



HAL
open science

Optimal ellipsometric parameter measurement strategies based on four intensity measurements in presence of additive Gaussian and Poisson noise

Xiaobo Li, François Goudail, Haofeng Hu, Qun Han, Zhenzhou Cheng, Tiegeng Liu

► **To cite this version:**

Xiaobo Li, François Goudail, Haofeng Hu, Qun Han, Zhenzhou Cheng, et al.. Optimal ellipsometric parameter measurement strategies based on four intensity measurements in presence of additive Gaussian and Poisson noise. *Optics Express*, 2018, 26 (26), pp.34529. 10.1364/OE.26.034529 . hal-04491913

HAL Id: hal-04491913

<https://hal.science/hal-04491913v1>

Submitted on 6 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimal ellipsometric parameter measurement strategies based on four intensity measurements in the presence of additive Gaussian and Poisson noise

XIAOBO LI,^{1,2,3} FRANÇOIS GOUDAIL,^{4,5} HAOFENG HU,^{1,2,3*} QUN HAN,^{1,2}
ZHENZHOU CHENG,^{1,2} AND TIEGEN LIU^{1,2}

¹ School of Precision Instrument and Opto-electronics Engineering, Tianjin University, Tianjin 300072, China

² Key Laboratory of Opto-electronics Information Technology, Ministry of Education, Tianjin 300072, China

³ Joint Laboratory for Ocean Observation and Detection, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266237, China

⁴ Laboratoire Charles Fabry, Institut d'Optique Graduate School, CNRS, Université Paris-Saclay, 91127 Palaiseau Cedex, France

⁵ francois.goudail@institutoptique.fr

* haofeng_hu@tju.edu.cn

Abstract: The two ellipsometric parameters of an isotropic sample can be measured with simplified polarimetry setups that acquire a minimum number of four intensity measurements. However, these measurements are perturbed by noise, and the measurement strategy has to be optimized in order to limit noise propagation. We determine two different measurement strategies that are optimal in the presence of both white Gaussian additive noise and Poisson shot noise. The first one involves a polarization state generator (PSG) with a single state of polarization and a polarization state analyzer (PSA) with four states, and in the second one, both PSG and PSA have two states. The total estimation variances obtained with both strategies are demonstrated to be minimal, of equal values and independent of the ellipsometric parameters to be measured. They are based on simple optical elements and have the potential to simplify and accelerate ellipsometric measurements.

© 2018 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Ellipsometry/polarimetry is a powerful tool for accurately determining the optical properties and geometric characteristics of materials in bulk or thin film form [1-3]. When observing smooth and isotropic samples, the ellipsometry consists of measuring two ellipsometric parameters [4-7]. Ellipsometric techniques can be broadly divided into two categories [5]. The first one consists in periodically varying the azimuth angle of a component of the polarimeter with time. The detected signal is then Fourier-analyzed in order to determine the ellipsometric parameters. Most commercial ellipsometers fall into this category [6,7]. The second consists of measuring the light intensity at predetermined azimuthal positions. This type of methods is called static ellipsometry [1,2,5].

In both categories of instruments, intensity measurements are performed, and then inverted by linear algorithms (Fourier transform or pseudo-inverse matrix) to yield the ellipsometric parameters [5-7]. However, the measurements are perturbed by noise, and this noise propagates to the estimates of ellipsometric parameters through the inversion process. It is of prime importance to limit this noise propagation, that is, to minimize the variance of the parameter estimates. The purpose of the present paper is to determine the static ellipsometer architectures that optimize estimation precision in the presence of the two main sources of noise corrupting the measurements: white Gaussian additive noise and Poisson shot noise. Since these noise sources are inevitable, the obtained results will constitute the fundamental

limits of measurement precision of ellipsometric parameters with any measurement strategy having the same integration time.

We will use the Stokes-Mueller formalism, which has proven its efficiency for finding optimal polarimeter configurations [2,8-11]. The Mueller matrix of isotropic materials is block-diagonal and only contains eight non-zero elements related to the ellipsometric parameters [5,6]. Moreover, these eight elements present some symmetry, so that they finally involve only four different unknown parameters. It should thus be possible to reduce the number of intensity measurements to four. We will describe two different ellipsometric measurement strategies that involve four measurements and that are optimal in the presence of both white Gaussian additive noise and Poisson shot noise. Both strategies minimize the total estimation variance while making it independent of the Mueller matrix elements, and thus of the ellipsometric parameters to be measured.

The paper is organized as follows. We define the model of the estimation problem we address in Section 2. In Section 3, we consider a first type of measurement strategy where the polarization state generator (PSG) has a single state of polarization and the polarization state analyzer (PSA) with four states. The PSG and PSA are optimized to yield minimal estimation error in the presence of additive Gaussian and Poisson shot noise. In Section 4, another type of setup where both the PSG and the PSA have two states is optimized in the same way. The obtained results are summarized in Section 5, and we draw conclusions and perspectives of this work in Section 6.

2. Polarimetry for measuring ellipsometric parameters

The ellipsometric parameters (ψ, Δ) are defined from the ratio of the amplitude reflection coefficients for p- and s-polarizations [5,6]:

$$\rho \equiv \tan \psi \exp(i\Delta) \equiv \frac{r_p}{r_s}. \quad (1)$$

where $\tan \psi = |r_p/r_s|$ refers to the amplitude ratio, and $\Delta = \delta_p - \delta_s$ refers to the difference in phase shift [5]. The two ellipsometric parameters are also related to the Mueller matrix of the sample. In particular, in the case of isotropic samples, only the upper left and lower right 2×2 sub-matrices do not vanish, and the Mueller matrix of the sample can thus be written in terms of the ellipsometric parameters (ψ, Δ) [5,6] as follows:

$$M = r \begin{bmatrix} 1 & a & 0 & 0 \\ a & 1 & 0 & 0 \\ 0 & 0 & b & c \\ 0 & 0 & -c & b \end{bmatrix}, \quad (2)$$

where $a = -\cos 2\psi$, $b = \sin 2\psi \cos \Delta$, $c = \sin 2\psi \sin \Delta$, and r is the surface power reflectance.

One can thus obtain the ellipsometric parameters (ψ, Δ) by measuring the Mueller matrix.

The typical system configuration of complete Mueller matrix polarimetry consists of a light source, a polarization state generator (PSG), a polarization state analyzer (PSA), and an intensity detector. In practice, there are various types of configurations for PSG and PSA [10-16], and the common system configuration is composed of a linear polarizer and a retarder in both the PSG and the PSA [1-5,10], as shown in Fig. 1. Complete Mueller matrix polarimetry requires 16 intensity measurements to estimate all the 16 elements of the Mueller matrix M , and thus both the PSG and PSA must be able to generate four different polarization states [1, 2].

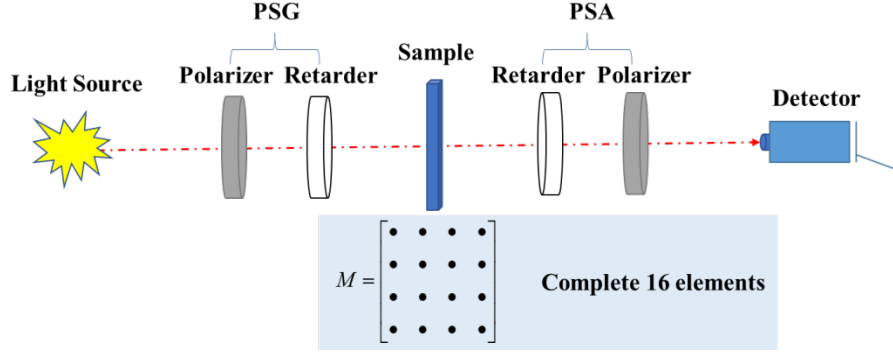


Fig. 1. Typical system configuration for complete Mueller matrix polarimetry.

However, it can be seen from Eq. (2) that the measurement of ellipsometric parameters depend on only four different parameters (r , a , b , c). Therefore, theoretically speaking, four intensity measurements should be sufficient to measure these four parameters. It is thus interesting to develop simplified strategies to measure these four parameters by taking only four intensity measurements. We will determine in the following two strategies based on four measurements that minimize the total estimation variance of Mueller parameters. They will differ by the number of polarization states generated by the PSG and the PSA.

