Estimating the Polarization Degree of Polarimetric Images in Coherent Illumination Using Maximum Likelihood Methods

Florent Chatelain,[1,*] Jean-Yves Tourneret,[2] Muriel Roche,[3] and Mehdi Alouini[4,5]

[1] *GIPSA-lab, Signal and Image Dept., Grenoble Inst. of Tech., BP 46, 38402 Saint Martin d'Hères, France*

[2] *University of Toulouse, IRIT/ENSEEIHT/TéSA, 2 rue Camichel, BP 7122, 31071 Toulouse, France*

[3] *Institut Fresnel, CNRS, Aix-Marseille Université, Ecole Centrale Marseille, Campus de Saint-Jérôme, 13013 Marseille, France*

[4] *Thales Research and Technology France, RD128, 91767 Palaiseau Cedex, France*

[5] *Institut de Physique de Rennes, CNRS-Université de Rennes I, 35042 Rennes Cedex, France*

[*] *Corresponding author: florent.chatelain@gipsa-lab.inpg.fr*

This paper addresses the problem of estimating the polarization degree of polarimetric images in coherent illumination. It has been recently shown that the degree of polarization associated to polarimetric images can be estimated by the method of moments applied to two or four images assuming fully developed speckle. This paper shows that the estimation can also be conducted by using maximum likelihood methods. The maximum likelihood estimators of the polarization degree are derived from the joint distribution of the image intensities. We show that the joint distribution of polarimetric images is a multivariate gamma distribution whose marginals are univariate, bivariate or trivariate gamma distributions. This property is used to derive maximum likelihood estimators of the polarization degree using two, three or four images. The proposed estimators provide better performance than the estimators of moments. These results are illustrated by estimations conducted on synthetic and real images.

1. Introduction

Polarimetric imagery allows one to analyze the polarimetric properties of the light backscattered or transmitted by a scene. It gives complementary information to standard intensity images. The materials which compose the scene are of different nature which induces different reflection or transmission of the received light. In particular,

the modification of the state of polarization of the received light can be related to the kind of material affected by this light. Polarimetric imaging has been used successfully in many image processing applications. For instance, in medical applications, polarimetry provides information about the loss of birefringence of the collagen when the skin is damaged, allowing one to classify normal and thermally damaged skins [1]. In military applications, polarimetry can be used to detect small targets embedded in a background clutter like minefield [2, 3]. Other examples of applications where polarimetry has shown interesting properties include computer vision [4–7], remote sensing [8], astrophysics [9], separation of diffuse and specular components [10], underwater imaging [11, 12], image dehazing [13].

Polarimetric imaging systems can be divided into two categories: passive [14, 15] and active [16, 17]. When active polarimetric imagery [18] is used, the scene is often illuminated with a coherent source of light, typically a laser beam. This type of imagery is interesting as it offers possible control of the emitted light allowing to obtain images by night and to facilitate the interpretation of the polarization state of the reflected light. The use of a laser beam has supplementary advantages. Its high directivity supports uniform illumination and its wavelenght improves the resolution for a given aperture. However when coherent light is used as a source of illumination, the polarimetric images are degraded by speckle fluctuations [19]. The speckle introduces a granulous structure in the images which deteriorates the analysis of the polarimetric properties of the light from the images. This phenomenon limits the use of a laser

when no statistical consideration of the speckle is introduced [20]. We propose in this paper to take advantage of the speckle noise in order to obtain informations on the polarimetric properties of the light backscattered by the objects. Other non statistical techniques can also be used to analyze the polarimetric contents of the light coming from the object (see, for instance, [21]), we focus here on statistical ones.

In order to analyze the polarimetric properties of the light coming from the scene, the covariance matrix of the Jones vector is usually estimated [22]. In particular, the knowledge of this covariance matrix allows one to compute the degree of polarization (DoP) of the light [19]. Some interesting polarization properties of the materials in the scene can be deduced from this scalar parameter. The DOP allows one, besides to emphasize the contrast between objects having different polarimetric properties, to quantify the way the light backscattered by the object is depolarized. As a consequence estimating the DoP is important in many field such as in optical fiber transmission for reflecting the degree of waveform degradation caused by polarization mode dispersion [23] and in quantum optics for characterizing polarimetric properties in this non classical optics [24]. The DoP is also used for analyzing the beam propagation in a turbulent atmosphere [25], for determining surface orientation of objects [26], as discrimination criterion by which to reject multiply scattered photons [27], and for medical applications [28].

In standard Stokes coherent imagery [12, 29], four images are needed to access to the DoP. These systems consist of a laser which illuminates the scene of interest, a

quarter-wave plate which introduces a phase shift to the electrical field and a polarizer which transmits the light in a particular direction. Four configurations of the couple quarter-wave plate/polarizer are used to obtain the different images.

In front of the complexity and the amount of polarimetric data to analyze, it seems interesting to develop methods that reduces the measurement time, the cost of the imagery systems as well as the complexity of the systems. Some methods have been proposed to reduce the number of acquired images. For example, the Orthogonal State Contrast Image (OSCI) [30, 31] gives access to polarimetric properties of the light using only two images. However, the OSCI is only an estimate of the degree of polarization in the particular cases of pure depolarizers. Recently, we have shown that when coherent illumination is considered the DoP can be estimated with moment based estimators with two images and assuming fully developped speckle [32] and with only a single image [33, 34] for a more general model of the speckle.

We propose in this paper maximum likelihood estimators of the DoP under coherent illumination assuming fully developped speckle that offers better performances as the moment based estimators. The study is carried out by decreasing the number of avalaible images from 4 to 2 images.

The paper is organized as follows. Section 2 recalls some important results regarding the statistical properties of the Stokes vector. Section 3 and 4 address the problem of estimating the DoP of polarimetric images using maximum likelihood (ML) and moment methods. Different situations are considered depending on the number of

available images. Estimation results conducted on synthetic and real data are presented in section 5 and 6 respectively. Conclusions and perspectives are reported in Section 7.

2. Background

The light can be described by a monochromatic electrical field propagating in the $\mathbf{e}_Z$ direction in an homogeneous and isotropic medium at a point $\mathbf{r}$, at time t

$$\mathbf{E}(\mathbf{r},t) = [A_X(\mathbf{r},t)\mathbf{e}_X + A_Y(\mathbf{r},t)\mathbf{e}_Y]e^{-i2\pi\nu t}, \tag{1}$$

where ν is the vibration central frequency, $\mathbf{e}_X$ and $\mathbf{e}_Y$ are orthonormal vector in the X and Y directions respectively, and $\mathbf{A}(\mathbf{r},t) = [A_X(\mathbf{r},t), A_Y(\mathbf{r},t)]^T$ is the Jones vector whose components are complex. For simplicity reasons, the notations $\mathbf{A} = \mathbf{A}(\mathbf{r},t)$, $A_X = A_X(\mathbf{r},t)$ and $A_Y = A_Y(\mathbf{r},t)$ will be used in the rest of the paper. The state of polarization of the light can be described by the random behavior of the Jones vector whose covariance matrix, called the polarization matrix, is

$$\Gamma = \begin{pmatrix} \mathsf{E}\left[A_X A_X^*\right] & \mathsf{E}\left[A_X A_Y^*\right] \\ \mathsf{E}\left[A_Y A_X^*\right] & \mathsf{E}\left[A_Y A_Y^*\right] \end{pmatrix} \triangleq \begin{pmatrix} a_1 & a_3 + ia_4 \\ a_3 - ia_4 & a_2 \end{pmatrix}, \tag{2}$$

where $\mathsf{E}[\cdot]$ and * denote the mathematical expectation and the complex conjugate, respectively. The covariance matrix Γ is a non negative hermitian matrix whose diagonal terms are the intensity components in the X and Y directions. The cross terms of Γ represents the correlations between the Jones components. If we assume a fully developed speckle, the phase of the light reflected by the elementary diffusers (which

constitute the diffusing surface of the scene) is uniformly distributed. In this case, the Jones vector $\mathbf{A}$ is distributed according to a complex Gaussian distribution with probability density function (pdf) [19]

$$p_{\mathbf{A}}(\mathbf{A}) = \frac{1}{\pi^2 |\Gamma|}\, e^{-\mathbf{A}^\dagger \Gamma^{-1} \mathbf{A}}, \tag{3}$$

where $|\Gamma|$ is the determinant of the matrix Γ and $\dagger$ denotes the conjugate transpose operator. The different components of the covariance matrix Γ can be classically estimated by using four intensity images, the scene being illuminated with a coherent polarized linear light. The two first images I_1 and I_2 are obtained by analyzing the light backscattered by the scene in two orthogonal states of polarization. This is done by introducing a polarizer between the scene and the camera, that is parallel or orthogonal to the incident light. The third intensity I_3 is obtained by recording the light backscattered in the direction oriented to 45° from the incident light, by modifying the orientation of the polarizer. Finally the last image I_4 is obtained by adding a quarter wave plate, before the polarizer, in order to introduce a phase difference of $\lambda/4$ in the previous configuration. As a consequence, the four intensities are related to the components of the Jones vector as follows:

$$\begin{aligned} I_1 &= |A_X|^2, \qquad I_2 = |A_Y|^2, \\ I_3 &= \frac{1}{2}|A_X|^2 + \frac{1}{2}|A_Y|^2 + \mathrm{Re}\left(A_X A_Y^*\right), \\ I_4 &= \frac{1}{2}|A_X|^2 + \frac{1}{2}|A_Y|^2 + \mathrm{Im}\left(A_X A_Y^*\right). \end{aligned} \tag{4}$$

The square DoP of a given pixel of the polarimetric image is defined by [19, p. 134]

$$P^2 = 1 - 4\frac{|\Gamma|}{\left[\text{trace}\,(\Gamma)\right]^2} = 1 - \frac{4\left[a_1 a_2 - (a_3^2 + a_4^2)\right]}{(a_1 + a_2)^2}, \tag{5}$$

where $\text{trace}\,(\Gamma)$ is the trace of the matrix Γ. The DoP characterizes the state of polarization of the light: the light is totally depolarized for $P = 0$, totally polarized for $P = 1$ and partially polarized when $P \in]0, 1[$. This paper studies different DoP estimators for polarimetric images at a given pixel. Since only one realization of the random vector $\boldsymbol{I} = (I_1, I_2, I_3, I_4)^T$ is available for this pixel, the images are supposed to be locally stationary and ergodic. These assumptions allow us to build estimates using several neighbor pixels belonging to a so-called estimation window. The estimators derived in this paper assume there is no spatial correlation between the pixels of the estimation window. This is a realistic hypothesis in many configurations of optical systems. Introducing spatial correlation between the different pixels of the estimation window would be interesting. However, it would increase considerably the computational complexity of the DoP estimators proposed in this paper. Note also that it could be interesting to consider temporal averaging on the data. This will lead to a more complex joint distribution of the image intensity.

