ANALYSIS OF SEEING-INDUCED POLARIZATION CROSS-TALK AND MODULATION SCHEME PERFORMANCE

R. CASINI, A. G. DE WIJN, AND P. G. JUDGE
High Altitude Observatory, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000, USA

Received 2012 May 8; accepted 2012 July 26; published 2012 September 4

ABSTRACT

We analyze the generation of polarization cross-talk in Stokes polarimeters by atmospheric seeing, and its effects on the noise statistics of spectropolarimetric measurements for both single-beam and dual-beam instruments. We investigate the time evolution of seeing-induced correlations between different states of one modulation cycle and compare the response to these correlations of two popular polarization modulation schemes in a dual-beam system. Extension of the formalism to encompass an arbitrary number of modulation cycles enables us to compare our results with earlier work. Even though we discuss examples pertinent to solar physics, the general treatment of the subject and its fundamental results might be useful to a wider community.

Key words: atmospheric effects – instrumentation: adaptive optics – instrumentation: polarimeters – methods: analytical – polarization – techniques: polarimetric

Online-only material: color figure

1. INTRODUCTION

Many important solar phenomena are driven by the interaction of turbulent plasma with magnetic fields. Measurements of the full state of polarization of solar spectral lines are routinely used to infer properties of solar vector magnetic fields from their imprints in the emergent spectra through the Zeeman and Hanle effects. All ground-based measurements are detrimentally affected by image distortion introduced by atmospheric seeing, which has significant power at frequencies up to at least 100 Hz, significantly higher than the frame rates of typical CCD cameras. The seeing-induced errors due to image distortion during an observation are expected to increase as the telescope aperture increases in size, since the telescope can then resolve finer structures and hence observe steeper gradients in images.

The Advanced Technology Solar Telescope (Rimmele et al. 2008), a 4 m aperture, off-axis telescope, will soon enter construction phase. The European Solar Telescope (Collados 2008) is also a 4 m class telescope currently under design. Spectropolarimetry is the prime mode of operation of this new generation of solar telescopes, but there is an urgent need to identify potential pitfalls before committing to a particular design for the modulation and detection of polarized light using such telescopes. Space-based spectropolarimetric instruments, like those on board the Hinode (Solar-B; Tsuneta et al. 2008) and Solar Dynamics Observatory (Scherrer et al. 2012) spacecrafts, are by necessity fed by smaller telescopes, and while their observations profit greatly from the absence of atmospheric disturbances, they are still affected by residual image motion due to spacecraft jitter.

The purpose of the present paper is to re-examine and extend earlier theoretical studies of polarization errors introduced by the effects of atmospheric seeing. Potentially, this can help refine the instrument requirements for the polarization systems of these large, multi-national projects, as well as of future solar space missions, such as Solar-C (Shimizu et al. 2011).

Two modes of operation are of particular interest for our study. Slit-based spectropolarimeters often use long integration times, including a number of modulation cycles that is typically of the order of a few tens. Because of the need to scan across the spectral line domain, tunable imaging spectropolarimeters typically implement much fewer measurements of the modulated intensities—often just one modulation cycle—for each wavelength position. The formalism developed in this paper is general, and can be applied to both types of instruments and observations. A substantial difference between the two types of observations is that post-processing techniques such as image shifting and de-stretching can be effectively adopted to partially correct for seeing-induced distortion in observations done with imaging polarimeters, whereas this option is not available in the case of slit-based spectropolarimeters.

A separate question—also of fundamental importance for observational spectropolarimetry—concerns the effects of polarization calibration errors on the degree of polarization cross-talk affecting an observation. This subject has been exhaustively treated by Asensio Ramos & Collados (2008), and therefore it will not be addressed here.

The plan of this paper is the following. In Section 2, we describe a model for the effects of atmospheric seeing on the polarization signals detected by an observer. In Section 3, we lay the foundation of the formalism. Using the point of view of the statistics of random processes we define the fundamental statistical observables for spectropolarimetry in the presence of atmospheric seeing. In Section 4, this formalism is extended to the case of dual-beam polarimetry, emphasizing those aspects of the problem where this extension is not trivial. In Section 5, we make use of the conceptual framework of the statistics of stationary random processes to gain a deeper understanding of the effects of seeing correlations on spectropolarimetric measurements. In particular, in that section we qualitatively study the effect of adaptive optics (AO) corrections on the temporal decay of seeing correlations. Those results are successively applied to study the performance of two popular modulation schemes in the presence of seeing. In Section 6, these results are further generalized to the case in which seeing correlations extend over the duration of one elemental observation (e.g., one slit position), and the effects of seeing on the performance of optimally efficient modulation schemes are studied both in the absence and in the presence of...
low-order (i.e., tip-tilt) AO corrections (based on observed data). Finally, we compare our results and conclusions with those from previous works on the subject.

2. POLARIZATION AND ATMOSPHERIC SEEING

We conventionally describe the polarization state of a radiation beam by its Stokes vector, \( \mathbf{S} = (S_1, S_2, S_3, S_4) \), where \( S_1 \) is the intensity of radiation, \( S_2 \) and \( S_3 \) describe the two independent states of linear polarization on the plane normal to the propagation direction, and finally \( S_4 \) is the parameter for circular polarization.

Because photon detectors are practically sensitive only to the intensity of the incoming radiation, the measurement of polarized radiation requires that its states of polarization be encoded into intensity signals. This encoding process is conventionally called modulation, and it is most often achieved through time-varying optical devices (modulators) that modify in a known way the polarization states of the incoming radiation.

Let \( \mathbf{m}(t) \) be the four-vector describing the polarization modulation process for the four Stokes parameters, so that \( \mathbf{m}(t) \cdot \mathbf{S} \) gives the detected signal of the modulated intensity at time \( t \) (see Equation (9) below). We will assume that this modulation vector is perfectly known (either by design or calibration), and that it is a periodic function of time, so that \( \mathbf{m}(t+\tau) = \mathbf{m}(t) \), where \( \tau \) is the period of the modulation cycle. Because of these properties, the time process of polarization modulation is both deterministic and stationary in a statistical sense. Evidently one needs at least four independent measurements during a modulation cycle to fully determine the Stokes vector. The actual number of measurements \( n \) taken during a given modulation cycle represents the number of modulation states for that cycle.

We also assume that the physical conditions of the light emitting region are stationary during the time interval \( T \) needed to perform a complete spectropolarimetric observation of a given element of spatial or spectral sampling, with the required sensitivity. In what follows, the term “complete observation” will always refer to such elemental operation, for instance, the integration of the light signal at one slit position of a spatial scan, in the case of a grating-based spectropolarimeter, or at one wavelength position of a spectral scan, in the case of a tunable, imaging spectropolarimeter. Evidently, such observation must always include at least one full modulation cycle.

In the presence of atmospheric seeing, a given pixel on the detector collects photons from different areas of the observed region, with a characteristic timescale \( t_0 \) (\( \sim 0.01 \) s). The cause for this “smearing” of the image can be a true displacement of the target by the seeing during the integration time or seeing-induced distortions of the effective point-spread function (PSF) at the detector, which determine a time-varying weighing of different elements of the observed region. This time-dependent smearing is in addition to the one determined by diffraction because of the finite aperture of the telescope. As a result, the combined effect of the atmosphere and the telescope is to smoothen the Stokes gradients present in the observed region, through a smearing length that is determined by the spatial resolution of the observation.

In order to demonstrate this intuitive result, we begin by considering the ideal case of a perfectly stigmatic optical system (e.g., Born & Wolf 1965) observing in the absence of atmospheric seeing. In this case, a 1–1 mapping can be established between the object and the image planes. We indicate with \( \tilde{S}_{ij}(t) \) the Stokes vector falling on the resolution element\(^2\) of coordinates \( (i, j) \) on the image plane, at a given time \( t \). Similarly, we indicate with \( S_{ij} \), the Stokes vector emitted by the element on the object plane (i.e., the solar surface) that corresponds to the element \( (i, j) \) on the detector, through the inverse mapping from the image plane onto the object plane. Note that the Stokes vector at the source is not given as a function of time, in agreement with the condition stated earlier that the time interval \( T \) to perform an elemental observation should be much shorter than the typical evolution time of the observed solar structure.

As customary, we assume that the transport of radiation from the object plane to the image plane is described by a linear operator. Hence, we can write

\[
\tilde{S}_{ij}(t) = \sum_{kl} T_{ij}^{kl}(t) S_{kl},
\]

where \( T_{ij}^{kl}(t) \) is the transfer (Mueller) matrix at time \( t \) of the imaging system telescope+atmosphere, which maps the element \( (k, l) \) on the object plane to the corresponding resolution element \( (i, j) \) on the image plane. In the ideal case of a perfectly stigmatic telescope and in the absence of atmospheric seeing, evidently

\[
T_{ij}^{kl}(t) = \delta_{ik} \delta_{jl} T_{ij}(t),
\]

where \( T_{ij}(t) \) is the transfer matrix of the imaging system, pertaining to the element \( (i, j) \) of the field of view.

