf{i}_+\| \\ &\leq \|\tilde{\mathcal{Q}}\|\|\mathbf{i}_+\| + \|q_+\|\|\mathbf{I}\| + \|q_+\|\|\mathbf{i}_+\| \\ &\leq \|\tilde{\mathcal{Q}}\|\|\delta\mathbf{I}\| + \|\delta\tilde{\mathcal{Q}}\|\|\mathbf{I}\| + \|\delta\tilde{\mathcal{Q}}\|\|\delta\mathbf{I}\|, \end{aligned} \tag{5}$$

where the last expression no longer depends on the optimization variables $\tilde{q}$ and **i**. So, there is a bound for $\|\delta\mathbf{S}\|$ as a function of $\delta\mathbf{I}$ and $\delta\tilde{\mathcal{Q}}$.

It is important to point out that in our analysis it is assumed that the norms of vectors and matrices are *consistent* [10] (see Appendix A).

It is well known that precision in the measurement of a quantity A is associated with the relative error $\varepsilon(A) \equiv \|\delta A\|/\|A\|$. In this way it is desirable to find an expression for $\varepsilon(\mathbf{S})$. Thus, if the norm is applied to both sides of Eq. (1), and by considering property (A4) of Appendix A, it leads to

$$\|\mathcal{Q}\|\|\mathbf{S}\| \geq \|\mathbf{I}\|. \tag{6}$$

Finally, if expression (5) is multiplied by expression (6), and by grouping conveniently, we obtain

$$\frac{\|\delta\mathbf{S}\|}{\|\mathbf{S}\|} \leq \left(\frac{\|\delta\mathbf{I}\|}{\|\mathbf{I}\|} + \frac{\|\delta\tilde{\mathcal{Q}}\|}{\|\tilde{\mathcal{Q}}\|} + \frac{\|\delta\mathbf{I}\|}{\|\mathbf{I}\|}\frac{\|\delta\tilde{\mathcal{Q}}\|}{\|\tilde{\mathcal{Q}}\|}\right)\|\tilde{\mathcal{Q}}\|\|\mathcal{Q}\|. \quad (7)$$

Additionally, keeping in mind that the CN is defined as [17]

$$\kappa(\mathcal{Q}) \equiv \|\tilde{\mathcal{Q}}\|\|\mathcal{Q}\|,$$

the expression for the relative error $\varepsilon(\mathbf{S})$ is given by

$$\varepsilon(\mathbf{S}) \leq (\varepsilon(\mathbf{I}) + \varepsilon(\tilde{\mathcal{Q}}) + \varepsilon(\mathbf{I})\varepsilon(\tilde{\mathcal{Q}}))\kappa(\mathcal{Q}). \quad (8)$$

Therefore, the CN naturally arises as an optimization parameter, given that it is a bounding factor that is independent of the $\mathbf{S}$ vector to be measured.

As a following step, it should be considered that Eq. (8) depends on the unknown parameter $\delta\tilde{\mathcal{Q}}$ (in fact, $\delta\mathcal{Q}$ is the known parameter); thus it is convenient to write $\delta\tilde{\mathcal{Q}}$ in terms of $\mathcal{Q}$, $\tilde{\mathcal{Q}}$, and $\delta\mathcal{Q}$. To do so, and by analogy with Eq. (3), let us write

$$\delta\tilde{\mathcal{Q}}(q) = [(\mathcal{Q}^T + q^T)(\mathcal{Q} + q)]^{-1}(\mathcal{Q}^T + q^T) - \tilde{\mathcal{Q}},$$

where $q \subset [-\delta\mathcal{Q}, \delta\mathcal{Q}]$ is a variable parameter.

Now, the matrix norm can be applied to both sides of the previous equation in order to get

$$\|\delta\tilde{\mathcal{Q}}(q)\| = \|[(\mathcal{Q}^T + q^T)(\mathcal{Q} + q)]^{-1}(\mathcal{Q}^T + q^T) - \tilde{\mathcal{Q}}\|.$$