3. DoP estimation using ML methods

This section studies DoP estimators based on several vectors $\boldsymbol{I}^1, \ldots, \boldsymbol{I}^n$ associated to the n pixels of the estimation window, where $\boldsymbol{I}^j$ denotes the intensity vector associated with the jth pixel. These estimators are constructed from estimates of the covariance

matrix elements a_i, $i = 1, ..., 4$. Different estimators are studied depending on the number of available polarimetric images, i.e. 4, 3 or 2 polarimetric images and allow us to study the loss induced by the simplification of the imagery system. The first step is to derive the joint distribution of the intensity vector $\boldsymbol{I}$. This joint distribution and its margins will be used to derive the different estimators.

3.A. Joint distribution of the intensity vector

To determine the distribution of a random vector $\boldsymbol{I}$, it is very usual to determine its Laplace transform (also denoted as moment generating function) defined as

$$L_{\boldsymbol{I}}(\boldsymbol{\theta}) = \mathsf{E}\left[\exp\left(-\sum_{j=1}^{4}\theta_j I_j\right)\right], \tag{6}$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4)^T \in \mathbb{R}^4$. The intensity vector $\boldsymbol{I}$ is related to the random hermitian matrix $\boldsymbol{S} = \boldsymbol{A}\boldsymbol{A}^\dagger$ where

$$\boldsymbol{S} = \begin{pmatrix} s_1 & s_3 + is_4 \\ s_3 - is_4 & s_2 \end{pmatrix} = \begin{pmatrix} |A_X|^2 & A_X A_Y^* \\ A_Y A_X^* & |A_Y|^2 \end{pmatrix}.$$

Indeed, by using (4) and by denoting $\boldsymbol{s} = (s_1, s_2, s_3, s_4)^T$, the following relation can be obtained

$$\boldsymbol{s} = \boldsymbol{M}\boldsymbol{I} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1/2 & -1/2 & 1 & 0 \\ -1/2 & -1/2 & 0 & 1 \end{pmatrix} \boldsymbol{I}. \tag{7}$$

As a consequence, the Laplace transform of $\boldsymbol{I}$ is defined as

$$L_{\boldsymbol{I}}(\boldsymbol{\theta}) = \mathsf{E}\left[\exp\left(-\boldsymbol{\theta}^T \boldsymbol{I}\right)\right] = \mathsf{E}\left[\exp\left(-\boldsymbol{\theta}^T \boldsymbol{M}^{-1} \boldsymbol{s}\right)\right] = L_{\boldsymbol{s}}\left(\boldsymbol{M}^{-T} \boldsymbol{\theta}\right). \tag{8}$$

On the other hand, the Laplace transform of the hermitian matrix $\boldsymbol{S}$ can be defined as (see [35])

$$L_{\boldsymbol{S}}(\boldsymbol{\Theta}) = \mathsf{E}\left[\exp\left[-\operatorname{trace}\left(\boldsymbol{S}\boldsymbol{\Theta}\right)\right]\right], \tag{9}$$

where $\boldsymbol{\Theta}$ is the following hermitian matrix

$$\boldsymbol{\Theta} = \begin{pmatrix} \theta_1 & \theta_3 + i\theta_4 \\ \theta_3 - i\theta_4 & \theta_2 \end{pmatrix}.$$

Straightforward computations allow one to show that

$$\operatorname{trace}\left(\boldsymbol{S}\boldsymbol{\Theta}\right) = \theta_1 s_1 + \theta_2 s_2 + 2\theta_3 s_3 + 2\theta_4 s_4,$$

hence

$$L_{\boldsymbol{S}}(\boldsymbol{\Theta}) = L_{\boldsymbol{s}}\left(\theta_1, \theta_2, 2\theta_3, 2\theta_4\right), \tag{10}$$

which allows one to relate the two Laplace transforms $L_{\boldsymbol{S}}$ and $L_{\boldsymbol{s}}$.

Using the complex Gaussian assumption for $\boldsymbol{A}$, it is well known that the hermitian matrix $\boldsymbol{S} = \boldsymbol{A}\boldsymbol{A}^{\dagger}$ is distributed according to a complex Wishart distribution [35] whose Laplace transform is

$$L_{\boldsymbol{S}}(\boldsymbol{\Theta}) = \mathsf{E}\left\{\exp\left[-\operatorname{trace}\left(\boldsymbol{S}\boldsymbol{\Theta}\right)\right]\right\} = \left|\mathbb{I}_2 + \Gamma\boldsymbol{\Theta}\right|^{-1}, \tag{11}$$

where $\mathbb{I}_2$ is the 2×2 identity matrix and $\boldsymbol{\Theta}$ is a 2×2 hermitian matrix ensuring existence of $L_{\boldsymbol{S}}(\boldsymbol{\Theta})$. This allows one to determine $L_{\boldsymbol{s}}(\boldsymbol{\theta})$ using (10), and $L_{\boldsymbol{I}}(\boldsymbol{\theta})$ using (8). The following result is obtained

$$L_{\boldsymbol{I}}(\boldsymbol{\theta}) = \frac{1}{P(\boldsymbol{\theta})}, \tag{12}$$

where $P(\boldsymbol{\theta})$ is an affine polynomial (A polynomial $P(\mathbf{z})$ where $\mathbf{z} = (z_1, \ldots, z_d)$ is *affine* if the one variable polynomial $z_j \mapsto P(\mathbf{z})$ can be written $Az_j + B$ (for any $j = 1, \ldots, d$), where A and B are polynomials with respect to the z_i's for $i \neq j$) defined as

$$P(\boldsymbol{\theta}) = 1 + \boldsymbol{\alpha}^T \boldsymbol{\theta} + k \left[2\theta_1\theta_2 + \theta_3\theta_4 + (\theta_1 + \theta_2)(\theta_3 + \theta_4) \right],$$

with $\boldsymbol{\alpha} = [a_1, a_2, (a_1 + a_2 + 2a_3)/2, (a_1 + a_2 + 2a_4)/2]^T$ and $k = \frac{1}{2}(a_1a_2 - a_3^2 - a_4^2)$. As a consequence, the intensity vector $\boldsymbol{I} = (I_1, I_2, I_3, I_4)^T$ is distributed according to a multivariate gamma distribution (MGD) as defined in [36,37]. Moreover, according to (12), the distribution of the intensity vector $\boldsymbol{I}$ is fully characterized by the parameter vector $\boldsymbol{\alpha}$, or equivalently by $\boldsymbol{a} = (a_1, a_2, a_3, a_4)^T$.

3.B. DoP estimation using 4 images

Using the properties of MGDs, it can be shown that the maximum likelihood estimator (MLE) of $\boldsymbol{\alpha} = \mathsf{E}[\boldsymbol{I}]$ is

$$\widehat{\boldsymbol{\alpha}}_{\mathrm{ML}} = \frac{1}{n} \sum_{j=1}^{n} \boldsymbol{I}^j.$$

This result is classical and was for instance mentioned in [38, p. 221] or [32]. The mean of the intensity vector is related to the vector $\boldsymbol{a}$ as follows

$$\alpha_1 = \mathsf{E}[I_1] = a_1, \quad \alpha_3 = \mathsf{E}[I_3] = (a_1 + a_2 + 2a_3)/2,$$

$$\alpha_2 = \mathsf{E}[I_2] = a_2, \quad \alpha_4 = \mathsf{E}[I_4] = (a_1 + a_2 + 2a_4)/2,$$

i.e., $\boldsymbol{a} = \boldsymbol{M}\boldsymbol{\alpha}$. The functional invariance principle [39, p. 176] can then be used to derive the MLE of $\boldsymbol{a}$

$$\widehat{\boldsymbol{a}}_{\mathrm{ML}} = \boldsymbol{M}\widehat{\boldsymbol{\alpha}}_{\mathrm{ML}}.$$

The MLEs of the parameters a_i, $i = 1, ..., 4$, are then plugged into (5) yielding an estimate of the polarization degree P^2 based on 4 polarimetric images. This yields the classical estimator of the DoP (see for instance [29, p. 340])

$$\widehat{P}_4^2 = 1 - \frac{4\left[\widehat{a}_1\widehat{a}_2 - (\widehat{a}_3^2 + \widehat{a}_4^2)\right]}{(\widehat{a}_1 + \widehat{a}_2)^2}. \tag{13}$$

Interestingly, the asymptotic variance of the estimator $\widehat{P}_4^2$ can be determined. Using the relation $\widehat{\boldsymbol{\alpha}}_{\mathrm{ML}} = \frac{1}{n}\sum_{j=1}^{n}\boldsymbol{I}^j$, it can be proved that the MLE of $\boldsymbol{a}$ is unbiased and efficient, providing an optimal estimation of $\boldsymbol{a}$. Moreover, the covariance matrix of $\widehat{\boldsymbol{a}}_{\mathrm{ML}}$ is expressed as

$$\mathrm{cov}\left(\widehat{\boldsymbol{a}}_{\mathrm{ML}}\right) = \frac{1}{n}\boldsymbol{M}\,\mathrm{cov}\left(\boldsymbol{I}\right)\boldsymbol{M}^T,$$

where $\mathrm{cov}\left(\boldsymbol{I}\right)$ is the covariance matrix of the intensity vector $\boldsymbol{I}$. The covariance matrix $\mathrm{cov}\left(\boldsymbol{I}\right)$ can be computed by using the moments of a bivariate gamma distribution. Indeed, by setting $\theta_i = 0$, for $i \notin \{k, l\}$ in (12), the Laplace transform of the vector

(I_k, I_l) is shown to be the Laplace transform of a bivariate gamma distribution. The second-order moments of the intensity vector $\boldsymbol{I}$ can then be computed, leading to

$$\operatorname{cov}\left(\widehat{\boldsymbol{a}}_{\mathrm{ML}}\right) = \frac{1}{n}\begin{pmatrix} a_1^2 & a_3^2+a_4^2 & a_1a_3 & a_1a_4 \\ a_3^2+a_4^2 & a_2^2 & a_2a_3 & a_2a_4 \\ a_1a_3 & a_2a_3 & c_{3,3} & a_3a_4 \\ a_1a_4 & a_2a_4 & a_3a_4 & c_{4,4} \end{pmatrix}, \tag{14}$$

with $c_{3,3} = (a_1a_2 + a_3^2 - a_4^2)/2$ and $c_{4,4} = (a_1a_2 - a_3^2 + a_4^2)/2$. The asymptotic variance of the estimator (13) is its Cramer-Rao bound which expresses as

$$\operatorname{var}_{\mathrm{A}}\left(\widehat{P}_4^2\right) = G_4^T \operatorname{cov}\left(\widehat{\boldsymbol{a}}_{\mathrm{ML}}\right) G_4 = \frac{2(1-P^2)^2P^2}{n}, \tag{15}$$

where

$$G_4(\boldsymbol{a}) = \left(\frac{4(a_1a_2 - a_2^2 - 2a_3^2 - 2a_4^2)}{(a_1+a_2)^3}, \frac{4(a_1a_2 - a_1^2 - 2a_3^2 - 2a_4^2)}{(a_1+a_2)^3}, \frac{8a_3}{(a_1+a_2)^2}, \frac{8a_4}{(a_1+a_2)^2}\right)^T \tag{16}$$

is the gradient of the function $g_4 : \boldsymbol{a} \mapsto g_4(\boldsymbol{a}) = P^2$ (see [39, p. 45] for details). Note that the expression (15) of the asymptotic variance of $\widehat{P}_4^2$ has already been obtained in [32] using a Taylor expansion of P^2 around its true value. However, the proposed methodology to determine $\operatorname{var}_{\mathrm{A}}\left(\widehat{P}_4^2\right)$ is interesting since it will be generalized to other DoP estimators. Note also that the asymptotic variance of $\widehat{P}_4^2$ (i.e. the asymptotic variance of the DoP estimator based on 4 images) only depends on the parameters a_i, $i = 1, ..., 4$ through P^2 and that this function of P^2 is maximum for $P^2 = 1/3$. This property will be confirmed in our simulation results.