3. Optimal strategy with one state of PSG and four states of PSA

In this section, we propose our first measurement strategy in which the PSG generates a single, fixed input polarization state, while the PSA analyzes with four different polarization states. We determine the optimal polarization states of the PSG and the PSA such that the total estimation variance of the Mueller matrix elements is minimal in the presence of white Gaussian additive noise and Poisson shot noise.

3.1 Modeling of the measurement strategy

In order to achieve four intensity measurements, a feasible strategy is that the PSG contains one polarization state characterized by eigenstate vector \mathbf{S} , while the PSA contains four different polarization states characterized by eigenstate vectors \mathbf{T}_i , $i \in \{1, 2, 3, 4\}$. We will call this strategy “1×4 type” in the following.

The eigenstate vectors \mathbf{S} and \mathbf{T}_i depend on the azimuths and the ellipticities (α^S, ε^S) and ($\alpha_i^T, \varepsilon_i^T$) of the polarization states used in PSG and PSA, respectively [2]:

$$\begin{aligned}
 \mathbf{S}(\alpha^S, \varepsilon^S) &= [1, s_1, s_2, s_3]^T \\
 &= [1, \cos(2\alpha^S)\cos(2\varepsilon^S), \sin(2\alpha^S)\cos(2\varepsilon^S), \sin(2\varepsilon^S)]^T, \\
 \mathbf{T}_i(\alpha_i^T, \varepsilon_i^T) &= [1, t_{i1}, t_{i2}, t_{i3}]^T \\
 &= [1, \cos(2\alpha_i^T)\cos(2\varepsilon_i^T), \sin(2\alpha_i^T)\cos(2\varepsilon_i^T), \sin(2\varepsilon_i^T)]^T,
 \end{aligned} \tag{3}$$

with $\alpha \in [-\pi/2, \pi/2]$, $\varepsilon \in [-\pi/4, \pi/4]$, $i \in \{1, 2, 3, 4\}$.

According to Eq. (3), the four intensities measured by the detector are thus equal to:

$$I_i = \frac{1}{2} \mathbf{T}_i^T \mathbf{M} \mathbf{S}, \quad i = \{1, 2, 3, 4\}, \tag{4}$$

where T denotes matrix transpose and M the product of the sample Mueller matrix by the source intensity. In order to facilitate the discussion, these four measured intensities are rewritten as a vector-matrix product:

$$I_i = \frac{1}{2} A_i^T \mathbf{V}_M \Leftrightarrow \mathbf{I} = \frac{1}{2} \mathbf{Q} \mathbf{V}_M, \quad \forall i \in \{1, 2, 3, 4\}, \quad (5)$$

where \mathbf{I} is the intensity vector of the four measured intensities I_i , \mathbf{V}_M is the 16-dimensional vector obtained by reading the measured Mueller matrix M of the sample in the lexicographic order, and $A_i = \mathbf{T}_i \otimes \mathbf{S}$ is the 16-dimensional vector computed from the Kronecker product \otimes between the eigenstate vector of the PSG and the i^{th} eigenstate vector of the PSA. Therefore, \mathbf{Q} is the 4×16 matrix obtained by stacking the vectors A_i^T row by row.

Taking into account the fact that only eight non-zero elements in Mueller matrix (vector) \mathbf{V}_M are to be measured, we pick out these eight elements in \mathbf{V}_M , and obtain an eight-dimensional vector \mathbf{V}_M^Ω containing all the non-zero elements of the Mueller matrix as:

$$\mathbf{V}_M^\Omega = [r, \quad ra, \quad ra, \quad r, \quad rb, \quad -rc, \quad rc, \quad rd]^T. \quad (6)$$

Then, Eq. (5) can be rewritten as:

$$\mathbf{I} = \frac{1}{2} \mathbf{Q}^\Omega \mathbf{V}_M^\Omega, \quad (7)$$

where \mathbf{Q}^Ω is a 4×8 matrix whose rows are the vectors A_i from which only the eight coefficients indexed by \mathbf{V}_M^Ω are kept. The intensity measured by the detector in Eq. (7) is thus given by:

$$\mathbf{I} = \frac{1}{2} \begin{bmatrix} 1 & t_{11} & s_1 & s_1 t_{11} & s_2 t_{12} & s_2 t_{13} & s_3 t_{12} & s_3 t_{13} \\ 1 & t_{21} & s_1 & s_1 t_{21} & s_2 t_{22} & s_2 t_{23} & s_3 t_{22} & s_3 t_{23} \\ 1 & t_{31} & s_1 & s_1 t_{31} & s_2 t_{32} & s_2 t_{33} & s_3 t_{32} & s_3 t_{33} \\ 1 & t_{41} & s_1 & s_1 t_{41} & s_2 t_{42} & s_2 t_{43} & s_3 t_{42} & s_3 t_{43} \end{bmatrix} \begin{bmatrix} r \\ ra \\ ra \\ r \\ rb \\ -rc \\ rc \\ rb \end{bmatrix}, \quad (8)$$

and by merging the terms that involve the same Mueller matrix parameters, Eq. (8) can be rewritten as:

$$\mathbf{I} = \frac{1}{2} \begin{bmatrix} 1 + s_1 t_{11} & t_{11} + s_1 & s_2 t_{12} + s_3 t_{13} & s_3 t_{12} - s_2 t_{13} \\ 1 + s_1 t_{21} & t_{21} + s_1 & s_2 t_{22} + s_3 t_{23} & s_3 t_{22} - s_2 t_{23} \\ 1 + s_1 t_{31} & t_{31} + s_1 & s_2 t_{32} + s_3 t_{33} & s_3 t_{32} - s_2 t_{33} \\ 1 + s_1 t_{41} & t_{41} + s_1 & s_2 t_{42} + s_3 t_{43} & s_3 t_{42} - s_2 t_{43} \end{bmatrix} \begin{bmatrix} r \\ ra \\ rb \\ rc \end{bmatrix}. \quad (9)$$

Let $\mathbf{V}_M^4 = [r, \quad ra, \quad rb, \quad rc]^T$ denote the four-dimensional vector of the Mueller matrix to be measured, which is called ‘‘Mueller vector’’ in the following, and let us define the matrix:

$$W = \begin{bmatrix} W_{11} & W_{12} & W_{13} & W_{14} \\ W_{21} & W_{22} & W_{23} & W_{24} \\ W_{31} & W_{32} & W_{33} & W_{34} \\ W_{41} & W_{42} & W_{43} & W_{44} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 + s_1 t_{11} & t_{11} + s_1 & s_2 t_{12} + s_3 t_{13} & s_3 t_{12} - s_2 t_{13} \\ 1 + s_1 t_{21} & t_{21} + s_1 & s_2 t_{22} + s_3 t_{23} & s_3 t_{22} - s_2 t_{23} \\ 1 + s_1 t_{31} & t_{31} + s_1 & s_2 t_{32} + s_3 t_{33} & s_3 t_{32} - s_2 t_{33} \\ 1 + s_1 t_{41} & t_{41} + s_1 & s_2 t_{42} + s_3 t_{43} & s_3 t_{42} - s_2 t_{43} \end{bmatrix}. \quad (10)$$

This matrix, which involves the polarization states of PSG and PSA, will be referred to as the “instrument matrix” in the following. Therefore, the vector of measured intensities $\mathbf{I} = [I_1, I_2, I_3, I_4]^T$ in Eq. (9) can be written as:

$$\mathbf{I} = W\mathbf{V}_M^4. \quad (11)$$

It has to be noted that $a^2 + b^2 + c^2 = 1$ [see in Eq. (2)], hence the Mueller vector \mathbf{V}_M^4 to be measured is structurally similar to a Stokes vector. Moreover, according to Eq. (10), the four elements of each row in W satisfy [the detailed demonstration is shown in Appendix I]:

$$W_{i1}^2 = W_{i2}^2 + W_{i3}^2 + W_{i4}^2, \quad \forall i \in \{1, 2, 3, 4\}. \quad (12)$$

It is thus noticed that the form in Eq. (11) is similar to that of Stokes vector polarimetry [16,17]. The only difference is that the elements W_{i1} of the first column of the instrument matrix W are not necessarily identical, while for Stokes vector polarimetry, they are all equal to one.

