

# Estimation of Mueller matrices using non-local means filtering

Sylvain Faisan,\* Christian Heinrich, Giorgos Sfikas, and Jihad Zallat

LSIIT - UMR 7005 CNRS/University of Strasbourg  
Pôle API, Bd Sébastien Brant BP 10413  
67412 Illkirch Cedex France

[\\*faisan@unistra.fr](mailto:*faisan@unistra.fr)

**Abstract:** This article addresses the estimation of polarization signatures in the Mueller imaging framework by non-local means filtering. This is an extension of previous work dealing with Stokes signatures. The extension is not straightforward because of the gap in complexity between the Mueller framework and the Stokes framework. The estimation procedure relies on the Cholesky decomposition of the coherency matrix, thereby ensuring the physical admissibility of the estimate. We propose an original parameterization of the boundary of the set of Mueller matrices, which makes our approach possible. The proposed method is fully unsupervised. It allows noise removal and the preservation of edges. Applications to synthetic as well as real data are presented.

© 2013 Optical Society of America

**OCIS codes:** (100.3020) Image reconstruction-restoration; (100.3190) Inverse problems; (120.5410) Polarimetry.

---

## References and links

1. J. Zallat, S. Anouz, and M. P. Stoll, "Optimal configurations for imaging polarimeters: impact of image noise and systematic errors," *J. Opt. A: Pure Appl. Opt.* **8**, 807–814 (2006).
2. J. Zallat, Ch. Collet, and Y. Takakura, "Clustering of polarization-encoded images," *Appl. Opt.* **43**, 283–292 (2004).
3. Ch. Collet, J. Zallat, and Y. Takakura, "Clustering of Mueller matrix images for skeletonized structure detection," *Opt. Express* **12**, 1271–1280 (2004).
4. J. Zallat, Ch. Heinrich, and M. Petremand, "A Bayesian approach for polarimetric data reduction: the Mueller imaging case," *Opt. Express* **16**, 7119–7133 (2008).
5. S. Faisan, Ch. Heinrich, F. Rousseau, A. Lallement, and J. Zallat, "Joint filtering-estimation of Stokes vector images based on a non-local means approach," *J. Opt. Soc. Am. A* **29**, 2028–2037 (2012).
6. A. Buades, B. Coll, and J. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.* **4**, 490–530 (2005).
7. Z. Xing, "On the deterministic and nondeterministic Mueller matrix," *J. Mod. Opt.* **39**, 461–484 (1992).
8. C. Givens and A. Kostinski, "A simple necessary and sufficient condition on physically realizable Mueller matrices," *J. Mod. Opt.* **40**, 471–481 (1993).
9. C. Vanderme, "An eigenvalue criterion for matrices transforming Stokes parameters," *J. Math. Phys.* **34**, 5072–5088 (1993).
10. D. Anderson and R. Barakat, "Necessary and sufficient conditions for a Mueller matrix to be derivable from a Jones matrix," *J. Opt. Soc. Am. A* **11**, 2305–2319 (1994).
11. A. Gopala Rao, K. Mallesh, and Sudha, "On the algebraic characterization of a Mueller matrix in polarization optics – I. Identifying a Mueller matrix from its N matrix," *J. Mod. Opt.* **45**, 955–987 (1998).
12. S. Cloude and E. Pottier, "Concept of polarization entropy in optical scattering: polarization analysis and measurement," *Opt. Eng.* **34**, 1599–1610 (1995).
13. A. Aiello, G. Puentes, D. Voigt, and J. P. Woerdman, "Maximum-likelihood estimation of Mueller matrices," *Opt. Lett.* **31**, 817–819 (2006).

14. C. Deledalle, L. Denis, and F. Tupin, "Iterative weighted maximum likelihood denoising with probabilistic patch-based weights," *IEEE Trans. Image Process.* **18**, 2661–2672 (2009).
  15. S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, 2004).
  16. G. Stewart, *Matrix algorithms. Vol II: eigensystems* (SIAM, 2001).
  17. T. Coleman and Y. Li, "On the convergence of reflective Newton methods for large-scale nonlinear minimization subject to bounds," *Math. Program.* **67**, 189–224 (1994).
  18. T. Coleman and Y. Li, "An interior, trust region approach for nonlinear minimization subject to bounds," *SIAM J. Optim.* **6**, 418–445 (1996).
  19. G. H. Golub and C. F. Van Loan, *Matrix Computations (3rd Edition)* (The Johns Hopkins University Press, 1996).
- 

## 1. Introduction

Mueller imaging is facing an ever growing scientific attention owing to the vast amount of information that it can reveal. In this framework, the Stokes-Mueller formalism linearly links polarization parameters to raw radiance acquisitions. This observation model is defined pixel-wise.

Pixelwise data reduction (inversion) approaches are classically used, the system being optimized so that the observation matrix (the polarization measurement matrix, PMM) reduces the impact of noise and system errors. The limitations of classical inversion procedures are known [1]. Even for a well-calibrated imaging polarimeter, errors and noise may be amplified, thus yielding unphysical Mueller matrices. Moreover, pixelwise inversion does not account for the spatial distribution of information. Exploiting the bidimensional structure of the information distribution has not been used systematically in the Mueller imaging framework. We note however that this was partially exploited in [2,3] and fully considered in [4] but in the particular case of classwise constant signatures.

A joint filtering-estimation procedure has been introduced in [5], to tackle the aforementioned shortcomings in the Stokes imaging modality. This procedure allows to estimate polarization signatures for Stokes images, while preserving sharp transitions. The procedure is based on non-local means (NLM) filtering, which is an efficient denoising algorithm that outperforms popular denoising methods regarding the preservation of sharp edges and fine texture details [6]. It is here extended to the Mueller imaging context: the noise is filtered while yielding physically admissible Mueller matrices at each pixel location. The proposed joint filtering-estimation procedure is expressed as a constrained optimization problem. Interestingly, we show that it can be equivalently seen as a two step method: a filtering stage based on the NLM approach followed by an estimation step ensuring physical admissibility. Our algorithm takes advantage of a nontrivial and original parameterization of the boundary of the set of Mueller matrices.

The article is organized as follows. Section 2 recalls the physical admissibility criteria related to Mueller matrices as well as the main lines of the approach developed in [5] for the Stokes imaging case. Section 3 extends the NLM-based polarimetric data reduction method to the Mueller imaging case. Section 4 deals with the application of the approach to synthetic and real data. Conclusions are drawn in section 5.

## 2. Related work

In this section, we briefly recall the physical criteria that must be met by any valid Mueller matrix. We summarize the approach that has been proposed in the Stokes imaging framework [5].

### 2.1. Physical admissibility of Mueller matrices

Whenever Mueller matrices are obtained from noisy polarized raw radiances, it is important to ensure that the estimation procedure provides physically admissible entities. A physical admissibility criterion for Mueller matrices has to be defined. This issue has been widely addressed in

the literature where several authors have dealt with the physical admissibility of experimentally measured Mueller matrices [7–11]. Many criteria have been proposed to assess the validity of such matrices.

Amongst them, the Jones criterion (also called the criterion for “physical” Mueller matrices) requires that a Mueller matrix is a linear combination, with non-negative coefficients, of at most four pure Mueller matrices (a pure Mueller matrix is obtained from a Jones matrix, see [11] for the complete definition). As stated in [11], the Jones criterion enables to represent a random assembly of Jones filters, thus covering perhaps every possible case of interest in polarization optics. Beyond any controversy about such a criterion, we use in the following the Jones criterion so that a Mueller matrix is defined hereafter as a matrix that verifies this criterion. With this definition, it has been established that a Mueller matrix is admissible if and only if the eigenvalues of the related coherency matrix are positive or null [12]. Moreover this criterion allows an interesting parameterization of the Mueller matrix in terms of Cholesky decomposition. This parameterization was first used by [13] in a probabilistic context and transferred efficiently to the Mueller imaging framework [4].

