Optimal Orientation Angle Configuration of Polarizers Exists in a 3 × 3 Mueller Matrix Polarimeter

Hanyue Wei 1,2,†, Yifu Zhou 1,2,†, Liyong Ren 1,2,* and Feiya Ma 1,2

1 School of Physics and Information Technology, Shaanxi Normal University, Xi’an 710119, China; why24@snnu.edu.cn (H.W.); zhouyf@snnu.edu.cn (Y.Z.); feiyam@snnu.edu.cn (F.M.)
2 Xi’an Key Laboratory of Optical Information Manipulation and Augmentation (OMA), Xi’an 710119, China
* Correspondence: renliy@snnu.edu.cn; Tel.: +86-(29)-8153-0793
† These authors contributed equally to this work.

Abstract: A 3 × 3 Mueller matrix polarimeter, as one of the primary polarization measurement tools, has attracted intensive interest in various fields of polarization optics. In this paper, from a novel viewpoint with more mathematical analysis of the basic configuration and the working performance of a 3 × 3 Mueller matrix polarimeter, we systematically study the relationship between Mueller matrix measurement errors with the configuration and characteristics of polarizers in the polarimeter. We figure out that, mathematically, there is an optimal orientation angle configuration (0°, 60°, and 120°) of polarizers, which inherently ensures a minimal relative measurement error in general. Moreover, we interpret the internal source of such a robust error resistance under this optimized configuration mathematically for the first time. Simulation and practical experiments show the effectiveness of our mathematical analyses, and compared to the frequently used orientation angle configuration, the average relative error is reduced by up to 32% under such an optimized configuration.

Keywords: Mueller matrix; perturbation analysis; polarimeter

1. Introduction

As a key dimension of optical information, polarization has received extensive attention. Optical measurement techniques based on polarization detection mainly focus on interactions between objects and polarized light. Because of the sensitivity of polarization states to morphological characteristics and physical and chemical properties of an object, polarimetry is applied as a precision measurement technology in various fields, such as cancer detection [1–3], characterization of biological tissue [4–7], defect inspection [8,9], etc.

In general, a polarization state can be represented by a Stokes vector in mathematics, which is a four-dimensional real vector. Linear transformation of polarization states caused by an object is described by a Mueller matrix (MM), which is a real matrix with a size of 4 × 4. There is a lot of useful information that could be extracted from the MM. For example, one can characterize the diattenuation, depolarization, and retardance of an object via Mueller matrix polar decomposition [10,11]. On the other hand, 3 × 3 MM as a submatrix of 4 × 4 MM has attracted intensive interest recently. Compared to 4 × 4 MM, 3 × 3 MM only characterizes the linear polarization characteristics of an object, which can be measured by a relatively simplified MM polarimeter that only has a pair of rotatable polarizers without waveplate. Therefore, 3 × 3 MM polarimetry is usually applied in scenarios where the complexity of optical setup is limited, such as endoscopy [12,13].

As the cornerstone of MM polarimetry, the MM polarimeter has become an essential instrument to measure the MMMs of samples. In general, a 3 × 3 MM polarimeter is simple and consists of a pair of rotatable polarizers. Therefore, the orientation angle and the...
extinction ratio of polarizers are two important parameters that could be optimized for minimizing the measurement errors of a polarimeter as much as possible. Although optimal designs of 4 × 4 polarimeters have been widely discussed by previous works [14–16], similar designs of 3 × 3 polarimeters are rarely discussed. In the past few years, there has been solid work [17] that discussed the calibration of a 3 × 3 polarimeter, where the method used to choose an optimal combination of calibration samples is similar to the method for finding the optimal parameters of polarimeters. In this paper, the fundamental principle of a canonical 3 × 3 MM polarimeter, the measurement errors, and the dependence of which on the configurations of polarizers are discussed. Based on rigorous mathematical analysis of the polarimeter, it is found that there always exists an optimal orientation angle configuration (0°, 60° and 120°) of polarizers in the 3 × 3 MM polarimeter, which inherently ensures obtaining a robust error resistance.

In order to verify the optimality of the suggested angle configuration and the validity of the mathematical analyses, some simulations and practical experiments are conducted. The simulations show that the mathematical analyses are rigorous, reasonable, and effective, and the polarimeter with the optimal orientation angle configuration has robust error resistance for three kinds of common errors. The result of the practical experiment is in accord with the results of the simulations. The average relative error of the polarimeter is reduced by up to 32% when such an optimal orientation angle configuration is adopted.

Moreover, we also discussed an optimization of the orientation angle configuration of a frequently used oversampling 3 × 3 MM polarimeter, and the performance of the canonical polarimeter is compared to the oversampling polarimeter. It is found that although the error resistances of both polarimeters are identical in theory, the practical performance of the oversampling 3 × 3 MM polarimeter is a little bit better than the canonical polarimeter.

Other chapters of this paper are organized according to the following structure. Section 2 introduces the 3 × 3 MM of polarizers and the general mathematical model of a 3 × 3 MM polarimeter. Section 3 discusses the optimization of the polarimeter, where the optimal orientation angle configuration is suggested. Section 4 discusses the validation of the optimality of the configuration and the validity of the mathematical analyses. The effectiveness of the error analysis, the optimal orientation angle configuration of the oversampling 3 × 3 MM polarimeter, and the performance comparison between the four-states-polarimeter and the three-states-polarimeter are also discussed in Section 5. The conclusion of the paper is given in Section 6.

2. Principle of a 3 × 3 MM Polarimeter

2.1. The 3 × 3. MMs of Polarizers

Polarizers are the most important elements in the 3 × 3 MM polarimetry. In this subsection, 3 × 3 MMs of polarizers are briefly introduced for future discussions. As examples of diattenuators [10], 3 × 3 MMs of polarizers, denoted by $M_D$, can be expressed in blocks by

$$M_D = m_{00} \begin{bmatrix} 1 & D^T \\ D & m_D \end{bmatrix},$$

where $m_{00}$ is the average intensity coefficient of $M_D$; $D$ is the diattenuation vector, which is a 2 × 1 real vector and satisfies $|D| \leq 1$; $T$ denotes the transpose operation; $m_D$ is a 2 × 2 real matrix that can be obtained by

$$m_D = \sqrt{1-D^2} I_{2 \times 2} + \left(1-\sqrt{1-D^2}\right) \hat{D} \hat{D}^T,$$
in which $I_{2 \times 2}$ is a $2 \times 2$ unit matrix; $D$ is the diattenuation that satisfies $D = |D|D$; $\hat{D}$ is the normalized diattenuation vector that can be obtained by $\hat{D} = D / |D|$. It can be deduced from Equations (1) and (2) that $M_D$ can be fully characterized by $D$ and $m_{00}$.

If $D$ can be parameterized by $D = D_0 \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$, $0 \leq \theta \leq \pi$, the diattenuator characterized by $M_D$ will be a linear diattenuator (or linear polarizer) with orientation angle $\theta$. Further, if $D = 1$, the linear polarizer is an ideal one, which can fully extinct a beam of linearly polarized light with a polarization angle of $\theta \pm (\pi/2)$.

The diattenuation of a polarizer can be calculated from its polarization extinction ratio ER in the following form:

$$D = \frac{ER - 1}{ER + 1}$$

As a special case, $D = 1$ when $ER = \infty$. Moreover, it can be seen in Equation (3) that $D$ is increasing with the increase of $ER$ when $1 \leq ER \leq \infty$, i.e., the higher the diattenuation, the higher the extinction ratio.

