

Noise minimization and equalization for Stokes polarimeters in the presence of signal-dependent Poisson shot noise

François Goudail

Laboratoire Charles Fabry de l'Institut d'Optique, CNRS, Univ. Paris-Sud, RD. 128, 91127 Palaiseau, France
(francois.goudail@institutoptique.fr)

Received November 6, 2008; revised January 14, 2009; accepted January 14, 2009;
posted January 23, 2009 (Doc. ID 103697); published February 25, 2009

We address the optimization of Stokes polarimeters in the presence of signal-dependent shot noise, which is the dominant type of noise in certain imaging systems. We show that in some precise sense, the polarimeters optimal for additive noise are also optimal for such noise and propose polarimeter architectures in which noise variances are equalized and independent of the input polarization state. © 2009 Optical Society of America

OCIS codes: 260.5430, 030.4280.

Active Stokes imaging consists of illuminating a scene with polarized light and measuring the Stokes vector of the light scattered by the scene. The design of Stokes imagers that minimize and/or equalize the noise power in the different Stokes channels has been widely studied [1–6]. It is usually assumed that the noise that affects the image is additive and independent of the level of the signal. However, in many cases, for example, photon counting systems or quantum detectors with a sufficient level of light, the shot noise due to the useful signal is dominant compared to the signal independent detector noise. It is thus important to determine which are the optimal Stokes polarimeter structures in the presence of signal dependent shot noise. This problem has already been addressed in [7,8] with different approaches.

We consider Stokes polarimeters that perform N intensity measurements to estimate the Stokes vector. Let us denote $\vec{S}=(S_0, S_1, S_2, S_3)^T$ the four-dimensional Stokes vector to estimate and $\vec{I}=(I_1, \dots, I_N)^T$ the N -dimensional vector representing the intensity measurements. The number of intensity measurement must be $N \geq 4$ for the estimation of \vec{S} to be possible. In the absence of noise, one has the relation

$$\vec{I} = W\vec{S}, \quad (1)$$

where W is the $N \times 4$ measurement (or synthesis) matrix. In its most general form, the i th line of W , with $i \in [1, N]$, is the first line of the Mueller matrix of a perfect polarizer whose totally transmitted polarization state is defined by the reduced Stokes vector \vec{d}_i , where \vec{d}_i is the three-dimensional unit-norm vector. This line thus has the following form: $(1, \vec{d}_i^T)/2$ [4]. In practice, polarimeter architectures are such that the degrees of freedom for choosing the vectors \vec{d}_i are limited. For example, variable retarder systems consist of two fixed retarders with variable delays and a fixed polarizer [4]. Rotating retarder fixed polarizer (RRFP) systems [3] consist of a fixed polarizer and a retarder with a fixed delay whose angle is variable.

In the presence of Poisson shot noise, \vec{I} is a random vector such that each of its elements I_i , $i \in [1, N]$ is a Poisson random variable of mean value $\langle I_i \rangle = \sum_{j=0}^3 W_{ij} S_j$, where $\langle \cdot \rangle$ denotes an ensemble average. The probability law of a Poisson random variable of mean λ is $P(n) = \exp[-\lambda] \lambda^n / n!$. Its mean and its variance are both equal to λ . In a vector form, one thus has $\langle \vec{I} \rangle = \text{VAR}[\vec{I}] = W\vec{S}$, where $\text{VAR}[\cdot]$ denotes the variance of a random vector. To estimate \vec{S} from \vec{I} , we will use the following estimator:

$$\hat{S} = W^\dagger \vec{I}, \quad (2)$$

where $W^\dagger = (W^T W)^{-1} W^T$ (T denotes matrix transposition) is the pseudoinverse of W . If $N=4$, W^\dagger is the classical inverse W^{-1} and is obviously the optimal estimate of \vec{S} . If $N > 4$, it is known that the pseudoinverse matrix is the maximum-likelihood estimator of \vec{S} in the presence of additive Gaussian noise [9]. In the presence of Poisson noise, it is a simple and closed-form algorithm that gives good results in practice. We shall thus use it in the following.