## 2.2. Joint filtering-estimation of Stokes vectors

In the context of the non-local means filtering of a one-channel noisy image  $I$ , the estimate of the denoised image  $I_{nlm}$  at pixel  $\mathbf{x}$  is a weighted average of all pixels in the image:

$$\begin{aligned} \forall \mathbf{x} \in \Omega, \quad I_{nlm}(\mathbf{x}) &= \sum_{\mathbf{y} \in \Omega} w(\mathbf{x}, \mathbf{y}) I(\mathbf{y}) \\ &= \arg \min_a \sum_{\mathbf{y} \in \Omega} w(\mathbf{x}, \mathbf{y}) (I(\mathbf{y}) - a)^2, \end{aligned} \quad (1)$$

where  $\Omega$  is the support of the digital image ( $\Omega \subset \mathbb{Z}^2$ ), and where  $w(\mathbf{x}, \mathbf{y})$  represents the similarity between pixels  $\mathbf{x}$  and  $\mathbf{y}$  with  $0 \leq w(\mathbf{x}, \mathbf{y}) \leq 1$ , and  $\sum_{\mathbf{y}} w(\mathbf{x}, \mathbf{y}) = 1$ . The similarity between two pixels derives from Euclidean distance between patches, a patch being a square window centered on  $\mathbf{x}$  or  $\mathbf{y}$ . More precisely, the similarity  $w(\mathbf{x}, \mathbf{y})$  is proportional to  $\exp\left(-\frac{1}{\beta^2} \|P_x - P_y\|^2\right)$ , where  $P_x$  (resp.  $P_y$ ) is the vectorized set of gray level intensities of the  $\mathbf{x}$ -patch (resp.  $\mathbf{y}$ -patch).

For Stokes imaging,  $K$  ( $K \geq 4$ ) measured intensities are available at each pixel location. They can be grouped in a  $K$ -component image  $\mathbf{I} = (I^{(1)}, I^{(2)}, I^{(3)}, \dots, I^{(K)})$ . Denoising  $\mathbf{I}$  can be performed for each channel independently as follows:

$$\begin{aligned} \forall \mathbf{x} \in \Omega, \quad \mathbf{I}_{nlm}(\mathbf{x}) &= \sum_{\mathbf{y} \in \Omega} \mathbf{D}_{\mathbf{x}}(\mathbf{y}) \cdot \mathbf{I}(\mathbf{y}) \\ &= \arg \min_{\mathbf{a}} \sum_{\mathbf{y} \in \Omega} (\mathbf{I}(\mathbf{y}) - \mathbf{a})^t \cdot \mathbf{D}_{\mathbf{x}}(\mathbf{y}) \cdot (\mathbf{I}(\mathbf{y}) - \mathbf{a}), \end{aligned} \quad (2)$$

where  $\mathbf{D}_{\mathbf{x}}(\mathbf{y})$  is a  $K \times K$  diagonal matrix. The diagonal element  $d_{ij}$  ( $i = 1 \dots K$ ) is the similarity (or weight) between pixel  $\mathbf{x}$  and pixel  $\mathbf{y}$  for the  $i^{th}$  channel,  $\sum_{\mathbf{y}} \mathbf{D}_{\mathbf{x}}(\mathbf{y})$  being the identity matrix. At each pixel location  $\mathbf{x}$ , the  $K$  intensity measurements  $\mathbf{I}(\mathbf{x})$  are related to the Stokes vector  $\mathbf{S}(\mathbf{x})$  by the linear equation  $\mathbf{I}(\mathbf{x}) \simeq \mathbf{P} \cdot \mathbf{S}(\mathbf{x})$ , where  $\mathbf{P}$  is the  $K \times 4$  PMM. Equation (2) becomes:

$$\forall \mathbf{x} \in \Omega, \quad \widehat{\mathbf{S}}(\mathbf{x}) = \arg \min_{\mathbf{S}} \sum_{\mathbf{y} \in \Omega} (\mathbf{I}(\mathbf{y}) - \mathbf{P} \cdot \mathbf{S})^t \cdot \mathbf{D}_{\mathbf{x}}(\mathbf{y}) \cdot (\mathbf{I}(\mathbf{y}) - \mathbf{P} \cdot \mathbf{S}). \quad (3)$$

The Stokes vector at pixel  $\mathbf{x}$  is thereby estimated using the whole set of pixels in the image, in a filtering-estimation procedure. Since the gradient of Eq. (3) is equal to the gradient of

$\|\mathbf{I}_{nlm}(\mathbf{x}) - \mathbf{P} \cdot \mathbf{S}\|^2$  (see [5]), the estimate writes:

$$\forall \mathbf{x} \in \Omega, \hat{\mathbf{S}}(\mathbf{x}) = \arg \min_{\mathbf{S}} \|\mathbf{I}_{nlm}(\mathbf{x}) - \mathbf{P} \cdot \mathbf{S}\|^2. \quad (4)$$

The estimate of Eq. (4) amounts to denoising each channel using a NLM approach, and then estimating the Stokes vector at each pixel separately using the denoised image. The criterion of Eq. (4) has a unique minimum which is the pseudo-inverse solution:  $\hat{\mathbf{V}}(\mathbf{x}) = (\mathbf{P}^t \cdot \mathbf{P})^{-1} \cdot \mathbf{P} \cdot \mathbf{I}_{nlm}(\mathbf{x})$ . However, this solution may not verify the admissibility constraints:  $\hat{\mathbf{V}}(\mathbf{x})$  may not be a Stokes vector. The following constrained criterion has been proposed:

$$\forall \mathbf{x} \in \Omega, \hat{\mathbf{S}}(\mathbf{x}) = \arg \min_{\mathbf{S} \in \mathbf{B}} \|\mathbf{I}_{nlm}(\mathbf{x}) - \mathbf{P} \cdot \mathbf{S}\|^2, \quad (5)$$

where  $\mathbf{B}$  is the set of admissible Stokes vectors. Note that optimizing the criterion of Eq. (5) is equivalent to optimizing the one of Eq. (3) under the physical admissibility constraint since both criteria are equal up to a constant.

Since the criterion of Eq. (5) is strictly convex, and since  $\mathbf{B}$  is a convex set, the criterion of Eq. (5) has a unique minimum. Instead of using a constrained optimization procedure, we use a much simpler procedure, which was defined in [5]. Convergence to the global minimum is ensured. It has been made possible thanks to the fact that the boundary  $\partial \mathbf{B}$  of  $\mathbf{B}$  can be easily parameterized.

### 3. Joint filtering-estimation of Mueller matrices

The filtering-estimation procedure which is suitable for Stokes vectors can be extended to the case of Mueller matrices provided (i) the criterion related to the estimation of the Mueller matrix can be written in the form of a weighted least squares problem (see Eq. (3) for the Stokes vector case), (ii) the Mueller matrix set is a convex set, and (iii) the boundary of the Mueller matrix set can be parameterized.

#### 3.1. Definition of the criterion

A Mueller imaging polarimeter allows the indirect measurement of the Mueller matrix of each pixel of the image. The scene is successively illuminated by  $p$  ( $p \geq 4$ ) independent polarization states through a polarization state generator (PSG). Each of these states interacts with the object, thus yielding an outgoing Stokes vector that is sensed through the polarization state analyzer (PSA) with  $q$  independent probing states ( $q \geq 4$ ). For each pixel location  $\mathbf{x}$  of the image, the  $p \times q$  intensity measurements  $\mathbf{I}$  are related to the Mueller matrix  $\mathbf{M}(\mathbf{x})$  at pixel  $\mathbf{x}$  by:

$$\mathbf{I}(\mathbf{x}) = \mathbf{A} \cdot \mathbf{M}(\mathbf{x}) \cdot \mathbf{W} (+ \text{noise}), \quad (6)$$

where  $\mathbf{A}$  and  $\mathbf{W}$  are the  $p \times 4$  PSA matrix, and the  $4 \times q$  PSG matrix, respectively. Equation (6) can be written as:

$$\begin{aligned} \underline{\mathbf{I}}(\mathbf{x}) &= (\mathbf{W}^t \otimes \mathbf{A}) \cdot \underline{\mathbf{M}}(\mathbf{x}) \quad (+ \text{noise}) \\ &= \mathbf{P} \cdot \underline{\mathbf{M}}(\mathbf{x}) \quad (+ \text{noise}), \end{aligned} \quad (7)$$

where the underline symbol is the operator which maps an  $m \times n$  into an  $mn \times 1$  vector formed by stacking up its columns, and  $\otimes$  is the Kronecker product. The known  $pq \times 16$  matrix  $\mathbf{P}$  is the PMM.