The general form of Equation (1) applies in the presence of smearing of the image, which may be caused by atmospheric seeing, by the instrument’s PSF, as well as by the optical aberrations of the imaging system.\(^3\) The Stokes vector measured by the detector is thus properly expressed as a function of the time, because of the variability of the atmosphere. Possible temporal variations of the instrument—for example, due to the changing optical configuration of the telescope during tracking of the target—are not a concern here, since their timescale is typically much larger than \( T \). Thus, in practice, the transfer matrix on the right-hand side of Equation (2) can be assumed to be constant.

When observations are affected by atmospheric seeing, the spatial resolution—and thus the size of the resolution element on the detector—depends critically on the particular observing conditions, such as the possible presence of AO corrections, and the temporal resolution of the observation (i.e., the exposure time), because of the effects of image smearing mentioned above. For example, in the limit of long time exposures, and in the absence of AO corrections, the spatial resolution is completely determined by the Fried parameter \( r_0 \) of the atmospheric seeing (Fried 1965). According to the Van Cittert–Zernike theorem (Van Cittert 1934; Zernike 1938; see also Mandel & Wolf 1995), distinct resolution elements can be regarded as incoherent light sources. The linear superposition of the source Stokes vectors expressed by Equation (1) is thus justified by this assumption of spatial and temporal incoherence of the radiation emitted by resolved structures on the object plane.

\(^2\) Typically, the pixel size in an optical system is matched to the critical sampling width (both spatial and spectral) determined in accordance to the Nyquist criterion (e.g., Goodman 1996), assuming diffraction-limit performance of the instrument. Thus, the resolution element will span at least \( 2 \times 2 \) pixels or a larger number for performance worse than the diffraction limit.

\(^3\) Of course, the discretization of the object plane implied by Equation (1) can only be an approximation, under these more general observing conditions.
Let us now consider a point \((p, q)\) on the image plane, located inside a sufficiently small neighborhood of \((i, j)\), such that we can approximate

\[
\tilde{S}_{pq}(t) \approx \tilde{S}_{ij}(t) + (x_{pq} - x_{ij}) \cdot \nabla \tilde{S}_{ij}(t). \tag{3}
\]

We recall that for an arbitrary vector \(v\),

\[
\nabla v_{ij} = (\partial_x v_{ij}) \hat{e}_x + (\partial_y v_{ij}) \hat{e}_y \equiv l_x^{-1} (v_{ij} - v_{ij-1}) \hat{e}_x + l_y^{-1} (v_{ij} - v_{ij-1}) \hat{e}_y,
\]

where \(l_x \equiv |x_{ij} - x_{i-1,j}|\) and \(l_y \equiv |x_{ij} - x_{i,j-1}|\) are expressed in pixel units. We then find, from Equation (1),

\[
\partial_x \tilde{S}_{ij}(t) = l_x^{-1} \sum_{kl} [T_{ij}^{kl}(t) - T_{ij-1}^{kl}(t)] \partial_x S_{kl}, \tag{4a}
\]

\[
\partial_y \tilde{S}_{ij}(t) = l_y^{-1} \sum_{kl} [T_{ij}^{kl}(t) - T_{ij-1}^{kl}(t)] \partial_y S_{kl}. \tag{4b}
\]

We now assume that the transfer matrix is shift-invariant over some neighborhood of the observed point. By definition, this condition is automatically satisfied when the neighborhood lies within an isoplanatic patch of the field of view (e.g., Goodman 1996). Then, Equations (4) can be rewritten in the following form:

\[
\partial_x \tilde{S}_{ij}(t) = -l_x^{-1} \sum_{kl} [T_{ij}^{kl}(t) - T_{ij-1}^{kl}(t)] S_{kl}, \tag{5a}
\]

\[
\partial_y \tilde{S}_{ij}(t) = -l_y^{-1} \sum_{kl} [T_{ij}^{kl}(t) - T_{ij-1}^{kl}(t)] S_{kl}. \tag{5b}
\]

Next we note that the smearing of the image at a given point is always bounded in practice, in the sense that, for every point \((i, j)\) on the image plane, there is always a positive integer \(M(i,j)\) such that

\[
T_{ij}^{k,m,j+l}(t) \equiv 0, \quad \text{if} \quad m, n \geq M(i,j). \tag{6}
\]

Under typical observing conditions, the isoplanatic patch is always larger than the contribution region to the element \((i, j)\) due to smearing, which is bounded according to Equation (6). For example, in the absence of AO corrections and for long time exposures, the smearing length is approximately given by the spatial resolution corresponding to \(r_0\), but the region of isoplanatism typically includes several of these resolution elements. Therefore, we can assume that the set of points \((k, l)\) contributing to Equations (5) always lies inside a region where the assumption of shift invariance of the transfer matrix is valid. This allows us to perform a summation by parts of Equations (5), neglecting the associated boundary terms because of Equation (6), to find

\[
\partial_x \tilde{S}_{ij}(t) = l_x^{-1} \sum_{kl} T_{ij}^{k+1,l}(t)(S_{k+1,l} - S_{kl}) \equiv \sum_{kl} T_{ij}^{kl}(t) \partial_x S_{kl}, \tag{7a}
\]

\[
\partial_y \tilde{S}_{ij}(t) = l_y^{-1} \sum_{kl} T_{ij}^{k,l+1}(t)(S_{k,l+1} - S_{kl}) \equiv \sum_{kl} T_{ij}^{kl}(t) \partial_y S_{kl}. \tag{7b}
\]

Equations (7) demonstrate the anticipated result that the Stokes gradients present at the object plane are smeared by the combined effect of the atmosphere and the telescope. The characteristic length for the variation of the Stokes vector on the image plane is determined by the convolution of the physical spatial scale of the observed structure, characterized by \(\nabla S_{kl}\), with the spatial resolution of the observation, corresponding to the domain of non-nullity of \(T_{ij}^{kl}(t)\) (cf. Equation (6)). The validity of the linear approximation expressed by Equation (3) is therefore limited to a range of distances \(|x_{pq} - x_{ij}|\) corresponding to this “smoothed” length scale, which is approximately given by \(|S_{ij}|/\|\nabla S_{ij}\|\).

Equation (3) provides a snapshot of the spatial distribution of polarization signals in a region around the element \((i, j)\) at the time \(t\). Because the phenomenon of atmospheric seeing can be regarded as a stationary and ergodic\(^4\) random process (e.g., Mandel & Wolf 1995), we can assume that all possible realizations \(S_{pq}(t)\) of Equation (3) will eventually manifest themselves at \(x_0 \equiv x_{ij}\) during the evolution of the atmospheric seeing. In other words, we can associate a stationary random process \(x(t)\) with the displacement vector of the image motion, which describes the effect of atmospheric seeing to the lowest order, such that (cf. Equation (3))

\[
\tilde{S}(x_0; t) \approx \tilde{S}(x_0) + (x(t) - x_0) \cdot \nabla \tilde{S}(x_0), \tag{8}
\]

where \(\tilde{S}(x_0)\) and \(\nabla \tilde{S}(x_0)\) must be interpreted as the ensemble (or long-time) averages of the quantities \(\tilde{S}_{ij}(t)\) and \(\nabla \tilde{S}_{ij}(t)\) appearing in Equation (3). Because of the random nature of atmospheric seeing, the long-time average of \(x(t) - x_0\) is zero, and therefore \(\tilde{S}(x_0)\) coincides with the Stokes vector that would be measured in the absence of atmospheric seeing.

For small telescopes with apertures of order \(r_0\), the detected signal at any given time is mostly affected by low-order aberrations of the wavefront, so that the seeing displacement vector \(x(t)\) can be effectively compensated just by the tip and tilt corrections. For large aperture telescopes, such that \((D/r_0)^2\) is of order 10 or larger, \(D\) being the telescope’s diameter, the distortions of the incoming wavefront are significantly more complex, resulting in the characteristic dynamic pattern of speckles on the image plane. Nonetheless, even in this more general case, the low-order Zernike terms corresponding to the tip-tilt wavefront correction still account for the largest part of the seeing-induced effects on the image (Noll 1976). We note that this more general case is still captured by the model of Equation (1). Thus, as long as a linear approximation of \(\tilde{S}(x; t)\) around \(x_0\) is justified, Equation (8) also describes the lowest order effects of seeing in observations performed with a large telescope.