### 2.2. Principle of $3 \times 3$ MM Polarimeter

A canonical $3 \times 3$ MM imaging polarimeter is shown in Figure 1 schematically. A fully unpolarized light as the light source of the system is modulated by a linear polarizer (P1) and then illuminates a sample. The transmitted light of the sample is modulated by another linear polarizer (P2). A camera following P2 is used to measure the intensity of the transmitted light of P2. Without loss of generality, assuming $3 \times 3$ MMs of P1 and P2 and the sample to be $M_{PSG}$, $M_{PSA}$, and $M$, respectively, then the three-dimensional Stokes vector of light transmitted through P2, denoted by $s_{out}$, can be expressed by Equation (4):

$$s_{out} = M_{PSA} M_{PSG} s_{in},$$

where $s_{in}$ is the three-dimensional Stokes vector of the light source. In a canonical $3 \times 3$ MM imaging polarimeter system shown in Figure 1, the light source is unpolarized; thus, $s_{in} = [I_{in} \ 0 \ 0]^T$ with $I_{in}$ is the intensity of light. Note that the MMs of P1 and P2, i.e., $M_{PSG}$ and $M_{PSA}$, are two known parameters, which have been described in Section 2.1. Moreover, $I_{in}$ can also be assumed to be known in this work, meaning that unknown parameters in Equation (4) are only $s_{out}$ and $M$.

![Figure 1. Diagram of a canonical $3 \times 3$ MM imaging polarimeter.](image)

As usual, P1 and P2 are known as a polarization state generator (PSG) and a polarization state analyzer (PSA), respectively. The PSG is used to generate different polarized light to illuminate the sample, while the PSA is used to analyze the polarization state of the transmitted light of the sample. For more detail, assuming that the three-dimensional Stokes vector of the transmitted light of the sample is $s = M_{PSG} s_{in}$, then Equation (4) can be rewritten as

$$s_{out} = M_{PSA} s.$$
The first Stokes parameter of $s_{\text{out}}$, denoted as $I_{\text{out}}$, is the intensity of the transmitted light of P2 and can be measured directly by a camera. Assuming the orientation angles of the PSA and the PSG are $\theta_{\text{PSA}}$ and $\theta_{\text{PSG}}$, respectively, it is obvious that $M_{\text{PSA}}$ is relevant to $\theta_{\text{PSA}}$, $s$ is relevant to $\theta_{\text{PSG}}$, and $I_{\text{out}}$ is relevant to both angles. According to Equations (1) and (5), $I_{\text{out}}$ satisfies the equation

$$m_{00} \begin{bmatrix} 1 & \mathbf{D}^T_{\text{PSA}}(\theta_{\text{PSA}}) \\ \mathbf{I}_{\text{PSG}}(\theta_{\text{PSG}}) \end{bmatrix} \mathbf{s}(\theta_{\text{PSG}}) = I_{\text{out}} \begin{bmatrix} \theta_{\text{PSA}} ), \theta_{\text{PSG}} \end{bmatrix},$$

where $\mathbf{D}_{\text{PSA}}(\theta_{\text{PSA}}) = \mathbf{D}_{\text{PSA}}(\theta_{\text{PSA}}) \mathbf{D}_{\text{PSA}}(\theta_{\text{PSG}})$ is the diattenuation vector of PSA and $\mathbf{D}_{\text{PSG}}$ is the corresponding diattenuation. Moreover, $s$ and $I_{\text{out}}$ in Equation (6) are the functions of $\theta_{\text{PSG}}$ and $(\theta_{\text{PSA}}, \theta_{\text{PSG}})$, respectively. If PSA (i.e., P2) was rotated three times and $I_{\text{out}}$ was captured by the camera for each time, then linear equations about $s$ can be built as

$$m_{00} \begin{bmatrix} 1 & \mathbf{D}^T_{\text{PSA}}(\theta_{\text{PSA}}) \\ \mathbf{I}_{\text{PSG}}(\theta_{\text{PSG}}) \end{bmatrix} \mathbf{s}(\theta_{\text{PSG}}) = I_{\text{out}} \begin{bmatrix} \theta_{\text{PSA}}, \theta_{\text{PSG}} \end{bmatrix},$$

or equivalently $\mathbf{A}_{\text{PSA}} s = I_{\text{out}},$

where $\theta_{\text{PSA}}, \theta_{\text{PSG}}$ and $\theta_{\text{PSG}}$ are the orientation angles of PSA for each rotation. If they were chosen to be appropriate angles, the coefficient matrix $\mathbf{A}_{\text{PSA}}$ can be invertible, meaning that $s(\theta_{\text{PSG}})$ can be solved from Equation (7). In a word, the Stokes vector of the light modulated by the camera can be measured by rotating the PSA and measuring the intensity of the transmitted light.

Further, if the PSG (i.e., P1) was rotated three times and each time the $s(\theta_{\text{PSG}})$ was measured accordingly by rotating PSA as mentioned above, based on $s = \mathbf{M} M_{\text{PSG}} s$ and the expression of $M_{\text{PSG}}$, a matrix equation about the $3 \times 3$ MM of the sample, i.e., $\mathbf{M}$, can be built as

$$\mathbf{M} \begin{bmatrix} \mathbf{I}_{\text{in}},m_{00} \end{bmatrix} \begin{bmatrix} \mathbf{D}_{\text{PSA}}(\theta_{\text{PSA}}) \\ \mathbf{D}_{\text{PSG}}(\theta_{\text{PSG}}) \\ \mathbf{D}_{\text{PSG}}(\theta_{\text{PSG}}) \end{bmatrix} \mathbf{s} = \begin{bmatrix} \mathbf{I}_{\text{out}}(\theta_{\text{PSA}}, \theta_{\text{PSG}}) \\ \mathbf{I}_{\text{out}}(\theta_{\text{PSA}}, \theta_{\text{PSG}}) \\ \mathbf{I}_{\text{out}}(\theta_{\text{PSA}}, \theta_{\text{PSG}}) \end{bmatrix},$$

or equivalently $\mathbf{A}_{\text{PSG}} = \mathbf{S},$

where $\mathbf{D}_{\text{PSG}}(\theta_{\text{PSG}}) = \mathbf{D}_{\text{PSG}}(\theta_{\text{PSG}}) \mathbf{D}_{\text{PSG}}(\theta_{\text{PSG}})$ is the diattenuation vector of PSG and $\mathbf{D}_{\text{PSG}}$ is the corresponding diattenuation; $\theta_{\text{PSG}}, \theta_{\text{PSG}},$ and $\theta_{\text{PSG}}$ are orientation angles of PSG for each rotation. If the three orientation angles of the PSG were appropriate, making $\mathbf{A}_{\text{PSG}}$ invertible, $\mathbf{M}$ can then be solved from Equation (8). In fact, according to Equation (7), the constant matrix $\mathbf{S}$ in Equation (8) can be obtained by

$$m_{00} \begin{bmatrix} 1 & \mathbf{D}^T_{\text{PSA}}(\theta_{\text{PSA}}) \\ \mathbf{D}_{\text{PSA}}(\theta_{\text{PSA}}) \\ \mathbf{D}_{\text{PSA}}(\theta_{\text{PSA}}) \end{bmatrix} \mathbf{s} = \begin{bmatrix} \mathbf{I}_{\text{out}}(\theta_{\text{PSA}}, \theta_{\text{PSG}}) \\ \mathbf{I}_{\text{out}}(\theta_{\text{PSA}}, \theta_{\text{PSG}}) \\ \mathbf{I}_{\text{out}}(\theta_{\text{PSA}}, \theta_{\text{PSG}}) \end{bmatrix},$$

or equivalently $\mathbf{A}_{\text{PSA}} \mathbf{S} = \mathbf{I}.$

Then, according to Equations (8) and (9), the intensity matrix $\mathbf{I}$ can be expressed as

$$\mathbf{I} = \mathbf{A}_{\text{PSA}} \mathbf{M} \mathbf{A}_{\text{PSG}}$$

(10)
It should be noted that Equations (7) and (8) imply the two polarizers, i.e., the PSA and PSG, having the same $m_{00}$, meaning that we do not distinguish $m_{00}$ of both. This is owing to the fact that the physical properties of the two polarizers in the polarimeter tend to be the same in general, and the value of $m_{00}$ is not important for a relative MM polarimeter that is used to measure the normalized MM of the sample instead of the precise value of the MM. The relative MM polarimeter is the most used type of polarimeter.