3.C. DoP estimation using 3 images

Straightforward computations using (4) show that the intensity vector $\boldsymbol{I}$ belongs to a cone whose equation is:

$$[I_3 - (I_1 + I_2)/2]^2 + [I_4 - (I_1 + I_2)/2]^2 = I_1 I_2.$$

Consequently, the distribution of $\boldsymbol{I}$ is singular and defined on this cone and one can think to estimate the unknown parameter vector $\boldsymbol{a}$ by using three images only. The analysis is conducted here with $\widetilde{\boldsymbol{I}} = (I_1, I_2, I_3)^T$. However, similar results could be obtained with $\widetilde{\boldsymbol{I}} = (I_1, I_2, I_4)^T$. The Laplace transform of $\widetilde{\boldsymbol{I}}$ is obtained by setting $\theta_4 = 0$ in (12)

$$L_{\widetilde{\boldsymbol{I}}}(\widetilde{\boldsymbol{\theta}}) = E\left[\exp\left(-\sum_{j=1}^{3}\theta_j I_j\right)\right] = \frac{1}{\widetilde{P}(\widetilde{\boldsymbol{\theta}})},$$

with $\widetilde{P}(\widetilde{\boldsymbol{\theta}}) = 1 + \sum_{i=1}^{3}\alpha_i\theta_i + k\left[2\theta_1\theta_2 + \theta_1\theta_3 + \theta_2\theta_3\right]$ and $\widetilde{\boldsymbol{\theta}} = (\theta_1, \theta_2, \theta_3)^T$. This expression shows that $\widetilde{\boldsymbol{I}}$ is distributed according to an MGD (since $\widetilde{P}(\widetilde{\boldsymbol{\theta}})$ is an affine polynomial). Interestingly, the density of the intensity vector $\widetilde{\boldsymbol{I}} = (I_1, I_2, I_3)^T$ can be determined. Letac and Wesolowski [40] recently derived the distributions whose Laplace transforms are

$$L_{\widetilde{\boldsymbol{I}}}(\widetilde{\boldsymbol{\theta}}) = \left[1 - 2c^T\widetilde{\boldsymbol{\theta}} + v(\widetilde{\boldsymbol{\theta}})\right]^{-p},$$

where $p > 0$, $c \in \mathbb{R}^N$ and $v(\widetilde{\boldsymbol{\theta}})$ is a quadratic form of $\widetilde{\boldsymbol{\theta}}$. By setting $N = 3$, $p = 1$, $c = \frac{1}{2}(\alpha_1, \alpha_2, \alpha_3)^T$ and $v(\widetilde{\boldsymbol{\theta}}) = k\left[2\theta_1\theta_2 + \theta_1\theta_3 + \theta_2\theta_3\right]$ in their results, the following

density for the intensity vector $\widetilde{\boldsymbol{I}}$ can be obtained:

$$p(\widetilde{\boldsymbol{I}}) = \frac{1}{k\sqrt{\pi v'(\widetilde{\boldsymbol{I}})}} \exp\left[-\frac{(a_2+a_3)I_1 + (a_1+a_3)I_2 - 2a_3I_3}{2k}\right] f_{\frac{1}{2}}\left(\frac{a_4^2 v'(\widetilde{\boldsymbol{I}})}{16k^2}\right) \mathbb{I}_\Omega(\widetilde{\boldsymbol{I}}), \tag{17}$$

where $v'(\widetilde{\boldsymbol{I}}) = (I_1+I_2)^2 - (I_1-I_2)^2 - (I_1+I_2-2I_3)^2$, $\Omega = \left\{\widetilde{\boldsymbol{I}} \in \mathbb{R}^3; v'(\widetilde{\boldsymbol{I}}) > 0\right\}$, $\mathbb{I}_\Omega(\widetilde{\boldsymbol{I}})$ is the indicator defined on Ω, and

$$f_q(z) = \sum_{m=0}^{\infty} \frac{z^m}{\Gamma(m+q)m!}, \qquad q > 0,$$

is related to the the confluent hypergeometric function [41, p. 374]. It is interesting to note here that a necessary and sufficient condition for $p(\widetilde{\boldsymbol{I}})$ to be the pdf of a probability distribution on $\mathbb{R}^3$ is $a_1a_2 > a_3^3 + a_4^2$. This condition is not restrictive since it is equivalent to assuming that the covariance matrix of the Jones vector defined in (2) is positive definite. The MLE of $\boldsymbol{a}$ based on three images can be obtained by differentiating $p(\widetilde{\boldsymbol{I}})$ with respect to each component of $\boldsymbol{a}$. By differentiating $p(\widetilde{\boldsymbol{I}})$ with respect to a_1, a_2 and a_3, the following relation can be obtained

$$\begin{pmatrix} \widehat{a}_1 \\ \widehat{a}_2 \\ \widehat{a}_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1/2 & 1 & -1/2 \end{pmatrix} \begin{pmatrix} \widehat{\alpha}_1 \\ \widehat{\alpha}_2 \\ \widehat{\alpha}_3 \end{pmatrix},$$

with $\widehat{\alpha}_l = \frac{1}{n}\sum_{j=1}^n \boldsymbol{I}_l^j$ for $l = 1, 2, 3$. After replacing $\widehat{a}_1$, $\widehat{a}_2$ and $\widehat{a}_3$ in $\frac{\partial p(\widetilde{\boldsymbol{I}})}{\partial a_4^2} = 0$, Appendix A shows that the MLE of parameter a_4^2 (denoted as $\widetilde{a_4^2}$) satisfies the following nonlinear

equation

$$\frac{1}{2n}\sum_{j=1}^{n}\sqrt{\frac{v'\left(\widetilde{\boldsymbol{I}}^{j}\right)}{\widetilde{a_4^2}}}\tanh\left(\frac{\sqrt{\widetilde{a_4^2}v'\left(\widetilde{\boldsymbol{I}}^{j}\right)}}{\widehat{d}-\widetilde{a_4^2}}\right)=1,$$

where $\widehat{d}=\widehat{a}_1\,\widehat{a}_2-\widehat{a}_3^2$. The estimates of (a_1,a_2,a_3) and a_4^2 are then plugged into (5) yielding a DoP estimate based on 3 polarimetric images

$$\widetilde{P}_3^2=1-\frac{4\left[\widehat{a}_1\widehat{a}_2-(\widehat{a}_3^2+\widetilde{a_4^2})\right]}{(\widehat{a}_1+\widehat{a}_2)^2}.$$

Interestingly, the asymptotic variance of the estimator $\widetilde{P}_3^2$ can be determined. The asymptotic covariance matrix of the estimator $(\widehat{a}_1,\widehat{a}_2,\widehat{a}_3,\widetilde{a_4})$ (for the unknown parameter vector $\boldsymbol{\eta}=(a_1,a_2,a_3,a_4^2)^T$) can be obtained from the asymptotic efficiency of the MLE. Indeed, the asymptotic covariance matrix of the MLE equals the Cramer Rao Lower Bound (CRLB), which is defined as the inverse of the following Fisher information matrix

$$F_3(\boldsymbol{\eta})=-E\left[\frac{\partial^2\log p(\widetilde{\boldsymbol{I}};\boldsymbol{\eta})}{\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}^T}\right].$$

However, the expectations appearing in this expression are difficult to compute analytically because of the term $\log f_{\frac{1}{2}}$ appearing in the log-density. In such situation, it is very usual to approximate the expectations by using Monte Carlo methods. More specifically, this approach consists of approximating the elements of the Fisher information matrix (FIM) $F_3(\boldsymbol{\eta})$ as follows

$$\left[F_3(\boldsymbol{\eta})\right]_{ij}\simeq-\frac{1}{N}\sum_{k=1}^{N}\frac{\partial^2\log p(\mathbf{x}_k)}{\partial\eta_i\partial\eta_j},$$

where $\mathbf{x}_k$ is distributed according to a trivariate gamma distribution whose pdf is defined in (17) and N is the number of Monte Carlo runs. Finally, the asymptotic variance of $\widetilde{P}_3^2$ can be determined as follows:

$$\text{var}\left(\widetilde{P}_3^2\right) = G_3^T F_3^{-1} G_3, \tag{18}$$

where G_3 is the gradient of the transformation from $\boldsymbol{\eta} = (a_1, a_2, a_3, a_4^2)$ to P^2, i.e.

$$G_3 = \left(\frac{4(a_1a_2 - a_2^2 - 2a_3^2 - 2a_4^2)}{(a_1+a_2)^3}, \frac{4(a_1a_2 - a_1^2 - 2a_3^2 - 2a_4^2)}{(a_1+a_2)^3}, \frac{8a_3}{(a_1+a_4)^2}\right)^T.$$

3.D. DoP estimation using 2 images

The Laplace transform of $\underline{\boldsymbol{I}} = (I_1, I_2)^T$ can be obtained by setting $\theta_3 = 0$ and $\theta_4 = 0$ in (12)

$$L_{\underline{\boldsymbol{I}}}(\underline{\boldsymbol{\theta}}) = E\left[\exp\left(-\sum_{j=1}^{2}\theta_j I_j\right)\right] = \frac{1}{\underline{P}(\underline{\boldsymbol{\theta}})}, \tag{19}$$

where $\underline{P}(\underline{\boldsymbol{\theta}}) = 1 + a_1\theta_1 + a_2\theta_2 + (a_1a_2 - a_3^2 - a_4^2)\theta_1\theta_2$ and $\underline{\boldsymbol{\theta}} = (\theta_1, \theta_2)^T$. As a consequence, the distribution of $\underline{\boldsymbol{I}}$ is a bivariate gamma distribution (since $\underline{P}(\underline{\boldsymbol{\theta}})$ is an affine polynomial). This distribution is parameterized by three parameters a_1, a_2 and $r = a_3^2 + a_4^2$. As a consequence, one can think of estimating these three parameters by using the ML method. The density of the bivariate gamma distribution having the Laplace transform $L_{\underline{\boldsymbol{I}}}(\underline{\boldsymbol{\theta}})$ has been defined in [37]:

$$p(\underline{\boldsymbol{I}}) = \frac{1}{2k}\exp\left(-\frac{a_2I_1 + a_1I_2}{2k}\right) f_1\left(cI_1I_2\right)\mathbb{I}_{\mathbb{R}_+^2}\left(\underline{\boldsymbol{I}}\right),$$

with $k = \frac{1}{2}(a_1 a_2 - a_3^2 - a_4^2)$ and $c = \frac{1}{4}(a_1 a_2 - 2k)k^{-2}$. By differentiating this density with respect to a_1, a_2, we obtain

$$\widehat{a}_1 = \widehat{\alpha}_1,\ \widehat{a}_2 = \widehat{\alpha}_2,$$

with $\widehat{\alpha}_l = \frac{1}{n}\sum_{j=1}^{n} \boldsymbol{I}_l^j$ for $l = 1, 2$. These two estimates are then replaced in $\frac{\partial p(\underline{\boldsymbol{I}})}{\partial r} = 0$, showing that the MLE of r satisfies the following nonlinear relation:

$$\widehat{a}_1\widehat{a}_2 - r - \frac{1}{n}\sum_{j=1}^{n} I_1^j I_2^j \frac{f_2\left(\frac{rI_1^jI_2^j}{(\widehat{a}_1\widehat{a}_2 - r)^2}\right)}{f_1\left(\frac{rI_1^jI_2^j}{(\widehat{a}_1\widehat{a}_2 - r)^2}\right)} = 0.$$

