Ensemble Estimation of Polarization Ellipse Parameters

By A.T. Walden and T. Medkour

Department of Mathematics, Imperial College London, 180 Queen's Gate, London SW7 2BZ, UK.

An ellipse describes the polarized part of a partially polarized quasi-monochromatic plane wave field. Its azimuth angle and aspect ratio are functions of the elements of the covariance matrix associated with the polarized part at a particular time instant. Given an ensemble of K independent samples at that time, the distributions of the estimators of these parameters are derived. The estimation is thus based on a sample ensemble at any time, and does not assume temporal stationarity. Additionally, the azimuth angle estimator has an angular distribution so that non-standard statistical methods are needed when deriving its mean and standard deviation.

Keywords: aspect ratio; azimuth; complex Gaussian distribution; polarization ellipse; Stokes parameters

1. Introduction

When a quasi-monochromatic (narrow-band) plane wave propagates in the $z-$ direction of a Cartesian coordinate system, it is found that, in the $x-y$ plane perpendicular to the direction of travel, the end point of the vector traces out an instantaneous ellipse, whose shape changes continuously with time, (e.g., Brosseau 1998); the wave is partially polarized. A wave whose ellipse maintains a constant azimuth angle and aspect ratio, but whose size varies, is called completely polarized, and if no regular pattern is exhibited the wave is unpolarized.

As well as being of fundamental importance in optics, polarization is of great interest in many other fields such as astronomy (Simmons and Stewart 1985), atmospheric science (Hayashi 1979), oceanography (Emery & Thomson 1998), geophysics (Park *et al* 1987) and elsewhere (Schreier 2007). Parameters describing the state of polarization can be calculated from observations but these are only estimates of the true parameter values (e.g., Simmons and Stewart 1985). Modern data-acquisition systems can often record multiple, (K, say), independent views of a propagating wave in the form of a set of discrete-time finite-length data series. For example, the presence of an ultra-low-frequency (ULF) wave in the solar magnetic field was captured simultaneously by $K = 4$ spacecraft in the Cluster mission, an international solar physics experiment, in February 2003 (Archer *et al* 2005); see Figure 1 and § 6. Assuming spatial homogeneity, these multiple samples can be used to reduce variance in the estimation of instantaneous polarization ellipse parameters, a procedure known as ensemble averaging. Considering the instantaneous ellipse describing the polarized portion of the signal, we shall look at estimators of its *azimuth* angle, the angle which the major axis of the ellipse makes with the x-direction, and its

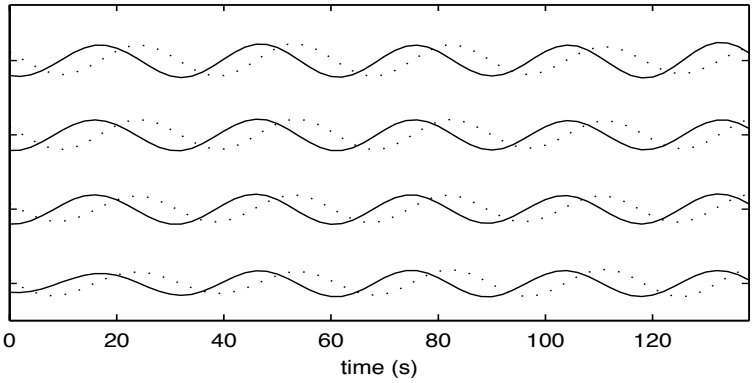

Figure 1. Narrowband observations of ULF wave, February 2003, by four spacecraft in the Cluster mission. In each case the solid line is the component $\mathcal{V}_1(t)$ measured parallel to the ecliptic plane, and the dotted line is the component $\mathcal{V}_2(t)$ measured perpendicular to the ecliptic plane.

aspect ratio, the ratio of length of the minor axis to that of the major axis, (signed according to right- or left-handedness). New statistical results are required for the estimation of polarization parameters from such data and are the subject of this paper.