Of course, the requirements for the validity of the linear approximation (8) become more stringent as the spatial resolution increases with the AO correction, because of the enhanced contrast of the image, and the expected increased complexity of the structure of the solar atmosphere at smaller spatial scales. In fact, as it has already been pointed out, the spatial range of applicability of the linear model of Equation (8) decreases like \(|S|/\|\nabla S\|\). On the other hand, the AO corrections effectively

\(^4\) Simply stated, ergodicity is the statistical property of a dynamical system that allows to use time averages in place of ensemble averages. In other words, it is characteristic of an ergodic process to attain, during its temporal evolution, all of its possible configurations, with a probability distribution identical to that of the ensemble.
reduce the rms displacement vector that can be associated with the residual image motion. Therefore, because of the overall scaling of the problem, we can expect that the linear model of Equation (8) will remain applicable even in the presence of AO corrections, at least as far as the effects of residual image motion are concerned.

Higher-order effects may be responsible for local fluctuations of the image intensity not associated with image motions, which would also be a source of polarization cross-talk, and they may even occur as image artifacts introduced by the same AO corrections. A quantitative study of these effects lies beyond the scope of this paper, but it could perhaps be pursued through an extension of the formalism presented in this work.

3. STATISTICAL DESCRIPTION OF THE EFFECTS OF ATMOSPHERIC SEEING

On the basis of the conclusions of the former section, in the following we will assume that, at any given time within the duration $T$ of a particular observation, the Stokes vector entering the polarization modulator of the telescope is expressed by Equation (8) (cf. Lites 1987; Judge et al. 2004). For notational convenience, in the following we always assume $x_0 = 0$, and also drop the "*m" from the quantities of Equation (8).

According to this model, the modulated intensity recorded by an ideal detector (i.e., neglecting bias and readout noise), for the $i$-th state of the modulation cycle, centered around the time $t_i$, and integrated over the exposure time $\Delta t$, is given by

$$\mathcal{I}_i = \kappa \sum_{j=1}^{4} \int_{-\Delta t/2}^{+\Delta t/2} m_j(t + t_i) \nabla S_j \cdot \mathbf{x}(t + t_i) \, dt + \delta \mathcal{I}_i, \quad i = 1, \ldots, n,$$

where $\kappa$ is a dimensional scaling constant for the detector, $n$ is the number of modulation states in the cycle, and $\delta \mathcal{I}_i$ is a random fluctuation due to photon-noise statistics.

It therefore makes sense to derive the expressions for the expectation value and variance of $\mathcal{I}_i$—respectively, $E(\mathcal{I}_i)$ and $\sigma^2(\mathcal{I}_i)$—as these provide important information about the quality of the observations (see also Asensio Ramos & Collados 2008). As we anticipated earlier in this section, $E(\mathbf{x}(t_i)) = 0$. Similarly, the photon-noise fluctuation of the measurement has zero expectation value. Therefore, using Equation (9), and the fact that the modulation process is assumed to be fully deterministic and stationary, we obtain

$$\bar{\mathcal{I}}_i \equiv E(\mathcal{I}_i) = \kappa \Delta t \sum_{j=1}^{4} m_j S_j$$

$$+ \frac{1}{\Delta t} \int_{-\Delta t/2}^{+\Delta t/2} m_j(t + t_i) \nabla S_j \cdot E(\mathbf{x}(t + t_i)) \, dt$$

$$\equiv \kappa \Delta t \sum_{j=1}^{4} m_{ij} S_j, \quad i = 1, \ldots, n,$$  \hfill (10)

where we have defined

$$m_{ij} \equiv \frac{1}{\Delta t} \int_{-\Delta t/2}^{+\Delta t/2} m_j(t + t_i) \, dt.$$  \hfill (11)

The $n \times 4$ matrix $\mathbf{M} \equiv (m_{ij})$ so constructed is conventionally called the modulation matrix.

Equation (10) shows that the expectation value of the $i$-th modulated intensity signal is determined exclusively by the true Stokes vector $\mathbf{S}$, regardless of the modulation scheme adopted. It is important to note that, in both Equations (10) and (11), the subscript $i$ no longer represents a specific instant $t_i$, like in Equation (9), but strictly represents only the corresponding step position of the $n$-state modulation cycle. In particular, for Equation (11), this is a direct consequence of the stationarity of the modulation process.

Similarly, the variance of $\mathcal{I}_i$, for $i = 1, \ldots, n$, is given by

$$\sigma^2(\mathcal{I}_i) \equiv E((\mathcal{I}_i - \bar{\mathcal{I}}_i)^2)$$

$$= E\left[\left(\kappa \sum_{j=1}^{4} \int_{-\Delta t/2}^{+\Delta t/2} m_j(t + t_i) \nabla S_j \cdot \mathbf{x}(t + t_i) \, dt + \delta \mathcal{I}_i\right)^2\right]$$

$$= \kappa^2 \sum_{j,k=1}^{4} \int_{-\Delta t/2}^{+\Delta t/2} m_j(t + t_i) m_k(t' + t_i)$$

$$\times E([\nabla S_j \cdot \mathbf{x}(t + t_i)][\nabla S_k \cdot \mathbf{x}(t' + t_i)]) \, dt \, dt' + \sigma^2_p(\mathcal{I}_i),$$

where we indicated with $\sigma_p(\mathcal{I}_i)$ the rms photon noise.

Equation (12) has no immediate applicability to Stokes data analysis, since it relies on physical parameters, such as $\nabla S_j$ and $\mathbf{x}(t)$, which are not directly available from typical observations. Instead, the variances defined by Equation (12) are determined in practice from the statistics of repeated intensity measurements over many modulation cycles (Lites 1987). Nonetheless, Equation (12) helps clarifying the physical origin of the noise in the measurement of the modulated intensities, and therefore can provide insights on optimal choices of modulation schemes and frequencies for reducing the final error.

It is useful to derive the expectation value and variance also in the case of $\mathcal{I}_i$. Using Equation (10), we find

$$\bar{S}_i = \sum_{j=1}^{n} d_{ij} \mathcal{I}_j, \quad i = 1, \ldots, 4.$$  \hfill (13)

This is in fact a general result that follows from error minimization in ordinary least-squares problems (e.g., Draper & Smith 1966).
\begin{equation}
\tilde{S}_i' \equiv E(S_i') = \sum_{j=1}^{n} d_{ij} \tilde{\mathcal{I}}_j = \kappa \Delta t \sum_{j=1}^{n} \left( \sum_{k=1}^{n} d_{jk} \right) S_k = \kappa \Delta t S_i, \quad i = 1, \ldots, 4,
\end{equation}

where we used the fact that \( \sum_j d_{ij} m_{jk} = \delta_{ik} \). Similarly, for the variance of \( S_i' \) we find, using Equations (13) and (14),

\begin{equation}
\sigma^2(S_i') = E(\{S_i' - \tilde{S}_i'\}^2) = E\left( \left[ \sum_{j=1}^{n} d_{ij} (\mathcal{I}_j - \tilde{\mathcal{I}}_j) \right]^2 \right) = \sum_{j,k=1}^{n} d_{ij} d_{ik} E(\{\mathcal{I}_j - \tilde{\mathcal{I}}_j\}[\mathcal{I}_k - \tilde{\mathcal{I}}_k]), \quad i = 1, \ldots, 4.
\end{equation}

Equations (10)–(12) and (14) and (15) provide the basic formulae through which we can evaluate the performance of different modulation schemes in the presence of atmospheric seeing. They will be related to earlier work in Section 7.

When the exposure time \( \Delta t \) is sufficiently large relative to the coherence time of the atmospheric seeing, like in the case of slow modulation cycles, then we can assume that the coherence terms \( E(\{\mathcal{I}_j - \tilde{\mathcal{I}}_j\}[\mathcal{I}_k - \tilde{\mathcal{I}}_k]) \) in Equation (15) are negligible for \( j \neq k \). In such case the variance of \( S_i' \) can be expressed directly as a diagonal quadratic form of the variances of the modulated intensity signals,

\begin{equation}
\sigma^2(S_i') = \sum_{j=1}^{n} d_{ij}^2 \sigma^2(\mathcal{I}_j), \quad i = 1, \ldots, 4.
\end{equation}

In Section 5, we provide a rigorous demonstration of the above statement. However, in many cases—certainly when fast cameras with frame rates \( \gtrsim 10 \text{ Hz} \) are employed—the time interval \( t_{n+1} - t_i \) (whose inverse conventionally defines the modulation frequency) is of the same order or less than the timescale \( t_0 \) of the atmospheric seeing. As a result, the intensity variations induced by seeing in different states of the modulation cycle are in general statistically correlated, so one must use the more general expression of Equation (15).

We note that Equation (16) also holds in the absence of seeing, in which case \( \sigma^2(\mathcal{I}_j) \) evidently reduces to just the contribution due to photon noise. Because the signals from different camera exposures are always statistically independent, no covariance terms arise in this case, and the photon noise only affects the diagonal terms of Equation (15). This provides a simple recipe to include photon noise in the results presented in this paper. Thus, for notational convenience, we will simply drop the photon-noise terms in all of the following treatment.