As in the case of three images, the practical determination of the MLE of r (denoted as $\underline{r}$) can be achieved by using a Newton-Raphson procedure. Note that the convergence of this numerical procedure has been proved in [42] for specific bivariate distributions. The MLEs of a_1, a_2 and r are then plugged into (5) yielding a DoP estimate based on 2 polarimetric images

$$\underline{P}_2^2 = 1 - \frac{4\left(\widehat{a}_1\widehat{a}_2 - \underline{r}\right)}{(\widehat{a}_1 + \widehat{a}_2)^2}.$$

The asymptotic variance of the estimator $\underline{P}_2^2$ can be computed similarly to the case of three images

$$\mathrm{var}_{\mathrm{A}}\left(\underline{P}_2^2\right) = G_2^T F_2^{-1} G_2, \tag{20}$$

where G_2 is the gradient of the transformation from (a_1, a_2, r) to P^2, i.e.

$$G_2 = \left(\frac{4(a_1a_2 - a_2^2 - 2r)}{(a_1 + a_2)^3}, \frac{4(a_1a_2 - a_1^2 - 2r)}{(a_1 + a_2)^3}\right)^T, \tag{21}$$

and F_2^{-1} is the inverse Fisher information matrix for the parameter vector $\boldsymbol{\eta} = (a_1, a_2, r)^T$.

4. DoP estimation using moment methods

In order to appreciate the performance of the DoP estimators derived above, this section studies estimators based on the classical method of moments.

4.A. 4 *images*

When four polarimetric images are available, the moment estimator of $\boldsymbol{a}$ based on the first order moments of the intensity vector $\boldsymbol{I} = (I_1, I_2, I_3, I_4)^T$ is also the estimator derived in section 3.B

$$\widehat{\boldsymbol{a}}_{\mathrm{Mo}} = M\widehat{\boldsymbol{\alpha}}_{\mathrm{Mo}},$$

where $\widehat{\boldsymbol{\alpha}}_{\mathrm{Mo}} = \frac{1}{n}\sum_{j=1}^{n} \boldsymbol{I}^j$. As a consequence, this estimator is unbiased and efficient providing an optimal estimation of $\widehat{\boldsymbol{a}}$.

4.B. 3 *images*

In the case of three observed intensities, the first order moments of $\widetilde{\boldsymbol{I}} = (I_1, I_2, I_3)^T$ are

$$\mathsf{E}[I_1] = a_1, \mathsf{E}[I_2] = a_2, \mathsf{E}[I_3] = \frac{1}{2}(a_1 + a_2) + a_3.$$

yielding the following moment estimators of a_1, a_2 and a_3:

$$(\widetilde{a}_1, \widetilde{a}_2, \widetilde{a}_3)^T = (\widehat{a}_1, \widehat{a}_2, \widehat{a}_3)^T.$$

Obviously, other moments of $\widetilde{\boldsymbol{I}}$ have to be considered to estimate a_4^2. This study focuses on the following second order moments

$$\mathsf{E}\left[I_1 I_2\right] = a_1 a_2 + a_3^2 + a_4^2, \tag{22}$$

$$\mathsf{E}\left[I_1 I_3\right] = a_1^2 + \left(\frac{a_2}{2} + 2a_3\right) a_1 + \frac{a_3^2 + a_4^2}{2}, \tag{23}$$

$$\mathsf{E}\left[I_2 I_3\right] = a_2^2 + \left(\frac{a_1}{2} + 2a_3\right) a_2 + \frac{a_3^2 + a_4^2}{2}. \tag{24}$$

The other second order moments $\mathsf{E}\left[I_1^2\right], \mathsf{E}\left[I_2^2\right]$ and $\mathsf{E}\left[I_3^2\right]$ have not been considered here since they do not depend on a_4. Of course, other higher order moments of the intensities could be considered as well. However, the estimation performance would not change significantly when using these moments. This indicates that most information regarding parameter a_4 is contained in the moments $\mathsf{E}\left[I_1 I_2\right], \mathsf{E}\left[I_1 I_3\right]$ and $\mathsf{E}\left[I_2 I_3\right]$. Based on these comments, the proposed method of moments estimates the parameter a_4^2 from (22), (23) and (24) using a non-linear least squares (NLLS) method. As the system is overdetermined, it is interesting to weight each equation by using the covariance matrix of the estimates. Indeed, suppose for example that the variance of $I_1 I_2$ is small with respect to the variances of $I_1 I_3$ and $I_2 I_3$. Using the same weight for (22), (23) and (24) would penalize the estimation with respect to a procedure based on (22) only. The NLLS method tackles this difficulty by minimizing an optimal weighted least-squares criterion. The NLLS method for the DoP estimation problem is briefly recalled in the end of this section (the interested reader is invited to consult [43] for more details).

Denote as $\boldsymbol{f}(\boldsymbol{\eta}) = \left(\mathsf{E}\left[I_1 I_2\right], \mathsf{E}\left[I_1 I_3\right], \mathsf{E}\left[I_2 I_3\right]\right)^T$, as $\boldsymbol{s_n}$ the empirical moments associated with $\boldsymbol{f}$ and as $C(\boldsymbol{\eta})$ the covariance matrix $C(\boldsymbol{\eta}) = n\mathrm{cov}(\boldsymbol{s_n})$, where $\boldsymbol{\eta} = (a_1, a_2, a_3, a_4^2)^T$ is the vector of the unknown parameters. The NLLS estimator of a_4^2 is obtained as the solution of the following optimization problem:

$$\widetilde{a}^2_{4\mathrm{Mo}} = \arg\min_{x>0} \left[\widetilde{\boldsymbol{f}}(x) - \boldsymbol{s_n}\right]^T \widetilde{C}(x)^{-1} \left[\widetilde{\boldsymbol{f}}(x) - \boldsymbol{s_n}\right],$$

with $\widetilde{\boldsymbol{f}}(x) = f(\widetilde{a}_1, \widetilde{a}_2, \widetilde{a}_3, x)$ and $\widetilde{C}(x) = C(\widetilde{a}_1, \widetilde{a}_2, \widetilde{a}_3, x)$. As there is no tractable expression for the criterion to be minimized, the computation of $\widetilde{a}^2_{4\mathrm{Mo}}$ is achieved by using a Levenberg-Marquardt gradient procedure. The DoP can then be estimated by replacing the values of a_1, a_2, a_3 and a_4^2 in (5) by their moment estimators

$$\widetilde{P}^2_{3\ \mathrm{Mo}} = 1 - \frac{4\left[\widehat{a}_1\widehat{a}_2 - (\widehat{a}_3^2 + \widetilde{a}^2_{4\mathrm{Mo}})\right]}{(\widehat{a}_1 + \widehat{a}_2)^2}.$$

A lower bound for the asymptotic variance of any estimator obtained from a method of moments can be obtained. This bound was first derived in the context of time series [43]. However, it can also be applied to the DoP estimation problem. Consider the following function $\boldsymbol{h}(\cdot) : \mathbb{R}^M \rightarrow \mathbb{R}^L$ composed of the first and second order moments of $\boldsymbol{I}$, i.e. such that $\boldsymbol{h}(I_1, I_2, I_3) = (I_1, I_2, I_3, I_1I_2, I_1I_3, I_2I_3)^T$ with $M = 3$ and $L = 6$. Consider also the following statistic (of size L) defined of empirical moments of the intensity vector

$$\boldsymbol{s}_n = \frac{1}{n}\sum_{j=1}^{n} \boldsymbol{h}(\boldsymbol{I}^j), \tag{25}$$

The first and second-order moments of $\boldsymbol{s}_n$ classically satisfy the following relations:

$$\mathsf{E}\left[\boldsymbol{s}_n\right] = \boldsymbol{f}(\boldsymbol{\eta}) = \mathsf{E}\left[\boldsymbol{h}(\boldsymbol{I}^1)\right], \tag{26}$$

$$n\mathsf{cov}[\boldsymbol{s}_n] = \boldsymbol{C}(\boldsymbol{\eta}) = \mathsf{cov}[\boldsymbol{h}(\boldsymbol{I}^1)]. \tag{27}$$

A major result in the framework of moment methods is that the covariance matrix of any estimator of moments based on $\boldsymbol{s}_n$ satisfies the following relation:

$$n\mathsf{cov}[\boldsymbol{s}_n] \geq \boldsymbol{B}(\boldsymbol{\eta}) = \left(\boldsymbol{H}(\boldsymbol{\eta})\boldsymbol{C}(\boldsymbol{\eta})^{-1}\boldsymbol{H}(\boldsymbol{\eta})^T\right)^{-1}, \tag{28}$$

where $\geq$ means that the difference between the two matrices is positive definite and $\boldsymbol{H}(\boldsymbol{\eta})$ is the Jacobian matrix of $\mathsf{E}\left[\boldsymbol{h}(I_1, I_2, I_3)\right]$ whose derivatives are computed with respect to the components of $\boldsymbol{\eta}$

$$\boldsymbol{H}(\boldsymbol{\eta}) = \begin{pmatrix} 1 & 0 & \frac{1}{2} & a_2 & 2a_1 + \frac{a_2}{2} + 2a_3 & \frac{a_2}{2} \\ 0 & 1 & \frac{1}{2} & a_1 & \frac{a_1}{2} & 2a_2 + \frac{a_1}{2} + 2a_3 \\ 0 & 0 & 1 & 2a_3 & 2a_1 + a_3 & 2a_2 + a_3 \\ 0 & 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \tag{29}$$

Since $\boldsymbol{B}(\boldsymbol{\eta})$ uses only the statistical properties of $\boldsymbol{s}_n$, it provides a lower bound on the asymptotic variances of all estimators constructed from the empirical moments contained in $\boldsymbol{s}_n$. The minimal bound $\boldsymbol{B}(\boldsymbol{\eta})$ can be used to obtain the following minimal asymptotic variance of any moment estimate of the DoP based on $\boldsymbol{s}_n$

$$\mathrm{var}_{\mathrm{A}}\left(\widetilde{P}_{3\ \mathrm{Mo}}^2\right) \geq G_3^T \boldsymbol{B}(\boldsymbol{\eta}) G_3. \tag{30}$$