A commonly-used model for a partially polarized plane wave is the complex representation whereby $Z_1(t)$ and $Z_2(t)$, say, are the analytical signal representations of the x and y components of the field (Barakat 1985): $Z_j(t) = \mathcal{V}_j(t) + \mathrm{i}\tilde{\mathcal{V}}_j(t),\ j = 1, 2$, where $\mathcal{V}_j(t)$ is the real field component and $\tilde{\mathcal{V}}_j(t)$ is its Hilbert transform. The real-valued stochastic processes $\mathcal{V}_1(t)$ and $\mathcal{V}_2(t)$ are taken to be zero mean and Gaussian distributed, justified by the central limit theorem (Brosseau 1998, sec. 3.3). Let $\boldsymbol{Z}_k = [Z_{1,k}, Z_{2,k}]^T, k = 0, \ldots, K-1$, denote the K independent complex-valued observations ('random samples') of $[Z_1(t), Z_2(t)]^T$ at an arbitrary time, where superscript 'T' denotes transpose. Associated with these observations is an instantaneous sample Hermitian covariance matrix of the form

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{K} \sum_{k=0}^{K-1} \begin{bmatrix} |Z_{1,k}|^2 & Z_{1,k} Z_{2,k}^* \\ Z_{1,k}^* Z_{2,k} & |Z_{2,k}|^2 \end{bmatrix} = (1/K) \sum_{k=0}^{K-1} \boldsymbol{Z}_k \boldsymbol{Z}_k^H, \tag{1.1}$$

where superscript '$*$' denotes complex conjugate, superscript 'H' denotes Hermitian (complex-conjugate) transpose, and the samples have zero mean and positive-definite covariance matrix $E\{\hat{\boldsymbol{\Sigma}}\} = \boldsymbol{\Sigma}$. Since $\hat{\boldsymbol{\Sigma}}$ is constructed from K independent complex-valued vectors, the parameter K is also the number of complex degrees of freedom in $\hat{\boldsymbol{\Sigma}}$, and $\lim_{K\to\infty} \hat{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}$.

Regarding the partially polarized quasi-monochromatic wave field as an incoherent superposition of a fully polarized monochromatic wave field and a completely unpolarized wave field, $\boldsymbol{\Sigma}$ can be uniquely decomposed as (Born & Wolf 1970),

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_P + \boldsymbol{\Sigma}_N, \tag{1.2}$$

respectively, where $\boldsymbol{\Sigma}_P$ has a determinant of zero. The estimators are based on the sample covariance matrix $\hat{\boldsymbol{\Sigma}}$ of (1.1) at any arbitrary time and could thus

be described as providing instantaneous estimates of the azimuth angle and aspect ratio. The results in Barakat (1985) are a special case of our results, since his results correspond to taking $K = 1$; also the azimuth angle estimator has an angular distribution so that we provide non-standard statistical methods for deriving its mean and standard deviation.

In § 2 we carefully define the azimuth angle and aspect ratio of the ellipse describing the polarized portion of the signal. Some fundamental statistical background is provided in § 3, and the forms of the parameter estimators are derived from the K-sample covariance matrix of (1.1). In § 4 we derive the probability density function (PDF), mean and standard deviation of the azimuth estimator, taking into account that it has an angular distribution, and a simulation experiment is carried out showing that the standard definitions of these quantities are, by comparison, quite unsatisfactory. The PDF of the aspect ratio is discussed in § 5. The ULF wave in the solar magnetic field is analysed in § 6. Our results are summarized in § 7.