According to the analyses mentioned above, in principle, the measurement of the MM of the sample can be divided into two parts; part one includes rotating the PSA, measuring the intensity, and solving Equation (7); part two includes rotating the PSG, repeating part one and solving Equation (8). Although, in practical applications, it may seem that the MM of the sample is obtained in other ways, not directly by Equations (7) and (8). For example, Ref. [12] calculates the MM by a formula that is deduced from the physical interpretation of the Stokes parameters instead of solving the equations. However, the formula they used can also be deduced from Equations (7) and (8) when $\theta^{(1)}_{\text{PSG}}, \theta^{(2)}_{\text{PSG}}, \theta^{(3)}_{\text{PSG}}$ are chosen to be $0^\circ, 45^\circ, 90^\circ$ and $\theta^{(1)}_{\text{PSA}}, \theta^{(2)}_{\text{PSA}}, \theta^{(3)}_{\text{PSA}}$ to be $0^\circ, 90^\circ, 135^\circ$, respectively. In fact, the measurement principle discussed above has generality, which provides a general mathematical model for a $3 \times 3$ MM polarimeter. All the following discussions will be based on the above principle.

Because there are nine unknown elements in $\mathbf{M}$, one needs to measure the intensity nine times, at least under different orientation angles of the PSA and the PSG. However, a measurement method that has been reported [13] needs to measure the intensity sixteen times for different PSG and PSA states in total to calculate the $3 \times 3$ MM of the sample. These methods add extra equations to Equations (7) and (8) by rotating the PSG and PSA and measuring the intensity more times. The number of equations is larger than the number of known parameters, which means the equations may be over-determined if there is noise in the measurement. In this case, one needs to use the pseudoinverse matrix to obtain $\mathbf{M}$, which will be discussed in Section 5.3. On the other hand, for $4 \times 4$ MM polarimetry, naturally, there is a need to measure the intensity at least sixteen times since there are sixteen unknown elements in $4 \times 4$ MMs. There are other works [18,19] that need more times (more than sixteen times) of rotations and intensity measurements to calculate a $4 \times 4$ MM. These methods, essentially, build over-determined equations via multiple intensity measurements of different polarization states and usually use some numerical methods (least square method or Fourier analysis) to obtain approximate solutions to the equations. The ideas of these $4 \times 4$ methods also can be applied on the $3 \times 3$ MM polarimetry. All the polarimeters that include extra rotations and intensity measurements are referred to as oversampling polarimeters. However, the oversampling polarimeters are not within the scope of this article (except for an oversampling polarimeter discussed in Section 5.3).

Fewer times of rotation and intensity measurement are conducive to reducing the cost of data processing and approving the measurement speed and timeliness.

### 3. Error Analysis of the MM Polarimeter

#### 3.1. Measurement Errors in MM Polarimeter

There are three kinds of sources of errors that should be considered in the polarimeter, i.e., the instrumental errors of the PSG, the PSA, and the camera. Mathematically, although the instrumental errors of the PSG and the PSA are reflected in the errors of $\mathbf{M}_{\text{PSG}}$ and $\mathbf{M}_{\text{PSA}}$, respectively, both of them, together with the inherent error of the camera, will be embodied in the error of the detected $\mathbf{I}$ simultaneously. As a result, such three kinds of errors will influence the measured MM of the sample.

It should be emphasized that all three kinds of errors will be converted to intensity errors finally. In this case, the solution of Equation (9), i.e., $\mathbf{S}$, also has errors, which thus leads to the errors in $\mathbf{M}$ based on Equation (8). In this paper, the matrix error is defined by using the 2-norm or the Frobenius norm (F-norm). As for $\mathbf{M}$, the error of which can be defined as
error = \| \mathbf{M} - \hat{\mathbf{M}} \|,
\quad (11)

where $\hat{\mathbf{M}}$ and $\mathbf{M}$ are the measurement value and the ground truth of MM of the sample; operator $\| \cdot \|$ stands for the 2-norm or the F-norm of the matrix. The reason for using the 2-norm or F-norm to define the matrix error is that the norms have some good properties for analysis and are compatible with the 2-norm of the vector, which is usually used to define the difference between two vectors (the difference is called the Euclid distance). The other reason for adopting these two norms is its invariance of orthogonal transformations: for any square matrix $\mathbf{X}$ and orthogonal matrix $\mathbf{Q}$ with the same size, the following equation always holds:

$$
\| \mathbf{X} \| = \| \mathbf{Q} \mathbf{X} \mathbf{Q}^\top \|.
\quad (12)
$$

That means the error defined by the 2-norm or F-norm is invariant when the reference coordinate system is rotated. This is because the rotation transformations are kinds of orthogonal transformations.

It should be noted that, in many applications [12,13,20], the PSG and PSA were treated as ideal linear polarizers, meaning that we are assuming $D_{\text{PSG}} = D_{\text{PSA}} = 1$. However, there is a work [21] suggesting a calibration method for MM polarimeters based on precise diattenuation of polarizers, in which polarizers are treated as general nonideal linear polarizers. Without loss of generality, the PSG and PSA will be treated as nonideal polarizers in the following analyses, in which the ideal polarizers are treated as a special case.

3.2. Error Analysis of the System Composed of the PSA and the Camera

According to Section 2.2, the measurement of $3 \times 3$ MM can be divided into two parts. The first part aims to measure the three-dimensional Stokes vectors of the transmitted lights of the sample, i.e., the $\mathbf{S}$, under three given states of PSG. Mathematically, this part is expressed by Equation (9), which, in this case, is only related to the system composed of the PSA and the camera. In fact, this system can be regarded as an independent linear Stokes polarimeter. This subsection analyzes the errors of the system composed of the PSA and the camera by analyzing the stability of Equation (9), where the impacts of the PSG are excluded.

If there are instrumental errors in the PSA, according to the discussion in Section 3.1, the intensity matrix $\mathbf{I}$ will involve some errors that will lead to errors in the solution of Equation (9). In this case, obviously, the error of the solution is relevant to the value of the coefficient matrix of Equation (9), i.e., $\mathbf{A}_{\text{PSA}}$. That means there may exist an optimal orientation angle configuration of the PSA to minimize the potential error of the solution, which can be obtained from the perturbation analysis for Equation (9). The main principle of perturbation analysis is introducing perturbations into equations and analyzing the corresponding error of the solution. In the case of Equation (9), only the perturbation of constant matrix $\mathbf{I}$ needs to be introduced if the calculation error can be ignored. Assuming the perturbative matrix of $\mathbf{I}$ is $\Delta \mathbf{I}$, and the corresponding error of the solution $\mathbf{S}$ is $\Delta \mathbf{S}$, then the equation containing errors can be expressed as

$$
\mathbf{A}_{\text{PSA}} (\mathbf{S} + \Delta \mathbf{S}) = (\mathbf{I} + \Delta \mathbf{I}),
\quad (13)
$$

where $\mathbf{S}$ is the precise solution of Equation (9) and satisfies $\mathbf{A}_{\text{PSA}} \mathbf{S} = \mathbf{I}$. As we mentioned before, $\mathbf{S}$ is a matrix composed of three Stokes vectors of the transmitted lights of the sample under different states of PSG. The perturbation of $\mathbf{I}$, i.e., $\Delta \mathbf{I}$, can be regarded as the intensity error caused by errors of the PSA and the camera. The perturbation of $\mathbf{S}$, i.e., $\Delta \mathbf{S}$, is the error of $\mathbf{S}$ caused by $\Delta \mathbf{I}$.

In fact, we can assume that $\mathbf{I} + \Delta \mathbf{I}$ is the measured value of $\mathbf{I}$, which satisfies
\[ (I + \Delta I) = (A_{PSA} + \Delta A_{PSA}) S + \Delta I_C, \]  

(14)

in which \( A_{PSA} + \Delta A_{PSA} \) is the ground truth of \( A_{PSA} \); \( \Delta I_C \) is the intensity perturbation caused by the instrumental error of the camera. In fact, \( \Delta A_{PSA} \) can be regarded as a perturbative matrix caused by the instrumental error of the PSA. In order to avoid ambiguity, it should be emphasized that there is not any perturbation added to the coefficient matrix of Equation (9), which can be seen in Equation (13). The \( \Delta A_{PSA} \) is the difference between \( A_{PSA} \) and its ground truth \( A_{PSA} + \Delta A_{PSA} \); the latter is correlated to the realistic physical state of PSA. The difference \( \Delta A_{PSA} \) is one of the reasons causing intensity error \( \Delta I \). The perturbation of the coefficient matrix in Equation (9) should be introduced only when the coefficient matrix used to solve the equations in practice may be different from its assumption (the difference is caused by calculating error in this case, but the calculating error has been ignored in this paper).