4.C. 2 images

When the intensity vector is $\underline{\boldsymbol{I}} = (I_1, I_2)^T$, the moment estimators of a_1, a_2 and $r = a_3^2 + a_4^2$ have been derived in [32] by using the following set of equations:

$$\mathsf{E}[I_1] = a_1, \ \mathsf{E}[I_2] = a_2, \tag{31}$$

$$\mathsf{E}[I_1 I_2] = a_1 a_2 + r. \tag{32}$$

The estimators of a_1 and a_2 are directly related to (31):

$$(\underline{a}_1, \underline{a}_2)^T = (\widehat{a}_1, \widehat{a}_2)^T,$$

whereas the estimator of r obtained from (32) is

$$\underline{r}_{\mathrm{Mo}} = \frac{1}{n} \sum_{j=1}^{n} I_1^j I_2^j - \widehat{a}_1 \widehat{a}_2.$$

The DoP can then be estimated by replacing the values of a_1, a_2 and r in (5) by their moment estimators

$$\underline{P}_{2\ \mathrm{Mo}}^2 = 1 - \frac{4\left[\widehat{a}_1 \widehat{a}_2 - \underline{r}_{\mathrm{Mo}}\right]}{(\widehat{a}_1 + \widehat{a}_2)^2}.$$

The asymptotic covariance matrix of the moment estimator vector $\underline{\boldsymbol{\eta}}_{2,\mathrm{Mo}} = (\widehat{a}_1, \widehat{a}_2, \underline{r}_{\mathrm{Mo}})^T$ can be easily computed as:

$$\mathrm{var}_{\mathrm{A}}\left(\underline{\boldsymbol{\eta}}_{2,\mathrm{Mo}}\right) = \frac{1}{n} \begin{pmatrix} a_1^2 & r & 2a_1 r \\ r & a_2^2 & 2a_2 r \\ 2a_1 & 2a_2 r & a_1^2 a_2^2 + 4a_1 a_2 r + 3r^2 \end{pmatrix}.$$

This expression can be used to derive the asymptotic variance of $\underline{P}_{2\ \mathrm{Mo}}^2$:

$$\mathrm{var}_{\mathrm{A}}\left(\underline{P}_{2\ \mathrm{Mo}}^2\right) = G_2^T \mathrm{var}_{\mathrm{A}}\left(\underline{\boldsymbol{\eta}}_{2,\mathrm{Mo}}\right) G_2, \tag{33}$$

where G_2 is the gradient of the transformation from (a_1, a_2, r) to P^2, which has been defined previously. Straightforward computations lead to the following result

$$\mathrm{var}_{\mathrm{A}}\left(\underline{P}^2_{2\ \mathrm{Mo}}\right) = \frac{2(1-P^2)^2(P^2+1/2)}{n} + \frac{64a_1a_2r}{n(a_1+a_2)^4}.$$

Note that this last result was also obtained in [32] by using a Taylor expansion of P^2 around its true value.

5. Estimation results on synthetic data

Several experiments have been conducted to evaluate the performance of the ML and moment estimators derived in this paper. The first simulations presented here have been obtained with polarimetric images with 9 different DoPs reported in Table 1 (inspired from [32]). The corresponding entries of the covariance matrices of the Jones vector, denoted as Γ_i for $i \in \{0, \ldots, 8\}$, are given in Table 2.

Figure 1 shows the log mean square errors (MSEs) of the square DoP estimates obtained with two images using the ML method (plus markers) and the method of moments (cross markers). These MSEs can be compared to those corresponding to 4 images (diamond markers) (note that the ML method and the method of moments coincide when 4 images are observed, as explained in section 4.A). The loss of performance obtained when using two polarimetric images instead of four can be clearly observed. The sample size is $n = 15 \times 15$ in these simulations. This corresponds to a square observation window containing 225 pixels. The theoretical asymptotic log MSEs associated to the ML estimators (dashed lines) and the asymptotic lower bound

for moment estimators (dotted line) (corresponding to Eq.'s (15), (20) and (33)) are also depicted in Fig 1. The asymptotic MSEs of the different estimators match perfectly with their estimates, except for the MLE associated to the matrices Γ_0, Γ_1 and Γ_3 for 2 images. This can be explained for matrices Γ_0 and Γ_3 by noting that the parameter $r = a_3^2 + a_4^2$ equals zero in these cases. In other words, r belongs to the boundary of its definition domain, preventing the use of its theoretical asymptotic variance [44, p. 8]. The difference between estimated and theoretical results regarding the matrix Γ_1 can be explained by noting that the parameter r is close to 0. In this case, the asymptotic MSE of the estimator is not reached for this sample size (a better match would be obtained for a larger sample size). A last comment resulting from Fig. 1 is that all estimators reach their best performance for large values of the DoP as expected.

Figure 2 shows the log MSEs between the true square DoP and the one estimated using three images by the ML method (plus markers) and the NLLS method of moments (cross markers). The theoretical asymptotic log MSEs of the different estimators are also depicted (they correspond to Eq.'s (18) and (30)). The asymptotic and estimated performance of all estimators match perfectly, except for the matrices Γ_0, Γ_1 and Γ_3 (same reason as above). The MLE clearly outperforms the moment estimator and its performance is better for large values of the DoP. The loss of performance obtained when observing three polarimetric images instead of four can be clearly observed by looking at the corresponding asymptotic and estimated log MSEs.

A comparison between Figs. 1 and 2 shows that the MLEs derived for two, three and four images have a similar global behavior, with a maximum near $P^2 = 1/3$ and a decreasing variance when P^2 tends to 1. Moreover the MLE for two images has roughly similar performance that the MLE for three images. This encouraging result indicates that the DoP of polarimetric images can be estimated with two images without significant loss of performance.

The next set of simulations studies the performance of the different estimators as a function of the sample size. Figures 3 and 4 show the log MSEs of the DoP estimates obtained for 2 and 3 images and for two particular matrices Γ_2 and Γ_7. These simulations allow us to appreciate the gain of performance obtained with the ML method when compared to the method of moments. The usual linear relation between $\log_{10}\mathrm{MSE}$ and $\log_{10}(n)$ can also be observed.

To appreciate the estimation performance on synthetic images, we have considered a synthetic polarimetric image of size 512×512 composed of three distinct objects located on an homogeneous background (according to the scheme depicted in Fig. 5 and inspired from [32]). The polarimetric properties of these objects and background (i.e. the covariance matrix of the Jones vector and the DoP) are reported in Table 3. Typical intensities associated to this polarimetric image are also represented in Fig. 6. The polarimetric contrast image, proposed in [21] and based on a non-statistical approach, is also displayed in Fig. 6(f). It shows that the object labeled as "3" in Table 3 is not visible on this contrast image. Moreover objects "1" and "2" exhibit quite

similar contrast levels despite their different polarimetric properties. These results emphasize the interest of statistical methods in order to estimate the DoP. The DoP of each pixel $\mathbf{x}^{(i,j)}$ (for $i, j = 1, \ldots, 512$) has been estimated from vectors belonging to windows of size $n = 15 \times 15$ centered around the pixel of coordinates (i, j) in the analyzed image. The estimated DoPs are depicted in Fig. 7 for the different estimation methods. The numbers appearing in each region of the image are given for indication. They represent the means of the DoP estimates for each object assuming that the region constituting objects are perfectly known (Note that these means are slightly different from the theoretical DoP since near the boundary of each object, the square estimation window is composed of both background and object pixels. Therefore, the intensities are not homogeneous on these estimation windows, and the estimation is biased.). Finally, these results confirm that the MLE performs better than the moment estimator for DoP estimation. They also show that the polarimetric properties of the image seem to be estimated efficiently with 3 or 2 intensity images only.

6. Estimation results on real data

6.A. Experimental framework

In order to confirm our simulation results, a simple but optimized imaging system has been designed. It enables to acquire real polarimetric images under coherent illumination. The emission part is a He:Ne laser, oscillating at 623nm and providing a 15mW output power. The corresponding output beam is linearly polarized along the

vertical direction and carefully reshaped in order to obtain a 5cm diameter uniform illumination spot. The reception part consists of a 12-bits Basler A312f camera giving images with a resolution of 782×582 pixels. The fore optics is a 50mm focal objective including an adjustable diaphragm. In order to analyze the polarimetric signature of the scene, a linear polarizer is placed in front of the camera. Actually, the intensity images corresponding to I_1, I_2 and I_3 are obtained by orienting this polarizer at respectively 0, 90 and 45 degrees with respect to the polarization of the illumination beam. Otherwise, the intensity image related to I_4 is obtained by inserting in front of the polarizer (oriented at 0°) a quarter wave whose fast axis is oriented at 45°. Finally, the scene, located 3m away, consists of two types of panels. The first one, intended to provide low DoP, is a grey diffuse plastic material (left object of Fig. 8), whereas the second one is made of a sand blasted aluminium panel (right object of Fig. 8) providing a high DoP value.

6.B. Estimation Results

The intensity images corresponding to I_1, I_2, I_3 and I_4 are depicted in Fig. 8. The total intensity $I_{\text{Tot}} = I_1 + I_2$, which is the intensity measured by a conventional imaging system of reflectivity, is represented in Fig. 8(e). This last image shows that the plastic and steel disks have a similar level of reflectivity and can hardly be distinguished without a polarimetric processing. It is important to note that the measured intensities are quite low due to the experimental conditions. Therefore, the noises

affecting these images is significant on this set of real data.

Figure 9 shows the estimated square DoPs P^2 for 2, 3 and 4 images. The estimation window size has been fixed to a quite small value $n = 9 \times 9$. This size allows one to mitigate the inhomogeneities on both plastic and steel disks and provides a good tradeoff between the estimation robustness and the expected resolution of the estimated polarimetric images. The DoP estimates obtained with the ML method for 2 and 3 images clearly provide better results than those obtained with the methods of moments, particularly for the plastic disk. Moreover, the expected values of these MLE are quite different on each object: the plastic disk appears highly depolarizing and the steel disk reveals itself to be not very depolarizing. This result is in good agreement with the theoretical properties of the studied materials and emphasizes the interest for polarimetric imaging systems.

7. Conclusions

Some recent studies have shown that the polarization degree of polarimetric images can be estimated by the method of moments using only two polarimetric images. This paper proved that the estimation can also be conducted by maximum likelihood methods using two, three or four images. The estimation results obtained on synthetic and real data were very encouraging. Future investigations include the generalization of these results to low flux images. Another degradation is added to the speckle noise when the intensity level of the light coming from the scene corresponds to

a small number of photons. This is for instance the case when the objects are far from the source of illumination or under low flux illumination. This combination of Poisson and speckle noises leads to a complex multivariate distribution for the images. Estimating the parameters of polarimetric images in this context using maximum likelihood methods is currently under investigation.