2. Stokes and ellipse parameters

We write the matrices in (1.2) as

$$\mathbf{\Sigma} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^* & \Sigma_{22} \end{bmatrix}; \qquad \mathbf{\Sigma}_P = \begin{bmatrix} \Sigma_B & \Sigma_D \\ \Sigma_D^* & \Sigma_C \end{bmatrix}; \qquad \mathbf{\Sigma}_N = \begin{bmatrix} \Sigma_A & 0 \\ 0 & \Sigma_A \end{bmatrix},$$

where $\Sigma_B, \Sigma_C, \Sigma_A \geq 0$, and $\det\{\mathbf{\Sigma}_P\} = \Sigma_B\Sigma_C - |\Sigma_D|^2 = 0$, where 'det' denotes determinant. The elements of these matrices are given by (Born & Wolf 1970)

$$\begin{aligned} \Sigma_A &= \tfrac{1}{2}(\Sigma_{11} + \Sigma_{22}) - \tfrac{1}{2}[(\Sigma_{11} - \Sigma_{22})^2 + 4|\Sigma_{12}|^2]^{1/2} \\ \Sigma_B &= \tfrac{1}{2}(\Sigma_{11} - \Sigma_{22}) + \tfrac{1}{2}[(\Sigma_{11} - \Sigma_{22})^2 + 4|\Sigma_{12}|^2]^{1/2} \\ \Sigma_C &= \tfrac{1}{2}(\Sigma_{22} - \Sigma_{11}) + \tfrac{1}{2}[(\Sigma_{11} - \Sigma_{22})^2 + 4|\Sigma_{12}|^2]^{1/2}; \quad \Sigma_D = \Sigma_{12}. \end{aligned}$$

If $\mathrm{Im}\{\Sigma_{12}\} \neq 0$, then $\mathrm{Im}\{\Sigma_D\} \neq 0$, and the polarized part $\mathbf{\Sigma}_P$ of $\mathbf{\Sigma}$ is associated with elliptical motion in the original coordinates. If, further, $\Sigma_{11} = \Sigma_{22}$, and $\mathrm{Re}\{\Sigma_{12}\} = 0$, then $\Sigma_B = \Sigma_C$, and the polarization component is circular (Hayashi 1979), a special case of elliptical polarization.

Associated with the *partial polarization* covariance matrix, $\mathbf{\Sigma}$, are the usual Stokes parameters (Eliyahu, 1994) given by,

$$\begin{aligned} s_0(\mathbf{\Sigma}) &= \Sigma_{11} + \Sigma_{22}; \quad s_1(\mathbf{\Sigma}) = \Sigma_{11} - \Sigma_{22} \\ s_2(\mathbf{\Sigma}) &= \Sigma_{12} + \Sigma_{21} = 2\mathrm{Re}\{\Sigma_{12}\}; \quad s_3(\mathbf{\Sigma}) = i[\Sigma_{21} - \Sigma_{12}] = 2\mathrm{Im}\{\Sigma_{12}\}. \end{aligned}$$

For notational simplicity we denote $s_0(\mathbf{\Sigma}), s_1(\mathbf{\Sigma}), s_2(\mathbf{\Sigma}), s_3(\mathbf{\Sigma})$ by s_0, s_1, s_2 and s_3.

The Stokes parameters associated with the *full polarization* covariance matrix, $\mathbf{\Sigma}_P$, are

$$\begin{aligned} s_0(\mathbf{\Sigma}_P) &= \Sigma_B + \Sigma_C = [(\Sigma_{11} - \Sigma_{22})^2 + 4|\Sigma_{12}|^2]^{1/2} \\ s_1(\mathbf{\Sigma}_P) &= \Sigma_B - \Sigma_C = \Sigma_{11} - \Sigma_{22} = s_1 \\ s_2(\mathbf{\Sigma}_P) &= 2\mathrm{Re}\{\Sigma_D\} = 2\mathrm{Re}\{\Sigma_{12}\} = s_2 \\ s_3(\mathbf{\Sigma}_P) &= 2\mathrm{Im}\{\Sigma_D\} = 2\mathrm{Im}\{\Sigma_{12}\} = s_3. \end{aligned}$$

So $s_0^2(\mathbf{\Sigma}_P) = s_1^2(\mathbf{\Sigma}_P) + s_2^2(\mathbf{\Sigma}_P) + s_3^2(\mathbf{\Sigma}_P) = s_1^2 + s_2^2 + s_3^2$, while the other Stokes parameters are the same for $\mathbf{\Sigma}$ and $\mathbf{\Sigma}_P$.