In the context of Mueller imaging, the criterion of Eq. (3) that is suitable for the Stokes case becomes:

$$\forall \mathbf{x} \in \Omega, \hat{\underline{\mathbf{M}}}(\mathbf{x}) = \arg \min_{\underline{\mathbf{M}}} \sum_{\mathbf{y} \in \Omega} (\underline{\mathbf{I}}(\mathbf{y}) - \mathbf{P} \cdot \underline{\mathbf{M}})^t \cdot \mathbf{D}_{\mathbf{x}}(\mathbf{y}) \cdot (\underline{\mathbf{I}}(\mathbf{y}) - \mathbf{P} \cdot \underline{\mathbf{M}}), \quad (8)$$

where  $\mathbf{D}_x(\mathbf{y})$  is a  $pq \times pq$  diagonal matrix,  $\sum_y \mathbf{D}_x(\mathbf{y})$  being the identity matrix. The diagonal element  $d_{ii}$  ( $i = 1 \dots pq$ ) is the similarity (or weight) between pixel  $\mathbf{x}$  and pixel  $\mathbf{y}$  for the  $i^{\text{th}}$  channel. As in the Stokes vector case, the estimate of Eq. (8) is equivalent to:

$$\forall \mathbf{x} \in \Omega, \quad \hat{\mathbf{M}}(\mathbf{x}) = \arg \min_{\mathbf{M}} \|\underline{\mathbf{I}}_{nlm}(\mathbf{x}) - \mathbf{P} \cdot \mathbf{M}\|^2, \quad (9)$$

where  $\underline{\mathbf{I}}_{nlm}$  is defined as:

$$\forall \mathbf{x} \in \Omega, \quad \underline{\mathbf{I}}_{nlm}(\mathbf{x}) = \sum_{\mathbf{y} \in \Omega} \mathbf{D}_x(\mathbf{y}) \cdot \underline{\mathbf{I}}(\mathbf{y}). \quad (10)$$

Finally, incorporating the physical admissibility constraints in Eq. (8) (or equivalently in Eq. (9)) leads to:

$$\forall \mathbf{x} \in \Omega, \quad \hat{\mathbf{M}}(\mathbf{x}) = \arg \min_{\mathbf{M} \in \mathbf{B}} \|\underline{\mathbf{I}}_{nlm}(\mathbf{x}) - \mathbf{P} \cdot \mathbf{M}\|^2, \quad (11)$$

where  $\mathbf{B}$  denotes in this part the set of admissible Mueller matrices.

The proposed criterion does not depend on the way the weights are computed. Consequently, an algorithm proposed for a one-channel image and computing the denoised value at a pixel by linearly combining pixel values of the original image (see Eq. (1)) can be extended to Mueller images using Eq. (8), or equivalently Eq. (9). Several authors have proposed strategies to determine the weights under various noise distributions (see for example [14]). Therefore, the proposed approach is well-suited for different kinds of noise distributions and may benefit from future advances in denoising methods.

### 3.2. Properties of the Mueller matrix set

There is a one-to-one correspondence between a Mueller matrix  $\mathbf{M}$  and the system coherency matrix  $\mathbf{H}$ , which reads:

$$\mathbf{M} = \mathbf{T} \cdot \mathbf{H}, \quad (12)$$

where  $\mathbf{T}$  is a  $16 \times 16$  invertible complex transformation matrix whose expression can be found for example in [4]. Given that a Mueller matrix  $\mathbf{M}$  is defined here as a matrix that verifies the Jones criterion (see Sec. 2.1),  $\mathbf{M}$  is physically acceptable *iff*  $\mathbf{H}$  is positive semidefinite ( $\mathbf{H} \in \mathcal{H}_+^4$ ) [12]. Since  $\mathcal{H}_+^4$  is a convex set, and since the image of a convex set under a linear transformation is also convex,  $\mathbf{B}$  is a convex set.

In order to extend the optimization algorithm that was proposed for Stokes vectors, we have to find a parameterization of the boundary  $\partial \mathbf{B}$  of  $\mathbf{B}$ . As  $\partial \mathbf{B}$  is the image of the boundary  $\partial \mathcal{H}_+^4$  of  $\mathcal{H}_+^4$  by the linear transformation  $\mathbf{T}$ , the key point is to define a parameterization for  $\partial \mathcal{H}_+^4$ . To this end, the following property is of great interest.

**Property 1:**  $\mathbf{H} \in \partial \mathcal{H}_+^4$  *iff*  $\mathbf{H} \in \mathcal{H}_+^4$  and  $\mathbf{H}$  is not invertible.

Property 1 is well-known in the case of the set  $\mathcal{S}_+^4$  of the symmetric positive semidefinite (real) matrices. The interior of  $\mathcal{S}_+^4$  consists of the positive definite (full-rank) matrices and all singular positive semidefinite matrices (having at least one null eigenvalue) reside on the boundary [15, p. 43]. This result can be extended to (complex) Hermitian positive semidefinite matrices because the eigenvalues of any matrix are continuous functions of the elements of the matrix [16, p. 36].

Moreover, every  $\mathbf{H} \in \mathcal{H}_+^4$  admits a Cholesky decomposition which reads:

$$\mathbf{H} = \mathbf{\Lambda} \cdot \mathbf{\Lambda}^t, \quad (13)$$

where  $\mathbf{\Lambda}^t$  is the conjugate transpose of  $\mathbf{\Lambda}$ . Matrix  $\mathbf{\Lambda}$  is a lower triangular matrix composed of 16 real parameters  $\{\lambda_i\}$ ,

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ \lambda_5 + i\lambda_6 & \lambda_2 & 0 & 0 \\ \lambda_7 + i\lambda_8 & \lambda_9 + i\lambda_{10} & \lambda_3 & 0 \\ \lambda_{11} + i\lambda_{12} & \lambda_{13} + i\lambda_{14} & \lambda_{15} + i\lambda_{16} & \lambda_4 \end{pmatrix}. \quad (14)$$

Property 1 and the fact that  $\det(\mathbf{\Lambda} \cdot \mathbf{\Lambda}^t) = (\lambda_1 \lambda_2 \lambda_3 \lambda_4)^2$  lead to property 2.

**Property 2:**  $\mathbf{\Lambda} \cdot \mathbf{\Lambda}^t \in \partial \mathcal{H}_+^4$  iff at least one value amongst  $\lambda_i$  ( $i = 1 \dots 4$ ) is null.

Property 2 enables to parameterize  $\partial \mathcal{H}_+^4$ , and consequently  $\partial \mathbf{B}$ , easily. However, the parameterization of the boundary  $\partial \mathcal{H}_+^4$  related to Property 2 is not practical for use in an optimization framework because a displacement on  $\partial \mathbf{B}$  may require to enforce a new  $\lambda_i$  ( $i = 1 \dots 4$ ) to 0, and eventually to relax the constraint for another one. A more useful property can be defined to parameterize  $\partial \mathcal{H}_+^4$ .

**Property 3:** if  $\mathbf{H} \in \partial \mathcal{H}_+^4$ , then there exists  $\mathbf{\Lambda}$  with  $\lambda_4 = 0$  such that  $\mathbf{H} = \mathbf{\Lambda} \cdot \mathbf{\Lambda}^t$ .