Appendix A: MLE of a_4^2 for three images

After removing the terms which do not depend on $\boldsymbol{a}$, the log-likelihood associated to the density of the intensity vector $\widetilde{\boldsymbol{I}} = (I_1, I_2, I_3)^T$ derived in section 3.C can be written

$$l_3\left(\widetilde{\boldsymbol{I}}^{(n)}; \boldsymbol{a}\right) = -\sum_{j=1}^{n}\left[\frac{(a_2+a_3)I_1 + (a_1+a_3)I_2 - 2a_3I_3}{2k} - \log f_{\frac{1}{2}}\left(\frac{a_4^2 v'(\widetilde{\boldsymbol{I}}^j)}{16k^2}\right)\right] - n\log(k),$$

with $k = (a_1a_2 - a_3^2 - a_4^2)/2$. By differentiating this log-likelihood with respect to a_1, a_2 and a_3 and by solving the resulting system, the following result can be obtained:

$$\begin{pmatrix} \widehat{a}_1 \\ \widehat{a}_2 \\ \widehat{a}_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1/2 & 1 & -1/2 \end{pmatrix} \begin{pmatrix} \widehat{\alpha}_1 \\ \widehat{\alpha}_2 \\ \widehat{\alpha}_3 \end{pmatrix},$$

with $\widehat{\alpha}_l = \frac{1}{n}\sum_{j=1}^{n} \boldsymbol{I}_l^j$ for $l = 1, 2, 3$. After differentiating the log-likelihood with respect to a_4^2, the following score function is obtained

$$\begin{aligned} g_3\left(\widetilde{\boldsymbol{I}}^{(n)}; \boldsymbol{a}\right) =& \frac{n}{4k^2}[2k - (a_2+a_3)\widehat{m}_1 - (a_1+a_3)\widehat{m}_2 + 2a_3\widehat{m}_3] \\ &+ \frac{(a_1a_2 - a_3^2 + a_4^2)}{32k^3}\sum_{j=1}^{n} v'(\widetilde{\boldsymbol{I}}^j)\frac{f_{\frac{3}{2}}\left(\frac{a_4^2 v'(\widetilde{\boldsymbol{I}}^j)}{16k^2}\right)}{f_{\frac{1}{2}}\left(\frac{a_4^2 v'(\widetilde{\boldsymbol{I}}^j)}{16k^2}\right)}. \end{aligned}$$

After replacing (a_1, a_2, a_3) by their MLEs in this score function, multiplying the result by $-8k^2[n(a_1a_2 - a_3^2 + a_4^2)]^{-1}$ and equating to zero, the following result can be obtained:

$$1 - \frac{1}{4n}\frac{1}{\widehat{d} - a_4^2}\sum_{j=1}^{n} v'\left(\widetilde{\boldsymbol{I}}^j\right)\frac{f_{\frac{3}{2}}\left(\frac{a_4^2 v'\left(\widetilde{\boldsymbol{I}}^j\right)}{4\left(\widehat{d} - a_4^2\right)^2}\right)}{f_{\frac{1}{2}}\left(\frac{a_4^2 v'\left(\widetilde{\boldsymbol{I}}^j\right)}{4\left(\widehat{d} - a_4^2\right)^2}\right)} = 0,$$

where $\widehat{d} = \widehat{a}_1 \widehat{a}_2 - \widehat{a}_3^2$. The relation $f_{\frac{3}{2}}(x)/f_{\frac{1}{2}}(x) = \tanh\left(2\sqrt{x}\right)/\sqrt{x}$ allows one to obtain the nonlinear relation used for the ML estimation of a_4^2:

$$1 - \frac{1}{2n}\sum_{j=1}^{n}\sqrt{\frac{v'\left(\widetilde{\boldsymbol{I}}^j\right)}{a_4^2}}\tanh\left(\frac{\sqrt{a_4^2 v'\left(\widetilde{\boldsymbol{I}}^j\right)}}{\widehat{d} - a_4^2}\right) = 0,$$

where $\tanh(x)$ is the hyperbolic tangent of x.

A. Acknowledgments

The authors would like to thank Gérard Letac for his fruitful comments regarding multivariate gamma distributions and Philippe Réfrégier and Julien Fade for their scientific contribution.

References

1. J. F. de Boer, T. E. Milner, M. J. C. van Gemert, and J. S. Nelson, "Two-dimensional birefringence imaging in biological tissue polarization-sensitive optical coherence tomography," Opt. Lett. **22**, 934–936 (1997).

2. A. F. Sadjadi and C. S. L. Chun, "Automatic detection of small objects from their infrared state-of-polarization vectors," Opt. Lett. **28**, 531–533 (2003).

3. J. W. Williams, J. S. Tee, and M. A. Poulter, "Image processing and classification for the uk remote minefield detection system infrared polarimetric camera," in "Proc. SPIE, Detection and Remediation Technologies for Mines and Minelike Targets VI," , vol. 4394, A. C. Dubey, J. F. Harvey, J. T. Broach, and V. George, eds. (2001), vol. 4394, pp. 139–152.

4. L. B. Wolff, "Polarization-based material classification from specular reflection," IEEE Trans. Pattern Anal. Mach. Intell. **12**, 1059–1071 (1990).

5. I. K. Miyazaki D., Kagesawa M., "Transparent surface modeling from a pair of polarization images," IEEE Trans. Pattern Anal. Mach. Intell. **26**, 73–82 (2004).

6. C. F. Tongbo Chen, Hendrik Lensch and H.-P. Seidel, "Polarization and phase-shifting for 3d scanning of translucent objects," in "IEEE Conference on Computer Vision and Pattern Recognition, 2007," (2007), pp. 1–8.

7. T. Shibata, T. Takahashi, D. Miyazaki, Y. Sato, and K. Ikeuchi, "Creating photorealistic virtual model with polarization-based vision system," in "Proc. SPIE,

Polarization Science and Remote Sensing II," , vol. 5888, J. A. Shaw and J. S. Tyo, eds. (2005), vol. 5888, pp. 25–35.

8. J. S. Tyo, D. L. Goldstein, D. B. Chenault, and J. A. Shaw, "Review of passive imaging polarimetry for remote sensing applications," Appl. Opt. **45**, 5453–5469 (2006).

9. E. S. Perlman, C. A. Padgett, M. Georganopoulos, W. B. Sparks, J. A. Biretta, C. P. O'Dea, S. A. Baum, M. Birkinshaw, D. M. Worrall, F. Dulwich, S. Jester, A. Martel, A. Capetti, and J. P. Leahy, "Optical polarimetry of the jets of nearby radio galaxies: I. the data," Astrophys. J. **651**, 735–748 (2006).

10. S. Umeyama and G. Godin, "Separation of diffuse and specular components of surface reflection by use of polarization and statistical analysis of images," IEEE Trans. Pattern Anal. Mach. Intell. **26**, 639–647 (2004).

11. N. Karpel and Y. Y. Schechner, "Portable polarimetric underwater imaging system with a linear response," in "Proc. SPIE, Polarization: Measurement, Analysis, and Remote Sensing VI," , vol. 5432 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, D. H. Goldstein and D. B. Chenault, eds. (2004), vol. 5432 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pp. 106–115.

12. P. C. Y. Chen, J. C. Flitton, K. I. Hopcraft, E. Jakeman, D. L. Jordan and J. G. Walker, "Improving visibility depth in passive underwater imaging by use

of polarization," Appl. Opt. **42**, 2794–2803 (2003).

13. Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, "Polarization-based vision through haze," Appl. Opt. **42**, 511–525 (2003).
14. Y. Y. Schechner, J. Shamir, and N. Kiryati, "Polarization and statistical analysis of scenes containing a semireflector," J. Opt. Soc. Am. A **17**, 276–284 (2000).
15. K. Koshikawa, "A polarimetric approach to shape understanding of glossy objects," in "Proc. of the 6th IJCAI," (Tokyo, Japan, 1979), pp. 493–495.
16. J. M. Bueno and P. Artal, "Double-pass imaging polarimetry in the human eye," Opt. Lett. **24**, 64–66 (1999).
17. J. S. Tyo, M. P. Rowe, E. N. Pugh, and N. Engheta, "Target detection in optical scattering media by polarization-difference imaging," Appl. Opt. **35**, 1855–1870 (1996).
18. S. Breugnot and P. Clémenceau, "Modeling and performances of a polarization active imager at $\lambda = 806$ nm," Opt. Eng. **39**, 2681–2688 (2000).
19. J. W. Goodman, *Statistical Optics* (Wiley, New York, 1985).
20. S. L. Jacques, J. C. Ramella-Roman, and K. Lee, "Imaging skin pathology with polarized light," J. Biomed. Opt. **7**, 329–340 (2002).
21. J. G. Walker, P. C. Y. Chang and K. I. Hopcraft, "Visibility depth improvement in active polarization imaging in scattering media," Appl. Opt. **39**, 4933–4941 (2000).

22. S. Huard, *Polarization of light* (Wiley, New York, 1997).

23. N. Kikuchi, "Analysis of Signal Degree of Polarization Degradation Used as Control Signal for Optical PolarizationMode Dispersion Compensation," J. Lightw. Technol. **19**, 480–486 (2001).

24. A. Luis, "Degree of polarization in quantum optics," Phys. Rev. A **66**, 013806–1 – 013806–8 (2002).

25. O. Korotkova, M. Salem, and E. Wolf, "The far-zone behavior of the degree of polarization of electromagnetic beams propagating through atmospheric turbulence," Opt. Commun. **233**, 225–230 (2004).

26. D. Miyazaki, M. Saito, Y. Sato, and K. Ikeuchi, "Determining surface orientations of transparent objects based on polarization degrees in visible and infrared wavelengths," J. Opt. Soc. Am. A **19**, 687–694 (2002).

27. V. Sankaran, M. J. Everett, D. J. Maitland, and J. T. Walsh, Jr., "Comparison of polarized-lightpropagation in biological tissue and phantoms," Opt. Lett. **24**, 1044–1046 (1999).

28. S. G. Demos, H. Savage, A. S. Heerdt, S. Schantz, and R. R. Alfano, "Time resolved degree of polarization for human breast tissue," Opt. Commun. **124**, 439–442 (1996).

29. C. Brosseau, *Fundamentals of polarized light - A statistical approach* (Wiley, New York, 1998).

30. F. Goudail and P. Réfrégier, "Target segmentation in active polarimetric images by use of statistical active contours," Appl. Opt. **41**, 874–883 (2002).

31. F. Goudail and P. Réfrégier, *Statistical image processing techniques for noisy images: an application oriented approach* (Kluwer, New York, 2004).

32. M. Roche, J. Fade, and P. Réfrégier, "Parametric estimation of the square degree of polarization from two intensity images degraded by fully developed speckle noise," J. Opt. Soc. Am. A **24**, 2719–2727 (2007).

33. P. Réfrégier, J. Fade, and M. Roche, "Estimation precision of the degree of polarization from a single speckle intensity image," Opt. Lett. **32**, 739–741 (2007).

34. J. Fade, M. Roche, and P. Réfrégier, "Precision of moment based estimation of the degree of polarization in coherent imagery without polarization device," J. Opt. Soc. Am. A **25**, 483–492 (2008).

35. P. Graczyk, G. Letac, and H. Massam, "The complex wishart distribution and the symmetric group," Ann. Statist. **31**, 287–309 (2003).

36. P. Bernardoff, "Which multivariate Gamma distributions are infinitely divisible?" Bernoulli **12**, 169–189 (2006).

37. F. Chatelain, J.-Y. Tourneret, A. Ferrari, and J. Inglada, "Bivariate gamma distributions for image registration and change detection," IEEE Trans. Image Process. **16**, 1796–1806 (2007).