The azimuth angle, ψ, satisfies (Born & Wolf 1970, pp. 31, 555)

$$\psi \;=\; \tfrac{1}{2}\tan^{-1}\left[\frac{s_2(\mathbf{\Sigma}_P)}{s_1(\mathbf{\Sigma}_P)}\right] = \tfrac{1}{2}\tan^{-1}\left[\frac{s_2}{s_1}\right] = \tfrac{1}{2}\tan^{-1}\left[\frac{2\mathrm{Re}\{\Sigma_{12}\}}{\Sigma_{11}-\Sigma_{22}}\right]. \quad (2.1)$$

Given s_2/s_1, no distinction can be made between a solution ψ_0 and a solution $\psi_0 \pm \pi/2$; there is not enough information to know which is which. Hence, as done by Eliyahu (1993) we in fact take ψ to be the angle of the major *or minor* axis of the ellipse from the x-direction, with $-\pi/4 \le \psi < \pi/4$; see Figure 2.

For circular polarization, both the numerator and denominator of the ratio in (2.1) are zero, and the angle is undefined, as it should be since the idea of the orientation of an ellipse axis is meaningless in the case of a circle.

Let $2a$ and $2b$ $(a \ge b)$ be the lengths of the major and minor axes of the ellipse, as in Figure 2, and $\chi, (-\pi/4 \le \chi < \pi/4)$, be the angle which characterizes the ellipticity and the sense in which the ellipse is being described. Then, (Born & Wolf 1970, p. 555), $\tan\chi = \pm b/a$ ($\chi \gtrless 0$ according as the polarization is right or left-handed) and in terms of the Stokes parameters

$$\sin 2\chi = s_3(\mathbf{\Sigma}_P)/s_0(\mathbf{\Sigma}_P) = s_3/[s_1^2 + s_2^2 + s_3^2]^{1/2}. \quad (2.2)$$

But $\cos 2\chi = [1-\sin^2 2\chi]^{1/2} = \left[(s_1^2+s_2^2)/(s_1^2+s_2^2+s_3^2)\right]^{1/2}$ for $(-\pi/4 \le \chi < \pi/4)$. If we define $y = \tan 2\chi$, and $\varepsilon = \tan\chi = \pm b/a$, then

$$y = \tan 2\chi = s_3/[s_1^2 + s_2^2]^{1/2}, \quad (2.3)$$

and

$$y = \tan 2\chi = (2\tan\chi)/(1-\tan^2\chi) = (2\varepsilon)/(1-\varepsilon^2), \quad -1 \le \varepsilon \le 1. \quad (2.4)$$

The only solution for ε in (2.4) for which $-1 \le \varepsilon \le 1$ is

$$\varepsilon = [-1 + (1+y^2)^{1/2}]/y, \quad (2.5)$$

a sigmoidal-shaped curve. Using (2.3) and (2.5) we see that the (signed) aspect ratio, ε, is given by

$$\varepsilon = [(s_1^2 + s_2^2 + s_3^2)^{1/2} - (s_1^2 + s_2^2)^{1/2}]/s_3. \quad (2.6)$$

The degree of polarization is (Born & Wolf 1970, p. 555)

$$P = (s_1^2 + s_2^2 + s_3^2)^{1/2}/s_0. \quad (2.7)$$

3. Statistical fundamentals

(a) *Complex Bivariate Gaussian distribution*

$\boldsymbol{Z}_0 = [Z_{1,0}, Z_{2,0}]^T$ is said to have the bivariate proper (or circular) complex Gaussian distribution with zero mean and covariance matrix $E\{\boldsymbol{Z}_0\boldsymbol{Z}_0^H\} = \mathbf{\Sigma}$,