In practice, Eq. (11) does not hold strictly since the intensity measurements are disturbed by noise, and thus \mathbf{I} is a random vector such that each of its element I_i is a random variable [16,17]. In order to estimate the Mueller vector \mathbf{V}_M^4 from the noisy intensity measurements vector \mathbf{I} , one just has to invert the instrument matrix W in Eq. (11):

$$\hat{\mathbf{V}}_M^4 = W^{-1}\mathbf{I}. \quad (13)$$

Once we get the value of $\hat{\mathbf{V}}_M^4$, we can calculate the two ellipsometric parameters (ψ, Δ) by [11,18]:

$$\psi = \frac{1}{2} \cos^{-1}[-a] = \frac{1}{2} \cos^{-1} \left\{ \frac{-[\hat{\mathbf{V}}_M^4]_2}{[\hat{\mathbf{V}}_M^4]_1} \right\}, \quad \Delta = \tan^{-1} \left[\frac{c}{b} \right] = \tan^{-1} \left\{ \frac{[\hat{\mathbf{V}}_M^4]_4}{[\hat{\mathbf{V}}_M^4]_3} \right\}, \quad (14)$$

where $[\hat{\mathbf{V}}_M^4]_i$ refers to the i^{th} element of $\hat{\mathbf{V}}_M^4$. Therefore, optimizing the estimation of Mueller vector is beneficial for achieving higher measurement precision of the ellipsometric parameters. It is clear that $\hat{\mathbf{V}}_M^4$ is an unbiased estimator, because $\langle \hat{\mathbf{V}}_M^4 \rangle = W^{-1} \langle \mathbf{I} \rangle = \mathbf{V}_M^4$, where $\langle \cdot \rangle$ denotes ensemble averaging. For an unbiased estimator, the measurement precision is determined by its variance [19]. The variance of $\hat{\mathbf{V}}_M^4$ can be characterized by its covariance matrix $\Gamma_{\hat{\mathbf{V}}_M^4}$, which is given by:

$$\Gamma_{\hat{\mathbf{V}}_M^4} = W^{-1} \Gamma_{\mathbf{I}} (W^{-1})^T, \quad (15)$$

where $\Gamma_{\mathbf{I}}$ denotes the covariance matrix of the intensity vector \mathbf{I} . The variances γ_i of each element of $\hat{\mathbf{V}}_M^4$ equal the diagonal element $[\Gamma_{\hat{\mathbf{V}}_M^4}]_{ii}$ of the covariance matrix, and thus the

total estimation variance of all the four elements in Mueller vector equals the trace of $\Gamma_{\hat{\mathbf{V}}_M^4}$, which is expressed as $\sum_{i=1}^4 \gamma_i = \text{trace}[\Gamma_{\hat{\mathbf{V}}_M^4}]$. Our goal will be to find the best association of polarization states of PSG and PSA yielding an instrument matrix W_{opt} with a minimal total estimation variance:

$$W_{opt} = \arg \min_W \left\{ \sum_{i=1}^4 \gamma_i \right\}. \quad (16)$$

3.2 Optimization in the presence of white Gaussian additive noise

In the presence of white Gaussian additive noise, \mathbf{I} is a random vector such that each of its elements I_i , $i \in \{1, 2, 3, 4\}$ is a Gaussian random variable with mean value $\langle I_i \rangle$ and variance σ^2 [19]. Since the noise is white, each intensity measurement is statistically independent from the other [20-22]. Therefore, the covariance matrix $\Gamma_{\mathbf{I}}$ is diagonal, and all the four diagonal elements are equal to σ^2 . By substituting $\Gamma_{\mathbf{I}}$ into Eq. (15), one can get the covariance matrix $\Gamma_{\hat{\mathbf{V}}_M^4}$ as follows:

$$\Gamma_{\hat{\mathbf{V}}_M^4} = \left[W^{-1} (W^{-1})^T \right] \sigma^2, \quad (17)$$

and the variances γ_i of each element of $\hat{\mathbf{V}}_M^4$ are thus given by:

$$\gamma_i = \left[\Gamma_{\hat{\mathbf{V}}_M^4} \right]_{ii} = \left[W^{-1} (W^{-1})^T \right]_{ii} \sigma^2. \quad (18)$$

It is seen that these variances depend only on the instrument matrix W , and not on the Mueller vector \mathbf{V}_M^4 to be measured. The total estimation variance is thus given by:

$$\sum_{i=1}^4 \gamma_i = \sigma^2 \text{trace} \left[W^{-1} (W^{-1})^T \right], \quad (19)$$

which is also independent of \mathbf{V}_M^4 . According to Eqs. (16) and (19), the optimal instrument matrix W_{opt}^{Gau} in the presence of white Gaussian additive noise should satisfy:

$$W_{opt}^{Gau} = \arg \min_W \left\{ \text{trace} \left[W^{-1} (W^{-1})^T \right] \right\}. \quad (20)$$

It is possible to obtain a closed-form solution for the optimization problem in Eq. (20), and the detailed demonstration is given in Appendix II. It is shown that the optimal association of polarization states of PSG and PSA has the following properties:

1. For the eigenstate vector of PSG given by Eq. (3), the second parameter $s_1 \equiv 0$;
2. The four eigenstate vectors of PSA have a regular tetrahedron structure [17,20].

These two properties induce that the instrument matrix W_{opt}^{Gau} corresponding to the optimal polarization states of PSG and PSA also has a regular tetrahedron structure. The variances γ_i of each element of the Mueller vector are:

$$\gamma_1 = \sigma^2, \gamma_2 = 3\sigma^2, \gamma_3 = 3\sigma^2, \gamma_4 = 3\sigma^2, \quad (21)$$

and the total variance is thus equal to $10\sigma^2$. It is interesting to notice that the optimal measurement matrix found in our case is the same as that which is optimal for Stokes vector estimation [17,20]. The reason for that is the structural similarity of our ellipsometric estimation problem with the Stokes vector estimation problem, which we have pointed out in Eqs (11) and (12).

The simplest eigenstate vector with $s_1 = 0$ is $\mathbf{S} = [1, 0, 1, 0]^T$, which corresponds to a linear polarizer at 45° for PSG [2]. Therefore, the optimal system configuration for polarimetry of “1×4 type” is shown in Fig. 2. Obviously, different from the typical system configuration shown in Fig. 1, the PSG of the proposed optimal strategy does not involve a retarder but only a linear polarizer at 45° angle.

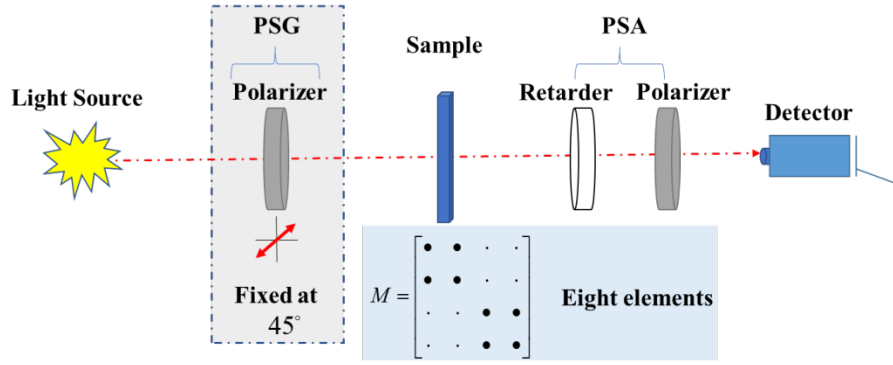


Fig. 2. Optimal system configuration for polarimetry of “1×4 type”.

In addition, for any given eigenstate vector \mathbf{S} and optimal instrument matrix W_{opt}^{Gau} with regular tetrahedron structure, the four eigenstate vectors of PSA can be calculated by Eq. (10).