It is clear that \hat{S} is an unbiased estimator, since $\langle \hat{S} \rangle = W^\dagger \langle \vec{I} \rangle = \vec{S}$. Its covariance matrix has the following expression: $\Gamma^{\hat{S}} = W^\dagger \Gamma^I [W^\dagger]^T$, where Γ^I is the covariance matrix of \vec{I} . From the properties of shot noise, the fluctuations are statistically independent from one intensity measurement to the other. The covariance matrix Γ^I of \vec{I} is thus a diagonal matrix whose i th element is the variance of the i th intensity measurement

$$\Gamma_{ij}^I = \begin{cases} \langle I_i \rangle = \sum_{k=0}^3 W_{ik} S_k, & \text{if } i=j \\ 0, & \text{otherwise} \end{cases}.$$

One thus has

$$\Gamma_{ij}^{\hat{S}} = \sum_{k=0}^3 S_k \sum_{n=1}^N W_{in}^{\dagger} W_{jn}^{\dagger} W_{nk}. \quad (3)$$

The variances on each component of \hat{S} are

$$\gamma_i = \Gamma_{ii}^{\hat{S}} = \sum_{k=0}^3 S_k \sum_{n=1}^N [W_{in}^{\dagger}]^2 W_{nk} = \sum_{k=0}^3 Q_{ik} S_k, \quad (4)$$

where the matrix Q is defined as

$$Q_{ij} = \sum_{n=1}^N [W_{in}^{\dagger}]^2 W_{nj}. \quad (5)$$

A standard performance criterion for a Stokes polarimeter is the total variance, which is the trace of $\Gamma^{\hat{S}}$

$$F(W, \vec{S}) = \text{trace}[\Gamma^{\hat{S}}] = \sum_{i=0}^3 \gamma_i = \sum_{k=0}^3 q_k S_k = \vec{q} \cdot \vec{S}, \quad (6)$$

where the vector \vec{q} is defined by $q_k = \sum_{i=0}^3 Q_{ik}$ and \cdot denotes the scalar product. Contrary to what happens in the presence of additive noise [3,4], this criterion does depend on the actual value of the Stokes vector. A possible way of defining a unique performance criterion is to consider a minimax optimization, that is, minimize with respect to W the following function:

$$\mathcal{F}(W) = \max_{\vec{S}} [F(W, \vec{S})]. \quad (7)$$

Any Stokes vector \vec{S} can be written as $\vec{S} = I_0(1, P\vec{s}^T)^T$, where I_0 is the intensity and P is the degree of polarization, which belongs to the interval $[0, 1]$. The normalized Stokes vector \vec{s} is three-dimensional and unit norm. It represents the principal state of polarization of \vec{S} . Using this notation, Eq. (6) becomes

$$F(W, \vec{S}) = I_0(q_0 + P\vec{v} \cdot \vec{s}), \quad (8)$$

where $\vec{v} = (q_1, q_2, q_3)^T$. Let us consider the two terms of this expression. The first one, $I_0 q_0$, is independent of \vec{s} . One has $q_0 = \sum_{n=1}^N W_{n0} \sum_{i=0}^3 [W_{in}^{\dagger}]^2$. Taking into account that $\forall_n, W_{n0} = 1/2$, one has

$$q_0 = \frac{1}{2} \text{trace}[(W^T W)^{-1}] = F_{\text{add}}(W). \quad (9)$$

The function $F_{\text{add}}(W)$ is well known in polarimeter optimization. Indeed, if the noise is signal independent and additive, the trace of the covariance matrix of \hat{S} is proportional to $F_{\text{add}}(W)$ [3,4]. The second term of Eq. (8), $I_0 P \vec{v} \cdot \vec{s}$, depends on \vec{s} and P . For a given intensity I_0 , it is obviously maximized by $P=1$ (totally polarized light) and $\vec{s}_{\text{max}} = \vec{v}/\|\vec{v}\|$. The function to minimize with respect to W is thus

$$\mathcal{F}(W) = I_0(F_{\text{add}}(W) + \|\vec{v}\|), \quad (10)$$

with $\|\vec{v}\| = (\sum_{i=1}^3 q_i)^{1/2}$. Both terms of the sum are positive.