This property is proved in Appendix A. From Property 3 and from Eq. (12), we derive:

$$\partial \mathbf{B} = \{\mathbf{M} : \mathbf{M} = \mathbf{T} \cdot \mathbf{\Lambda} \cdot \mathbf{\Lambda}^t \text{ with } \lambda_4 = 0\}. \quad (15)$$

Since  $\partial \mathbf{B}$  can be easily parameterized (Eq. (15)), the optimization algorithm that has been proposed for Stokes vectors can be extended to Mueller matrices. Optimization of Eq. (11) is briefly described in the next section.

### 3.3. Optimization algorithm

Since the criterion of Eq. (11) is strictly convex and  $\mathbf{B}$  is convex, Eq. (11) has a unique minimum denoted  $\mathbf{M}^*(\mathbf{x})$ .

The optimization algorithm of Eq. (11) proceeds as follows. If the pseudo-inverse solution  $\widehat{\mathbf{V}}(\mathbf{x}) \in \mathbf{B}$ , then  $\mathbf{M}^*(\mathbf{x}) = \widehat{\mathbf{V}}(\mathbf{x})$ . Otherwise,  $\mathbf{M}^*(\mathbf{x})$  belongs to  $\partial \mathbf{B}$  and can be estimated by using these two strategies iteratively:

(i) local descent on the boundary  $\partial \mathbf{B}$  (see property 3):

$$\widehat{\mathbf{M}}(\mathbf{x}) = \mathbf{T} \cdot \widehat{\mathbf{\Lambda}} \cdot \widehat{\mathbf{\Lambda}}^t \text{ with } \widehat{\mathbf{\Lambda}} = \arg \min_{\mathbf{\Lambda}, \lambda_4=0} \|\mathbf{I}_{nlm}(\mathbf{x}) - \mathbf{P} \cdot \mathbf{T} \cdot \mathbf{\Lambda} \cdot \mathbf{\Lambda}^t\|^2 \quad (16)$$

(ii) local descent on the interior  $\check{\mathbf{B}}$  of  $\mathbf{B}$ .

The algorithm is initialized on  $\partial \mathbf{B}$ , with the orthogonal projection of  $\widehat{\mathbf{V}}(\mathbf{x})$  onto  $\mathbf{B}$ , and strategy (i) is considered. When a minimum is obtained, computing the gradient of criterion of Eq. (11) enables to determine if this minimum is the global one or a local one: the gradient gives the direction of the interior of  $\mathbf{B}$  iff the minimum is the global one. If this minimum is a local one, the criterion can be decreased by entering into  $\mathbf{B}$ . The descent is then continued with strategy (ii). In this strategy, the boundary is ensured to be met since there is no local minimum in  $\check{\mathbf{B}}$ . Then, strategy (i) is used again. This procedure is repeated until the global minimum is reached with strategy (i).

The optimization algorithm for strategy (i) is a subspace trust region method that is based on the interior-reflective Newton method described in [17, 18]. For strategy (ii), a simple gradient approach is used: the maximal admissible step constraining  $\hat{\mathbf{M}}(\mathbf{x})$  to be a Mueller matrix is computed at each iteration. The proposed algorithm is finally much simpler than a standard constrained optimization procedure, and it is ensured to converge to the unique minimum.

Note however that one major problem arises during the implementation of the optimization algorithm. Strategy (i) is used starting from a Mueller matrix  $\mathbf{M}$  that belongs to  $\partial\mathbf{B}$ . This matrix has been obtained either from the orthogonal projection of  $\hat{\mathbf{V}}(\mathbf{x})$  onto  $\mathbf{B}$ , or from strategy (ii) that is ensured to converge to a Mueller matrix of  $\partial\mathbf{B}$ . In both cases, starting from  $\mathbf{M} \in \partial\mathbf{B}$ , initialization of strategy (i) requires the determination of  $\mathbf{\Lambda}$  such that  $\lambda_4 = 0$  and  $\mathbf{M} = \mathbf{T} \cdot \mathbf{\Lambda} \cdot \mathbf{\Lambda}^t$ . Due to numerical problems, this computation has to be done carefully (see Appendix B).

Finally, if  $\hat{\mathbf{M}}$  denotes the estimated Mueller matrix image, the image  $\hat{\mathbf{I}}$  defined as  $\hat{\mathbf{I}}(\mathbf{x}) = \mathbf{P} \cdot \hat{\mathbf{M}}(\mathbf{x})$  for each  $\mathbf{x}$ , can be considered as the denoised version of  $\mathbf{I}$ . For  $pq = 16$ , the images  $\hat{\mathbf{I}}$  and  $\mathbf{I}_{nlm}$  differ only for the pixels for which the pseudo-inverse solution does not satisfy the admissibility constraints.

## 4. Results

### 4.1. Results on synthetic data

We synthesized a  $256 \times 256$  pixel Mueller image  $\mathbf{M}^{gt}$  composed of two distinct regions: (i) a background with a uniform polarization signature  $\mathbf{M}$ , and (ii) a 100 pixel radius circle with a smoothly varying polarized Mueller signature placed in the center of the image (see Fig. 1). The Mueller matrices of the synthetic target lie on  $\partial\mathbf{B}$  and even a small noise level may lead to non physical solutions if one uses the pseudo-inverse approach.

The Mueller matrices are defined randomly based on the  $\lambda$  parameterization with  $\lambda_4 = 0$  as follows:

- The Mueller matrix associated to the background has been defined by drawing each component of  $\mathbf{\Lambda}$  (except  $\lambda_4$  that is set to 0) according to a normal distribution of standard deviation 1.
- For the circle, 15 images of size  $256 \times 256$  (an image for each  $\lambda_i, i \neq 4$ ) are computed by drawing the value at each pixel according to a normal distribution of standard deviation 1. Each image is then filtered using an isotropic Gaussian filter of standard deviation of 5 pixels, thereby creating 15 smooth images. A constant is then added to each image (each constant is drawn according to a normal distribution of standard deviation 1).

The Mueller image  $\mathbf{M}^{gt}$  can then be computed. A standard observation model was finally used to generate intensity images  $I_i^{gt}$  ( $i = 1 \dots 16$ ) that were degraded by adding white Gaussian noise of variance  $\sigma^2$ .

Since the ground truth is known, estimation accuracy can be evaluated by comparing the estimated values (the Mueller matrices  $\hat{\mathbf{M}}$  and the associated intensity values  $\hat{\mathbf{I}}$ ) with the original ones ( $\mathbf{M}^{gt}$  and  $\mathbf{I}^{gt}$ ). The method is first evaluated by comparing the original image  $\mathbf{I}^{gt}$  (noise-free intensity image) with its estimation  $\hat{\mathbf{I}}$  using the Peak Signal-to-Noise Ratio (PSNR):

$$PSNR(\mathbf{I}^{gt}, \hat{\mathbf{I}}) = 10 \log_{10} \left( \frac{d^2}{\frac{1}{16 \cdot P} \sum_{j=1}^{16} \alpha_j^2 \sum_{\mathbf{x}} \left( I_j^{gt}(\mathbf{x}) - \hat{I}_j(\mathbf{x}) \right)^2} \right), \quad (17)$$

where  $\alpha_j$  is computed so that the dynamic of  $\alpha_j \cdot I_j^{gt}$  is  $d$  (for example 255), and where  $P$  is the number of pixels. Complementary to the PSNR, a Mueller matrix dedicated measurement is