Finally, we observe that the diagonality of \( \sigma^2(S_i') \), as expressed by Equation (16), is a fundamental assumption in the derivation of the optimal form of demodulation matrices presented by del Toro Iniesta & Collados (2000). In other words, the usual definition of the efficiency of polarization modulation schemes relies on the fact that seeing-induced noise be negligible with respect to photon noise, or at least that the seeing-induced covariances of the type appearing in Equation (15) be vanishing. Since that approach is based on the minimization of the noise of the Stokes measurements in the absence of systematic errors, which is expressed by a relation formally identical to Equation (16), there remains an unanswered question: How would the condition of optimality of the demodulation matrix change, if the more general expression of Equation (15)—which takes into account the systematic errors due to the seeing—were adopted. This is a subject of research in its own right, which we are not going to address further in this paper.

### 4. Dual-Beam Polarimetry

In the case of dual-beam polarimetry, one has two independent sets of \( n \) measurements that can be combined into \( 2n \) new intensity signals of the form \( \mathcal{I}_j^\pm \equiv \mathcal{I}_j \pm \mathcal{I}_k \), where \( \mathcal{I}_j \) and \( \mathcal{I}_k \) refer to the two beams. For an ideal polarimeter, the modulation vector for the two beams satisfies the simple relations \( m_i^a(t) = m_i^b(t) = 1 \) and \( m_i^a(t) = -m_i^b(t) \) for \( i = 2, 3, 4 \). However, in practical cases the two beams are never perfectly balanced. We take into account beam imbalance through the detector’s gain factor in front of Equation (9), and introduce new modulation vectors for the dual-beam system, \( m_i^a(t) \) and \( m_i^b(t) \), according to

\begin{equation}
\tilde{\mathcal{I}}_i^\pm = \mathcal{I}_i^+ \mathcal{I}_i^- = \tilde{\mathcal{I}}_i^+, \quad \tilde{\mathcal{I}}_i^\pm(t) = (m_i^a(t), m_i^b(t))^T.
\end{equation}

Through these new vectors, Equations (10)–(12) can be directly extended to take into account the dual-beam redundancy. It is convenient to introduce a new \( 2n \) intensity vector with the corresponding modulation vector,

\begin{equation}
\tilde{\mathcal{I}}_i^\pm = \mathcal{I}_i^+ \mathcal{I}_i^- = \tilde{\mathcal{I}}_i^+, \quad \tilde{\mathcal{I}}_i^\pm(t) = (m_i^a(t), m_i^b(t))^T.
\end{equation}

We then find, for \( i = 1, \ldots, 2n \),

\begin{equation}
\tilde{\mathcal{I}}_i^\pm = E(\mathcal{I}_i^\pm) = \tilde{\mathcal{I}}_i^+ \tilde{\mathcal{I}}_i^- = \kappa \Delta t \sum_{j=1}^{4} m_i^j S_j,
\end{equation}

\begin{equation}
\sigma^2(\mathcal{I}_i^\pm) = E(\{\mathcal{I}_i^\pm - \tilde{\mathcal{I}}_i^\pm\}^2) = \kappa^2 \sum_{j,k=1}^{4} \int_{-\Delta t/2}^{\Delta t/2} m_i^j(t + t_{i,j}) m_i^k(t_{i,k}) \times E(\{\mathcal{I}_j^\pm - \tilde{\mathcal{I}}_j^\pm\}[\mathcal{I}_k^\pm - \tilde{\mathcal{I}}_k^\pm]) dt dt'.
\end{equation}

Note that we dropped the time shift \( t_i \) from both arguments of \( \mathcal{I}_j(t) \), relying on the stationarity of seeing (see the discussion after Equation (12), and Section 5). The index of \( t_i \) in the argument of the modulation functions \( m_i^\pm(t) \) must be modulo \( n \), since it is related to the actual time stamp during the modulation cycle.

In the usual way, we can associate with \( \mathbf{M}^\pm \equiv (m_i^\pm) \) an “optimal,” dual-beam demodulation matrix, \( \mathbf{D}^\pm \). Then, the extensions of Equations (14) and (15) to the case of dual-beam polarimetry become, respectively,

\begin{equation}
\delta_i^\pm = \sum_{j=1}^{2n} d_{ij}^\pm \tilde{\delta}_j^\pm, \quad i = 1, \ldots, 4,
\end{equation}

\begin{equation}
\sigma^2(S_i^\pm) = \sum_{j,k=1}^{2n} d_{ij}^\pm d_{ik}^\pm \times E(\{\mathcal{I}_j^\pm - \tilde{\mathcal{I}}_j^\pm\}[\mathcal{I}_k^\pm - \tilde{\mathcal{I}}_k^\pm]) = 1, \ldots, 4.
\end{equation}

As before, off-diagonal covariances in Equation (21) can be neglected only in the limit of exposure times much larger than the seeing coherence time (\( \Delta t \gg t_0 \)).
In order to illustrate the effect of beam imbalance, we consider the ideal modulation matrix for a general modulation scheme with \( n \) states,

\[
M = \begin{pmatrix}
1 & m_{12} & m_{13} & m_{14} \\
& 1 & m_{23} & m_{24} \\
& & 1 & m_{34} \\
& & & 1
\end{pmatrix}.
\]

(22)

If we introduce the degree of beam imbalance, \( \rho = (k_a - k_b)/(k_a + k_b) \), the modulation matrix \( M^\pm \) appearing in Equation (18) is accordingly given by

\[
M^\pm = \begin{pmatrix}
1 & \rho m_{12} & \rho m_{13} & \rho m_{14} \\
1 & \rho m_{22} & \rho m_{23} & \rho m_{24} \\
\rho & m_{12} & m_{13} & m_{14} \\
\rho & m_{22} & m_{23} & m_{24}
\end{pmatrix}.
\]

(23)

where the top and bottom halves correspond respectively to the “+” and “−” linear combinations of the \( n \) measurements from the two beams.

From the form of \( M^\pm \) we immediately see that the seeing variations on the intensity and those on the polarization parameters of the incoming Stokes vector are decoupled when \( \rho = 0 \). Such removal of cross-talk between intensity and polarization is in fact the rationale behind dual-beam polarimetry. However, this ideal goal is attained only if the intensity signals of the two beams can be balanced (either by camera gain adjustment or by rescaling of the signals during data reduction) with an error that must be smaller than the target polarimetric sensitivity. To clarify this point, let us assume that the gain factors \( k_a \) and \( k_b \) have been experimentally determined with some error by the polarization calibration procedure, so that \( k_{a,b} = k'_{a,b} + \delta k'_{a,b} \), where \( k'_{a,b} \) is the measured value of \( k_{a,b} \). From the definition of the dual-beam modulation vector, Equation (17), after rescaling, we have

\[
\tilde{k} m^\pm_{ij} \equiv \frac{k_a}{k'_{a}} m_{ij} \pm \frac{k_b}{k'_{b}} m'_{ij}, \quad \tilde{k} \equiv \frac{k_a}{k'_{a}} + \frac{k_b}{k'_{b}}.
\]

Correspondingly, the degree of beam imbalance becomes

\[
\rho = \left( \frac{\delta k'_{a} - \delta k'_{b}}{k'_a} \right) \left( 2 + \frac{\delta k'_a}{k'_a} + \frac{\delta k'_b}{k'_b} \right) \approx \frac{1}{2} \left( \frac{\delta k'_{a}}{k'_a} - \frac{\delta k'_{b}}{k'_b} \right),
\]

so that \( |\rho| \lesssim (1/2)(|\delta k'_{a}|/k'_a + |\delta k'_{b}|/k'_b) \) provides a sensible bound on the maximum error acceptable in order to achieve a given polarimetric precision in the presence of residual intensity-to-polarization cross-talk.

Since the two beams can in principle always be rescaled post facto, we will assume in the following that we are always dealing with perfectly balanced beams. It is understood that the possible rescaling of the two beams has to be taken into account for the proper determination of the noise on the combined beams, according to the usual formula \( \sigma^2(aX + bY) = a^2\sigma^2(X) + b^2\sigma^2(Y) \), where \( a \) and \( b \) are real numbers. This has an effect on the derivation of the quantities \( \sigma^2(\rho a x) \) in Equation (19) from the noise statistics of the individual beams. In Appendix A, we illustrate this problem for the particular case where the beam imbalance is produced by the differential efficiency of a diffraction grating in the \( p \) and \( s \) polarizations.