By analyzing Equations (13) and (14), an inequality can be deduced, which can be expressed as

\[ \frac{\| \Delta S \|}{\| S \|} \leq \text{cond.}(A_{PSA}) \left( \frac{\| \Delta A_{PSA} \|}{\| A_{PSA} \|} + \frac{\| \Delta I_C \|}{\| I \|} \right) \]

(15)

where \( \text{cond.}(A_{PSA}) \) is the condition number of \( A_{PSA} \), which is defined based on the 2-norm or the F-norm and satisfies

\[ \text{cond.}(A_{PSA}) = \| A_{PSA} \| \| A_{PSA}^{-1} \|. \]

(16)

In Equation (15), the ratio \( \frac{\| \Delta A_{PSA} \|}{\| A_{PSA} \|} \) can be treated as the relative error of \( A_{PSA} \). The ratios \( \frac{\| \Delta I_C \|}{\| I \|} \) and \( \frac{\| \Delta S \|}{\| S \|} \) can be interpreted in the same way. Equation (15) gives the upper boundary of the relative error of the solution of Equation (9) under a given amount of the relative instrumental errors of the PSA and the camera. It is obvious that if the relative instrumental errors of the PSA and the camera were given, the upper boundary of the relative error of the solution \( S \) is only relative to \( A_{PSA} \).

At the end of this subsection, it should be noted that values of \( \Delta A_{PSA} \) and \( \Delta I_C \) are arbitrary in Equation (15), meaning that our analysis is applicable for any error sources of the PSA and the camera. For example, for PSA, both the mechanical rotation error and the depolarization error are applicable despite the fact that the latter will lead to the MM of the PSA being depolarized.

3.3. Error Analysis of the PSG

The MM of the sample can be solved from Equation (8) if the constant matrix \( S \) in Equation (8) had been solved from Equation (9). Assuming \( S \) is a known parameter, then the solution of Equation (8) is only dependent on \( A_{PSG} \). Similar to the method that the error analysis of the system composed of the PSA and the camera is based on the perturbation analysis of Equation (9), the error analysis of the PSG can be fulfilled based on the perturbation analysis of Equation (8).

As we mentioned before, if there exists an error in the solution of Equation (9), i.e., \( S + \Delta S \), the MM of the sample solved from Equation (8) is also in error. Assuming that the error of the MM of the sample is \( \Delta M \), then Equation (8) in error can be expressed as

\[ (M + \Delta M) A_{PSG} = (S + \Delta S). \]

(17)

In this case, \( \Delta S \) can be divided into two parts: the first part is the solution error of Equation (9), and the second part is the error caused by the instrumental error of the PSG. Therefore, \( S + \Delta S \) can be expressed as
\[(S + \Delta S) = M (A_{PSG} + \Delta A_{PSG}) + \Delta S_c, \quad (18)\]

in which \(A_{PSG} + \Delta A_{PSG}\) is the ground truth of \(A_{PSG}\); \(\Delta S_c\) is the solution error of Equation (9).

Similar to the analysis of the system composed of the PSA and the camera, an inequality can be deduced from Equations (17) and (18):

\[
\frac{\|\Delta M\|}{\|M\|} \leq \text{cond}_{2}(A_{PSG}) \left( \frac{\|\Delta A_{PSG}\|}{\|A_{PSG}\|} + \frac{\|\Delta S_c\|}{\|S\|} \right). \quad (19)
\]

Equation (19) gives the upper boundary of the relative error of the solution of Equation (8) under a given amount of the relative instrumental errors of PSG and the relative solution error of Equation (9).

### 3.4. Error Estimation of the MM Polarimeter

Having known that both the error of the PSG and the error of the system composed of PSA and camera will affect the final solution of \(M\), i.e., the MM of the sample, it is necessary to discuss how these errors propagate to the MM of the sample.

By analyzing Equation (10), one can obtain the following inequality:

\[
\frac{\|\Delta M\|}{\|M\|} \leq \text{cond}_{2}(A_{PSA}) \frac{\|\Delta A_{PSA}\|}{\|A_{PSA}\|} + \text{cond}_{2}(A_{PSG}) \frac{\|\Delta A_{PSG}\|}{\|A_{PSG}\|} + \text{cond}(A_{PSA}) \text{cond}(A_{PSG}) \left( \frac{\|\Delta A_{PSA}\|}{\|A_{PSA}\|} + \frac{\|\Delta A_{PSG}\|}{\|A_{PSG}\|} + \frac{\|\Delta I_{C}\|}{\|I\|} \right). \quad (20)
\]

A proof of Equation (20) is given in the Appendix A. Similar to Equations (15) and (19), Equation (20) gives an upper limitation of the relative MM error (RMME, i.e., \(\|\Delta M\|/\|M\|\)) under given error matrices, i.e., \(\Delta A_{PSA}, \Delta A_{PSG}\), and \(\Delta I_{C}\). The three error matrices can be regarded as the mathematical expressions of the three error resources introduced in Section 3.1. Therefore, Equation (20) gives an error estimation to evaluate the range of RMME of the 3 × 3 MM polarimeter.

### 4. Optimal Design of Polarizers in MM Polarimeter

As mentioned above, it is easy to see from Equation (20) that one possible way to reduce the RMME of the 3 × 3 MM polarimeter is to minimize \(\text{cond}_{2}(A_{PSA})\) and \(\text{cond}_{2}(A_{PSG})\) as much as possible. Therefore, Equation (20), as the error estimation of the MM polarimeter, actually gives us a way to make an optimal design of it. A similar method is also applied in a relevant work [17] that discussed the calibration of the 3 × 3 MM polarimeter. In [17], the minimization of the condition number is used to find an optimal combination of calibration samples, where the combination of calibration samples with a minimal condition number corresponds to the minimal calibration errors. Therefore, the method has been proven effective.

Let us consider the minimization of the condition number of \(A_{PSA}\). As mentioned in Section 3.1, we will only consider matrix errors defined on the 2-norm and the F-norm. The condition number of \(A_{PSA}\) defined on the 2-norm, denoted by \(\text{cond}_{2}(A_{PSA})\), can be calculated from the following equation:

\[
\text{cond}_{2}(A_{PSA}) = \frac{\sigma_{\text{max}}}{\sigma_{\text{min}}}, \quad (21)
\]
where σ_{\text{max}} and σ_{\text{min}} are the maximal and the minimal singular values of A_{\text{PSA}}, respectively. Singular values of A_{\text{PSA}} are square roots of eigenvalues of a positive-definite matrix that is A_{\text{PSA}}A_{\text{PSA}}^T, which satisfies

\[
A_{\text{PSA}}A_{\text{PSA}}^T = m_{00}^2 \begin{bmatrix}
D_{\text{PSA}}^2 + 1 & D_{\text{PSA}}^2 \cos(2x) + 1 & D_{\text{PSA}}^2 \cos(2y) + 1 \\
D_{\text{PSA}}^2 \cos(2x) + 1 & D_{\text{PSA}}^2 + 1 & D_{\text{PSA}}^2 \cos\left[2(y - x)\right] + 1 \\
D_{\text{PSA}}^2 \cos(2y) + 1 & D_{\text{PSA}}^2 \cos\left[2(y - x)\right] + 1 & D_{\text{PSA}}^2 + 1
\end{bmatrix}
\] (22)

where x = \theta_{(2)}^{(1)} - \theta_{(1)}^{(1)} and y = \theta_{(3)}^{(1)} - \theta_{(1)}^{(1)}. Equation (22) shows that \text{cond}2(A_{\text{PSA}}) is relevant to x and y, i.e., the angle differences of \theta_{(2)}^{(1)} and \theta_{(3)}^{(1)} with respect to \theta_{(1)}^{(1)}, instead of the three angles themselves. In fact, this property is caused by the invariance of orthogonal transformations of the 2-norm, which has been discussed in Section 3.1 as shown in Equation (12). Moreover, a well-known fact is that the condition number of a matrix is invariant while the matrix times a scalar coefficient. Therefore, \text{cond}2 cannot affect \text{cond}2(A_{\text{PSA}}), meaning that \text{cond}2(A_{\text{PSA}}) is only relevant to x, y, and D_{\text{PSA}}.