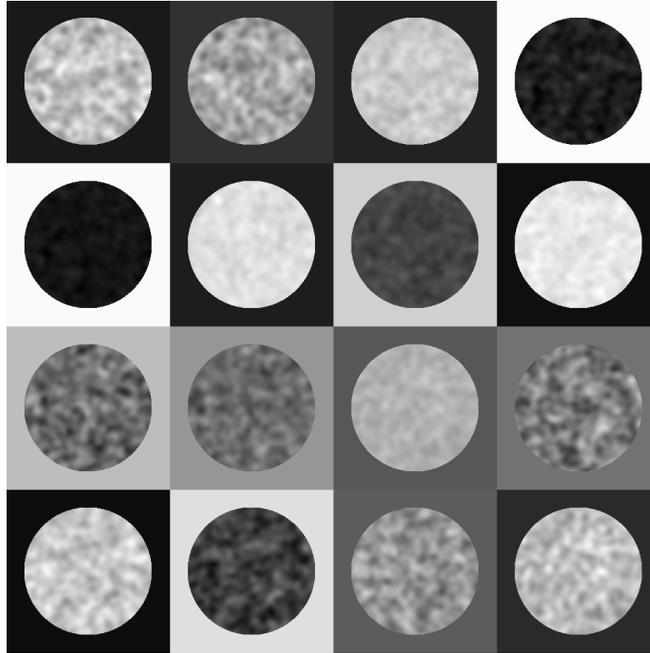


Fig. 1. Mueller image of the synthetic scene. For convenience, 16 images (one for each channel of the Mueller matrix) are represented on a  $4 \times 4$  grid. Row  $i$  column  $j$  image corresponds to the  $m_{ij}$  Mueller matrix element image.

also used to evaluate the estimation accuracy:

$$e(\mathbf{M}^{gt}, \hat{\mathbf{M}}) = 100 \sqrt{\frac{1}{P} \sum_{\mathbf{x}} \frac{\|\mathbf{M}^{gt}(\mathbf{x}) - \hat{\mathbf{M}}(\mathbf{x})\|^2}{\|\mathbf{M}^{gt}(\mathbf{x})\|^2}}. \quad (18)$$

Four methods,  $M_{PI}$ ,  $M_{PIortho}$ ,  $M_{PIproj}$ , and  $M_P$  are evaluated:

- for  $M_{PI}$ :  $\hat{\mathbf{M}}(\mathbf{x})$  is the pseudo-inverse solution;
- for  $M_{PIortho}$ :  $\hat{\mathbf{M}}(\mathbf{x})$  is the pseudo-inverse solution, further orthogonally projected onto the Mueller matrix set if it does not verify the admissibility constraints;
- for  $M_{PIproj}$ :  $\hat{\mathbf{M}}(\mathbf{x})$  is the pseudo-inverse solution, further projected onto the Mueller matrix set if it does not verify the admissibility constraints (the projection is performed with the proposed optimization algorithm so as to reduce at most the error reconstruction between the observed measurements and the predicted ones);
- $M_P$  is the proposed approach.

Note that the methods  $M_{PI}$ ,  $M_{PIortho}$ ,  $M_{PIproj}$  do not use any spatial filtering.

The PSNR (Eq. (17)) and the Mueller matrix estimation error (Eq. (18)) corresponding to  $M_{PI}$ ,  $M_{PIortho}$ ,  $M_{PIproj}$  and  $M_P$  are given in Tab. 1 for different values of  $\sigma$  (from  $\sigma = 0.05$  to 1).

Table 1 shows that the proposed approach outperforms the other methods in the particular context of this experiment. This illustrates the benefit of accounting for spatial information for

Table 1. PSNR (left) and Mueller matrix estimation error (right) obtained with the four different methods, and for different values of  $\sigma$ . Bold values correspond to the best results.

$\sigma$	$M_{PI}$	$M_{PIortho}$	$M_{PIproj}$	$M_P$	$\sigma$	$M_{PI}$	$M_{PIortho}$	$M_{PIproj}$	$M_P$
0.05	33.09	32.62	33.44	<b>45.35</b>	0.05	2.75	2.71	2.57	<b>0.65</b>
0.1	27.05	26.28	27.50	<b>41.50</b>	0.1	5.50	5.36	4.93	<b>0.89</b>
0.5	13.08	10.66	14.19	<b>30.61</b>	0.5	27.48	24.83	19.34	<b>3.06</b>
1	7.04	3.80	8.78	<b>25.67</b>	1	54.94	46.97	32.71	<b>5.67</b>

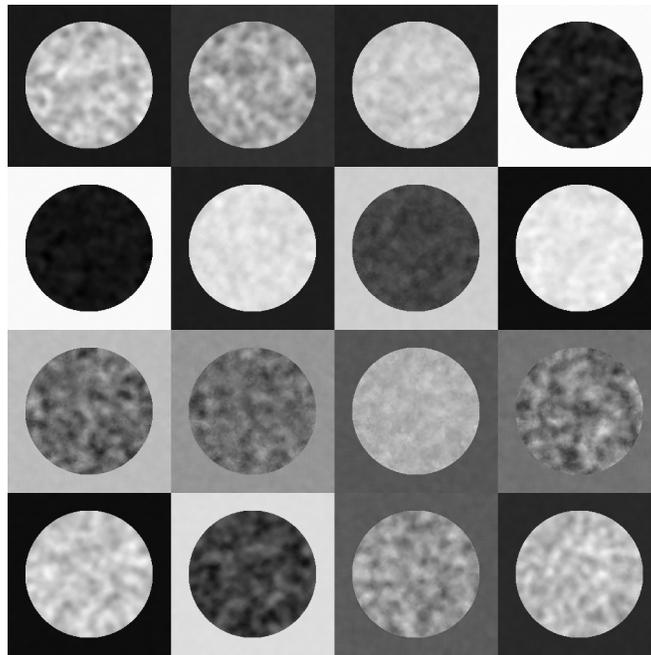
the estimation, and in particular, the efficiency of the NLM approach. The PSNR (left part of Tab. 1) obtained with the proposed approach is increased of at least 10 dB, and the estimation error (right part of Tab. 1) is decreased of a factor varying approximately from 4 to 10. The same conclusion can be drawn from Fig. 2 which presents the Mueller matrix images obtained from the noisy images ( $\sigma = 0.1$ ) with the NLM approach (a), and with  $M_{PI}$  (b). Since results obtained with the methods  $M_{PIproj}$  and  $M_{PIortho}$  are visually very similar to Fig. 2(b), they are not presented here. The channels associated to the third line of the Mueller matrix are for example noisy with  $M_{PI}$  but not with the NLM approach. Besides performing noise reduction, we can observe that edges have been preserved. This nice property is inherited from NLM filtering: NLM filtering preserves sharp edges and fine texture details [6].

Finally, amongst the methods which do not consider spatial information,  $M_{PIproj}$  is the one providing the best results. This shows that the traditional way of projecting onto the Mueller matrix set (method  $M_{PIortho}$ , orthogonal projection) is not an efficient approach, compared to the proposed method which reduces the most the error reconstruction between the observed measurements and the predicted ones, while constraining the solution to be physically acceptable. Note also that  $M_{PIortho}$  provides better results than  $M_{PI}$  in terms of accuracy of the Mueller matrix estimation, but less satisfactory results in terms of PSNR. This highlights the fact that the choice of orthogonal projection is arbitrary.