3.3 Optimization in the presence of Poisson shot noise

Let us now consider that the measurements are perturbed by Poisson shot noise. In this case, each element I_i of the intensity vector \mathbf{I} is a Poisson random variable with mean and variance both equal to $\langle I_i \rangle$ [19]. Because of the statistical independence of Poisson shot noise, the covariance matrix $\Gamma_{\mathbf{I}}$ is also diagonal with elements:

$$[\Gamma_{\mathbf{I}}]_{ii} = \langle I_i \rangle = \sum_{k=1}^4 W_{ik} [\mathbf{V}_M^4]_k, \quad i \in \{1, 2, 3, 4\}, \quad (22)$$

and thus, according to Eq. (15), one can get the variances γ_i of each element $[\mathbf{V}_M^4]_i$ as:

$$\gamma_i = \sum_{j=1}^4 [\mathbf{V}_M^4]_j \sum_{k=1}^4 [W_{ik}^{-1}]^2 W_{kj}, \quad \forall i \in \{1, 2, 3, 4\}. \quad (23)$$

It is noticed that contrary to the case of white Gaussian additive noise, the criterion $\sum_{i=1}^4 \gamma_i$ does depend on the true value of the Mueller vector \mathbf{V}_M^4 to be measured. In this case, a proper strategy is to consider a minmax optimization [17], that is, minimize the following function:

$$F(W) = \max_{\mathbf{V}_M^4} \left\{ \sum_{i=1}^4 \gamma_i \right\} = \max_{\mathbf{V}_M^4} \left\{ \sum_{j=1}^4 u_j [\mathbf{V}_M^4]_j \right\}, \quad (24)$$

where $u_j = \sum_{i=1}^4 \sum_{k=1}^4 [W_{ik}^{-1}]^2 W_{kj}$. To facilitate the discussion, let us define the vectors $\mathbf{U} = [1, \mathbf{u}^T]^T$ and $\mathbf{V}_M^4 = r [1, \mathbf{v}^T]^T$, where $\mathbf{u} = [u_2, u_3, u_4]^T$ and $\mathbf{v} = [a, b, c]^T$ is the normalized Mueller vector. Therefore, the function in Eq. (24) can be rewritten as:

$$F(W) = \max_{\mathbf{u}} \{r(u_1 + \mathbf{u} \cdot \mathbf{v})\}, \quad (25)$$

where “ \cdot ” denotes the scalar product. Considering the two terms of $F(W)$ in Eq. (25), it can be seen that the first term ru_1 is independent of the normalized Mueller vector \mathbf{v} to be measured, while the second term $r\mathbf{u} \cdot \mathbf{v}$ depends on \mathbf{v} . For a given r , this second term is obviously maximized by $\mathbf{v} = \mathbf{u}/\|\mathbf{u}\|$, and thus the function to be minimized with respect to W is:

$$F(W) = r(u_1 + \|\mathbf{u}\|), \quad (26)$$

where $\|\mathbf{u}\| = \left(\sum_{i=2}^4 u_i^2\right)^{1/2}$. Thus, the instrument matrix W_{opt}^{Poi} in the presence of Poisson shot noise should minimize $F(W)$ as:

$$W_{opt}^{Poi} = \arg \min_W F(W). \quad (27)$$

It is difficult to obtain a closed-form solution of this optimization problem. However, one can apply a numerical optimization algorithm to find the solution. Indeed, according to Eq. (3), each eigenstate vector of PSG and PSA is defined by two parameters (azimuth α and ellipticity ε), and thus the numerical optimization consists in optimizing 10 parameters of the PSG (two parameters: α^S, ε^S) and the PSA (eight parameters: $\alpha_i^T, \varepsilon_i^T, i \in \{1, 2, 3, 4\}$). We apply the shuffled complex evolution (SCE) method [23], which is robust to the presence of local maxima when there are multiple parameters to be optimized. For the optimization problem in Eq. (27), we have verified that this algorithm can converge rapidly to the global minimum.

In addition, we repeat SCE method numerous times with different starting points, and all the numerical results show that:

1. The minimal total estimation variance is equal to $5r$;
2. The second parameter of the optimal eigenstate vector \mathbf{S} of PSG $s_1 \equiv 0$, and thus, from Eq. (10), $\forall k \in \{1, 2, 3, 4\} \quad W_{k1} = 1/2$.

Indeed, considering $W_{k1} = 1/2$, one has:

$$u_1 = \frac{1}{2} \sum_{i=1}^4 \sum_{k=1}^4 [W_{ik}^{-1}]^2 = \frac{1}{2} \text{trace} \left[W^{-1} (W^{-1})^T \right] = \frac{1}{2} F_{Gau}(W), \quad (28)$$

where $F_{Gau}(W)$ denotes the total estimation variance in the presence of white Gaussian additive noise, and therefore, $F(W)$ in Eq. (26) can be rewritten as:

$$F(W) = r \left(\frac{1}{2} F_{Gau}(W) + \|\mathbf{u}\| \right), \quad (29)$$

According to the optimization in the presence of white Gaussian additive noise in section 3.2, $F_{Gau}(W)$ is minimal only when the instrument matrix has a regular tetrahedron structure. Moreover, all the regular tetrahedra lead to a second term $\|\mathbf{u}\|=0$ [17,21], and thus to the minimal total estimation variance of $5r$. According to the analysis above, it is reasonable to consider a solution as optimal if it has the following properties:

1. For the eigenstate vector of PSG, the second parameter $s_1 \equiv 0$;
2. The instrument matrix W and thus the four eigenstate vectors of PSA have a regular tetrahedron structure.

It should be noted that all the instrument matrices with regular tetrahedron structure lead to minimization of total estimation variance of $5r$, and this total estimation variance does not depend on \mathbf{V}_M^4 . In the general case, however, the variances γ_i of each element of \mathbf{V}_M^4 may vary with the true value of \mathbf{V}_M^4 to be measured. An attractive property for instrument matrix would be that the variances γ_i do not depend on \mathbf{V}_M^4 . The only two “standard” regular tetrahedron matrices that have this property are [17]:

$$W_{opt}^{Poi} = \frac{1}{2} \begin{bmatrix} 1 & 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1 & -1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \\ 1 & -1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \\ 1 & 1/\sqrt{3} & -1/\sqrt{3} & -1/\sqrt{3} \end{bmatrix}, \quad (30)$$

and the matrix W_{opt-1}^{Poi} obtained from W_{opt}^{Poi} by reversing the signs of all the elements of the last three columns. By substituting this optimal matrix into Eq. (23), it is seen that the variances γ_i of each element of \mathbf{V}_M^4 have the following expressions:

$$\gamma_1 = \frac{1}{2}r, \quad \gamma_i = \frac{3}{2}r \quad \forall i \in \{2,3,4\}, \quad (31)$$

and thus the total estimation variance is equal to $5r$. It is verified that they are all independent of \mathbf{V}_M^4 . Besides, the variances of the last three elements of \mathbf{V}_M^4 are “equalized” and equal to three times the variance of the first element. However, it should be noted that, in the presence of Poisson shot noise, the “equalization” property is not valid for all instrument matrices with regular tetrahedron structure, but only for the “standard” regular tetrahedron instrument matrix W_{opt}^{Poi} and W_{opt-1}^{Poi} given by Eq. (30).

Table 1. Optimal parameters and eigenstate vectors of PSG/PSA for “1×4 type”.

Parameters/ Eigenstate vectors	PSG	PSA
Parameters (α, ε)	$(\alpha^s, \varepsilon^s) = (45^\circ, 0^\circ)$	$(\alpha_1^T, \varepsilon_1^T) = (22.50^\circ, 17.63^\circ); (\alpha_2^T, \varepsilon_2^T) = (-67.50^\circ, 17.63^\circ);$ $(\alpha_3^T, \varepsilon_3^T) = (67.50^\circ, -17.63^\circ); (\alpha_4^T, \varepsilon_4^T) = (-22.50^\circ, -17.63^\circ)$
Eigenstate vectors ($\mathbf{S} / \mathbf{T}_i$)	$\mathbf{S} = (1, 0, 1, 0)$	$\mathbf{T}_1 = [1, 1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}]; \mathbf{T}_2 = [1, -1/\sqrt{3}, -1/\sqrt{3}, 1/\sqrt{3}];$ $\mathbf{T}_3 = [1, -1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3}]; \mathbf{T}_4 = [1, 1/\sqrt{3}, -1/\sqrt{3}, -1/\sqrt{3}]$

In addition, since the second parameter of the optimal eigenstate vector \mathbf{S} of PSG $s_1=0$, the simplest PSG setting is one linear polarizer at 45° . Therefore, the optimal system configuration for the case of Poisson shot noise is the same as that for white Gaussian additive noise shown in Fig. 2. The corresponding parameters (azimuth and ellipticity) of the PSA can be calculated according to Eqs. (3), (10) and (30), and are shown in Table 1. In addition, the eigenstate vectors corresponding to PSG and PSA are also presented in Table 1.