By considering the same arguments explained at the beginning of this section, there exists a q_+ value that maximizes the error (or that makes the matrix norm tend to infinity), and as a consequence

$$\|\delta\tilde{\mathcal{Q}}\| = \|[\mathcal{Q}^T\mathcal{Q}(I_n + \Gamma_+)]^{-1}(\mathcal{Q}^T + q_+^T) - \tilde{\mathcal{Q}}\|,$$

where $\Gamma_+ = (\mathcal{Q}^T\mathcal{Q})^{-1}[(\mathcal{Q}^T + q_+^T)q_+ + q_+^T\mathcal{Q}]$ and I_n represents the identity matrix of order n. From the inverse matrix properties and those related to the inverse of the product of two matrices, we obtain

$$\begin{aligned}\|\delta\tilde{\mathcal{Q}}\| &= \|(I_n + \Gamma_+)^{-1}(\mathcal{Q}^T\mathcal{Q})^{-1}(\mathcal{Q}^T + q_+^T) - \tilde{\mathcal{Q}}\| \\ &= \|(I_n + \Gamma_+)^{-1}[\tilde{\mathcal{Q}} + (\mathcal{Q}^T\mathcal{Q})^{-1}q_+^T] - \tilde{\mathcal{Q}}\|.\end{aligned}$$

Taking out $(I_n + \Gamma_+)^{-1}$ as a common factor and distributing the remaining brackets, the previous equation reduces to

$$\begin{aligned}\|\delta\tilde{\mathcal{Q}}\| &= \|(I_n + \Gamma_+)^{-1}[(\mathcal{Q}^T\mathcal{Q})^{-1}q_+^T - \Gamma_+\tilde{\mathcal{Q}}]\| \\ &\leq \|(I_n + \Gamma_+)^{-1}\|[\|(\mathcal{Q}^T\mathcal{Q})^{-1}q_+^T\| + \|\Gamma_+\tilde{\mathcal{Q}}\|] \\ &\leq \|(I_n + \Gamma_+)^{-1}\|[\|(\mathcal{Q}^T\mathcal{Q})^{-1}\|\|q_+^T\| + \|\Gamma_+\|\|\tilde{\mathcal{Q}}\|].\end{aligned}$$

From the singular value decomposition [10] it can be shown that $\|(\mathcal{Q}^T\mathcal{Q})^{-1}\| \leq \|\tilde{\mathcal{Q}}\|^2$ (see Appendix B), and consequently

$$\begin{aligned}\|\delta\tilde{\mathcal{Q}}\| &\leq \|(I_n + \Gamma_+)^{-1}\|(\|\tilde{\mathcal{Q}}\|^2\|q_+^T\| + \|\Gamma_+\|\|\tilde{\mathcal{Q}}\|), \\ \varepsilon(\tilde{\mathcal{Q}}) &\leq \|(I_n + \Gamma_+)^{-1}\|(\|\tilde{\mathcal{Q}}\|\|q_+^T\| + \|\Gamma_+\|).\end{aligned}$$

Additionally, the application of the norm on the definition of Γ_+ leads to

$$\begin{aligned}\|\Gamma_+\| &= \|(\mathcal{Q}^T\mathcal{Q})^{-1}[(\mathcal{Q}^T + q_+^T)q_+ + q_+^T\mathcal{Q}]\| \\ &\leq \|(\mathcal{Q}^T\mathcal{Q})^{-1}\|\|[(\mathcal{Q}^T + q_+^T)q_+ + q_+^T\mathcal{Q}]\| \\ &\leq \|(\mathcal{Q}^T\mathcal{Q})^{-1}\|[\|\mathcal{Q}^T + q_+^T\|\|q_+\| + \|q_+^T\|\|\mathcal{Q}\|] \\ &\leq \|\tilde{\mathcal{Q}}\|^2\|q_+\|(2\|\mathcal{Q}\| + \|q_+\|) \\ &\leq \|\tilde{\mathcal{Q}}\|^2\|\mathcal{Q}\|\|\delta\mathcal{Q}\|[2 + \varepsilon(\mathcal{Q})] \\ &\leq \kappa(\mathcal{Q})^2\varepsilon(\mathcal{Q})[2 + \varepsilon(\mathcal{Q})], \end{aligned} \quad (9)$$

where it was considered that $\|q_+\| = \|q_+^T\|$ (see Appendix A) and that $\|q_+\| \leq \|\delta\mathcal{Q}\|$ (this last relationship is evident from the definition of q).