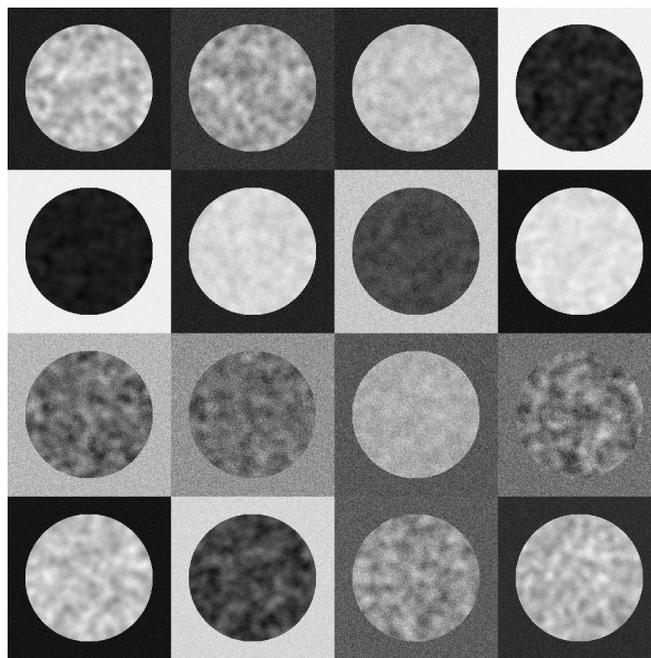
#### 4.2. Results on real data

To evaluate the potential of the method to handle real images, we used a well-calibrated Mueller polarimeter [1] to image two real scenes with different kinds of polarization signatures: (i) the first scene (Fig. 3) corresponds to two shapes made from two different transparent thin layers (cellophane for wrapping food) sandwiched between two glass sections. This object is expected to show class-wise constant polarization responses; (ii) the second one consists of an ensemble of objects (Fig. 4) leading to an image with various polarization responses and complex geometrical properties. Observations were carried out through a narrow band interferential filter to ensure that the PMM is known with high accuracy. The Mueller image was reconstructed from the raw intensity images using  $M_P$  (the proposed approach) and  $M_{PIortho}$  (the classical pseudo-inverse approach except that the solution is further projected orthogonally onto the Mueller matrix set if it does not verify the admissibility constraints).

Figure 3 and Fig. 4 present the results for the first scene and second scene respectively. For both scenes, the images reconstructed by our approach are less noisy than the pseudo-inverse images. As an example, for the first scene, the information content of some of the Mueller channels are lost in the projected pseudo-inverse solution (e.g.  $m_{11}$ ,  $m_{12}$ ,  $m_{21}$ , and  $m_{22}$  element images). In particular, a lot of fine details are definitely lost (channels  $m_{12}$  and  $m_{21}$ ). All channels estimated with  $M_{PIortho}$  are particularly noisy for the second scene which is not the case with the proposed approach. For the sake of conciseness, only channels  $m_{22}$  (Fig. 4(a)) and  $m_{44}$  (Fig. 4(c)) are presented. The  $M_{PIortho}$  approach propagates intensity noises to the Mueller channels leading to less workable Mueller images. Results obtained with the second

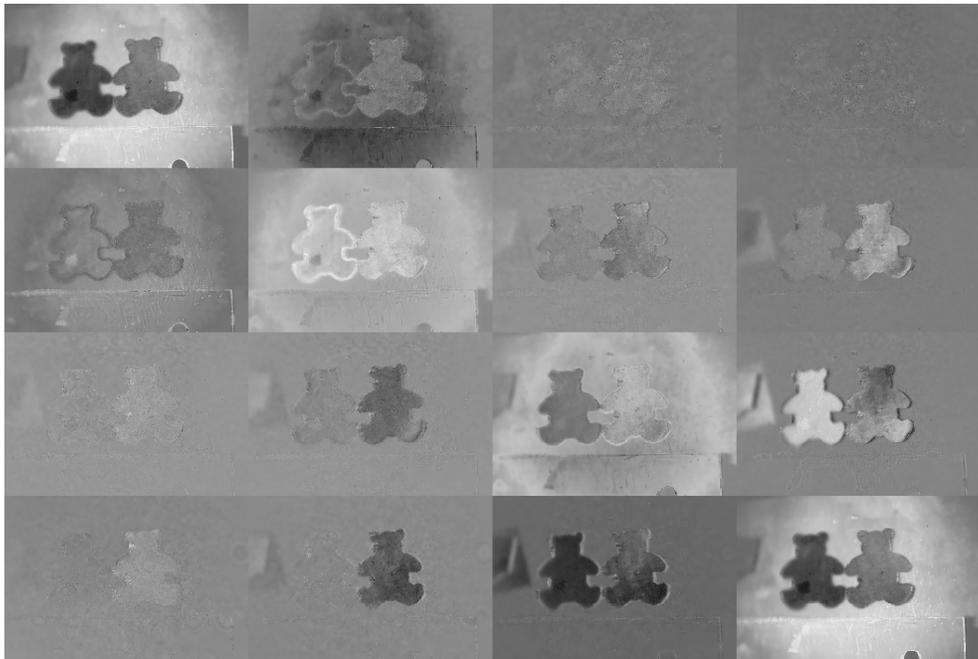


(a)

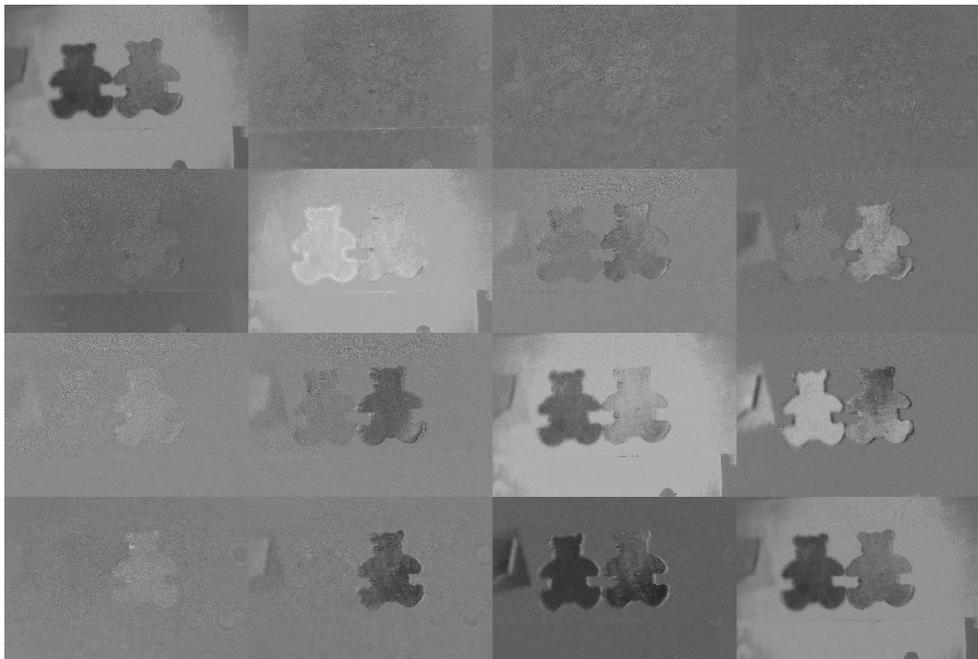


(b)

Fig. 2. Mueller image of the synthetic scene estimated with the proposed (NLM) approach (a) and pseudo-inverse (b) from the noisy intensity images ( $\sigma = 0.1$ ). Data are presented with the convention used in Fig. 1.

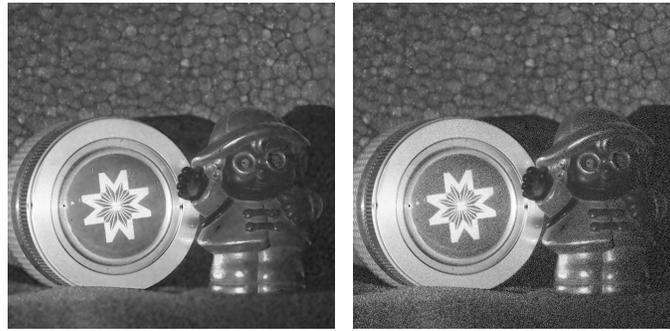


(a)

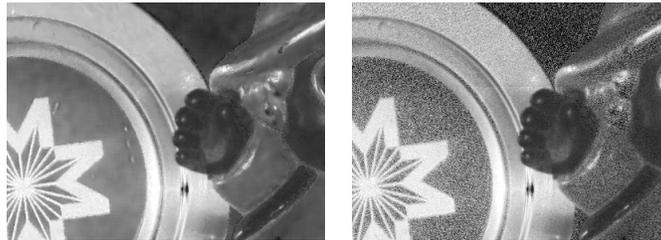


(b)

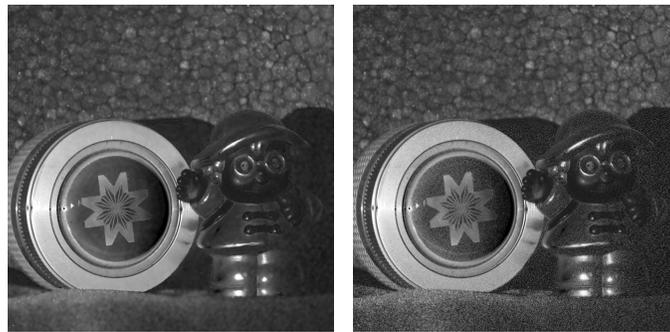
Fig. 3. Mueller image of the first scene estimated with the proposed (NLM) approach (a) and projected pseudo-inverse (b). Data are presented with the convention used in Fig. 1 except that all channels but  $m_{11}$  have been pixelwise normalized with respect to  $m_{11}$  ( $m_{ij}(\mathbf{x}) = m_{ij}(\mathbf{x})/m_{11}(\mathbf{x})$ ).