It has to be noted that since the optimal instrument matrix for Poisson shot noise also minimizes the total estimation variance for white Gaussian additive noise, the optimal parameters and eigenstate vectors of the PSG and the PSA in Table 1 are also optimal in the presence of white Gaussian additive noise. Therefore, by performing four intensity measurements with the simplified system configuration shown in Fig. 2 and the optimal parameters shown in Table 1, we can estimate the Mueller vector \mathbf{V}_M^4 with minimal total estimation variance in the presence of both white Gaussian additive noise and Poisson shot noise.

4. Optimal strategy with two states of both PSG and PSA

In this section, we will consider another type of strategy to achieve four intensity measurements, where both PSG and PSA implement two polarization states characterized by the eigenstate vectors $\mathbf{S}_i, \mathbf{T}_i, i \in \{1, 2\}$, respectively. This strategy will be called “2×2 type” in the following.

Applying the same reasoning that led to Eq. (10) in the case of the “1×4 type” strategy, it is easily shown that the instrument matrix for the “2×2 type” strategy is:

$$W_{2 \times 2} = \frac{1}{2} \begin{bmatrix} 1 + s_{11}t_{11} & t_{11} + s_{11} & s_{12}t_{12} + s_{13}t_{13} & s_{13}t_{12} - s_{12}t_{13} \\ 1 + s_{11}t_{21} & t_{21} + s_{11} & s_{12}t_{22} + s_{13}t_{23} & s_{13}t_{22} - s_{12}t_{23} \\ 1 + s_{21}t_{11} & t_{11} + s_{21} & s_{22}t_{12} + s_{23}t_{13} & s_{23}t_{12} - s_{22}t_{13} \\ 1 + s_{21}t_{21} & t_{21} + s_{21} & s_{22}t_{22} + s_{23}t_{23} & s_{23}t_{22} - s_{22}t_{23} \end{bmatrix}, \quad (32)$$

where $\mathbf{S}_i = [1, s_{i1}, s_{i2}, s_{i3}]$, $\mathbf{T}_i = [1, t_{i1}, t_{i2}, t_{i3}]$, $i \in \{1, 2\}$ refer to the two eigenstate vectors of the polarization states of PSG and PSA, respectively. Then the four elements in \mathbf{V}_M^4 can be extracted by inverting the instrument matrix $W_{2 \times 2}$ as $\hat{\mathbf{V}}_M^4 = W_{2 \times 2}^{-1} \mathbf{I}$.

4.1 Optimization in the presence of white Gaussian additive noise

In the case of “2×2 type”, the optimal association of the polarization states of PSG and PSA that yields the optimal instrument matrix $W_{2 \times 2}^{Gau}$ in the presence of white Gaussian additive noise should minimize the total estimation variance of the four elements in \mathbf{V}_M^4 :

$$W_{2 \times 2}^{Gau} = \arg \min_{W_{2 \times 2}} \left\{ \text{trace} \left[W_{2 \times 2}^{-1} \left(W_{2 \times 2}^{-1} \right)^T \right] \right\}. \quad (33)$$

It is difficult to find a closed-form solution of this optimization problem. One can get the optimal solution numerically by applying an optimization algorithm with SCE method [23], which involves optimizing the eight parameters of PSG and PSA. The numerical optimization result shows that, the minimal total estimation variance is $10\sigma^2$, and the second parameter of the optimal eigenstate vectors of PSG fulfils $s_{i1} = 0, i \in \{1, 2\}$. In this case, $W_{i1} = 1/2 \forall i \in \{1, 2, 3, 4\}$ in Eq. (32), and thus the matrices W with regular tetrahedron structure lead to

the minimal total estimation variance of $10\sigma^2$. It is thus reasonable to consider a solution as optimal if it has the following properties:

1. For the two eigenstate vectors \mathbf{S}_i of PSG, the second parameter $s_{i1} = 0$, $i \in \{1, 2\}$;
2. The instrument matrix W has the regular tetrahedron structure.

With the optimal instrument matrix $W_{2 \times 2-opt}^{Gau}$, the variances γ_i of each element of $\hat{\mathbf{V}}_M^4$ and the total estimation variance are also calculated by $\gamma_i = \left\{ \left[W_{2 \times 2-opt}^{Gau} \right]^{-1} \left[\left[W_{2 \times 2-opt}^{Gau} \right]^{-1} \right]^T \right\}_{ii} \sigma^2$ and $\sum_{i=1}^4 \gamma_i$, respectively. One obtains the same values as for the “1×4 type” strategy:

$$\gamma_1 = \sigma^2, \gamma_2 = \gamma_3 = \gamma_4 = 3\sigma^2, \sum_{i=1}^4 \gamma_i = 10\sigma^2, \quad (34)$$

The simplest couple of eigenstate vectors of PSG with $s_{i1} = 0$, $i \in \{1, 2\}$ is $\mathbf{S}_1 = [1, 0, 1, 0]^T$ and $\mathbf{S}_2 = [1, 0, -1, 0]^T$, which corresponds to linear $\pm 45^\circ$ polarized light. Therefore, different from the “1×4 type” strategy, the linear polarizer in PSG is not just fixed at 45° but must be able to implement two different angles (45° and -45°). The system configuration of “2×2 type” is shown in Fig. 3. Comparing with the configuration shown in Fig. 1, the PSG in Fig. 3 does not involve a retarder but only a linear polarizer, which is same to the one in Fig. 2. In addition, in Fig. 2, the polarizer of the PSA is set to a single fixed angle of 45° , while that in Fig. 3 has to be set to two different angles of $\pm 45^\circ$.

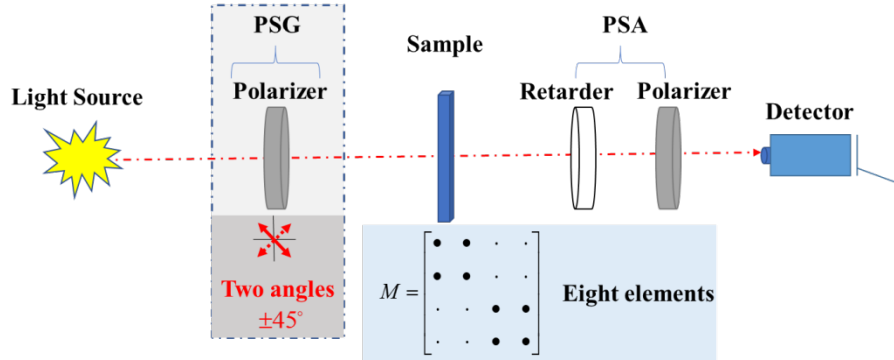


Fig. 3. Optimal system configuration for polarimetry of “2×2 type”.

In addition, as soon as the two eigenstate vectors \mathbf{S}_i and the optimal instrument matrix $W_{2 \times 2-opt}^{Gau}$ with regular tetrahedron structure are decided, the two eigenstate vectors of PSA can be calculated by Eq. (32).

4.2 Optimization in the presence of Poisson shot noise

For the case of “2×2 type” strategy in the presence of Poisson shot noise, the optimal instrument matrix $W_{2 \times 2-opt}^{Poi}$ is such that:

$$W_{2 \times 2-opt}^{Poi} = \arg \min_{W_{2 \times 2}} \left\{ r(\mathbf{u}_i + \|\mathbf{u}\|) \right\}. \quad (35)$$

where $\underline{u}_j = \sum_{i=1}^4 \sum_{k=1}^4 \left[[W_{2 \times 2}]_{ik}^{-1} \right]^2 [W_{2 \times 2}]_{kj}$. We obtain the optimal solution of the optimization problem in Eq. (35) numerically by using SCE method. The numerical optimization results show that the second parameter of the eigenstate vectors of PSG fulfills $s_{i1} = 0$, $i \in \{1, 2\}$, and the minimal total variance equals $5r$. Similar to the case of “1×4 type” strategy, the instrument matrix with regular tetrahedron structure is also optimal for “2×2 type” strategy. It indicates that the optimal instrument matrix for Poisson shot noise also minimizes the total variance for white Gaussian additive noise. However, only the “standard” matrix given by Eq. (30) minimizes and equalizes the variances of each Mueller parameters to be measured, and makes all these variances independent of the value of this Mueller vector.