Figure 2a shows a mesh surface plot of a two-dimensional distribution of \text{cond}2(A_{\text{PSA}}) under D_{\text{PSA}} = 1 and different values of x and y in degree. Values of \text{cond}2(A_{\text{PSA}}) that correspond to x = 0, y = 0, and x = y were not calculated and shown since A_{\text{PSA}} is singular in these cases. Moreover, Figure 2b shows one-dimensional distributions of \text{cond}2(A_{\text{PSA}}) from x = 58° to x = 62° under D_{\text{PSA}} = 1 and different y from 100° to 140°. According to the results shown in Figure 2, \text{cond}2(A_{\text{PSA}}) is minimized at (x, y) = (60°, 120°) under D_{\text{PSA}} = 1. It is worth noting that values of \text{cond}2(A_{\text{PSA}}) at (x, y) are always equal to values at (y, x); therefore, \text{cond}2(A_{\text{PSA}}) is also minimized at (x, y) = (120°, 60°) under D_{\text{PSA}} = 1. In fact, (x, y) and (y, x) are physically equivalent, which corresponds to the same orientation angles configuration of the PSA.

Figure 2. (a) Mesh surface plot of the two-dimensional distribution of \text{cond}2(A_{\text{PSA}}) under D_{\text{PSA}} = 1 and different (x, y); (b) one-dimensional distributions of \text{cond}2(A_{\text{PSA}}) from x = 58° to x = 62° under D_{\text{PSA}} = 1 and different y.

The results shown in Figure 2 are achieved by numerical calculations, which only give distributions of \text{cond}2(A_{\text{PSA}}) under D_{\text{PSA}} = 1. For further exploring the effects of D_{\text{PSA}} on \text{cond}2(A_{\text{PSA}}), we used a symbolic computation system to solve the minimum point of \text{cond}2(A_{\text{PSA}}). The result of the symbolic computation shows two important properties of \text{cond}2(A_{\text{PSA}}):

1. Partial derivatives of \text{cond}2(A_{\text{PSA}}) with respect to x and y at (x, y) = (60°, 120°) are always zeros for arbitrary D_{\text{PSA}} ≠ 0.

2. Values of \text{cond}2(A_{\text{PSA}}) at (x, y) = (60°, 120°) equal to \frac{\sqrt{2}}{|D_{\text{PSA}}|}.
The two properties imply that, under a physical constrain of \(0 \leq D_{PSA} \leq 1\), a minimum point of \(\text{cond}_2(A_{PSA})\) occurs at \((x, y, D_{PSA}) = (60^\circ, 120^\circ, 1)\), which corresponds to a minimum value of \(2^{1/2}\).

On the other hand, the condition number of \(A_{PSA}\) defined on the F-norm, denoted by \(\text{cond}_F(A_{PSA})\), satisfies

\[
\text{cond}_F(A_{PSA}) = \frac{\sqrt{6\left[(D_{PSA}^4 + D_{PSA}^2)(3 - P_{x,y})^2 + 4(D_{PSA}^2 + 1)(3 - P_{x,y})\right]}}{2\left|D_{PSA}Q_{x,y}\right|},
\]

where

\[
P_{x,y} = \cos(2x) + \cos(2y) + \cos[2(y - x)],
\]
\[
Q_{x,y} = \sin(2y) - \sin(2x) - \sin[2(y - x)].
\]

Similar to \(\text{cond}_2(A_{PSA})\), \(\text{cond}_F(A_{PSA})\) is also a function of \(x = \theta_{PSA}^{(2)} - \theta_{PSA}^{(1)}, y = \theta_{PSA}^{(3)} - \theta_{PSA}^{(1)}, D_{PSA}\), owing to the fact that the F-norm also has the invariance of orthogonal transformations. According to Equation (23), for arbitrary \(D_{PSA} \neq 0\), minimum points of \(\text{cond}_F(A_{PSA})\) with respect to \((x, y)\) are always at \((x, y) = (60^\circ, 120^\circ)\). The corresponding minimum values of \(\text{cond}_F(A_{PSA})\), denoted by \(\min[\text{cond}_F(A_{PSA}), D_{PSA}]\), can be expressed as

\[
\min[\text{cond}_F(A_{PSA}), D_{PSA}] = \frac{\sqrt{(D_{PSA}^2 + 1)(D_{PSA}^2 + 4)}}{D_{PSA}}.
\]

According to Equation (24), \(\text{cond}_F(A_{PSA})\) at \((x, y) = (60^\circ, 120^\circ)\) is monotonically decreasing with the increase of \(D_{PSA}\) from 0 to 1, meaning that a minimum point of \(\text{cond}_F(A_{PSA})\) under the physical limitation of \(0 \leq D_{PSA} \leq 1\) occurs at \((x, y, D_{PSA}) = (60^\circ, 120^\circ, 1)\), which corresponds to a minimum \(\text{cond}_F(A_{PSA})\) of \(10^{1/2}\). This result about \(\text{cond}_F(A_{PSA})\) is similar to the case of \(\text{cond}_2(A_{PSA})\).

The above discussion is about \(\text{cond}_*(A_{PSA})\). On the other hand, the analysis of \(\text{cond}_*(A_{PSG})\) is similar to that of \(\text{cond}_*(A_{PSA})\). Therefore, we directly give conclusions here: \(\text{cond}_*(A_{PSG})\) is a function of \(x = \theta_{PSG}^{(2)} - \theta_{PSG}^{(1)}, y = \theta_{PSG}^{(3)} - \theta_{PSG}^{(1)}, D_{PSG}\); the minimum points of \(\text{cond}_2(A_{PSG})\) and \(\text{cond}_F(A_{PSG})\) are both occurring at \((x, y, D_{PSG}) = (60^\circ, 120^\circ, 1)\), and the corresponding minimum values are \(2^{1/2}\) and \(10^{1/2}\), respectively.

According to the discussions above, an optimal criterion for the designation of polarizers in the \(3 \times 3\) MM polarimeter is that

1. The orientation angle configurations of the PSA and the PSG, i.e., \((\theta_{PSA}^{(1)}, \theta_{PSA}^{(2)}, \theta_{PSA}^{(3)})\) and \((\theta_{PSG}^{(1)}, \theta_{PSG}^{(2)}, \theta_{PSG}^{(3)})\), need to both satisfy the form of \((c^\circ, c^\circ + 45^\circ, c^\circ + 90^\circ)\), where \(c\) is an arbitrary real constant.
2. Extinction ratios of the PSA and the PSG need to be as large as possible.

The criterion can minimize \(\text{cond}_*(A_{PSA})\) and \(\text{cond}_*(A_{PSG})\) at the same time, meaning that the MM polarimeter will have an optimal error resistance when such a criterion is fulfilled.

At the end of this section, we calculate \(\text{cond}_*(A_{PSA})\) and \(\text{cond}_*(A_{PSG})\) under some other orientation angle configurations and extinction ratios for reference, which are shown in Table 1, where \(x = \theta_{PSA}^{(2)} - \theta_{PSA}^{(1)} = \theta_{PSG}^{(2)} - \theta_{PSG}^{(1)}, y = \theta_{PSA}^{(3)} - \theta_{PSA}^{(1)} = \theta_{PSG}^{(3)} - \theta_{PSG}^{(1)}\). In Table 1, the cases of \((x, y) = (45^\circ, 90^\circ)\) are corresponding to orientation angle configurations of \((c^\circ, c^\circ + 45^\circ, c^\circ + 90^\circ)\), which are frequently used configurations [12,20].
Table 1. Values of $\text{cond}^2(\text{APSA})$ and $\text{cond}^2(\text{APSG})$ under different orientation angle configurations and extinction ratios of the PSA and the PSG.