Likewise, we find that

$$\|\tilde{\mathcal{Q}}\|\|q_+^T\| \leq \|\tilde{\mathcal{Q}}\|\|\delta\mathcal{Q}\| = \kappa(\mathcal{Q})\varepsilon(\mathcal{Q}) \leq \kappa(\mathcal{Q})^2\varepsilon(\mathcal{Q}).$$

In this way, from these bounding relations, it follows that

$$\varepsilon(\tilde{\mathcal{Q}}) \leq \kappa(\mathcal{Q})^2\varepsilon(\mathcal{Q})[3 + \varepsilon(\mathcal{Q})]\|(I_n + \Gamma_+)^{-1}\|.$$

The last factor on the second member of the equation above can be expanded as a power series of Γ_+. From the theory of matrix power series it is known that the radius of convergence is less than unity, i.e., the Taylor series will be convergent if it satisfies the condition $\|\Gamma\| \leq 1$ [18]. However, even if the power series is not convergent, it will be majorant with respect to the original function. This fact implies that in both cases the relation

$$\|(I_n + \Gamma)^{-1}\| \leq \sum_{i=0}^{\infty}\|\Gamma\|^i$$

holds, and as a consequence it is possible to write, in general,

$$\varepsilon(\tilde{\mathcal{Q}}) \leq \kappa(\mathcal{Q})^2\varepsilon(\mathcal{Q})[3 + \varepsilon(\mathcal{Q})]\sum_{i=0}^{\infty}\|\Gamma_+\|^i. \quad (10)$$

In addition, Eq. (9) leads to

$$\sum_{i=0}^{\infty}\|\Gamma_+\|^i \leq \sum_{i=0}^{\infty}\{\kappa(\mathcal{Q})^2\varepsilon(\mathcal{Q})[2 + \varepsilon(\mathcal{Q})]\}^i,$$

and by combining the above two equations, we obtain

$$\varepsilon(\tilde{\mathcal{Q}}) \leq \sum_{i=0}^{\infty}\{[\kappa(\mathcal{Q})^2\varepsilon(\mathcal{Q})]^{i+1}[3 + \varepsilon(\mathcal{Q})][2 + \varepsilon(\mathcal{Q})]^i\}.$$

It is apparent that

$$[3 + \varepsilon(\mathcal{Q})][2 + \varepsilon(\mathcal{Q})]^i \leq [3 + \varepsilon(\mathcal{Q})]^{i+1},$$

so it follows that

$$\varepsilon(\tilde{\mathcal{Q}}) \leq \sum_{i=1}^{\infty}\{\kappa(\mathcal{Q})^2\varepsilon(\mathcal{Q})[3 + \varepsilon(\mathcal{Q})]\}^i,$$

where the initial value of index i was shifted in order to introduce the change $i + 1 \rightarrow i$ in the sum exponent.

The factor between braces is a monotonous function of $\kappa(\mathcal{Q})$ and $\varepsilon(\mathcal{Q})$. Therefore, the complete series behaves monotonously, in relation to both parameters.

Then, let us define the functions

$$\xi(\mathcal{Q}) \equiv \kappa(\mathcal{Q})^2 \varepsilon(\mathcal{Q})[3 + \varepsilon(\mathcal{Q})] \tag{11}$$

and

$$\Xi(\mathcal{Q}) \equiv \sum_{i=1}^{\infty} [\xi(\mathcal{Q})]^i.$$

In this way, Eq. (8) can be written as

$$\varepsilon(\mathbf{S}) \leq \{\varepsilon(\mathbf{I}) + [1 + \varepsilon(\mathbf{I})]\Xi(\mathcal{Q})\}\kappa(\mathcal{Q}), \tag{12}$$

where the relation $\varepsilon(\tilde{\mathcal{Q}}) \leq \Xi(\mathcal{Q})$ has been used.

Equation (12) shows that the parameters that impose a limit on the precision of a polarimeter are $\kappa(\mathcal{Q})$, $\varepsilon(\mathbf{I})$, and $\Xi(\mathcal{Q})$. Also, it is derived from this equation that

$$\lim_{\varepsilon(\mathcal{Q}),\varepsilon(\mathbf{I}) \to 0} \varepsilon(\mathbf{S}) = 0,$$

which means that the relative error associated with **S** approaches zero as the uncertainties related to the intensity measurements and the matrix of analyzers decrease, as was expected.