Table 2. Optimal parameters and eigenstate vectors of PSG/PSA for “2×2 type”.

Parameters/ Eigenstate vectors	PSG	PSA
Parameters	$(\alpha_1^s, \varepsilon_1^s) = (45^\circ, 0^\circ)$; (α, ε)	$(\alpha_1^T, \varepsilon_1^T) = (22.50^\circ, 17.63^\circ)$; $(\alpha_2^T, \varepsilon_2^T) = (67.50^\circ, -17.63^\circ)$
Eigenstate vectors (\mathbf{S}_i / \mathbf{T}_i)	$\mathbf{S}_1 = [1, 0, 1, 0]$; $\mathbf{S}_2 = [1, 0, -1, 0]$	$\mathbf{T}_1 = [1, 1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}]$; $\mathbf{T}_2 = [1, -1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3}]$

Therefore, in this case again, the system configuration optimal for Poisson shot noise is identical to the configuration optimal for white Gaussian noise shown in Fig. 3, in which only one linear polarizer is required for the PSG. The four intensity measurements can be achieved with this polarizer at two angles ($+45^\circ$ and -45°). The corresponding optimal parameters (azimuth α and ellipticity ε) and the eigenstate vectors of PSG and PSA can be calculated according to Eqs. (3), (30) and (32), and are shown in Table 2.

By comparing this table with Table 1, it is clearly seen that the “2×2 type” strategy has the same optimal performance as “1×4 type” strategy:

1. The variances of each element of the Mueller vector $\hat{\mathbf{V}}_M^4$ are $\gamma_1 = 1/2r$, $\gamma_i = 3/2r$, $\forall i \in \{2, 3, 4\}$ and the total estimation variance is equal to $5r$;
2. The noise variances are equalized and independent of the Mueller vector, and thus of the ellipsometric parameters to be measured.

Therefore, by performing four intensity measurements with the optimal configuration shown in Fig. 3 and the optimal parameters shown in Table 2, one can estimate the Mueller vector \mathbf{V}_M^4 with minimal total estimation variance for both white Gaussian additive noise and Poisson shot noise. In addition, the total estimation variances obtained with the “2×2 type” and “1×4 type” strategies are identical.

5. Summary

In the sections above, based on the Stokes-Mueller formalism, we have proposed two types of ellipsometric parameters measurement strategies that are optimal in the presence of both white Gaussian additive noise and Poisson shot noise. In order to present the results clearly, we summarize them in Table 3.

It should be noted that in the case of white Gaussian additive noise, all the instrument matrices W with regular tetrahedron structure are optimal. Therefore, the optimal set of eigenstate vectors of PSA is not unique: any set of four measurement vectors obtained by 3D rotation of the standard regular tetrahedron matrix given by Eq. (30) is optimal [21]. For the purpose of illustration, we presented in Table 3 the optimal eigenstate vectors of PSA

corresponding to the standard regular tetrahedron instrument matrix given by Eq. (30) in the presence of white Gaussian additive noise.

In practice, the intensity detection is disturbed by both additive Gaussian noise and Poisson shot noise. The quantum efficiency (QE) η influences the relative weights of the Gaussian and Poisson noise components, since it has a quadratic influence on the variance of Gaussian noise σ_{Gau}^2 , but a linear influence on the variance of Poisson noise σ_{Poi}^2 . Therefore, the variance of the detected intensity I is [9,15,17]:

$$\text{VAR}[I] = \eta\sigma_{Poi}^2 + \eta^2\sigma_{Gau}^2, \quad (36)$$

However, as shown in Table 3, the regular tetrahedron matrix given by Eq. (30) is optimal for both Gaussian and Poisson noise. Therefore, this regular tetrahedron matrix is also optimal when Gaussian and Poisson noise are simultaneously present.

Table 3. The performance for the two types of optimal strategies in the presence of white Gaussian additive noise and Poisson shot noise.

Type	Noise	Optimal instrument matrix (W)	Optimal eigenstate vectors		Total variance
			PSG (S)	PSA (T)	
1×4 type	Gaussian	All the matrices with regular tetrahedron structure.	(1,0,1,0)	$\mathbf{T}_1 = [1, 1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}]$; $\mathbf{T}_2 = [1, -1/\sqrt{3}, -1/\sqrt{3}, 1/\sqrt{3}]$; $\mathbf{T}_3 = [1, -1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3}]$; $\mathbf{T}_4 = [1, 1/\sqrt{3}, -1/\sqrt{3}, -1/\sqrt{3}]$	$10\sigma^2$
	Poisson	$\frac{1}{2} \begin{bmatrix} 1 & 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1 & -1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \\ 1 & -1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \\ 1 & 1/\sqrt{3} & -1/\sqrt{3} & -1/\sqrt{3} \end{bmatrix}$		Not unique, and can be obtained by 3D rotation. $\mathbf{T}_1 = [1, 1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}]$; $\mathbf{T}_2 = [1, -1/\sqrt{3}, -1/\sqrt{3}, 1/\sqrt{3}]$; $\mathbf{T}_3 = [1, -1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3}]$; $\mathbf{T}_4 = [1, 1/\sqrt{3}, -1/\sqrt{3}, -1/\sqrt{3}]$	
2×2 type	Gaussian	All the matrices with regular tetrahedron structure.	(1,0,1,0)	$\mathbf{T}_1 = [1, 1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}]$; $\mathbf{T}_2 = [1, -1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3}]$	$10\sigma^2$
	Poisson	$\frac{1}{2} \begin{bmatrix} 1 & 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1 & -1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \\ 1 & -1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \\ 1 & 1/\sqrt{3} & -1/\sqrt{3} & -1/\sqrt{3} \end{bmatrix}$	(1,0,-1,0)	$\mathbf{T}_1 = [1, 1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}]$; $\mathbf{T}_2 = [1, -1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3}]$	

6. Conclusion

In this paper, we have described two types of optimal strategies for estimating the ellipsometric parameters of isotropic samples with only four intensity measurements. These two strategies minimize the total estimation variance of the Mueller parameters in the presence of both white Gaussian additive noise and Poisson shot noise. They yield the same value of the total estimation variance, and this value is independent of the ellipsometric parameters to be measured. Interestingly, in both strategies, the optimal instrument matrix is found to have a regular tetrahedron structure, which is well known to also minimize noise propagation in Stokes vector estimation problems [17,20]. This is because we have been able to express the problem of ellipsometric parameters estimation in a mathematical form that is similar – albeit not identical – to Stokes vector estimation. In addition, these optimal strategies only require a polarizer in the PSG, which simplifies the system and reduces its cost.

In short, we have proven, for the first time to our knowledge, that ellipsometric parameters estimation with minimal error propagation can be performed using very simple optical setups.

Besides, our proposed method can also be implemented on spectroscopic ellipsometers. In this case, a key issue is taking into account that the phase retardance of the compensator (retarder) depends on the wavelength [26]. Therefore, the optical implementation of the optimal measurement matrix should be adapted to the wavelength. On the other hand, we can also use achromatic compensators, which can provide almost the same phase retardance over a wide spectral range [26], to implement our method for spectroscopic ellipsometers.

Compared to the existing common commercial ellipsometers, which use continuously rotating motors and spread the total integration time to a large number of data points (for example, ~40 points [24,25] or ~80 points [26] per one optical cycle), the outstanding advantage of our method is that it only requires four optimal intensity measurements, and the whole available integration time can be devoted to four points, which further reduces the impact of noise. However, a key issue for our method is the “dead time” between each measurement, since the system must wait for the motor to move to the next location. During this time, intensity is not collected, contrary to what happens in ellipsometers with continuous data collection. A step-and-stare system might thus not be efficient from the standpoint of intensity collection. In order to reduce the “dead time” to the order of microseconds, rotating optical elements could for example be replaced by electro-optical or photoelastic modulation devices with no moving parts (such as photoelastic modulator (PEM) [27]). More globally speaking, comparison between step-and-stare and continuous ellipsometers amounts to a tradeoff between using only optimal measurement points with dead time in between, or having no dead time but a large set of measurement points. Optimization of this tradeoff in terms of estimation accuracy will depend on the optical precision and noise parameters of the two compared systems. Studying this tradeoff in concrete cases is a very interesting perspective, for which our present work on performance optimization of step-and-stare systems will constitute an important element.