5. THE BEHAVIOR OF SEEING-INDUCED CORRELATIONS BETWEEN MODULATION STATES

We must evaluate covariance terms \( E(\{J_f - \bar{J}_f\} [\bar{J}_k - \bar{J}_k]) \) in Equation (15) (or Equation (21)) and study their behavior as a function of the modulation frequency. Using the definition (9),

\[
E(\{J_f - \bar{J}_f\} [\bar{J}_k - \bar{J}_k])
\]

\[
= \kappa^2 \sum_{p,q=1}^{4} \int_{-\Delta t/2}^{+\Delta t/2} m_p(t + t_j) \nabla S_p \cdot x(t + t_j) dt 
\times \int_{-\Delta t/2}^{+\Delta t/2} m_q(t' + t_k) \nabla S_q \cdot x(t' + t_k) dt' 
\]

\[
= \kappa^2 \int_{-\Delta t/2}^{+\Delta t/2} \int_{-\Delta t/2}^{+\Delta t/2} m_p(t + t_j) m_q(t' + t_k) 
\times E(x_a(t + t_j) x_b(t' + t_k)) dt dt',
\]

(24)

where in the last line a double summation over the coordinate indices \( \alpha, \beta = 1, 2 \) is implicit.

Because seeing can be considered a stationary random process, the expectation value in the last line of Equation (24) can be written in terms of the two-time correlation matrix,

\[
\Gamma_{\alpha\beta}(t' - t) \equiv E(x_\alpha(t)x_\beta(t')) , \quad \alpha, \beta = 1, 2,
\]

(25)

and because of the isotropy of the seeing motion, we also have\(^6\)

\[
\Gamma_{12}(t) = \Gamma_{22}(t) \equiv \Gamma(t),
\]

(26a)

\[
\Gamma_{12}(t) = \Gamma_{21}(t),
\]

(26b)

since the two components of such a motion cannot be distinguishable. In addition, the two component motions \( x_1(t) \) and \( x_2(t) \) are orthogonal, and therefore they can be assumed to be independent random processes. Hence,

\[
E(x_\alpha(t)x_\beta(t')) = E(x_\alpha(t))E(x_\beta(t')) = 0, \quad \alpha \neq \beta,
\]

since \( x_1(t) \) and \( x_2(t) \) are random processes with zero average. Thus, the correlation matrix is diagonal and proportional to the unit matrix.\(^7\)

\[
\Gamma_{\alpha\beta}(t) = \delta_{\alpha\beta} \Gamma(t).
\]

Equation (24) then becomes

\[
E(\{J_f - \bar{J}_f\} [\bar{J}_k - \bar{J}_k])
\]

\[
= \kappa^2 \int_{-\Delta t/2}^{+\Delta t/2} \int_{-\Delta t/2}^{+\Delta t/2} m_p(t + t_j) m_q(t' + t_k) 
\times \Gamma(t' - t + t_j) dt dt'.
\]

(27)

\(^6\) The cross-correlation function of a real stationary random process satisfies the relation \( \Gamma_{\alpha\beta}(t) = \Gamma_{\beta\alpha}(-t) \). The additional symmetry constraint provided by Equation (26b) can be viewed as a consequence of the time-reversal symmetry of the seeing motion.

\(^7\) Another way to state this result is to consider that since the correlation matrix (26) is symmetric, it can be diagonalized via a similarity transformation involving standard rotation matrices in \( \Omega(2) \). On the other hand, because of the isotropy of the seeing motion, there cannot be any preferential direction to attain such a diagonal form, and so the correlation matrix for the seeing displacement must always be diagonal.
Following the formalism of Section 4, for an n-state modulation scheme in dual-beam configuration, j and k vary from 1 to 2n. However, the indices of $t_j$ and $t_k$ must be taken modulo n, because these refer to the actual steps of the modulation cycle.

No further simplifications can be made at this point in the case of a continuously modulating device. In the remaining part of this section, we will then restrict ourselves to the case of stepped modulators, for which the following relation holds:

$$m_j(t+t_j) = m_i(t_j), \quad \forall t \in (-\Delta t/2, +\Delta t/2).$$

The last equivalence follows directly from the definition of the modulation matrix, Equation (11), and so we can simply operate the substitution $m_p(t+t_j)m_q(t'+t_k) \rightarrow m_{jp}m_{kq}$ in Equation (27). Using standard manipulations (e.g., Mandel & Wolf 1995), the double integral in Equation (27) can then be transformed into a single integral. In fact, noting that $t_k-t_j = (k-j)\Delta t/r$, where r is the duty cycle of the camera ($0 < r \leq 1$), we have

$$F_i(\Delta t) = \frac{1}{\Delta t^2} \int_{-\Delta t/2}^{+\Delta t/2} \Gamma(t'-t+s\Delta t/r) \, dt \, dt'$$

$$= \frac{1}{\Delta t} \int_{-\Delta t}^{\Delta t} \left( 1 - \frac{|\tau|}{\Delta t} \right) \Gamma(\tau + s\Delta t/r) \, d\tau. \quad (28)$$

For the sake of demonstration, in the following we assume for $\Gamma(t)$ a functional dependence typical of a Gauss–Markov random process,

$$\Gamma(t) = \Gamma(0) e^{-\chi|t|}, \quad \chi > 0. \quad (29)$$

It is known (Tatarski 1961) that the Kolmogorov description of atmospheric turbulence leads instead to an auto-correlation function of the form $e^{-\chi|\tau|^2}$. However, its use in place of Equation (29) would not qualitatively change the conclusions of this section, since Equation (29) already contains the essential features of the seeing power spectrum that we are going to analyze. Using the auto-correlation function (29), $F_i(\Delta t)$ can be integrated analytically:

$$F_i(\Delta t) = \begin{cases} 
\Gamma(0) \frac{(e^{\chi \Delta t} - 1)^2}{\chi^2 \Delta t^2} e^{-\chi \Delta t (1+|\tau|/r)}, & s \neq 0, \\
\Gamma(0) \frac{2(\chi \Delta t + e^{-\chi \Delta t} - 1)}{\chi^2 \Delta t^2}, & s = 0.
\end{cases}$$

We then see that $F_i(\Delta t) \rightarrow 0$ when $\Delta t \gg \chi^{-1}$ as expected, because of the random nature of atmospheric seeing. In particular, for very long exposure times, $F_i(\Delta t)$ tends to zero at least as $(\chi \Delta t)^{-2}$ for $s \neq 0$, whereas $F_0(\Delta t) \sim (\chi \Delta t)^{-1}$. For typical atmospheric conditions, $\chi^{-1} \sim 10^{-2} \text{s}$, which is of the same order of magnitude of typical exposure times. Hence, the terms $E(\mathbf{S}_j - \mathbf{S}_j || \mathbf{S}_k - \mathbf{S}_k)$ must in general be taken into account for a proper determination of the measurement errors on the Stokes vector, even for $j \neq k$. Figure 1 shows how fast these off-diagonal terms drop as $|j-k|$ increases. A camera duty cycle with $r = 1$ was assumed for that figure.

In order to illustrate the effects of the covariances (27) on the seeing noise on the measured Stokes parameters, we introduce the Stokes gradient matrix, $G_{ij} = \nabla S_i \cdot \nabla S_j$, which allows us to recast Equation (27) in matrix form,

$$\mathbf{Cov}_{jk}(\Delta t) = E(\mathbf{S}_{j} - \mathbf{S}_j || \mathbf{S}_k - \mathbf{S}_k) = [\kappa^2 \Delta t^2 (\text{MGM})^T]_{jk} F_{ij}(\Delta t). \quad (30)$$

Here we used Equation (28), and the fact that $F_{j-k}(\Delta t) = F_{-j-k}(\Delta t)$, as indicated by the explicit functional form of those integrals in practical cases. Consequently, Equation (15) can also be written in matrix form as

$$\sigma^2(S'_i) = (\text{D Cov}(\Delta t) \text{D}^T)_{ii}. \quad (31)$$

We conclude this section by studying modifications of covariances (27) arising from the use of AO. The general effect is an important reduction of the low-frequency part of the seeing power spectrum. Figure 2 shows an example based on the model of seeing correlations described by Equation (29). Those curves
should be compared qualitatively with the models of the seeing used by Judge et al. (2004). Here, we define the power spectrum $S(\omega)$ as the Fourier transform of the auto-correlation function (e.g., Mandel & Wolf 1995):

$$S(\omega) = \int_{-\infty}^{+\infty} \Gamma(t) e^{-i\omega t} dt = \Gamma(0) \int_{-\infty}^{+\infty} \gamma(t) e^{-i\omega t} dt,$$  \hspace{1cm} (32)

where $\gamma(t) = \Gamma(t)/\Gamma(0)$. The thin curve in Figure 2 represents the power spectrum of the uncorrected seeing motion as described by the model of Equation (29), whereas the thick curve shows a simple analytic modification of this spectrum with effects that could be representative of the action of AO. Note how the high-frequency part of the spectrum is not modified by the AO correction. These covariances, however, change sign, so one does not converge to zero any faster than in the absence of AO correction. On the other hand, the covariances do not converge to zero any faster than in the absence of AO correction. These covariances, however, change sign, so one could in principle adopt modulation schemes and frequencies such that some of the covariances are either vanishing, or even contributing a negative term to the expression of $\sigma^2(S')$, thus possibly reducing the final error.