38. P. Réfrégier, *Noise Theory and Application to Physics: From Fluctuations to In-*

formation (Springer, New-York, 2004).

39. S. M. Kay, *Fundamentals of statistical signal processing: estimation theory* (Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993).

40. G. Letac and J. Wesolowski, "Laplace transforms which are negative powers of quadratic polynomials," Trans. Amer. Math. Soc. **360**, 6475–6496 (2008).

41. M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).

42. F. Chatelain, G. Letac, and J.-Y. Tourneret, "Estimating the correlation coefficient of bivariate gamma distributions using the maximum likelihood principle and the inference for margins," Tech. rep., IRIT/ENSEEIHT/TéSA (2007).

43. B. Porat and B. Friedlander, "Performance analysis of parameter estimation algorithms based on high-order moments," Int. J. Adapt. Control Signal Process. **3**, 191–229 (1989).

44. M. G. Kendall and A. Stuart, *The advanced theory of statistics*, vol. 2 (Griffin, London, 1961).

Table 1. Polarimetric image DoPs.

P_0^2	P_1^2	P_2^2	P_3^2	P_4^2	P_5^2	P_6^2	P_7^2	P_8^2
0	0.2	0.3	0.4	0.5	0.6	0.8	0.9	0.99

Table 2. Covariance matrices of the Jones vector.

	Γ_0	Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_6	Γ_7	Γ_8
a_1	2	15	1	16	82	18	30	2	1.25
a_2	2	6	1	3.6	17	11	14	2	26
a_3	0	0.2	0.4	0	0	7	16	0.6	0
a_4	0	0.5	$\sqrt{0.14}$	0	13	8	8	1.8	5.5

Table 3. Polarimetric properties of image DoPs.

Object	Polarization matrix Γ	P^2	Remarks
Background	$\begin{pmatrix} 0.79 & 0 \\ 0 & 0.98 \end{pmatrix}$	0.0115	very depolarizing and dark background
1	$\begin{pmatrix} 3.6 & 0 \\ 0 & 0.22 \end{pmatrix}$	0.783	very bright and weakly depolarizing object (typically steel)
2	$\begin{pmatrix} 3 & 0.1 \\ 0.1 & 0.6 \end{pmatrix}$	0.447	bright object quite depolarizing
3	$\begin{pmatrix} 0.7 & 0.5+0.2\,i \\ 0.5-0.2\,i & 1.07 \end{pmatrix}$	0.414	dark object whose mean total intensity is the same that the background

List of Figures

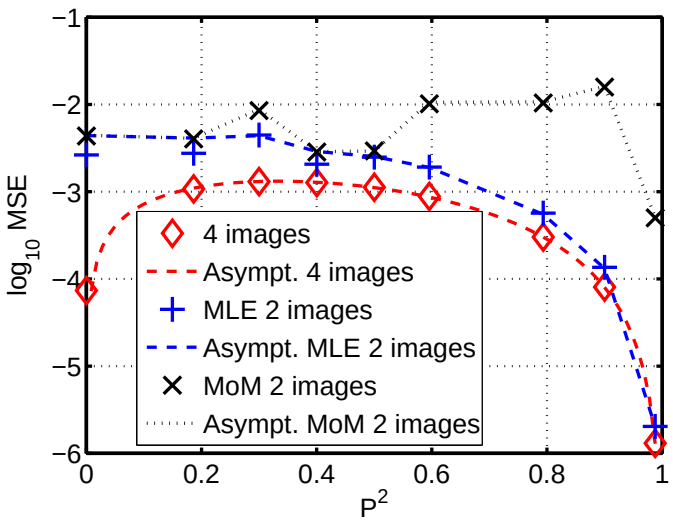


Fig. 1. log MSEs of the square DoP estimates using 2 and 4 images vs P^2 for the set of polarization matrices defined in Tab. 2 ($n = 15 \times 15$, "MoM": method of moments estimators, "MLE": maximum likelihood estimators, "Asympt.": theoretical asymptotic value of the log MSE for a given estimator).

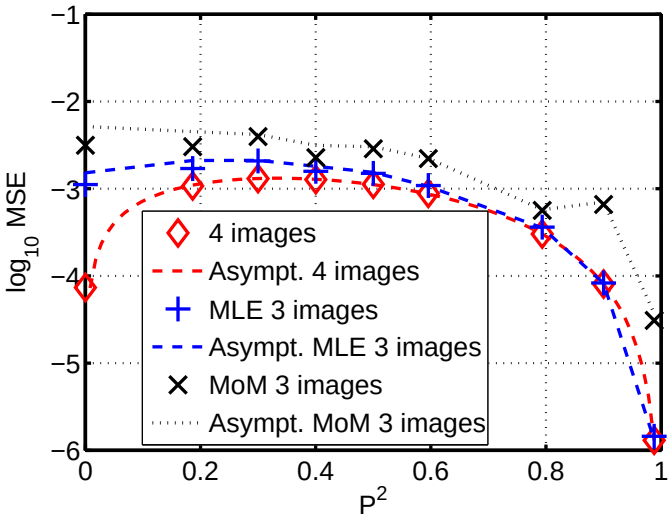
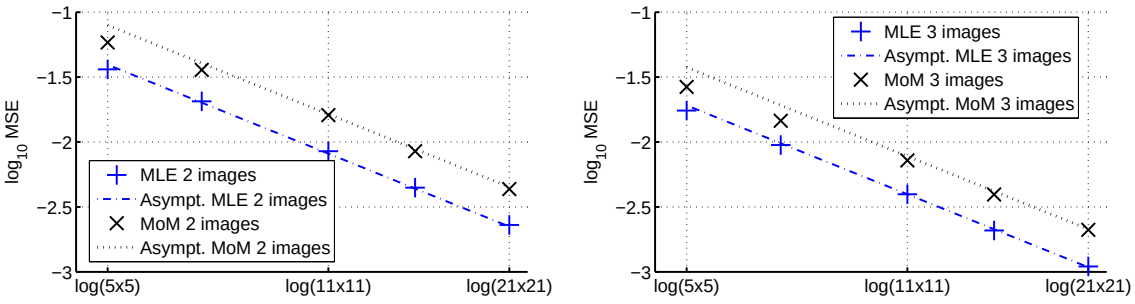
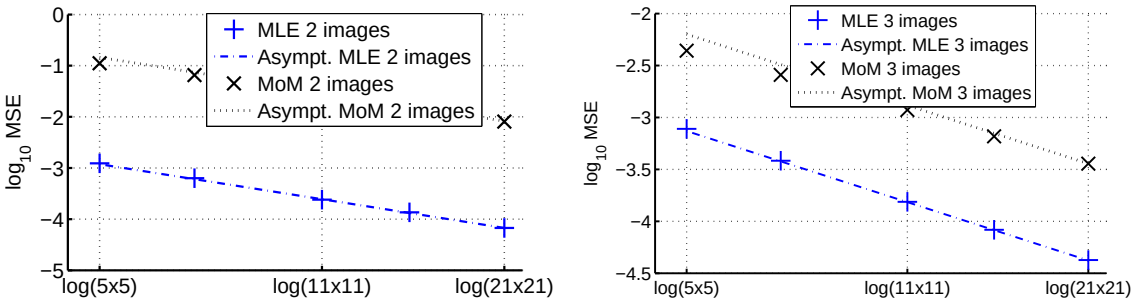


Fig. 2. log MSEs of the square DoP estimates using 3 and 4 images vs P^2 for the set of polarization matrices defined in Tab. 2 ($n = 15 \times 15$, "MoM": method of moments estimators, "MLE": maximum likelihood estimators, "Asympt.": theoretical asymptotic value of the log MSE for a given estimator).



(a) 2 images (b) 3 images

Fig. 3. log MSE of the estimated square DoP P^2 using 2 or 3 intensity images versus the logarithm of the sample size for the matrix Γ_2 ("MoM": method of moments estimators, "MLE": maximum likelihood estimators, "Asympt.": theoretical asymptotic value of the log MSE for a given estimator).



(a) 2 images (b) 3 images

Fig. 4. log MSE of the estimated square DoP P^2 using 2 or 3 intensity images versus the logarithm of the sample size for the matrix Γ_7 ("MoM": method of moments estimators, "MLE": maximum likelihood estimators, "Asympt.": theoretical asymptotic value of the log MSE for a given estimator).

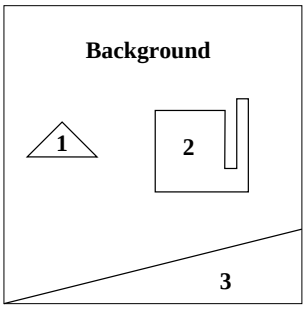
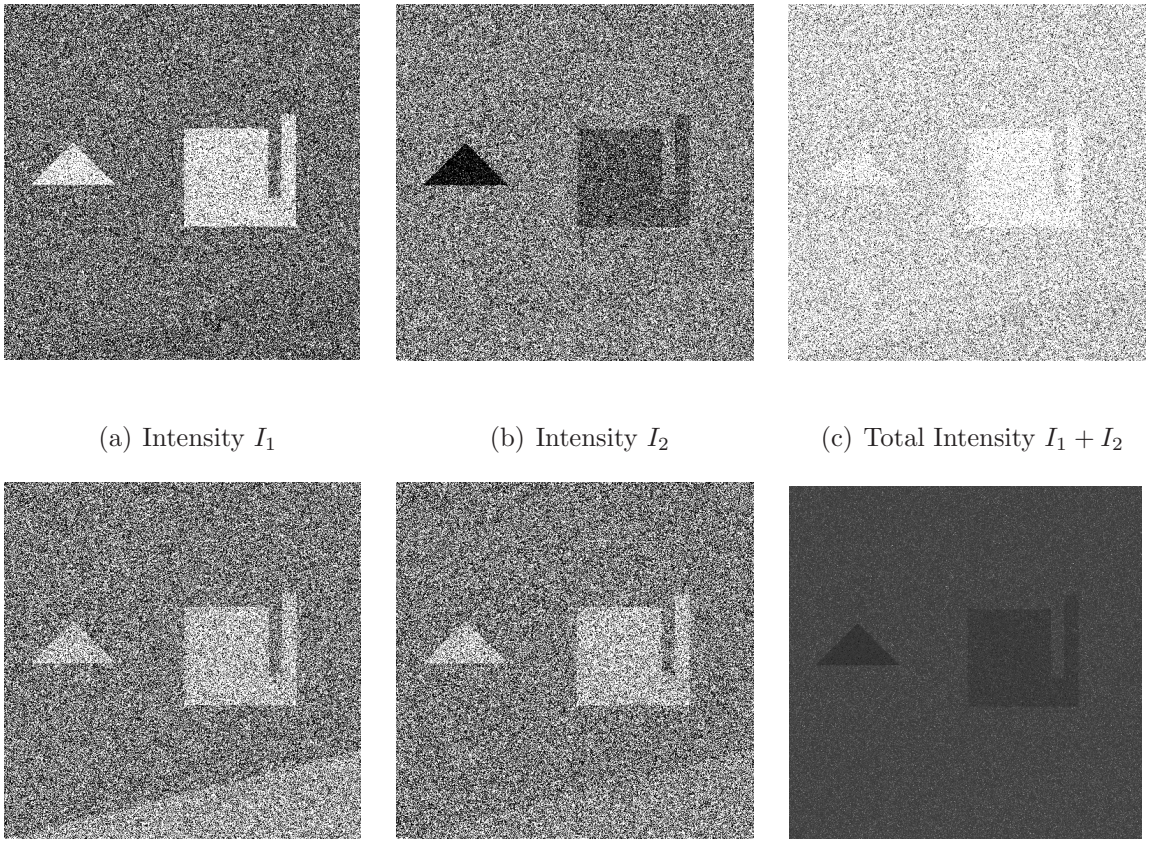