So far, it has been shown that the figure of merit CN arises naturally as an optimization parameter. Let us prove that the same is valid for the EWV. In order to do that, Eq. (11) can be rewritten as

$$\xi(\mathcal{Q}) \equiv \|\tilde{\mathcal{Q}}\| \|\delta\mathcal{Q}\| [3\kappa(\mathcal{Q}) + \|\tilde{\mathcal{Q}}\| \|\delta\mathcal{Q}\|],$$

where the explicit dependence on $\|\delta\mathcal{Q}\|$ and $\|\tilde{\mathcal{Q}}\|$ appears. Consequently, Eq. (12) becomes

$$\varepsilon(\mathbf{S}) \leq \{\varepsilon(\mathbf{I}) + [1 + \varepsilon(\mathbf{I})]\|\tilde{\mathcal{Q}}\| \|\delta\mathcal{Q}\| \Xi'(\mathcal{Q})\}\kappa(\mathcal{Q}), \tag{13}$$

where, in order to shorten it, we have defined

$$\Xi'(\mathcal{Q}) \equiv \kappa(\mathcal{Q})[3 + \varepsilon(\mathcal{Q})] \sum_{i=0}^{\infty} [\xi(\mathcal{Q})]^i.$$

It is not difficult to verify that EWV $= \|\tilde{\mathcal{Q}}\|^2$ when $\|\cdot\|$ is the Frobenius norm [8,17] (see Appendix C). Therefore, Eq. (13) shows that the CN and the EWV emerge as optimization parameters when errors associated with Stokes polarimeters are bounded in terms of the norm of the vectors and matrices related to the measurement process.

It is important that both magnitudes are present in this equation, because, as was shown in previous papers [8,9], the EWV is the figure of merit that takes into account the improvement in the accuracy of the results as the number of measurements N increases.

Additionally, in Eq. (13) a third factor $\|\delta\mathcal{Q}\|$ arises. If $\delta\mathcal{Q}$ is a function of the same variables as $\mathcal{Q}$ (as happens with alignment errors), then the optimal solution should minimize not only the CN and the EWV but also this parameter. In that case, the optimal solution will be a trade-off among the values of the involved figures of merit, and, in general, this solution will differ from that obtained by considering only the CN and the EWV. This fact has been experimentally found in a previous work [8].

The results achieved suggest that a more general function should be considered. Let us define the *overall precision factor* (Λ) as

$$\Lambda(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) \equiv \|\tilde{\mathcal{Q}}\| \|\delta\mathcal{Q}\| \kappa(\mathcal{Q}),$$

(where each $\mathbf{x}_i$ represents a possible configuration of the system). This is a global figure of merit, because it takes into account the dependence of the system on all the involved parameters. This function includes error amplification through matrix product (CN), redundant information (EWV), and errors associated with matrix $\mathcal{Q}$ that depend on the selected configurations of analyzers ($\|\delta\mathcal{Q}\|$). The functional form of each factor will be dependent on the optical elements that conform the polarimeter, its architecture, and degrees of freedom.

3. APPLICATION EXAMPLE

Let us show the use of Λ in an optimization process of a polarimeter based on a fixed linear polarizer and a rotating quarter-wave plate, as is shown in Fig. 1.

In order to compute Λ, we will consider only those uncertainties in alignment of polarizers and quarter-wave plates, because it will simplify the calculation of $\delta\mathcal{Q}$. If a matrix $\mathcal{Q}$ is determined by a set of analyzers $\boldsymbol{X} \equiv (\mathbf{x}_1, \ldots, \mathbf{x}_N)$ that have uncertainties $\delta\boldsymbol{X} \equiv (\delta\mathbf{x}_1, \ldots, \delta\mathbf{x}_N)$, let us call $\mathcal{D} \equiv [\boldsymbol{X} - \delta\boldsymbol{X}, \boldsymbol{X} + \delta\boldsymbol{X}]$. Then we can define

$$\mathcal{Q}_+ \equiv \mathcal{Q}(\boldsymbol{Y}), \tag{14}$$

$$\mathcal{Q}_- \equiv \mathcal{Q}(\boldsymbol{Z}), \tag{15}$$

with $\boldsymbol{Y}$, $\boldsymbol{Z} \subset \mathcal{D}$, such that

$$\mathcal{Q}_{ij}(\boldsymbol{Y}) \geq \mathcal{Q}_{ij}(\boldsymbol{X}'), \quad \mathcal{Q}_{ij}(\boldsymbol{Z}) \leq \mathcal{Q}_{ij}(\boldsymbol{X}'),$$

for all i, j and $\boldsymbol{X}' \subset \mathcal{D}$. Therefore, it is possible to assert that the real value of $\mathcal{Q}$ belongs to $[\mathcal{Q}_-, \mathcal{Q}_+]$.