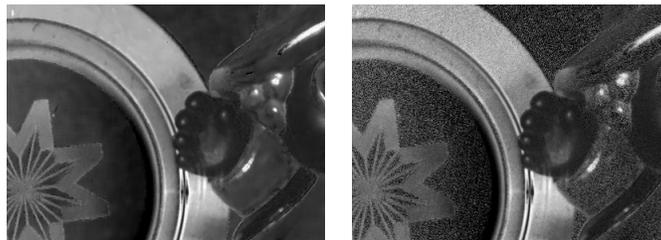
(a)



(b)



(c)



(d)

Fig. 4. Mueller image of the second scene estimated with the proposed (NLM) approach (left column) and projected pseudo-inverse (right column). For convenience, only channels  $m_{22}$  (a) and  $m_{44}$  (c) are presented. Images (b) and (d) correspond to a zoom in of (a) and (c), respectively.

scene illustrate also the efficiency of the NLM filtering approach to preserve edges and thin structures (see Fig. 4(b) and (d)).

Since the NLM solution is much less noisy than the projected pseudo-inverse one, it provides also a better estimation of physical characteristics (e.g. diattenuation, depolarization). Figure 5 shows the mean diattenuation calculated from the  $M_P$  and the  $M_{PIortho}$  solutions for the first scene. The poor quality of the result obtained with the  $M_{PIortho}$  approach can be explained by the fact that the upper left  $2 \times 2$  block of the Mueller matrix is in this case particularly noisy (see Fig. 3(b)).

More generally, we found that, for all cases considered here, our approach performed better than the pseudo-inverse solution that is widely used in Mueller data reduction phase. Our approach is fully unsupervised and requires no parameter tuning.

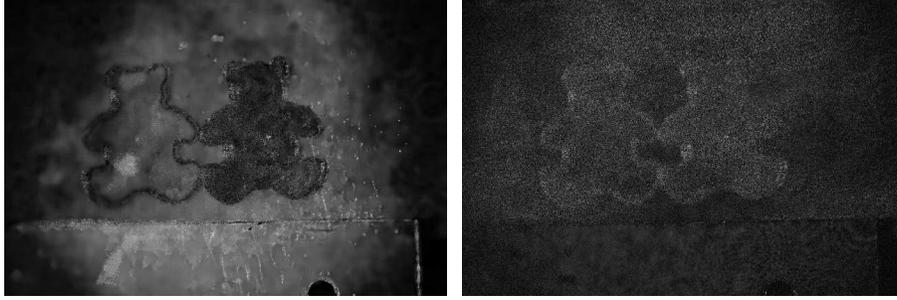


Fig. 5. Mean diattenuation of the first scene obtained by the polar decomposition applied to the NLM (left column) and to the  $M_{PIortho}$  solutions (right column).

## 5. Conclusion

We introduced a new reconstruction-estimation method that allows obtaining Mueller images from raw measured intensities. The proposed approach performs better than state-of-the-art methods, while yielding admissible Mueller signatures. The major interest of this study comes from the association of up-to-date denoising algorithms with physical admissibility criteria of Mueller channels that allow better exploitation and interpretation of the final images. This approach turns out to be of great benefit since the reconstructed signature images are little affected by measurement noise while sharp transitions and thin structures remain preserved.

### A. Characterization of the boundary $\partial \mathcal{H}_+^4$ of the set of H-type matrices

The objective of this appendix is to prove property 3 (see page 6), which reads: if  $\mathbf{H} \in \partial \mathcal{H}_+^4$ , then there exists  $\mathbf{\Lambda}$  with  $\lambda_4 = 0$  such that  $\mathbf{H} = \mathbf{\Lambda} \cdot \mathbf{\Lambda}^t$ .

Let us write

$$\mathbf{H} = \begin{pmatrix} h_1 & h_5 - ih_6 & h_7 - ih_8 & h_{11} + ih_{12} \\ h_5 + ih_6 & h_2 & h_9 - ih_{10} & h_{13} - ih_{14} \\ h_7 + ih_8 & h_9 + ih_{10} & h_3 & h_{15} - ih_{16} \\ h_{11} - ih_{12} & h_{13} + ih_{14} & h_{15} + ih_{16} & h_4 \end{pmatrix}. \quad (19)$$

By substituting Eq. (14) and Eq. (19) into Eq. (13), we get a system of equations in  $\lambda$ 's that has at least one solution if  $\mathbf{H} \in \mathcal{H}_+^4$ . In practice, there are several solutions when  $\mathbf{H} \in \mathcal{H}_+^4$ . As an example, if  $\mathbf{H} \in \mathcal{H}_+^4$  and if  $\mathbf{H}$  is invertible, then the signs of  $\lambda_1, \lambda_2, \lambda_3$ , and  $\lambda_4$  can be set arbitrarily leading to  $2^4$  different solutions. In the general case where  $\mathbf{H} \in \mathcal{H}_+^4$ , one possible

solution in  $\lambda$ 's writes:

$$\left\{ \begin{array}{ll} \lambda_1 = \sqrt{h_1} & \\ \lambda_i = \frac{h_i}{\lambda_1} \text{ (for } i = 5 \text{ to } 8 \text{ and for } i = 11 \text{ to } 12) & \text{if } \lambda_1 > 0, \text{ and } 0 \text{ otherwise} \\ \lambda_2 = \sqrt{h_2 - \lambda_5^2 - \lambda_6^2} & \\ \lambda_9 = \frac{1}{\lambda_2} (h_9 - \lambda_5 \lambda_7 - \lambda_6 \lambda_8) & \text{if } \lambda_2 > 0, \text{ and } 0 \text{ otherwise} \\ \lambda_{10} = \frac{1}{\lambda_2} (h_{10} - \lambda_6 \lambda_7 - \lambda_5 \lambda_8) & \text{if } \lambda_2 > 0, \text{ and } 0 \text{ otherwise} \\ \lambda_{13} = \frac{1}{\lambda_2} (h_{13} - \lambda_5 \lambda_{11} - \lambda_6 \lambda_{12}) & \text{if } \lambda_2 > 0, \text{ and } 0 \text{ otherwise} \\ \lambda_{14} = \frac{1}{\lambda_2} (h_{14} + \lambda_6 \lambda_{11} - \lambda_5 \lambda_{12}) & \text{if } \lambda_2 > 0, \text{ and } 0 \text{ otherwise} \\ \lambda_3 = \sqrt{h_3 - \lambda_7^2 - \lambda_8^2 - \lambda_9^2 - \lambda_{10}^2} & \\ \lambda_{15} = \frac{1}{\lambda_3} (h_{15} - \lambda_7 \lambda_{11} - \lambda_8 \lambda_{12} - \lambda_9 \lambda_{13} - \lambda_{10} \lambda_{14}) & \text{if } \lambda_3 > 0, \text{ and } 0 \text{ otherwise} \\ \lambda_{16} = \frac{1}{\lambda_3} (h_{16} + \lambda_8 \lambda_{11} - \lambda_7 \lambda_{12} + \lambda_{10} \lambda_{13} - \lambda_9 \lambda_{14}) & \text{if } \lambda_3 > 0, \text{ and } 0 \text{ otherwise} \\ \lambda_4 = \sqrt{h_4 - \lambda_{11}^2 - \lambda_{12}^2 - \lambda_{13}^2 - \lambda_{14}^2 - \lambda_{15}^2 - \lambda_{16}^2} & \end{array} \right. \quad (20)$$

If  $\mathbf{H}$  is invertible ( $\lambda_i > 0$  for  $i = 1 \dots 4$ ), Eq. (20) can be either obtained by solving the system of equations in  $\lambda$ 's directly or by using the algorithm of [19, p. 145]. In [19], the algorithm is described for symmetric positive definite (real) matrices but the extension to (complex) Hermitian positive definite matrices is straightforward.