Finally, we have considered in this paper only the simplest and most fundamental noise sources. Our results thus represent the fundamental upper limit on the performance that can be reached with a static ellipsometer. Of course, in practice, other sources of perturbations such as instrumental errors may be present or even dominant [28]. Implementation of the optimization approach described in this paper on more precise measurement models is indeed an interesting perspective to the present work. Nonetheless, we think that the results obtained in this paper provide a solid basis for performing these future works and interpreting their results.

Appendix I: Demonstration of $W_{i1}^2 = W_{i2}^2 + W_{i3}^2 + W_{i4}^2$, $\forall i \in \{1, 2, 3, 4\}$ in Eq. (12).

According to Eq. (10), one has:

$$\begin{aligned} W_{i2}^2 + W_{i3}^2 + W_{i4}^2 &= \frac{1}{4} \left[(t_{i1} + s_1)^2 + (s_2 t_{i2} + s_3 t_{i3})^2 + (s_3 t_{i2} - s_2 t_{i3})^2 \right] \\ &= \frac{1}{4} \left[t_{i1}^2 + s_1^2 + 2t_{i1}s_1 + s_2^2 t_{i2}^2 + s_3^2 t_{i3}^2 + s_3^2 t_{i2}^2 + s_2^2 t_{i3}^2 \right] \\ &= \frac{1}{4} \left[t_{i1}^2 + s_1^2 + 2t_{i1}s_1 + (s_2^2 + s_3^2)(t_{i2}^2 + t_{i3}^2) \right]. \end{aligned} \quad (37)$$

Since $s_1^2 + s_2^2 + s_3^2 = 1$ and $t_{i1}^2 + t_{i2}^2 + t_{i3}^2 = 1$, $\forall i \in \{1, 2, 3, 4\}$, thus

$$\begin{aligned}
W_{i2}^2 + W_{i3}^2 + W_{i4}^2 &= \frac{1}{4} \left[t_{i1}^2 + s_1^2 + 2t_{i1}s_1 + (1-s_1^2)(1-t_{i1}^2) \right] \\
&= \frac{1}{4} \left[1 + 2t_{i1}s_1 + s_1^2 t_{i1}^2 \right] \\
&= \frac{1}{4} (t_{i1} + s_1)^2 \\
&= W_{i1}^2.
\end{aligned} \tag{38}$$

Therefore, we have $W_{i1}^2 = W_{i2}^2 + W_{i3}^2 + W_{i4}^2$, $\forall i \in \{1, 2, 3, 4\}$.

Appendix II: Demonstration of the optimal solution in case of “1×4 Type”.

According to Eqs. (7), (10) and (11), the instrument matrix W can be rewritten as the product of two matrices T and Q :

$$W = TQ = \frac{1}{2} \begin{bmatrix} 1 & t_{11} & t_{12} & t_{13} \\ 1 & t_{21} & t_{22} & t_{23} \\ 1 & t_{31} & t_{32} & t_{33} \\ 1 & t_{41} & t_{42} & t_{43} \end{bmatrix} \begin{bmatrix} 1 & s_1 & & & & \\ & s_1 & 1 & & & \\ & & & s_2 & s_3 & \\ & & & s_3 & -s_2 & \end{bmatrix}. \tag{39}$$

Therefore, the total estimation variance in Eq. (19) is given by [29]:

$$\text{trace} \left[(TQ)^T (TQ) \right]^{-1} = \text{trace} \left[(T^T T)^{-1} (Q^T Q)^{-1} \right]. \tag{40}$$

Since Q and T are invertible, this criterion can be written as:

$$F(A, B) = \text{trace}(A^{-1}B^{-1}), \tag{41}$$

with $A = T^T T$ and $B = Q^T Q$ which are symmetric, positive definite matrices. Our goal will be to find the best association of A and B that minimizes the total estimation variance:

$$(A, B) = \arg \min_{A, B} \{F(A, B)\}, \tag{42}$$

with the following constraints:

$$\text{trace}(A) = 2 \quad \text{and} \quad \|\mathbf{s}\| = 1, \tag{43}$$

where $\mathbf{s} = (s_1, s_2, s_2)$ is the reduced eigenstate vector of the PSG. Since A and B are both positive definite, it can be shown that [29]:

$$\text{trace}(A^{-1}B^{-1}) \leq \text{trace}(A^{-1})\text{trace}(B^{-1}). \tag{44}$$

Let us denote $G(A, B) = \text{trace}(A^{-1})\text{trace}(B^{-1})$. Since the function $\text{trace}(\cdot)$ is decreasing and convex, and the constraints in Eq. (43) define convex sets, the functions $F(A, B)$ and $G(A, B)$ thus have a single extremum in the domain defined by the constraints, and this is thus a minimum [30].

It is clear that,

$$\begin{aligned}
(A, B) &= \arg \min_{A, B} \{G(A, B)\} \\
&= \arg \min_A \{\text{trace}(A^{-1})\} \cdot \arg \min_B \{\text{trace}(B^{-1})\}.
\end{aligned} \tag{45}$$

According to previous works [17,20], it is known that under the constraint $\text{trace}(A) = 2$, the matrix A that minimizes the value of $\text{trace}(A^{-1})$ is:

$$A_0^{-1} = \arg \min_A \{ \text{trace}(A^{-1}) \} = \begin{bmatrix} 1 & & & \\ & 3 & & \\ & & 3 & \\ & & & 3 \end{bmatrix}, \quad (46)$$

and it corresponds to the matrix T in Eq. (39) having a regular tetrahedron structure on the Poincaré sphere [20].

Besides, with the constraint $\|\mathbf{s}\| = 1$, it is easy to find that:

$$\text{trace}(B^{-1}) = \frac{4}{(1-s_1^2)^2}, \quad (47)$$

which has a minimal value whenever $s_1 = 0$. Therefore, the optimal matrix B that minimizes the value of $\text{trace}(B^{-1})$ is:

$$B_0^{-1} = \arg \min_B \{ \text{trace}(B^{-1}) \} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}. \quad (48)$$

According to Eqs. (46) and (48), it can be noticed that:

$$\text{trace}(A_0^{-1} B_0^{-1}) = \text{trace}(A_0^{-1}) \text{trace}(B_0^{-1}). \quad (49)$$

Therefore, $F(A_0, B_0) = G(A_0, B_0)$.

According to the above, we thus have:

1. $(A_0, B_0) = \arg \min_{A,B} \{ G(A, B) \}$;
2. $F(A_0, B_0) = G(A_0, B_0)$;
3. $F(A, B) \leq G(A, B) \quad \forall A, B$;
4. Since $F(A, B)$ and $G(A, B)$ are convex, they have a single minimum [30].