<table>
<thead>
<tr>
<th>(x, y)</th>
<th>Extinction Ratio</th>
<th>$\text{cond}^2(\text{APSA})$ and $\text{cond}^2(\text{APSG})$</th>
<th>$\text{cond}^2(\text{APSA})$ and $\text{cond}^2(\text{APSG})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(60°, 120°)</td>
<td>$\infty$</td>
<td>1.4142</td>
<td>3.1623</td>
</tr>
<tr>
<td>(60°, 120°)</td>
<td>100:1</td>
<td>1.4428</td>
<td>3.1818</td>
</tr>
<tr>
<td>(60°, 120°)</td>
<td>10:1</td>
<td>1.7285</td>
<td>3.4124</td>
</tr>
<tr>
<td>(45°, 90°)</td>
<td>$\infty$</td>
<td>2.4142</td>
<td>3.8730</td>
</tr>
<tr>
<td>(45°, 90°)</td>
<td>100:1</td>
<td>2.4489</td>
<td>3.8969</td>
</tr>
<tr>
<td>(45°, 90°)</td>
<td>10:1</td>
<td>2.8162</td>
<td>4.1794</td>
</tr>
</tbody>
</table>

5. Discussion
5.1. Effectiveness of the Error Analysis: A Simulation

The optimal criterion of the designation of polarizers in the 3 × 3 MM polarimeter is based on the error estimation given by Equation (20) in Section 3.4. To validate the optimality of this criterion, one needs to validate the effectiveness of the error estimation.

Equation (20) gives an upper boundary of the RMME under given amounts of the relative errors of PSG, PSA, and camera. It should be emphasized that the higher upper boundary of the RMME deduced from the perturbation analysis does not necessarily correspond to the higher RMME. Therefore, to validate the effectiveness of the error estimation, we need to confirm whether the lower upper boundary of RMME is really corresponding to the lower RMME statistically.

An intuitive method to verify the effectiveness of the error estimation is to calculate the precise RMME and its upper boundary under given relative errors of PSG, PSA, and the camera. By comparing the precise RMME to its upper boundary predicted via Equation (20), the effectiveness of the perturbation analysis can be verified: if the precise error is positively correlated with its upper limitation, then the error estimation will be effective. Considering the errors of PSG, PSA, and camera cannot be controlled precisely in practice, a numerical simulation will be conducted as an alternative.

In the simulation, we calculated the RMMEs and their upper boundaries under given samples, instrumental errors, and different orientation angle configurations to verify the effectiveness of the error estimation. The experimental samples are air and a quarter-wave plate (QWP) with an orientation angle of 45°, which are frequently used calibration samples in the calibration of the MM polarimeter [17, 22]. The orientation angle configurations of the PSA and the PSG, i.e., $\theta_{\text{PSA}}$, $\theta_{\text{PSG}}$, $\theta_{\text{PSA}}$, and $\theta_{\text{PSG}}$, are set to $\theta_1$, $\theta_2$, $\theta_3$, where $\theta_1 = 0°$ and the range of $\theta_2$ and $\theta_3$ are $[20°, 160°]$. The cases of $|\theta_2 - \theta_3| \leq 3°$ are not calculated since, in these cases, $A_{\text{PSA}}$ and $A_{\text{PSG}}$ are singular or nearly singular, meaning that the solutions of equations are unstable. The polarizers in the MM polarimeter are regarded as ideal polarizers, meaning that both $D_{\text{PSA}}$ and $D_{\text{PSG}}$ are equal to 1. For simplicity, the intensity of the light source (i.e., $I_0$) and the average intensity coefficient of polarizers (i.e., $m_{00}$) are set to 1 and 1/2, respectively. Instrumental errors are added into the orientation angles (i.e., $\theta_{\text{PSA}}$, $\theta_{\text{PSG}}$, $\theta_{\text{PSA}}$, and $\theta_{\text{PSG}}$), the diattenuations (i.e., $D_{\text{PSA}}$ and $D_{\text{PSG}}$) and the intensities measured by the camera (i.e., $I_{\text{out}}(\theta_{\text{PSG}}, \theta_{\text{PSA}})$), respectively. In this case, an RMME is a relative matrix error between the MM solved from equations with errors and the MM matrix of the sample. The orientation angle errors are independent and satisfy a uniform distribution on $[-3°, 3°]$. The diattenuation errors are independent and satisfy a uniform distribution on $[-2/101, 0]$ (corresponding to the extinction ratio range from 100:1 to $\infty$). The intensity errors are also independent and satisfy Gaussian distributions with a mean of 0 and a variance of 0.002. In the simulation, the RMMEs and their upper boundaries are calculated 5000 times under different random additional instrumental errors and averaged for each orientation angle configuration and each sample. It should be noted that the statistical distributions of the errors have been considered when choosing the above parameters to simulate the practical condition. In fact, for a high-precision MM...
imaging system, the angle error of a polarizer driven by an electrical motor is no more than 1°; the extinction ratio of a polarizer is usually larger than 100:1; the Gaussian distribution of the intensity error is chosen to simulate a Gaussian noise, and a variance of 0.002 can balance the impact of the above three errors.

The two-dimensional distributions of average RMMEs and corresponding average upper boundaries with respect to different orientation angle configurations and different samples are shown in Figure 3. The distributions of average RMMEs and corresponding average upper boundaries based on the 2-norm for air and the QWP are shown in Figures 3a and 3c, respectively. The distributions of RMMEs and corresponding average upper boundaries based on the F-norm for air and the QWP are shown in Figures 3b and 3d, respectively. It can be seen in Figure 3 that the average RMME distribution is similar to the corresponding average upper boundary distributions, meaning that the higher upper boundary of RMME predicted by Equation (20) does correspond to the larger RMME statistically. That means the error estimation is effective. In addition, the minimal average RMME is located near \((\theta_1, \theta_2, \theta_3) = (0°, 60°, 120°)\) and \((\theta_1, \theta_2, \theta_3) = (0°, 120°, 60°)\) for both norms and samples, where the two configurations satisfy the optimal criterion. Therefore, based on the simulation result, a conclusion can be made that Equation (20) can be used as an effective estimation of the RMME of the MM polarimeter, and the optimal criterion derived from the estimation is effective, which inherently ensures a minimal RMME. Moreover, it can be seen in Figure 3 that the RMME distributions for the MMs of air and QWP are similar, meaning that the conclusion has a certain universality.

In detail, for comparison, Figure 4 shows the histograms of 2-norm-defined and the F-norm-defined average RMMEs for the MM of air and the corresponding average upper boundaries of four orientation angle configurations, i.e., the optimal one \((0°, 60°, 120°)\), frequently used ones which include \((0°, 45°, 90°)\) and \((0°, 45°, 135°)\), and \((0°, 30°, 90°)\). Figure 4a shows the histogram with error bars of RMMEs. Figure 4b shows the histogram of the upper boundaries of the RMMEs. It can be seen in Figure 4 that the higher relative error emerged with a higher upper boundary deduced from Equation (20), which again implies that the error estimation is effective. According to error bars shown in Figure 4a, the degree of dispersion of RMME is large when its average upper boundary is large, meaning that the MM polarimeter satisfies the optimal criterion has optimal stability. Moreover, according to Figure 4, the degree of dispersion of relative error is large when the average upper boundary of relative error is large, meaning that the measurement result under the optimal configuration, which corresponds to the minimal average upper boundary of relative error, has excellent numerical stability. On the other hand, the 2-norm-defined average RMMEs on \((0°, 60°, 120°)\) is about 0.2613, and the cases on \((0°, 45°, 90°)\) and \((0°, 45°, 135°)\) are about 0.3821 and 0.3818, respectively. Therefore, under the optimal configuration, the average RMME was reduced by 32% compared to the frequently used configuration. Moreover, for F-norm, the F-norm-defined average RMMEs on \((0°, 60°, 120°)\) is about 0.1729, and the cases on \((0°, 45°, 90°)\) and \((0°, 45°, 135°)\) are about 0.2392 and 0.2393, respectively, meaning that there exists a reduction of 28%.