From now on, we shall consider the particular case where $N=4$ measurements are performed and thus $W^{\dagger} = W^{-1}$. In this case, it has been shown that the ma-

trices W_{opt} that minimize $F_{\text{add}}(W)$ are such that the points defined by the vectors \vec{d}_i of each row of matrix \vec{W} form a regular tetrahedron [3,4,6]. If W has such a structure, one has the two following properties: $\forall j > 0, \sum_{i=0}^3 W_{ij} = 0$ and $\forall j, \sum_{i=0}^3 [W_{ij}^{-1}]^2 = A$. These two properties lead to $\forall_k, q_k = 0$. Consequently, when W has this peculiar structure, $\|\vec{v}\| = 0$, it is obviously the minimal possible value of $\|\vec{v}\|$. This demonstrates that the polarimeter structures that are optimal for additive signal independent noise are also optimal for signal dependent Poisson noise in the sense of the minimax criterion defined in Eq. (10). Indeed, if W has such a structure, F_{add} is minimal and $\|\vec{v}\| = 0$. A second interesting result is that since $\vec{v} = 0$, the total variance $F(W, \vec{S}) = \sum_{i=0}^3 \gamma_i$ is independent of the input principal polarization state \vec{s} [see Eq. (8)]. However, it must be noted that contrary to the case of additive noise, the variances γ_i on each Stokes channel may vary with \vec{s} . Indeed, Eq. (4) yields

$$\gamma_i = \sum_{j=0}^3 Q_{ij} S_j = I_0(Q_{i0} + P\vec{u}^i \cdot \vec{s}), \quad (11)$$

with $\vec{u}^i = (Q_{i1}, Q_{i2}, Q_{i3})^T$. The polarization state that maximizes γ_i is $\vec{s}_{\text{max}}^i = \vec{u}^i / \|\vec{u}^i\|$ and that which minimizes it is $\vec{s}_{\text{min}}^i = -\vec{s}_{\text{max}}^i$, which represents the state orthogonal to \vec{s}_{max}^i . The maximal and minimal values of the variance are thus given by

$$\gamma_i^{\text{max}} = I_0(Q_{i0} + P\|\vec{u}^i\|), \quad \gamma_i^{\text{min}} = I_0(Q_{i0} - P\|\vec{u}^i\|). \quad (12)$$

If P is close to zero, the variances become independent of the polarization state. The situation is then equivalent to signal independent additive noise whose variance is controlled by the mean value of S_0 , that is, $I_0 Q_{i0}$, as was already noticed in [3]. In the general case however, the estimation variance of a given Stokes parameter depends on the true values of all Stokes parameters.

If W has the regular tetrahedron structure, it has the following properties:

- $\forall i \in [0, 3], W_{0i}^{-1} = W_{i0}$,
- $\forall i \in [1, 3], j \in [0, 3], W_{ji}^{-1} = 3W_{ji}$.

Consequently, one has $u_j^i = 9 \sum_{k=0}^3 [W_{ki}]^2 W_{kj}$, where u_j^i denotes the j th coordinate of vector \vec{u}^i , and the matrix Q has the following structure:

$$Q = \begin{bmatrix} 1/2 & 0 & 0 & 0 \\ 3/2 & u_1^1 & u_2^1 & u_3^1 \\ 3/2 & u_1^2 & u_2^2 & u_3^2 \\ 3/2 & u_1^3 & u_2^3 & u_3^3 \end{bmatrix}, \quad (13)$$

and thus one always has $q_0 = 5$ and $\gamma_0 = I_0/2$ independently of \vec{s} .

For illustration purposes, let us consider the two following measurement matrices:

$$W_a = \begin{bmatrix} 0.5 & 0.2895 & 0.3624 & -0.1867 \\ 0.5 & 0.2895 & -0.3624 & 0.1867 \\ 0.5 & -0.2895 & -0.1885 & -0.3615 \\ 0.5 & -0.2895 & 0.1885 & 0.3615 \end{bmatrix},$$

$$W_b = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & -0.1667 & -0.4080 & 0.2355 \\ 0.5 & -0.1667 & 0 & -0.4710 \\ 0.5 & -0.1667 & 0.4080 & 0.2355 \end{bmatrix}.$$

The first one is an optimal RRF configuration as given by Sabatke *et al.* [3]. It consists of a plate of retardation $\delta=132^\circ$ that is rotated by four different angles $\pm 51.7^\circ$ and $\pm 15.1^\circ$. The second one is given by Savenkov [6]. In the presence of additive noise, these two matrices are totally equivalent; they both lead to a noise covariance matrix on \vec{S} , $(W^T W)^{-1}$, which is diagonal with eigenvalues (1, 3, 3, 3) [4]. We will see that they have different behaviors in the presence of Poisson shot noise.