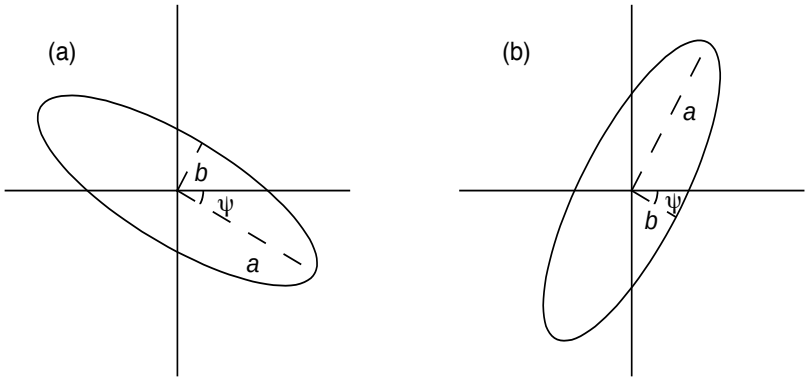

Figure 2. Illustration of the azimuth angle. ψ is the angle of (a) the major, and (b) the minor, axis of the ellipse from the x-direction. ψ is the same in both cases and $-\pi/4 \leq \psi < \pi/4$. The semi-major and semi-minor axes of lengths a and b are shown as dashed lines.

$\det\{\boldsymbol{\Sigma}\} > 0$, written $\boldsymbol{Z}_0 \sim \mathcal{N}_2^C(\boldsymbol{0}, \boldsymbol{\Sigma})$, if $\boldsymbol{Z}_0$ has PDF (Goodman 1963, Picinbono 1993, p. 122)

$$f_{\boldsymbol{Z}_0}(\boldsymbol{z}) \;=\; \frac{1}{\pi^2 \det(\boldsymbol{\Sigma})} \exp\left[-\frac{\Sigma_{22}|Z_1|^2 + \Sigma_{11}|Z_2|^2 - 2\mathrm{Re}(\Sigma_{12} Z_1^* Z_2)}{\det(\boldsymbol{\Sigma})}\right]. \qquad (3.1)$$

Under the Gaussian asumption on the real-valued stochastic processes $\mathcal{V}_1(t)$ and $\mathcal{V}_2(t)$, the random vector $\boldsymbol{Z}_k$ is obviously a bivariate complex Gaussian vector, but it is only proper and hence has probability density (3.1), if (Picinbono 1993, p. 120) $E\{\boldsymbol{Z}_k \boldsymbol{Z}_k^T\} = \boldsymbol{0}$. For this to be true we need to show that $E\{Z_{1,k}^2\} = E\{Z_{2,k}^2\} = 0$ and $E\{Z_{1,k} Z_{2,k}\} = E\{Z_{2,k} Z_{1,k}\} = 0$. These follow immediately from results on Hilbert transform relationships given in Bendat & Piersol (1986, Table 13.2, p. 499) and thus we conclude that $\boldsymbol{Z}_k$ has the bivariate proper (or circular) complex Gaussian distribution with PDF (3.1).

(*b*) *Complex Wishart distribution*

Let $\boldsymbol{Z} = [\boldsymbol{Z}_0, \boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{K-1}]$, where $\boldsymbol{Z}_0, \ldots, \boldsymbol{Z}_{K-1}$ are K independent bivariate proper complex Gaussian random vectors, each with the $\mathcal{N}_2^C(\boldsymbol{0}, \boldsymbol{\Sigma})$ distribution. When $K \geq 2$, the random 2×2 matrix $\boldsymbol{Z}\boldsymbol{Z}^H = \sum_{k=0}^{K-1} \boldsymbol{Z}_k \boldsymbol{Z}_k^H$ is full-rank and