5.1. An Illustrative Example

In order to illustrate the effect of seeing on polarization measurements, based on the results derived above, we will consider two popular, step-wise modulation schemes: a “Stokes-definition” scheme and a “balanced” scheme, with ideal modulation matrices given respectively by

$$\mathbf{M}_{\text{Def}} \equiv \left( \begin{array}{cccc} 1 & +1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & +1 & 0 \\ 1 & 0 & 0 & -1 \end{array} \right),$$

$$\mathbf{M}_{\text{Bal}} \equiv \left( \begin{array}{cccc} 1 & +\frac{1}{\sqrt{3}} & +\frac{1}{\sqrt{3}} & +\frac{1}{\sqrt{3}} \\ 1 & +\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ 1 & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & +\frac{1}{\sqrt{3}} \\ 1 & -\frac{1}{\sqrt{3}} & +\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{array} \right).$$  \hspace{1cm} (35)

Both schemes provide maximum modulation efficiencies for all Stokes parameters ($1$ for $S_1$, and $1/\sqrt{3}$ for $S_i$ with $i \geq 2$). The respective modulation matrices in the dual-beam case are derived according to Equation (23). We consider here the case of perfect balancing of the two beams.
In the case of the Stokes-definition scheme, by design, only one of the Stokes parameters contributes at any time to any given intensity signal combined from the two beams, so that all cross-talk terms between different Stokes parameters are eliminated. In addition, the seeing-induced Stokes variations enter the Stokes-definition scheme only through terms that are diagonal in $\nabla S_j$. In the case of a balanced modulation scheme, instead, one must take into account general covariance terms that depend on both $\nabla S_j$ and $\nabla S_k$, for $j, k = 2, 3, 4$, which complicate the expressions of the seeing-induced polarization cross-talk.

For illustration consider a particular case with $\nabla S_2 = \nabla S_3 = 0$, and $\nabla S_1$ parallel to $\nabla S_4$. This example might represent realistic distributions of magnetic fields on the solar surface, such as those of magnetic bright points associated with emerging flux, observed near disk center. Introducing then the quantities $g_1 = |\nabla S_1|$ and $g_4 = |\nabla S_4|$ in Equation (31), we find, for the Stokes-definition scheme,

$$\sigma^2(S'_1) = \frac{1}{6} \kappa^2 \Delta t^2 g_1^2 \left[ \mathcal{T}_0(\Delta t) + \frac{1}{3} \sum_{s=1}^{5} (6 - s) \mathcal{T}_s(\Delta t) \right] \tag{36a}$$

$$\sigma^2(S'_2) = \sigma^2(S'_1) = 0, \tag{36b}$$

$$\sigma^2(S'_4) = \frac{1}{2} \kappa^2 \Delta t^2 g_4^2 \left[ \mathcal{T}_0(\Delta t) + \frac{1}{2} \mathcal{T}_1(\Delta t) \right] \tag{36c}$$

while for the balanced modulation scheme,

$$\sigma^2(S'_1) = \frac{1}{4} \kappa^2 \Delta t^2 g_1^2 \left[ \mathcal{T}_0(\Delta t) + \frac{1}{2} \sum_{s=1}^{3} (4 - s) \mathcal{T}_s(\Delta t) \right] \tag{37a}$$

$$\sigma^2(S'_2) = \frac{1}{4} \kappa^2 \Delta t^2 g_4^2 \left[ \mathcal{T}_0(\Delta t) - \mathcal{T}_2(\Delta t) - \frac{1}{2} \mathcal{T}_3(\Delta t) \right] \tag{37b}$$

$$\sigma^2(S'_4) = \frac{1}{4} \kappa^2 \Delta t^2 g_4^2 \left[ \mathcal{T}_0(\Delta t) + \frac{1}{2} \sum_{s=1}^{3} (4 - s) \mathcal{T}_s(\Delta t) \right] \tag{37c}$$

$$\sigma^2(S'_2) = \frac{1}{4} \kappa^2 \Delta t^2 g_4^2 \left[ \mathcal{T}_0(\Delta t) - \mathcal{T}_2(\Delta t) - \frac{1}{2} \mathcal{T}_3(\Delta t) \right] \tag{37d}$$

If we assume that the exposure time, $\Delta t$, is sufficiently large compared with the characteristic time of the seeing, $t_0$, then we can neglect all $\mathcal{T}_i(\Delta t)$ terms with $s \neq 0$ in the above equations. In such case, we see that the two schemes are affected by the same total error on the inferred Stokes vector, under identical conditions of camera exposure and time duration of the observation. This is because there are $6/4 = 1.5$ more modulation cycles for the balanced scheme than for the Stokes-definition scheme during the same time interval. Equations (36) also show that the seeing-induced error in the Stokes-definition scheme only affects the Stokes parameters that have non-vanishing gradients at the entrance of the modulator. In particular, $g_4$ only induces an error on $S_4$. In contrast, in the balanced scheme, the same error is evenly distributed among all of $S_2, S_3,$ and $S_4$.

It is important to remark that the variances expressed by Equations (36) and (37) strictly apply to Stokes vectors entering the modulator of a polarization-free telescope. In general, instead, typical telescopes’ Mueller matrices map the gradient vector $(\nabla S_1, 0, 0, \nabla S_4)$ at the entrance to the telescope onto a new gradient vector $(\nabla S'_1, \nabla S'_2, \nabla S'_3, \nabla S'_4)$ at the entrance to the modulator, so that all inferred Stokes parameters are affected by seeing-induced errors. In such a case, there is no evident advantage in adopting the Stokes-definition scheme over a balanced modulation scheme.

6. THE EFFECT OF SEEING DURING A FULL OBSERVATIONAL SEQUENCE

The results of Section 5 apply to elemental observations that consist of a single modulation cycle. The typical mode
of operation of slit-based spectropolarimeters (and of some imaging polarimeters, such as the IMaX instrument on board the Sunrise balloon mission; Martínez Pillet et al. 2011) is instead of operation of slit-based spectropolarimeters (and of some

The Astrophysical Journal

analysis to derive ˜

We then introduce the box function

To this purpose, we first note that the elements

expression of the Stokes variances, Equation (15), and extend it to the case of a series N of modulation cycles. Substituting Equation (27), we find

σ2(S′ i ) = \frac{\kappa^2}{2\pi} \sum_{j,k=1}^{nN} \nabla S_p \cdot \nabla S_q \int_{-\Delta t/2}^{+\Delta t/2} m_p(t + t_j) m_q(t' + t_k) dt dt',

\times m_q(t' + t_k) \Gamma(t' - t + t_k - t_j) dt dt'.

We then introduce the box function

\Pi_a(t) = \begin{cases} 1, & |t| \leq a/2 \\ 0, & |t| > a/2, \end{cases}

substitute Equation (33), and finally operate the changes of variable \( \xi = t + t_j \) and \( \xi' = t' + t_k \). We find

σ2(S′ i ) = \frac{\kappa^2}{2\pi} \sum_{p,q=1}^{4} \nabla S_p \cdot \nabla S_q \int_{-\Delta t/2}^{+\Delta t/2} m_p(t + t_j) \int_{-\Delta t/2}^{+\Delta t/2} m_q(t' + t_k) dt dt',

\times m_q(t' + t_k) \Gamma(t' - t + t_k - t_j) dt dt'.

To make this formula applicable to specific modulation schemes and seeing realizations, we must evaluate the Fourier transform \( \widetilde{H}_{ij}(\omega) \) of the functions

\widetilde{H}_{ij}(\xi) = \frac{1}{\Delta t} \sum_{k=1}^{nN} d_{ik} \Pi_\Delta(\xi - t_k) m_j(\xi).

To this purpose, we first note that the elements \( d_{ik} \) are periodic in \( k \) with period \( n \), so we can rewrite

\widetilde{H}_{ij}(\xi) = \frac{1}{\Delta t} \sum_{k=1}^{n} \sum_{l=0}^{N-1} \Pi_\Delta(\xi - t_k - l\Delta t/r) m_j(\xi)

\times \int_{-\Delta t/2}^{+\Delta t/2} m_p(t + t_j) \frac{e^{-i\omega \xi}}{\omega} d\xi

\times \int_{-\Delta t/2}^{+\Delta t/2} m_q(t' + t_k) \frac{e^{-i\omega \xi'}}{\omega} d\xi'.