Let us assume that the input light has an intensity of $I_0=100$ photoelectrons and a degree of polarization equal to $P=1$. Since $q_0=5$, the total noise variance is $\mathcal{F}(W)=500$. We have represented in Table 1, for both matrices W_a and W_b , the values of γ_i^{\min} and γ_i^{\max} for $i \in [1, 3]$ (γ_0 is constant and equal to 50) and the azimuth α_i^{\max} and ellipticity ϵ_i^{\max} of the input polarization states \vec{s} for which γ_i^{\max} is reached. For matrix W_a , the variance on S_1 is independent of the input polarization state. On the other hand, γ_2 and γ_3 vary with \vec{s} in the interval [100, 200]. It is seen in Table 1 that for this particular architecture, γ_2 is maximal when the input polarization is linear with azimuth 0° and γ_3 is maximal when it has azimuth 90° . Since minimal and maximal variances are reached for orthogonal input states, γ_2 is maximal when γ_3 is minimal and vice versa. When the measurement matrix W_b is

Table 1. Minimal, Maximal Variance, α_i^{\max} , and ϵ_i^{\max} for Each Element of the Stokes Vector

		γ_i^{\min}	γ_i^{\max}	α_i^{\max} ($^\circ$)	ϵ_i^{\max} ($^\circ$)
Matrix W_a	S_1	150	150	–	–
	S_2	100	200	0	0
	S_3	100	200	90	0
Matrix W_b	S_1	50	250	0	0
	S_2	64	237	90	27
	S_3	64	237	90	-27

used, the variances on the three Stokes parameters depend on the input polarization state and reach their maxima for three different input states.

An attractive property for a measurement matrix would be that all variances γ_i do not depend on \vec{s} . The only two matrices that have this property are (within arbitrary row permutations)

$$W_c = 1/2 \times \begin{bmatrix} 1 & 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1 & -1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \\ 1 & -1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \\ 1 & 1/\sqrt{3} & -1/\sqrt{3} & -1/\sqrt{3} \end{bmatrix}, \quad (14)$$

and the matrix W'_c obtained from W_c by reversing the signs of all the elements of the last three columns. It is easily verified that $\forall i, j \in [1, 3] \times [1, 3]$, $u_j^i=0$. The noise variances are thus independent of \vec{s} and equal to $\gamma_0=I_0/2$ and $\forall i \in [1, 3]$, $\gamma_i=3/2I_0$. The situation is very similar to the case of additive noise; the variances on the three last elements of the Stokes vector are “equalized,” equal to three times the variance on S_0 , and they are independent of the input polarization state. However, in the case of Poisson noise, these properties are not obtained for all polarimeter structures based on regular tetrahedra but only in the case of measurement matrices W_c and W'_c .

The results presented in this Letter make it possible to optimize Stokes polarimeters in the presence of Poisson shot noise. In particular, we have shown that polarimeter structures that minimize and equalize the noise variance on each Stokes parameter exist. An interesting perspective to this Letter is to determine the practical polarimeter architectures that make it possible to easily generate a measurement matrix, such as Eq. (14).

The author thanks Arnaud B eni ere for fruitful discussions.

References

1. R. M. A. Azzam, I. M. Elminyawi, and A. M. El-Saba, *J. Opt. Soc. Am. A* **5**, 681 (1988).
2. A. Ambirajan and D. C. Look, *Opt. Eng.* **34**, 1651 (1995).
3. D. S. Sabatke, M. R. Descour, E. L. Dereniak, W. C. Sweatt, S. A. Kemme, and G. S. Phipps, *Opt. Lett.* **25**, 802 (2000).
4. J. S. Tyo, *Opt. Lett.* **25**, 1198 (2000).
5. J. S. Tyo, *Appl. Opt.* **41**, 619 (2002).
6. S. N. Savenkov, *Opt. Eng.* **41**, 965 (2002).
7. V. L. Gamiz and J. F. Belsher, *Opt. Eng.* **41**, 973 (2002).
8. M. R. Foreman, C. M. Romero, and P. T or ok, *Opt. Express* **16**, 15212 (2008).
9. S. M. Kay, *Fundamentals of Statistical Signal Processing—Volume I: Estimation Theory* (Prentice-Hall, 1993).