$$\boldsymbol{W} \;=\; \boldsymbol{Z}\boldsymbol{Z}^H = \sum_{k=0}^{K-1} \boldsymbol{Z}_k \boldsymbol{Z}_k^H = \begin{bmatrix} w_{11} & w_{12} \\ w_{12}^* & w_{22} \end{bmatrix} \sim \mathcal{W}_2^C\{K, \boldsymbol{\Sigma}\}, \qquad (3.2)$$

i.e., $\boldsymbol{W}$ has the non-singular 2-dimensional complex central Wishart distribution with K complex degrees of freedom and mean $K\boldsymbol{\Sigma}$. Goodman (1963) pointed out that $\hat{\boldsymbol{\Sigma}}$ in (1.1) is the maximum likelihood estimator of the covariance matrix $\boldsymbol{\Sigma}$.

(*c*) *Parameter estimators*

From the sample covariance matrix $\hat{\boldsymbol{\Sigma}}$ in (1.1) can be defined the random variables,

$$\begin{array}{rclrcl} X_0 = s_0(\hat{\boldsymbol{\Sigma}}) & = & \hat{\Sigma}_{11} + \hat{\Sigma}_{22}; & X_1 = s_1(\hat{\boldsymbol{\Sigma}}) & = & \hat{\Sigma}_{11} - \hat{\Sigma}_{22}; \\ X_2 = s_2(\hat{\boldsymbol{\Sigma}}) & = & 2\mathrm{Re}\{\hat{\Sigma}_{12}\}; & X_3 = s_3(\hat{\boldsymbol{\Sigma}}) & = & 2\mathrm{Im}\{\hat{\Sigma}_{12}\}. \end{array} \tag{3.3}$$

From (2.1) the estimator for the azimuth ψ is the random variable

$$\Psi = \tfrac{1}{2}\tan^{-1}(X_2/X_1). \tag{3.4}$$

The estimator of the aspect ratio ε is, from (2.6), the random variable $\mathcal{E}$ given by

$$\mathcal{E} = \left[(X_1^2 + X_2^2 + X_3^2)^{1/2} - (X_1^2 + X_2^2)^{1/2}\right]/X_3. \tag{3.5}$$

Hence we seek the PDFs of (3.4) and (3.5).

4. Statistical properties of azimuth estimator

(*a*) *Distribution of* Ψ

The PDF can be derived via the unitary transformation matrix

$$\boldsymbol{A} = \tfrac{1}{\sqrt{2}}\begin{bmatrix} 1 & \mathrm{i} \\ 1 & -\mathrm{i} \end{bmatrix}.$$

Applying it to $\boldsymbol{Z}_0$ gives $\boldsymbol{Y}_0 \equiv \boldsymbol{A}\boldsymbol{Z}_0 \sim \mathcal{N}_2^C(\boldsymbol{0}, \boldsymbol{\Upsilon})$, where the Hermitian covariance matrix $\boldsymbol{\Upsilon}$ is given by $\boldsymbol{\Upsilon} = \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^H$ with elements

$$\begin{array}{rcl} \Upsilon_{11} & = & \tfrac{1}{2}(\Sigma_{11} + \Sigma_{22}) + \mathrm{Im}(\Sigma_{12}) = \tfrac{1}{2}[s_0 + s_3] \\ \Upsilon_{12} = \Upsilon_{21}^* & = & \tfrac{1}{2}(\Sigma_{11} - \Sigma_{22}) + \mathrm{iRe}(\Sigma_{12}) = \tfrac{1}{2}[s_1 + \mathrm{i}s_2] \\ \Upsilon_{22} & = & \tfrac{1}{2}(\Sigma_{11} + \Sigma_{22}) - \mathrm{Im}(\Sigma_{12}) = \tfrac{1}{2}[s_0 - s_3]. \end{array}$$

Then $|\Upsilon_{12}|^2/[\Upsilon_{11}\Upsilon_{22}] = (s_1^2 + s_2^2)/(s_0^2 - s_3^2) = \rho^2$, while $\arg\Upsilon_{12} = \tan^{-1}(s_2/s_1) = 2\psi$. We note also that ρ^2 is the squared correlation coefficient between the variables $(Z_1 + \mathrm{i}Z_2)/\sqrt{2}$ and $(Z_1 - \mathrm{i}Z_2)/\sqrt{2}$.