In the end, it also can be seen in Figure 4 that the 2-norm-defined average RMMEs are higher than the F-norm-defined ones, but the 2-norm-defined upper boundaries are lower than the F-norm-defined ones. That means the error estimation based on the 2-norm is more accurate than the F-norm-based one.
Figure 3. Two-dimensional distributions of average RMMEs and corresponding average upper boundaries with respect to different orientation angle configurations of the MM polarimeter: (a) 2-norm-defined distributions for air; (b) F-norm-defined distributions for air; (c) 2-norm-defined distributions for QWP; (d) F-norm-defined distributions for QWP.
5.2. Effectiveness of the Optimal Criterion: A Practical Experiment

The optimal criterion proposed in Section 4 points out that the optimal orientation angle configuration of polarizers is \((c^o, c^o + 60^o, c^o + 120^o)\). As a supplement to the simulation experiment, we conducted a simple experiment to verify the effectiveness of the optimal criterion.

In the experiment, images of MMs of air are measured by our 3 × 3 MM imaging system under different orientation angle configurations. We can verify the effectiveness of the optimal criterion via a direct comparison of RMMEs under different configurations. The experiment setup is shown in Figure 5a. The basic architecture of the 3 × 3 MM imaging system is the same as the MM polarimeter shown in Figure 1, where the polarizers are driven by electrical motors, and the light source is an integrating sphere. The system is set to be uncalibrated, of which polarizers in the system are regarded as ideal polarizers in calculations. The orientation angle configurations are set to the optimal one \((0^o, 60^o, 120^o)\), the frequently used ones including \((0^o, 45^o, 90^o)\) and \((0^o, 45^o, 135^o)\), and the random one \((0^o, 30^o, 90^o)\). The resolution of the camera is 2048 × 3072, and the size of the region of interest is 501 × 501, so there are 251,001 MMs in one MM image for comparison. For each orientation angle configuration, we capture one MM image and calculate the RMME for each MM of the image.

Figure 5b shows a boxplot of 2-norm-defined and the F-norm-defined RMMEs for the MMs of air under the four orientation angle configurations, where Conf. 1, Conf. 2, Conf. 3 and Conf. 4 are \((0^o, 60^o, 120^o)\), \((0^o, 45^o, 90^o)\), \((0^o, 45^o, 135^o)\) and \((0^o, 30^o, 90^o)\), respectively. It can be seen in Figure 5b that the RMMEs of the optimal configuration are smaller than the others in terms of the minimum value, the 25th percentile, the 50th percentile, the 75th percentile, and the maximum value. Again, the experimental result is in accord with the theoretical analysis and simulation.
5.3. Three States vs. Four States

As we mentioned before, there are 3 × 3 MM polarimeters that need to measure the intensity of the transmitted light of the sample sixteen times (different combinations of four states of the PSA and the PSG, respectively) to determine the MM of the sample. This subsection will discuss the optimal orientation angle configuration of the MM polarimeter with PSA and PSG with four states (i.e., the four-states-polarimeter, FSP) briefly and compare the performance of the FSP to the polarimeter with PSA and PSG with three states (i.e., the three-states-polarimeter, TSP).

The mathematical model of the FSP can be expressed as

\[
\mathbf{A}^{(4 \times 3)}_{PSA} \mathbf{M}^{(3 \times 4)}_{PSG} = \mathbf{I}^{(4 \times 4)},
\]

where \( \mathbf{M} \) is the 3 × 3 MM of the sample; \( \mathbf{A}^{(4 \times 3)}_{PSA}, \mathbf{A}^{(4 \times 4)}_{PSG}, \mathbf{I}^{(4 \times 4)} \) are matrices with size of 4 × 3, 4 × 4, respectively. The matrices of \( \mathbf{A}^{(4 \times 3)}_{PSA} \) and \( \mathbf{A}^{(4 \times 4)}_{PSG} \) satisfy
\[
\mathbf{A}_{\text{PSA}}^{(4 \times 3)} = m_{\text{in}} \begin{bmatrix}
1 & \mathbf{D}_{\text{PSA}}^T \left( \theta_{\text{PSA}}^{(1)} \right) & \\
1 & \mathbf{D}_{\text{PSA}}^T \left( \theta_{\text{PSA}}^{(2)} \right) & \\
1 & \mathbf{D}_{\text{PSA}}^T \left( \theta_{\text{PSA}}^{(3)} \right) & \\
1 & \mathbf{D}_{\text{PSA}}^T \left( \theta_{\text{PSA}}^{(4)} \right)
\end{bmatrix},
\]
\[
\mathbf{A}_{\text{PSG}}^{(3 \times 4)} = \mathbf{I}_{\text{in}} m_{\text{in}} \begin{bmatrix}
1 & \mathbf{D}_{\text{PSG}}^T \left( \theta_{\text{PSG}}^{(1)} \right) & \\
1 & \mathbf{D}_{\text{PSG}}^T \left( \theta_{\text{PSG}}^{(2)} \right) & \\
1 & \mathbf{D}_{\text{PSG}}^T \left( \theta_{\text{PSG}}^{(3)} \right) & \\
1 & \mathbf{D}_{\text{PSG}}^T \left( \theta_{\text{PSG}}^{(4)} \right)
\end{bmatrix}^T,
\]  

and \( \mathbf{I}^{(4 \times 4)} \) satisfies
\[
\mathbf{I}^{(4 \times 4)} = \begin{bmatrix}
\mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(1)} , \theta_{\text{PSG}}^{(1)} \right) & \mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(1)} , \theta_{\text{PSG}}^{(2)} \right) & \mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(1)} , \theta_{\text{PSG}}^{(3)} \right) & \mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(1)} , \theta_{\text{PSG}}^{(4)} \right) \\
\mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(2)} , \theta_{\text{PSG}}^{(1)} \right) & \mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(2)} , \theta_{\text{PSG}}^{(2)} \right) & \mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(2)} , \theta_{\text{PSG}}^{(3)} \right) & \mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(2)} , \theta_{\text{PSG}}^{(4)} \right) \\
\mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(3)} , \theta_{\text{PSG}}^{(1)} \right) & \mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(3)} , \theta_{\text{PSG}}^{(2)} \right) & \mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(3)} , \theta_{\text{PSG}}^{(3)} \right) & \mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(3)} , \theta_{\text{PSG}}^{(4)} \right) \\
\mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(4)} , \theta_{\text{PSG}}^{(1)} \right) & \mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(4)} , \theta_{\text{PSG}}^{(2)} \right) & \mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(4)} , \theta_{\text{PSG}}^{(3)} \right) & \mathbf{I}_{\text{out}} \left( \theta_{\text{PSA}}^{(4)} , \theta_{\text{PSG}}^{(4)} \right)
\end{bmatrix},
\]  

where \( m_{\text{in}} \) is the average intensity coefficient of the PSA and the PSG; \( \mathbf{I}_{\text{in}} \) is the intensity of the light source; \( \theta_{\text{PSA}}^{(i)}, \theta_{\text{PSG}}^{(i)} \) are four orientation angles of the PSA and the PSG, respectively; \( \mathbf{D}_{\text{PSA}}^{(\theta_{\text{PSA}})} \) and \( \mathbf{D}_{\text{PSG}}^{(\theta_{\text{PSG}})} \) are the light source of the sample under the PSA angle of \( \theta_{\text{PSA}} \) and the PSG angle of \( \theta_{\text{PSG}} \).