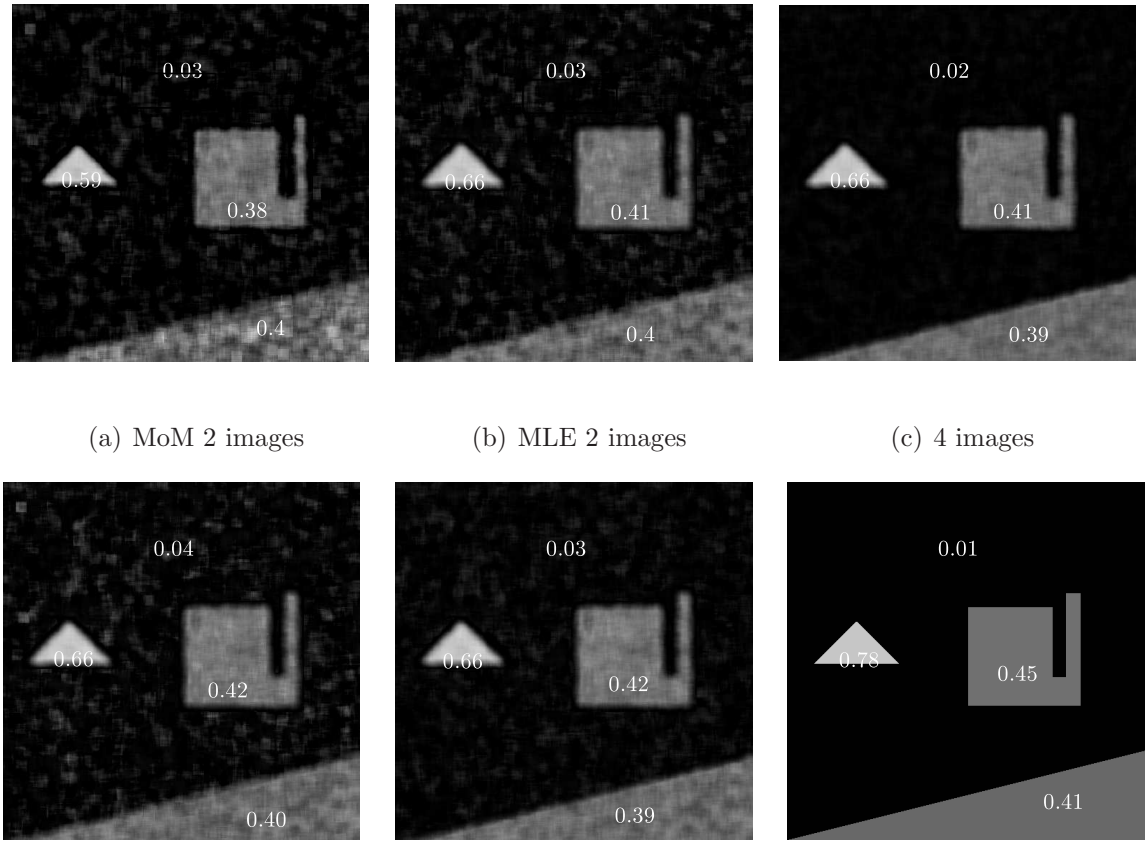
Fig. 5. Composition of the scene used to generate synthetic polarimetric intensity images.



(a) Intensity I_1 (b) Intensity I_2 (c) Total Intensity $I_1 + I_2$

(d) Intensity I_3 (e) Intensity I_4 (f) Image contrast C

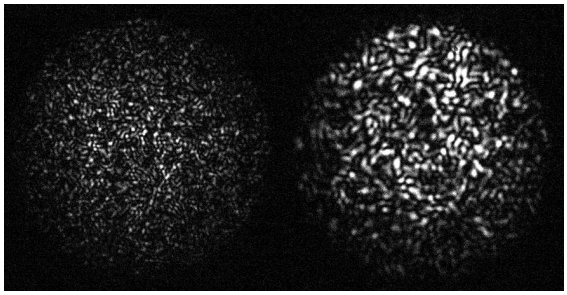
Fig. 6. Synthetic intensity images and polarimetric contrast image $C = I_2 - \alpha I_1$ [21] (with a typical value of $\alpha = 12\%$) for the scene depicted in Fig. 5 and described in Tab. 3.



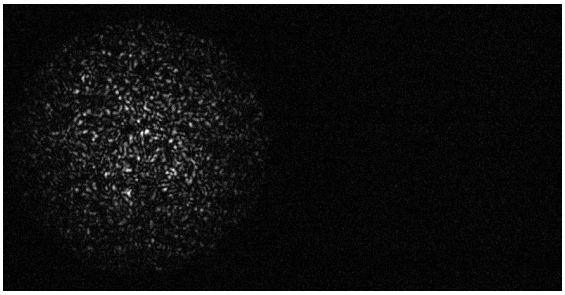
(a) MoM 2 images (b) MLE 2 images (c) 4 images

(d) MoM 3 images (e) MLE 3 images (f) Theoretical DoP

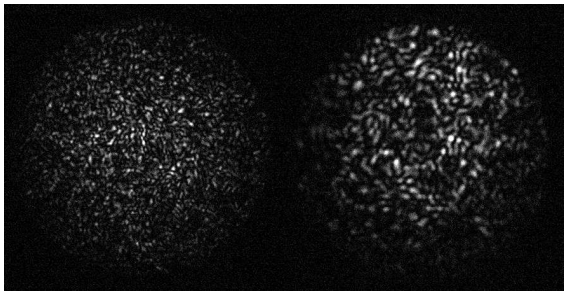
Fig. 7. Estimates of P^2 using 2, 3 or 4 intensity images for the synthetic polarimetric images (the numbers appearing in each region of the image are given for indication and represent the means of the estimates for each object assuming that the region constituting each object is perfectly known) for an estimation window of size $n = 15 \times 15$. "MoM": method of moments estimators, "MLE": maximum likelihood estimators.



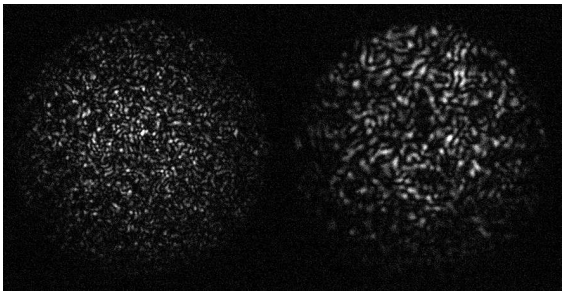
(a) Intensity I_1



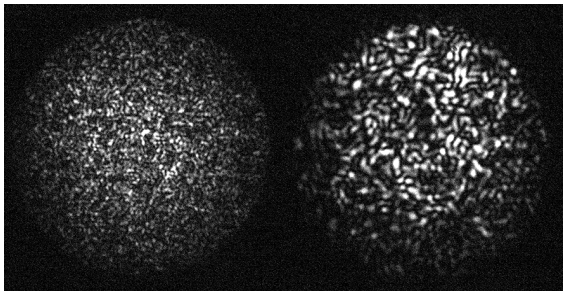
(b) Intensity I_2



(c) Intensity I_3

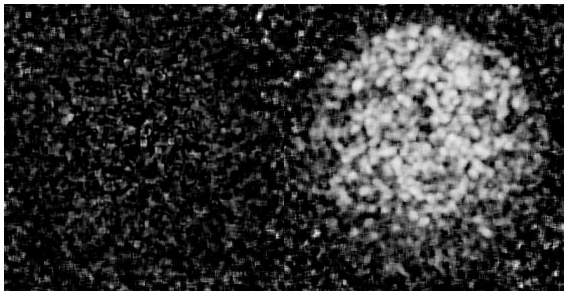


(d) Intensity I_4

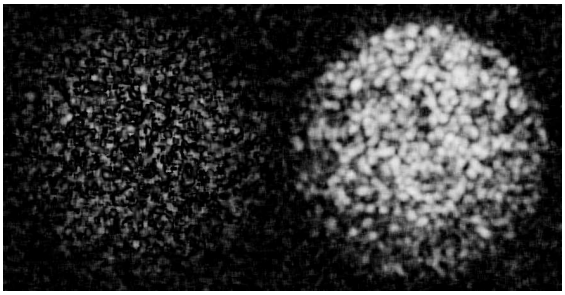


(e) Total Intensity $I_{\text{Total}} = I_1 + I_2$

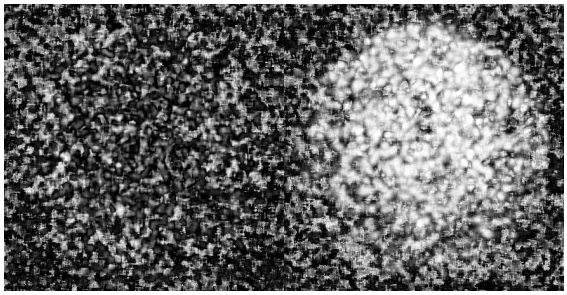
Fig. 8. Real polarimetric intensity images of a scene composed of a plastic disk (left) and a steel disk (right).



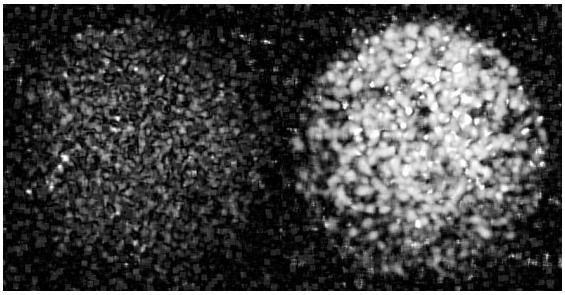
(a) 2 images: MoM



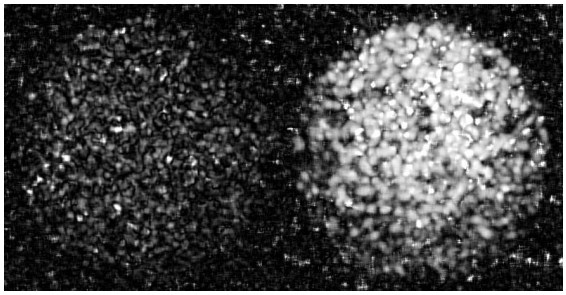
(b) 2 images: MLE



(c) 3 images: MoM



(d) 3 images: MLE



(e) 4 images

Fig. 9. Estimates of P^2 using 2, 3 or 4 intensity images for the real polarimetric images (size of the estimation window: $n = 9 \times 9$, "MoM": method of moments estimators, "MLE": maximum likelihood estimators).