Now let $\boldsymbol{Y}_0, \ldots, \boldsymbol{Y}_{K-1}$ be K independent bivariate proper complex Gaussian random vectors, each with the $\mathcal{N}_2^C(\boldsymbol{0}, \boldsymbol{\Upsilon})$ distribution. The maximum likelihood estimator of the true covariance matrix $\boldsymbol{\Upsilon}$ is given by $\hat{\boldsymbol{\Upsilon}} = (1/K)\sum_{k=0}^{K-1}\boldsymbol{Y}_k\boldsymbol{Y}_k^H$. The random variable $\hat{\omega} = \arg\hat{\Upsilon}_{12}, -\pi/2 \le \hat{\omega} < \pi/2$, has PDF (Miller 1980, p. 91)

$$f_\Omega(\hat{\omega}) = \frac{(1-\rho^2)^K}{\pi}{}_2F_1\left(1, K; \tfrac{1}{2}; \rho^2\cos^2(\hat{\omega} - \omega)\right), \tag{4.1}$$

where $\omega = \arg\Upsilon_{12}, \rho^2 < 1, K \ge 2$. Here ${}_2F_1(\alpha_1, \alpha_2; \beta_1; z)$ is the hypergeometric function with 2 and 1 parameters, α_1, α_2 and β_1, and scalar argument z, which may be written explicitly as

$$\sum_{m=0}^{\infty}\frac{\Gamma(\alpha_1 + m)\Gamma(\alpha_2 + m)\Gamma(\beta_1)z^m}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\beta_1 + m)m!}. \tag{4.2}$$

(It is a special case of ${}_pF_q(\alpha_1, \ldots, \alpha_p; \beta_1, \ldots, \beta_q; z)$, the generalized hypergeometric series, defined by (Gradshteyn & Ryzhik 1980, p. 1045)

$$\sum_{m=0}^{\infty} \frac{(\alpha_1)_m \ldots (\alpha_p)_m}{(\beta_1)_m \ldots (\beta_q)_m} \frac{z^m}{m!}, \tag{4.3}$$

where $(y)_m$ is defined in terms of the Gamma function $\Gamma(\cdot)$ as $(y)_m = y(y+1)\ldots(y+m-1) = \Gamma(y+m)/\Gamma(y)$ with $(y)_0 = 1$.)

On transforming (4.1) with $\hat{\psi} = \hat{\omega}/2$ and $\psi = \omega/2$, we get

$$f_\Psi(\hat{\psi}) = 2f_\Omega(2\hat{\psi}) = C\ {}_2F_1\left(1, K; \tfrac{1}{2}; \rho^2\ \cos^2 2(\hat{\psi} - \psi)\right), \tag{4.4}$$

$-\pi/4 \le \hat{\psi} < \pi/4$, where $C = \frac{2}{\pi}(1-\rho^2)^K$ and

$$\rho^2 = (s_1^2 + s_2^2)/(s_0^2 - s_3^2) < 1. \tag{4.5}$$

The PDF in (4.4) is valid for $K = 1$ as well as $K \ge 2$. For $K = 1$, $f_\Psi(\hat{\psi}) = f^B(\hat{\psi}) + f^B(\hat{\psi} \pm \pi/2)$, where $f^B(\cdot)$ denotes Barakat's PDF, (Barakat 1985, eqn. 5.27), defined over $-\pi/2 \le \psi < \pi/2$; this is exactly what we should obtain (Eliyahu 1993, p. 2885). Of course the importance and novelty of (4.4) is its validity for $K \ge 2$. Note that the PDF in (4.4) depends on ρ^2 and the true value, ψ, of the azimuth. If we define