Thus,

$$\delta\mathcal{Q} \leq \mathcal{Q}_+ - \mathcal{Q}_-, \tag{16}$$

so $\mathcal{Q}_+ - \mathcal{Q}_-$ is a reasonable bounding expression for $\delta\mathcal{Q}$.

For a Stokes polarimeter consisting of a fixed linear polarizer and a rotating quarter-wave plate, each analyzer will be characterized by the angles of rotation of both optical

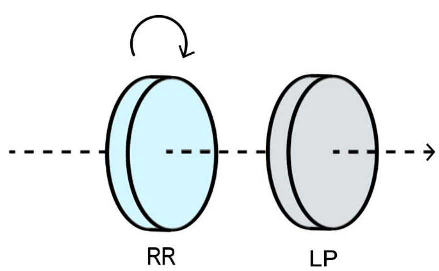

Fig. 1. Basic scheme of the polarimeter used as an example.

Table 1. Numerical Values of the Four Figures of Merit Calculated for Each of the Optimization Processes

	No. 1	No. 2	No. 3	No. 4
Λ	3.96	4.04	4.13	3.79×10^{17}
κ	6.49	6.42	6.42	3.05×10^{9}
$\|\tilde{\mathcal{Q}}\|$	4.59	4.54	4.54	2.16×10^{9}
$\|\delta\mathcal{Q}\|$	0.133	0.139	0.142	0.0576

elements. The value of the retardance of the wave plate is supposed to be exact. But given that the polarizer is fixed, its angle of rotation will be the same for all the analyzers, while the angle of rotation of the wave plate will be different for each one. Consequently, for N analyzers, $N + 1$ parameters should be considered.

In order to find the set of $N + 1$ optimum parameters, for $N = 4$, four different computational simulation processes were performed. Four figures of merit were used as optimization parameters, one for each process: Λ for process no. 1, κ for no. 2, $\|\tilde{\mathcal{Q}}\|$ for no. 3, and $\|\delta\mathcal{Q}\|$ for no. 4. For each optimum configuration obtained, the values of the other figures of merit were calculated. It is important to note that all the calculations were performed by using the Frobenius norm. We computed $\delta\mathcal{Q}$ assuming that all the δX_{ij} were the same and that the experimental uncertainty in all angles was 0.5°, so $\delta X_{ij} = \delta X = 0.5°$. The results of all processes are given in Table 1.

The first remarkable result is that the parameters shown in the first three columns have similar values. In fact, κ and $\|\tilde{\mathcal{Q}}\|$ barely differ in the third significant digit. This fact explains why previously reported optimization processes, in which CN or EWV was minimized, led to a good performance of the polarimeter. Second, it should be noted that the $\|\delta\mathcal{Q}\|$ values are all of the same order of magnitude in all the table columns, but the other parameters are greater by more than eight orders of magnitude in the last column. This means that the $\|\delta\mathcal{Q}\|$ is a very inefficient optimization figure of merit.

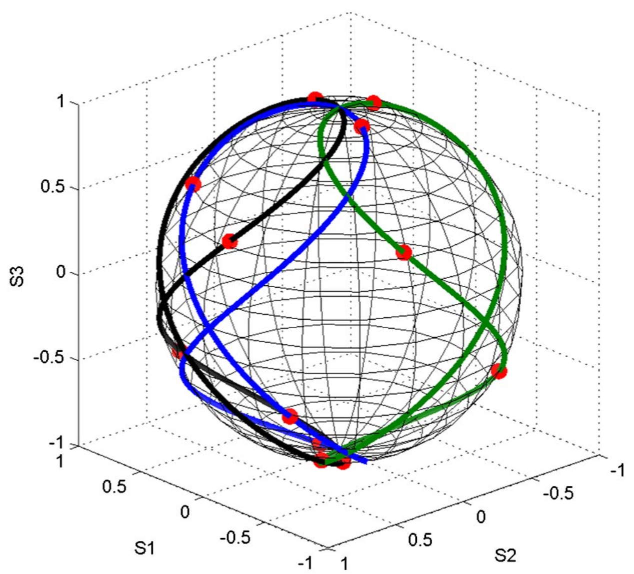

Fig. 2. Optimum analyzers obtained in optimization processes nos. 1, 2, and 3.