(38)

Note that \( d_{ik} \) is an element of the demodulation matrix corresponding to the extended measurement of \( N \) cycles. Such matrix contains a factor \( 1/N \) with respect to the analogous matrix for one cycle. Therefore, we can replace \( d_{ik} \) with the standard (one-cycle) demodulation matrix, and drop the factor \( N \) from front of Equation (41). Because the matrix \( \tilde{H}(\omega) \) always appears in a product with \( \tilde{H}^\dagger(\omega) \), we can drop all common phase factors from its definition, so we can rewrite

\widetilde{H}_{ij}(\omega) = \frac{\sin(\omega \Delta t/2)}{\sin(\omega T/2)} \sum_{k=1}^{n} d_{ik} m_j e^{-i(k-1)\omega T/N}.

(42)

where we also used the fact that \( Nn\Delta t/r = T \), i.e., the duration of one elemental observation. In the case of dual-beam polarimetry, the summation in Equation (42) is extended to \( 2n \), while at the same time \( (k - 1) \) in the exponential must be taken modulo \( n \).

We note that \( \tilde{H}_{ip}(\omega)\tilde{H}_{jp}^\dagger(\omega) \rightarrow \delta_{ip} \delta_{jq} \) for vanishing \( \Delta t \). Therefore, only \( G_{il} = |\nabla S_i|^2 \) contributes to \( \sigma^2(S'_i) \) in Equation (40) for \( \Delta t \rightarrow 0 \). This is in complete agreement with the result that can be derived from Equations (30) and (31) under the same limit conditions. The results presented by Lites (1987) and Judge et al. (2004) correspond evidently to the diagonal case \( p = q \).
Figure 5. Plots of the Stokes errors of Equation (40) normalized to the incoming intensity, as a function of the modulation frequency. The case shown corresponds to the balanced scheme of Equation (35) in single-beam configuration, with non-zero gradients of the Stokes parameters, such that $g_1 = g_4 = S_1$ arcsec$^{-1}$ and $g_2 = g_3 = 0.1 g_1$. Measured power spectra of seeing-induced image motions both with and without AO correction were used to produce these plots (Judge et al. 2004). This figure shows the clear benefit of AO correction in terms of a reduction of the seeing-induced errors by more than an order of magnitude for $S_1$ and $S_2$. For modulation periods larger than the seeing correlation time, i.e., for long exposures, the variances of $S_2$, $S_3$, and $S_4$ are dominated by cross-talk from gradients in $S_1$. When the modulation frequency is increased to the point that seeing-induced displacements are practically frozen for the duration of a modulation cycle, the cross-talk terms in each error become negligible compared with the diagonal terms. The use of AO implies shorter correlation times of the seeing-induced displacements, and hence a higher modulation frequency is required to satisfy this condition.

For the dual-beam case (Figure 6) with perfectly balanced beams, there is no cross-talk from $S_1$ to $S_2$, $S_3$, and $S_4$ (see Section 4). This fact is illustrated by the drop in the error curves for $S_2$, $S_3$, and $S_4$ with respect to the single-beam case of Figure 5. However, the cross-talk among $S_2$, $S_3$, and $S_4$ remains. Analogously to the single-beam case, at high modulation frequencies, the seeing-induced error on each element of the Stokes vector tends to the contribution from the diagonal term only. Comparing the plots of Figures 5 and 6, we see that for very high modulation frequencies (100 Hz without AO or 1 kHz with AO) the performance of single- and dual-beam modulation schemes becomes comparable, and the only benefit of a dual-beam polarimeter in such case is the redundancy of polarimetric information, leading to a reduction of the photon noise by a factor of $\sqrt{2}$ with respect to the single-beam case.

Figure 6. Same as Figure 5, but for the dual-beam configuration with perfectly balanced beams.

Figure 7. Same as Figure 6, but for the case of a stepped modulator (with eight states) consisting of a waveplate with 150° retardance, identical to the one considered in Lites (1987). The same demodulation scheme of Lites (1987) was adopted for this calculation. These plots are for a dual-beam configuration with perfectly balanced beams. The gradients of the Stokes parameters are $g_2 = g_3 = S_1$ arcsec$^{-1}$ and $g_1 = g_4 = 0$. The red curves (see online version of the figure) show the predicted cross-talk level neglecting terms of the form $g_2 g_3$, which are missing in the treatment of Lites (1987) and Judge et al. (2004). (A color version of this figure is available in the online journal.)
In the work of Lites (1987) and Judge et al. (2004), the “variances” there defined depend explicitly on the total observation time and the modulation frequency, through the integral of the product of the seeing power spectrum with a real function of frequency that depends implicitly on both (cf. Equation (15) of Lites 1987, and points 1–5 of Section 2 of Judge et al. 2004). In our formalism, the total integration time does not appear explicitly because Equation (12) represents the variance on a single measurement of \( \tilde{I} \). However, if a number \( N \) of such measurements are made, and those measurements can be considered statistically uncorrelated, then the variance on the average signal \( \tilde{I} \) is reduced by a factor \( N \). In such case, for a fixed modulation frequency, but increasing the total observation time, we expect the same qualitative scaling of variance with the integration time as derived in previous work. Second, the dependence on the modulation frequency and the seeing power spectrum is also apparent within our formalism—already in the case of a single measurement—when we consider Equation (9).

Incidentally, we note how the power spectrum of the seeing appears most naturally in our approach as the Fourier transform of the two-time correlation function of the seeing displacement vector (see Section 5), whereas in previous work it is identified instead with the modulus square of the Fourier transform of the seeing displacement.\(^8\) The correspondence between these two approaches to the definition of the power spectrum of a stationary random process is clearly described by Mandel & Wolf (1995).

The importance of the covariances \( E(\{\tilde{I}_j - \tilde{I}\}_j|\{\tilde{I}_k - \tilde{I}\}_k\}) \) in typical cases has fundamental implications for the concept of polarimetric measurements. Through our analysis we are able to quantify the significance of seeing-induced correlations between measurements corresponding to different modulation steps, and how these correlations decay for decreasing modulation frequencies. Based on those results, it must be expected that seeing-induced correlations between different modulation states typically extend beyond the time interval of one modulation cycle. In other words, the Stokes vector measurements corresponding to different modulation cycles during an elemental observation are in general statistically correlated. Under this condition, we must expect that the variance of the average signal \( \tilde{I} \) will obey the ordinary scaling law by the total number \( N \) of cycles of the elemental observation only approximately. Formally, one should consider instead such elemental observation as a single measurement that is realized through the totality of the \( N \) modulation cycles, and thus characterized by a corresponding \( nN \times 4 \) modulation matrix. The variance of this measurement can then be determined through the usual equations, adopting such extended definition of the modulation and demodulation matrices (see Section 6). A similar approach is taken also in the earlier work (Lites 1987; Judge et al. 2004), where these correlations are made manifest by performing a spectral analysis of the demodulated signal over the entire time interval of the elemental observation.

Our formalism reveals however that, in the work of Lites (1987) and Judge et al. (2004), the covariance terms corresponding to contributions proportional to \( \nabla \cdot \mathbf{S} \), with \( i \neq j \) (see Equation (27)), are not accounted for. In fact, Lites (1987) derives the variances as Equation (12) of that paper directly from Equation (11), which are explicitly of diagonal form. In other words, looking at our Equations (30) and (31), previous work has computed the seeing-induced noise always under the assumption that the gradient matrix \( \mathbf{G} \) was diagonal. The additional off-diagonal components would arise in the development by Lites (1987) and Judge et al. (2004) when properly evaluating the expectation value of \( (\hat{O}_i - \bar{O}_i)^2 \) in the notation of Lites (1987). Equations (30) and (31) also clarify the vanishing of the cross-talk terms in the limit of large modulation frequencies \( (\chi \Delta t \rightarrow 0) \). Because all the integrals \( \tilde{I}_{j=k}((\Delta t)) \) tend to the same value in such a limit, the vector of the Stokes variances becomes simply proportional to the diagonal of the gradient matrix, \( \mathbf{G} \).

We are grateful to our HAO colleague B. Lites for a careful reading of the manuscript and helpful comments. We thank F. Wöger (NSO) for insightful discussion on the phenomenology of atmospheric seeing and of AO corrections, and for a critical reading of Section 2. We also acknowledge the valuable criticism and helpful comments by the referees of a previous version of this manuscript.

APPENDIX A

BEAM IMBALANCE INTRODUCED BY A POLARIZING GRATING

We consider a beam of unpolarized light incident on a diffraction grating. Typically the diffraction efficiency of a grating is different in the \( p \) and \( s \) directions, so the diffracted beam consists of a mixture of two orthogonally polarized beams with generally different intensities \( S_p \) and \( S_s \). Let us consider the case in which these two beams get separated by a perfectly polarizing beam splitter placed after the grating. The emerging beams will then have intensity \( S_p \) and \( S_s \), respectively.