If  $\mathbf{H}$  is not invertible (i.e.  $\exists i \in [1, 4]$  such that  $\lambda_i = 0$ ), the  $i$ -th column of  $\mathbf{A}$  is set to 0 (see Eq. (20)). This is justified in [19, p. 148] for real matrices, but the same reasoning applies to complex ones.

We suppose that  $\mathbf{H} \in \partial \mathcal{H}_+^4$ , which implies that  $\lambda_1 = 0$ , or  $\lambda_2 = 0$ , or  $\lambda_3 = 0$ , or  $\lambda_4 = 0$ . In the following,  $\mathbf{A}$  denotes the lower triangular matrix whose elements are defined in Eq. (20). By construction,  $\mathbf{A}$  verifies  $\mathbf{H} = \mathbf{A} \cdot \mathbf{A}^t$  but  $\lambda_4$  may be greater than 0. To obtain a solution with  $\lambda_4 = 0$ , an equivalence class is defined on the set of the lower triangular matrices (see Eq. (14)):  $\mathbf{A}_a$  and  $\mathbf{A}_b$  are equivalent if there exists an orthonormal matrix  $\mathbf{R}$  such that  $\mathbf{A}_a = \mathbf{A}_b \cdot \mathbf{R}$ . Matrices  $\mathbf{A}_a$  and  $\mathbf{A}_b$  correspond then to the same underlying coherency matrix  $\mathbf{A}_a \cdot \mathbf{A}_a^t = \mathbf{A}_b \cdot \mathbf{A}_b^t$ , and consequently to the same Mueller matrix. If  $\mathbf{H} \in \partial \mathcal{H}_+^4$ , we show hereafter that we can find in the class of  $\mathbf{A}$  a matrix with  $\lambda_4 = 0$ . Three cases have to be considered, depending on which  $\lambda_i$  vanishes:

1. if  $\mathbf{H}$  is such that  $\lambda_1 = 0$  then from Eq. (20), we have:

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & \lambda_9 + i\lambda_{10} & \lambda_3 & 0 \\ 0 & \lambda_{13} + i\lambda_{14} & \lambda_{15} + i\lambda_{16} & \lambda_4 \end{pmatrix}, \quad (21)$$

with  $\mathbf{H} = \mathbf{A} \cdot \mathbf{A}^t$ . A matrix equivalent to  $\mathbf{A}$  can be derived from  $\mathbf{A}$  by swapping its first and last columns. This can be formally defined as follows:  $\mathbf{A}$  is equivalent to  $\mathbf{A} \cdot \mathbf{R}$ , with

$$\mathbf{R} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \quad (22)$$

In the class of  $\mathbf{A}$ , we thereby have a matrix  $(\mathbf{A} \cdot \mathbf{R})$  with  $\lambda_4 = 0$ . This means that  $\mathbf{A} \cdot \mathbf{R}$  is a lower triangular matrix with  $\lambda_4 = 0$  that verifies  $\mathbf{H} = (\mathbf{A} \cdot \mathbf{R}) \cdot (\mathbf{A} \cdot \mathbf{R})^t$ ;

2. if  $\mathbf{H}$  is such that  $\lambda_2 = 0$  in Eq. (20), then the second column of  $\mathbf{A}$  vanishes and the reasoning above applies with

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}; \quad (23)$$

3. if  $\mathbf{H}$  is such that  $\lambda_3 = 0$  in Eq. (20), then the third column of  $\mathbf{A}$  vanishes and the reasoning above applies with

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (24)$$

This proves that every  $\mathbf{H} \in \partial \mathcal{H}_+^4$  admits a decomposition  $\mathbf{H} = \mathbf{A} \cdot \mathbf{A}^t$ , with  $\lambda_4 = 0$ .

Note that when  $\lambda_i = 0$  ( $i = 1 \dots 3$ ), the elements of the  $i$ -th column cannot be uniquely determined. In this case, these elements have been set to 0 (see Eq. (20)). This has provided us a way to easily find a matrix equivalent to  $\mathbf{A}$  that verifies  $\lambda_4 = 0$ . If the undetermined elements had not been set to 0, it would not have been possible to use such a simple procedure, because swapping columns would have led to a non-triangular matrix.

## B. Algorithm for the estimation of $\mathbf{A}$ from $\mathbf{M} \in \partial \mathbf{B}$

Let  $\mathbf{M}$  be a Mueller matrix that belongs to the boundary  $\partial \mathbf{B}$  of the Mueller matrix set. Let  $\mathbf{H}$  be the associated coherency matrix ( $\mathbf{H} \in \partial \mathcal{H}_+^4$ ). It has been shown in Appendix A that there exists  $\mathbf{A}$  with  $\lambda_4 = 0$  such that  $\mathbf{H} = \mathbf{A} \cdot \mathbf{A}^t$ . The goal of this Appendix is to explain how to determine  $\mathbf{A}$ . At first sight, an algorithm can be easily derived from Appendix A:

1. Determine  $\mathbf{A}$  using Eq. (20).
2. Let  $i_0$  be an integer in  $[1, 4]$  verifying  $\lambda_{i_0} = 0$ . Such an integer exists since at least one value of  $\lambda_i$  is null ( $i = 1 \dots 4$ ).
3. Swap the  $i_0$ -th column of  $\mathbf{A}$  with its last one to obtain the desired result.

However, such an algorithm cannot be used to properly estimate  $\mathbf{A}$ . Indeed, due to numerical problems, it may happen that  $\lambda_1, \lambda_2, \lambda_3$  and  $\lambda_4$  are all strictly positive. We could then define a threshold beyond which the values of  $\lambda_i$  would be set to 0 ( $i = 1 \dots 4$ ). However, there is in this case a risk that  $\lambda_i$  is set to 0 while it should not be, thereby leading to a matrix  $\mathbf{A}$  with  $\mathbf{A} \cdot \mathbf{A}^t$  highly different from  $\mathbf{H}$ .

To solve this problem, several matrices  $\mathbf{A}$  are computed by setting some values of  $\lambda_i$  to 0 ( $i = 1 \dots 4$ ). This can be formally described as follows. Let  $\mathcal{A}$  denote a combination of  $p$  elements of  $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$  ( $1 \leq p \leq 4$ ). Each  $\lambda_i$  of  $\mathcal{A}$  is set to 0, and the other values of  $\mathbf{A}$  are computed from Eq. (20). All combinations are tested, leading to 15 different matrices  $\mathbf{A}$ . The matrix  $\mathbf{A}$  that minimizes the quadratic error between  $\mathbf{H}$  and  $\mathbf{A} \cdot \mathbf{A}^t$  is chosen. The equivalence class defined in Appendix A enables finally to derive a lower triangular matrix  $\mathbf{A}_0$  from  $\mathbf{A}$ , with  $\lambda_4 = 0$  that verifies  $\mathbf{H} \simeq \mathbf{A}_0 \cdot \mathbf{A}_0^t$ .

## Acknowledgment

Giorgos Sfikas was supported by a grant from Région Alsace (France).