In order to analyze the obtained configurations, we plotted the optimum analyzers found through optimization processes nos. 1, 2, and 3 over the Poincaré sphere (see Fig. 2). Each of the continuous curves represents the position of all the analyzers that have the same value for the polarizer angle θ. The blue, green, and black curves correspond to processes nos. 1, 2, and 3, respectively. Note that the blue curve is obtained by rotating the green one at an angle of 50° to the left, and the black one at an angle of 12° to the right.

From the results in Table 1 and Fig. 2, we conclude that optimization processes nos. 1, 2 and 3 lead to configurations that have a similar geometric distribution over the Poincaré sphere and similar conditioning properties (the values of κ and $\|\tilde{\mathcal{Q}}\|$ are very close). However, an optimization process that minimizes Λ provides more complete information, because it takes into account the influence of the error matrix $\delta\mathcal{Q}$ and yields a lower value of its norm. The fact that in this particular case the influence of $\|\delta\mathcal{Q}\|$ seems to be low does not imply that this condition holds for any other polarimeter design.

4. CONCLUSIONS

It has been demonstrated that from a bounding process of the relative errors associated with a complete Stokes polarimeter, the CN and EWV parameters naturally appear as adequate figures of merit for optimization. Furthermore, a third optimization parameter emerges and explains why the inclusion of calibration errors into the optimization process leads to an optimal configuration that does not agree with the configuration obtained when the matrix calibration errors are not considered.

On the other hand, in order to derive these results, no assumptions were made here about the statistical behavior of the involved uncertainties, and therefore the analysis is completely general.

Finally, a global figure of merit was proposed as an optimization parameter. This overall precision factor includes the CN, the EWV, and errors associated with the matrix of analyzers $\mathcal{Q}$.

APPENDIX A

A matrix norm of $\mathcal{A} \in \mathbb{C}^{m\times n}$, denoted by $\|\mathcal{A}\|$, is defined as a function $f: \mathbb{C}^{m\times n} \to \mathbb{R}$ such that

$$\|\mathcal{A}\| \geq 0, \qquad \|\mathcal{A}\| = 0 \Leftrightarrow \mathcal{A} = \mathbf{0}, \tag{A1}$$

$$\|\alpha\mathcal{A}\| = |\alpha|\|\mathcal{A}\|, \tag{A2}$$

$$\|\mathcal{A} + \mathcal{B}\| \leq \|\mathcal{A}\| + \|\mathcal{B}\|, \tag{A3}$$

$$\|\mathcal{A}\mathcal{B}\| \leq \|\mathcal{A}\|\|\mathcal{B}\|, \tag{A4}$$

for every $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{m\times n}$ and $\alpha \in \mathbb{C}$.

Additionally, the norm satisfies

$$\|\mathcal{A}\| = \|\mathcal{A}^T\|. \tag{A5}$$

A given matrix norm may have an important connection with some vector norms. This connection is called *consistency*, and it is said that a matrix norm $\|\cdot\|_M$ and a vector norm $\|\cdot\|_V$ are *consistent* if and only if for any matrix $\mathcal{A} \in \mathbb{C}^{m\times n}$ and any vector $x \in \mathbb{C}^n$

$$\|\mathcal{A}x\|_V \leq \|\mathcal{A}\|_M \|x\|_V. \tag{A6}$$

For example, the Frobenius norm and the usual vector norm are consistent.

APPENDIX B

Let us represent the $\mathcal{Q}$ matrix and its pseudoinverse through their singular value decomposition

$$\mathcal{Q} = U\Sigma V^T \Rightarrow \tilde{\mathcal{Q}} = V\tilde{\Sigma}U^T,$$

where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal matrices.

From this expression, and taking into account the property of *unitary invariance* [10] of the 2-norm and Frobenius norm (which are the commonly used in polarimetry), it follows that

$$\|\tilde{\mathcal{Q}}\| = \|V\tilde{\Sigma}U^T\| = \|\tilde{\Sigma}\|.$$

In addition, from the singular value decomposition, it is found that

$$\|(\mathcal{Q}^T\mathcal{Q})^{-1}\| = \|(V\Sigma U^T U\Sigma V^T)^{-1}\| = \|(V\Sigma^2 V^T)^{-1}\|$$
$$= \|(V^T\Sigma^{-2}V)\| = \|\Sigma^{-2}\| \leq \|\tilde{\Sigma}\|^2 = \|\tilde{\mathcal{Q}}\|^2.$$

Therefore,

$$\|(\mathcal{Q}^T\mathcal{Q})^{-1}\| \leq \|\tilde{\mathcal{Q}}\|^2.$$

APPENDIX C

The EWV is usually defined in the literature as

$$\text{EWV}(\mathcal{Q}) \equiv \sum_{j=1}^{r} \frac{1}{\sigma_j^2},$$

where σ_j represents the singular values of matrix $\mathcal{Q}$.