Let us consider the combined signal of intensity \( S \) so defined,

\[
\tilde{k} S = k_p S_p + k_s S_s, \quad \tilde{k} = k_p + k_s,
\]

where \( k_p \) and \( k_s \) are the two gain factors applied to the detector to balance the two beams. The ratio \( k_p/k_s \) then corresponds to the ratio of the two orthogonal efficiencies of the grating, \( \epsilon_p \) and \( \epsilon_s \), according to

\[
k_p/k_s = \epsilon_s/\epsilon_p. \tag{A1}\]

The photon noise associated with the combined signal \( S \) is given by

\[
\sigma^2(S) = \left( \frac{\partial S}{\partial S_p} \right)^2 \sigma^2(S_p) + \left( \frac{\partial S}{\partial S_s} \right)^2 \sigma^2(S_s)
\]

\[
= \left( \frac{k_p}{k} \right)^2 \sigma^2(S_p) + \left( \frac{k_s}{k} \right)^2 \sigma^2(S_s).
\]

Correspondingly, the relative error is given by

\[
\frac{\sigma^2(S)}{S^2} = \left( \frac{k_p S_p}{k S} \right)^2 \frac{\sigma^2(S_p)}{S_p^2} + \left( \frac{k_s S_s}{k S} \right)^2 \frac{\sigma^2(S_s)}{S_s^2}.
\]

Evidently, for balanced beams

\[
\frac{\sigma^2(S)}{S^2} = \left( \frac{k_p}{k} \right)^2 \frac{\sigma^2(S_p)}{S_p^2} + \left( \frac{k_s}{k} \right)^2 \frac{\sigma^2(S_s)}{S_s^2}.
\]

\(\text{References}\)

\(8\) We observe that the ordinary Fourier transform is not well defined for the seeing displacement vector, because the associated random process is not limited in time. One can get around this problem by defining a generalized Fourier transform of the seeing, as the limit of ordinary Fourier transforms of finite samples of the seeing for increasing duration of those samples (e.g., Mandel & Wolf 1995).
If we express the signals in terms of photon flux, assuming Poisson’s statistics for the photon counts, and indicating with $r_N$ the readout noise of the camera, we can rewrite Equation (A3) as

$$\sigma^2(S) = \frac{1}{S^2} \left[ \frac{\sigma^2(S_p)}{S_p^2} + \frac{\sigma^2(S_c)}{S_c^2} \right].$$

(A3)

If we indicate with $N_{in}$ the photon count before the grating, evidently $N_p = (1/2)\kappa_p N_{in}$ and $N_s = (1/2)\kappa_s N_{in}$, so we find

$$\sigma^2(S) = \frac{1}{2N_{in}} \left[ \frac{\epsilon_p + \epsilon_s}{\epsilon_p \epsilon_s} \right] + \frac{r_N^2}{N_{in}^2} \left[ \frac{\epsilon_p^2 + \epsilon_s^2}{\epsilon_p \epsilon_s} \right].$$

(A4)

where we also introduced the grating’s average efficiency $\bar{\epsilon} = (1/2)(\epsilon_p + \epsilon_s)$. We note that the applicability of the above treatment in practical cases relies on the previous knowledge of the grating efficiencies, $\epsilon_p$ and $\epsilon_s$. These must be determined as a function of wavelength from flat fielding, ideally using unpolarized radiation. If the incident beam is instead weakly polarized (as it is typically the case for sunlight flat fields), the above treatment is still valid, but Equation (A1) can only provide an approximate estimate of the appropriate gain factors, $\kappa_s$ and $\kappa_p$, if the measured efficiencies of the grating are adopted in that equation.

**APPENDIX B**

**ROTATING MODULATOR**

We consider the case of a retarding device, which is continuously rotating with angular frequency $\Omega$, and which at $0^\circ$ position is described by the Mueller matrix

$$\mathbf{\mu}_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \mu_{22} & \mu_{23} & \mu_{24} \\ 0 & \mu_{32} & \mu_{33} & \mu_{34} \\ 0 & \mu_{42} & \mu_{43} & \mu_{44} \end{pmatrix},$$

where the following norm conditions must be satisfied,

$$\sum_i \mu_{ij}^2 \leq 1, \quad \sum_j \mu_{ij}^2 \leq 1.$$

We note that a full modulation cycle corresponds to only half rotation of the modulator, because of the characteristic 180° periodicity of polarization modulation.

In order to determine the cross-talk terms (40) for such a device, we need to compute the Fourier transforms of the functions (see Equation (39))

$$Z_{jk}(\xi) = \Pi_k(\xi) m_j(\xi + t_k), \quad j = 1, \ldots, 4, \quad k = 1, \ldots, n,$$

where $n$ is the number of modulation states (camera exposures) in the modulation cycle. These Fourier transforms are given by the following expressions:

$$\tilde{Z}_{1k}(\omega) = \Delta t \sin(\omega \Delta t/2),$$

$$\tilde{Z}_{2k}(\omega) = \frac{1}{2} (\mu_{22} + \mu_{33}) \Delta t \sin(\omega \Delta t/2) + \frac{1}{4} [\mu_{22} - \mu_{33} + i(\mu_{23} + \mu_{32})] e^{i(\Omega t + \pi(k-1)/n)} \Delta t \sin((\omega - 4\Omega) \Delta t/2) + \frac{1}{4} [\mu_{22} - \mu_{33} - i(\mu_{23} + \mu_{32})] e^{-i(\Omega t + \pi(k-1)/n)} \Delta t \sin((\omega + 4\Omega) \Delta t/2),$$

$$\tilde{Z}_{3k}(\omega) = \frac{1}{2} (\mu_{32} - \mu_{23}) \Delta t \sin(\omega \Delta t/2) + \frac{1}{4} [\mu_{22} - \mu_{33} + i(\mu_{23} + \mu_{32})] e^{i(\Omega t + \pi(k-1)/n)} \Delta t \sin((\omega - 4\Omega) \Delta t/2) - \frac{i}{4} [\mu_{22} - \mu_{33} + i(\mu_{23} + \mu_{32})] e^{-i(\Omega t + \pi(k-1)/n)} \Delta t \sin((\omega + 4\Omega) \Delta t/2),$$

$$\tilde{Z}_{4k}(\omega) = \frac{1}{2} (\mu_{24} + i\mu_{34}) e^{i(\Omega t + \pi(k-1)/n)} \Delta t \sin((\omega - 2\Omega) \Delta t/2) + \frac{1}{2} (\mu_{24} - i\mu_{34}) e^{-i(\Omega t + \pi(k-1)/n)} \Delta t \sin((\omega + 2\Omega) \Delta t/2),$$

where $\Delta t$ is the exposure time for each modulation state, and $t_k$ is the time at which the modulator is found in the $0^\circ$ position.

**REFERENCES**

Asensio Ramos, A., & Collados, M. 2008, Appl. Opt., 47, 2541
Born, M., & Wolf, E. 1965, Principles of Optics (3rd ed.; Pergamon: Oxford)
Collados, M. 2008, Proc. SPIE, 7012, 17
Draper, N. R., & Smith, H. 1966, Applied Regression Analysis (Wiley: New York)
Fried, D. L. 1965, J. Opt. Soc. Am., 55, 1427
Goodman, J. W. 1996, Introduction to Fourier Optics (2nd ed.; McGraw-Hill: New York)
Judge, P. G., Elmore, D. F., Lites, B. W., Keller, C. U., & Rimmelle, T. 2004, Appl. Opt., 43, 3817
Lites, B. W. 1987, Appl. Opt., 26, 3838
Mandel, L., & Wolf, E. 1995, Optical Coherence and Quantum Optics (Cambridge: Cambridge Univ. Press)
Martinez Pillet, V., del Toro Iniesta, J. C., Alvarez-Herrero, A., et al. 2011, Sol. Phys., 268, 57
Noll, R. J. 1976, J. Opt. Soc. Am., 66, 207
Rimmelle, T., & the ATST Team 2008, Adv. Space Res., 42, 78
Scherrer, P. H., Schou, J., Bush, R. I., Kosovichev, A. G., & Bogart, R. S. 2012, Sol. Phys., 275, 207
Shimizu, T., Tsuneta, S., Hara, H., et al. 2011, Proc. SPIE, 8148, 10
Tatarski, V. I. 1961, Wave Propagation in a Turbulent Medium (Dover: New York)
Tsuneta, S., Ichimoto, K., Katsukawa, Y., et al. 2008, Sol. Phys., 249, 167
Van Cittert, P. H. 1934, Physics, 1, 201
Zernike, F. 1938, Physics, 5, 785