Let us denote $X = (A, B)$ the set of variables, and $X_0 = (A_0, B_0)$ a minimum of $G(A, B)$. The constraints in Eq. (43) are rewritten as $g(A, B) = 0$. Thus, by using the Lagrange multiplier method [31], one gets:

$$\frac{\partial G(X_0)}{\partial X} - \lambda \frac{\partial g(X_0)}{\partial A} - \mu \frac{\partial g(X_0)}{\partial B} = 0, \quad (50)$$

where λ , μ refer to the Lagrange parameters. For any prescribed small interval around the minimum $(X_0 + \delta X)$, one has:

$$\begin{aligned}
F(X_0 + \delta X) &= F(X_0) + \left(\frac{\partial F}{\partial X}\right)^T(X_0)\delta X + o_1(\delta X), \\
G(X_0 + \delta X) &= G(X_0) + \left(\frac{\partial G}{\partial X}\right)^T(X_0)\delta X + o_2(\delta X),
\end{aligned} \tag{51}$$

where $o_i(\delta X)$, $i \in \{1, 2\}$ refers to the higher order terms of Taylor series of $F(A, B)$ and $G(A, B)$, respectively. Since $F(X_0) = G(X_0)$, and let us denote:

$$\Delta(\delta X) = G(X + \delta X) - F(X + \delta X), \tag{52}$$

One can get, by substituting Eq. (51) into Eq. (52), and then using Eq. (50), that:

$$\begin{aligned}
\Delta(\delta X) &= \left(\frac{\partial G}{\partial X}\right)^T(X_0)\delta X - \left(\frac{\partial F}{\partial X}\right)^T(X_0)\delta X + o(\delta X) \\
&= \left[\lambda \frac{\partial g(X_0)}{\partial A} + \mu \frac{\partial g(X_0)}{\partial B} - \frac{\partial F}{\partial X}(X_0) \right]^T \delta X + o(\delta X),
\end{aligned} \tag{53}$$

where $o(\delta X) = o_2(\delta X) - o_1(\delta X)$ is also a higher order term, which can be neglected and considered as zero. Since $\forall A, B \ F(A, B) \leq G(A, B)$, one must have $\forall \delta X, \Delta(\delta X) \geq 0$. Therefore, the following equality must be satisfied:

$$\frac{\partial F}{\partial X}(X_0) - \lambda \frac{\partial g(X_0)}{\partial A} - \mu \frac{\partial g(X_0)}{\partial B} = 0. \tag{54}$$

According to Eq. (54), it is seen that X_0 is also an extremum of $F(A, B)$ under the constraint $g(A, B) = 0$ [30]. Since $F(A, B)$ is a convex function, and the constraint $g(A, B)$ also defines a convex set, this extremum is unique. It means that X_0 is the single minimum of $F(A, B)$ [31]. Therefore, the only minimum of $F(A, B)$ must also be (A_0, B_0) .

According to the analyses above, the optimal closed-form solution of the optimization problem in Eq. (42) and thus that in Eq. (20) should satisfy:

1. For the eigenstate vector of PSG given by Eq. (3), the second parameter $s_1 \equiv 0$;
2. The four eigenstate vectors of PSA and thus the matrix T in Eq. (39) has the regular tetrahedron structure [17,20].

Funding

National Natural Science Foundation of China (No. 61775163), National Instrumentation Program (No. 2013YQ030915), Young Elite Scientists Sponsorship Program by CAST (2017QNRC001), Director Fund of Qingdao National Laboratory for Marine Science and Technology (QNL201717), China Postdoctoral Science Foundation (No. 2016M601260).

Acknowledgments

The authors thank Enric Garcia-Caurel and Razvigor Ossikovski for fruitful discussions.

References

1. D. Goldstein, *Polarized Light* (Dekker, 2003).
2. R. M. A. Azzam, "Stokes-vector and Mueller-matrix polarimetry," *J. Opt. Soc. Am. A* **33**(7), 1396-1408 (2016).

3. X. Li, H. Hu, T. Liu, B. Huang, and Z. Song, "Optimal distribution of integration time for intensity measurements in degree of linear polarization polarimetry," *Opt. Express* **24**(7), 7191-7200 (2016).
4. W. Li, C. Zhang, H. Jiang, X. Chen, and S. Liu, "Depolarization artifacts in dual rotating-compensator Mueller matrix ellipsometry," *J. Opt.* **18**, 055701 (2016).
5. R. M. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light* (Elsevier Science Publishing, 1987).
6. E. Garcia-Caurel, R. Ossikovski, M. Foldyna, A. Pierangelo, B. Drévilion, and A. D. Martino, "Advanced Mueller ellipsometry instrumentation and data analysis," in *Ellipsometry at the Nanoscale*, M. Losurdo and K. Hingerl, ed. (Springer, 2013), Chap. 2.
7. H. Gu, X. Chen, H. Jiang, C. Zhang, W. Li, and S. Liu, "Accurate alignment of optical axes of a biplate using a spectroscopic Mueller matrix ellipsometer," *Appl. Opt.* **55**(15), 3935-3941 (2016).
8. K. Wang, J. Zhu, H. Liu, and B. Du, "Expression of the degree of polarization based on the geometrical optics pBRDF model," *J. Opt. Soc. Am. A* **34**(2), 259-263 (2017).
9. N. Hagen and Y. Otani, "Stokes polarimeter performance: general noise model and analysis," *Appl. Opt.* **57**(15), 4283-4296 (2018).
10. G. Anna, H. Sauer, F. Goudail, and D. Dolfi, "Fully tunable active polarization imager for contrast enhancement and partial polarimetry," *Appl. Opt.* **51**(21), 5302-5309 (2012).
11. X. Li, H. Hu, L. Wu, T. Liu, "Optimization of instrument matrix for Mueller matrix ellipsometry based on partial elements analysis of the Mueller matrix," *Opt. Express* **25**(16), 18872-18885 (2017).
12. J. S. Tyo and H. Wei, "Optimizing imaging polarimeters constructed with imperfect optics," *Appl. Opt.* **45**(22), 5497-5503 (2006).
13. H. Hu, R. Ossikovski, and F. Goudail, "Performance of maximum likelihood estimation of Mueller matrices taking into account physical realizability and Gaussian or Poisson noise statistics," *Opt. Express* **21**(4), 5117-5129 (2013).
14. J. S. Tyo, Z. Wang, S. J. Johnson, and B. G. Hoover, "Design and optimization of partial Mueller matrix polarimeters," *Appl. Opt.* **49**(12), 2326-2332 (2010).
15. S. Roussel, M. Boffety, and F. Goudail, "Polarimetric precision of micropolarizer grid-based camera in the presence of additive and Poisson shot noise," *Opt. Express* **26**(23), 29968-29982 (2018).
16. X. Li, T. Liu, B. Huang, Z. Song, and H. Hu, "Optimal distribution of integration time for intensity measurements in Stokes polarimetry," *Opt. Express* **23**(21), 27690-27699 (2015).
17. F. Goudail, "Noise minimization and equalization for Stokes polarimeters in the presence of signal-dependent Poisson shot noise," *Opt. Lett.* **34**(5), 647-649 (2009).
18. S. Krishnan, P. C. Nordine, "Mueller-matrix ellipsometry using the division-of-amplitude photopolarimeter: a study of depolarization effects," *Appl. Opt.* **33**(19), 4184-4192 (1994).
19. A. Papoulis, *Probability, Random Variables and Stochastic Processes* (McGraw-Hill, 1991).
20. M. R. Foreman, A. Favaro, and A. Aiello, "Optimal frames for polarization state reconstruction," *Phys. Rev. Lett.* **115**(26), 263901 (2015).
21. G. Anna, F. Goudail, "Optimal Mueller matrix estimation in the presence of Poisson shot noise," *Opt. Express* **20**(19), 21331-21340 (2012).
22. X. Li, H. Hu, H. Wang, L. Wu, and T. Liu, "Influence of noise statistics on optimizing the distribution of integration time for degree of linear polarization polarimetry," *Optical Engineering* **57**(6), 064110 (2018).
23. Q. Y. Duan, V. K. Gupta, and S. Sorooshian, "Shuffled complex evolution approach for effective and efficient global minimization," *J. Optim. Theory Appl.* **76**(3), 501-521 (1993).
24. D. E. Aspnes and A. A. Studna, "High precision scanning ellipsometer," *Appl. Opt.* **14**(1), 220-228 (1975).
25. H. Fujiwara, *Spectroscopic Ellipsometry: Principles and Applications* (Wiley, 2007).
26. H. Tompkins and E. Irene, *Handbook of Ellipsometry* (William Andrew, 2005).
27. S. Alali, A. Gribble, and I. A. Vitkin, "Rapid wide-field Mueller matrix polarimetry imaging based on four photoelastic modulators with no moving parts," *Opt. Lett.* **41**(5), 1038-1041 (2016).
28. F. Goudail and A. Bènière, "Estimation precision of the degree of linear polarization and of the angle of polarization in the presence of different sources of noise," *Appl. Opt.* **49**(4), 683-693 (2010).
29. C. D. Meyer, *Matrix Analysis and Applied Linear Algebra* (SIAM, 2000).
30. S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University, 2004).
31. D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods* (Academic, 1982).