Taking into account the unitary invariance of the Frobenius norm, it is possible to write

$$\|\tilde{\mathcal{Q}}\| = \|\tilde{\Sigma}\|.$$

Additionally, by definition of the Frobenius norm,

$$\|\tilde{\Sigma}\| \equiv [\text{tr}(\tilde{\Sigma}^T\tilde{\Sigma})]^{1/2} = \left(\sum_{j=1}^{r} \frac{1}{\sigma_j^2}\right)^{1/2},$$

and as a consequence,

$$\|\tilde{\mathcal{Q}}\|_F^2 = \|\tilde{\Sigma}\|^2 = \sum_{j=1}^{r} \frac{1}{\sigma_j^2} = \text{EWV}.$$

REFERENCES

1. N. Ghosh and I. A. Vitkin, "Tissue polarimetry: concepts, challenges, applications and outlook," J. Biomed. Opt. **16**, 110801 (2011).
2. P. Shukla and A. Pradhan, "Mueller decomposition images for cervical tissue: potential for discriminating normal and dysplastic states," Opt. Express **17**, 1600–1609 (2009).
3. M. F. G. Wood, N. Ghosh, X. Guo, and I. A. Vitkin, "Towards noninvasive glucose sensing using polarization analysis of multiply scattered light," in *Handbook of Optical Sensing of Glucose in Biological Fluids and Tissues*, V. V. Tuchin, ed., Vol. **12** of Series in Medical Physics and Biomedical Engineering (Taylor & Francis, 2008), Chap. 17.
4. D. Côté and I. A. Vitkin, "Robust concentration determination of optically active molecules in turbid media with validated three-dimensional polarization sensitive Monte Carlo calculations," Opt. Express **13**, 148–163 (2005).
5. S. Tan and R. M. Narayanan, "Design and performance of a multiwavelength airborne polarimetric lidar for vegetation remote sensing," Appl. Opt. **43**, 2360–2368 (2004).
6. G. Horvath, A. Barta, J. Gal, B. Suhai, and O. Haiman, "Ground-based full-sky imaging polarimetry of rapidly changing skies and its use for polarimetric cloud detection," Appl. Opt. **41**, 543–559 (2002).
7. 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).
8. J. S. Tyo, "Design of optimal polarimeters: maximization of signal-to-noise ratio and minimization of systematic error," Appl. Opt. **41**, 619–630 (2002).
9. A. Peinado, A. Lizana, J. Vidal, C. Iemmi, and J. Campos, "Optimization and performance criteria of a Stokes polarimeter based on two variable retarders," Opt. Express **18**, 9815–9830 (2010).
10. A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications* (Springer, 2003).
11. A. Peinado, A. Lizana, J. Vidal, C. Iemmi, and J. Campos, "Optimized Stokes polarimeters based on a single twisted nematic liquid-crystal device for the minimization of noise propagation," Appl. Opt. **50**, 5437–5445 (2011).
12. J. M. Daniels, "Optimum design of a polarimeter for visible light," Rev. Sci. Instrum. **57**, 1570–1573 (1986).
13. K. M. Twietmeyer and R. A. Chipman, "Optimization of Mueller matrix polarimeters in the presence of error sources," Opt. Express **16**, 11589–11603 (2008).
14. I. J. Vaughn and B. G. Hoover, "Noise reduction in a laser polarimeter based on discrete waveplate rotations," Opt. Express **16**, 2091–2108 (2008).
15. D. Layden, M. F. G. Wood, and I. A. Vitkin, "Optimum selection of input polarization states in determining the sample Mueller matrix: a dual photoelastic polarimeter approach," Opt. Express **20**, 20466–20481 (2012).
16. T. M. Apostol, *Calculus*, *Volume II: Multi-Variable Calculus and Linear Algebra, with Applications to Differential Equations and Probability* (Wiley, 1967).
17. G. H. Golub and C. F. Van Loan, *Matrix Computations* (Johns Hopkins University, 1996).
18. J. T. Schwartz, *Introduction to Matrices and Vectors* (Dover, 2001).