According to Equation (25), the measured MM of the sample, denoted by \( \hat{\mathbf{M}} \), can be expressed as
\[
\hat{\mathbf{M}} = \left[ \mathbf{A}_{\text{PSA}}^{(4 \times 3)} \right]^T \mathbf{I}^{(4 \times 4)} \left[ \mathbf{A}_{\text{PSG}}^{(3 \times 4)} \right]^T,
\]  

where \( \mathbf{A}_{\text{PSA}}^{(4 \times 3)} \) and \( \mathbf{A}_{\text{PSG}}^{(3 \times 4)} \) are Moore–Penrose generalized inverse matrices of \( \mathbf{A}_{\text{PSA}}^{(4 \times 3)} \) and \( \mathbf{A}_{\text{PSG}}^{(3 \times 4)} \), respectively. If there is not any instrumental error, and \( \mathbf{A}_{\text{PSA}}^{(4 \times 3)} \) and \( \mathbf{A}_{\text{PSG}}^{(3 \times 4)} \) are full column rank and full row rank, respectively, then \( \hat{\mathbf{M}} \) will equal to \( \mathbf{M} \). Conversely, if there are instrumental errors so that Equation (25) is over-determined, \( \hat{\mathbf{M}} \) is a least squares solution with minimum norm.

There is a work [23] points out that 2-norm-defined condition numbers of \( \mathbf{A}_{\text{PSA}}^{(4 \times 3)} \) and \( \mathbf{A}_{\text{PSG}}^{(3 \times 4)} \) are minimized by an orientation angle configuration that is \( (0^\circ, 45^\circ, 90^\circ, 135^\circ) \). Moreover, according to our analysis, the F-norm-defined condition numbers of \( \mathbf{A}_{\text{PSA}}^{(4 \times 3)} \) and \( \mathbf{A}_{\text{PSG}}^{(3 \times 4)} \) are also minimized by such an orientation angle configuration. A coincidence is that the minimized condition numbers of \( \mathbf{A}_{\text{PSA}}^{(4 \times 3)} \) and \( \mathbf{A}_{\text{PSG}}^{(3 \times 4)} \) are fully the same as the minimized condition numbers of \( \mathbf{A}_{\text{PSA}}^{(4 \times 3)} \) and \( \mathbf{A}_{\text{PSG}}^{(3 \times 4)} \). That implies error resistances of the FSP and the TSP are the same under the frame of the perturbation analysis [23].

To compare the performances of the TSP and the FSP, we conducted a supplemental simulation. The simulation is similar to one introduced in Section 5.1, where instrumental errors were added into the orientation angle, the diattenuations and the intensities measured by the camera for both the schemes, and RMMEs defined by the 2-norm and the F-norm are calculated. Angle errors and intensity errors were added to the four angles and the sixteen intensities for the FSP. The statistical distributions of random errors in the FSP and the TSP are the same. In the simulation, the FSP and the TSP are optimized, where the diattenuation of the PSAs and the PSGs equal one, and orientation angle configurations are optimal.
A boxplot of RMMEs of 5000 times of simulation is shown in Figure 6, where the maximum values of RMMEs of different polarimeters and different norms are outliers, which are marked by “+”. It can be seen from Figure 6 that, for both the 2-norm and the F-norm, the RMME of FSP is always lower than that of TSP in terms of the minimum value, the 25th percentile, the 50th percentile, the 75th percentile, the upper adjacent value, and even the abnormal maximum value. Therefore, the error resistance of FSP is better than that of TSP significantly. This conclusion is reasonable since the TSP needs more time for measurement. In fact, in 4 × 4 MM polarimetry, an effective way to improve the accuracy of measurement is to increase the number of measurements, such as dual-rotating retarder MM polarimeter [16,19]. However, the cost of time and the accuracy need to be balanced. Therefore, how to reach a high accuracy as much as possible at the minimum times of measurements is still a valuable problem.

Figure 6. Boxplot of 2-norm-defined and F-norm-defined RMMEs of TSP and FSP. The maximum values of RMMEs are marked by “+”.

6. Conclusions

In this paper, based on the rigorous mathematical analysis, an optimal criterion of the designation of polarizers in a canonical 3 × 3 Mueller matrix polarimeter is given. The mathematical analysis reveals the fact that the polarimeter with the optimal orientation angle configurations (0°, 60°, and 120°) of polarizers has a robust optimal error resistance, which inherently ensures a minimal relative measurement error statistically. Moreover, to ensure the effectiveness of the mathematical analysis, a simulation is conducted. The simulation experiment shows that the mathematical analysis is effective, and the average relative error reaches a minimal value under the optimal orientation angle configuration. In contrast to the frequently used orientation angle configuration, the RMME reduced by up to 32% under the condition of our simulation experiment when the optimal design criterion was adopted.

Author Contributions: Conceptualization, methodology, H.W., Y.Z. and L.R.; investigation, writing—original draft preparation; writing—review and editing; H.W. and Y.Z.; supervision, L.R. and F.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by National Natural Science Foundation of China (No. 61535015); Natural Science Foundation of Shaanxi Province (2021JM-204); Xi’an Scientific and Technological Projects (2020KJRC0013).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.
Appendix A

To prove Equation (20), we need to introduce two fundamental inequalities about the 2-norm and the F-norm. For any square matrix $A$ and $B$ with same size, the following inequalities always hold:

$$ \|A + B\| \leq \|A\| + \|B\|, \quad (A1) $$

and

$$ \|AB\| \leq \|A\| \cdot \|B\|. \quad (A2) $$

The first inequality is known as the Triangle Inequality, the second is the Cauchy–Schwarz Inequality.

The intensity matrix measured with all the three kinds of instrumental errors, denoted by $I + \Delta I$, can be expressed as

$$ I + \Delta I = (A_{PSA} + \Delta A_{PSG})M(A_{PSG} + \Delta A_{PSG}) + \Delta I_C. \quad (A3) $$

The solved MM of the sample, denoted by $M + \Delta M$, can be expressed as

$$ M + \Delta M = A_{PSA}^{-1} (I + \Delta I) A_{PSG}. \quad (A4) $$

Note that the precise MM of the sample, i.e., $M$, satisfies Equation (10). Therefore, according to Equations (10), (A3) and (A4), we can obtain

$$ \Delta M = A_{PSA}^{-1} \Delta A_{PSA} M + M \Delta A_{PSG} A_{PSG}^{-1} + A_{PSA}^{-1} \Delta A_{PSA} M \Delta A_{PSG} A_{PSG}^{-1} + A_{PSA}^{-1} \Delta I_C A_{PSG}^{-1}. \quad (A5) $$

Now, according to the Triangle Inequality and the Cauchy–Schwarz Inequality, it can be deduced that

$$ \|\Delta M\| \leq \|A_{PSA}^{-1}\| \cdot \|\Delta A_{PSA}\| \cdot \|A_{PSG}\|, \quad (A6) $$

$$ + \|A_{PSA}^{-1}\| \cdot \|A_{PSG}^{-1}\| \cdot \|\Delta A_{PSA}\| \cdot \|\Delta A_{PSG}\| \cdot \|A_{PSG}\| \cdot \|A_{PSG}\| \cdot \|\Delta I_C\| \cdot \|M\|. $$

Equation (34) can be rewritten as

$$ \|\Delta M\| \leq \text{cond.} (A_{PSA}) \cdot \|\Delta A_{PSA}\| \cdot \|A_{PSG}\| + \text{cond.} (A_{PSG}) \cdot \|\Delta A_{PSG}\| \cdot \|A_{PSG}\| $$

$$ + \text{cond.} (A_{PSA}) \cdot \text{cond.} (A_{PSG}) \left( \|A_{PSA}\| \cdot \|\Delta A_{PSG}\| \cdot \|A_{PSG}\| + \|\Delta I_C\| \cdot \|M\| \right). \quad (A7) $$

Note that according to Equation (10) and the two fundamental inequalities, we can obtain

$$ \frac{1}{\|M\|} \leq \frac{\|A_{PSA}\| \cdot \|A_{PSG}\|}{\|I\|}. \quad (A8) $$

Substituting Equation (A8) into Equation (A7), the Equation (20) can be deduced.
References


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.