$$\delta = s_3/s_0 = (2P\varepsilon)/(1+\varepsilon^2), \tag{4.6}$$

then

$$\rho^2 = \left\{[(s_1^2 + s_2^2 + s_3^2)/s_0^2] - [s_3^2/s_0^2]\right\} / \left\{(s_0^2 - s_3^2)/s_0^2\right\} = (P^2 - \delta^2)/(1-\delta^2),$$

so that the degree of polarization, P, and the aspect ratio, ε, together determine ρ^2, and so the PDF could alternatively, but not so neatly, be directly parameterized in terms of these quantities.

The PDF $f_\Psi(\hat{\psi})$ is plotted in Figures 3(a)-(c) for three values of K, ψ and ρ^2, respectively. We see, that as K increases, then as expected, the distribution becomes concentrated around the true value of $\psi = 0.2$ in Figure 3(a). The same effect can be seen when ρ^2 increases for a fixed K in Figure 3(c). It is also easy to see that when $\psi = \pi/5$, in Figure 3(b), that the distribution shows its circular nature, clearly wrapping around in a circular manner. This is an important observation when thinking about the mean and variance of such a distribution, and we look at two specific examples to illustrate the more complicated approach required.

We consider the two covariance matrices, $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$, where

$$\mathbf{\Sigma}_1 = \begin{bmatrix} 6 & 7+i \\ 7-i & 10 \end{bmatrix}; \qquad \mathbf{\Sigma}_2 = \begin{bmatrix} 10 & -1+2i \\ -1-2i & 3 \end{bmatrix}. \tag{4.7}$$

The PDF $f_\Psi(\hat{\psi})$ is shown in Figure 4 for both matrices.

(*b*) *Mean and standard deviation for the angular distribution*

Since $f_\Psi(\hat{\psi}) = f_\Psi(\hat{\psi} + [\pi/2])$ the PDF corresponds to an *angular* distribution on the range $[-\pi/4, \pi/4)$. As a result, special techniques are required to sensibly

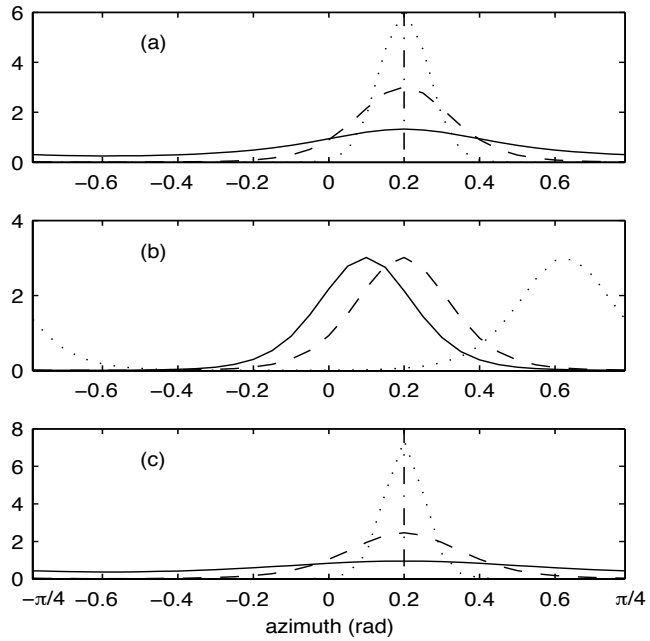

Figure 3. PDFs $f_\Psi(\hat{\psi})$. (a) $K = 1, 5, 20$ (solid, dashed, dotted) with $\rho^2 = 0.6, \psi = 0.2$. (b) $\psi = \pi/32, \pi/16, \pi/5$ (solid, dashed, dotted) with $\rho^2 = 0.6, K = 5$. (c) $\rho^2 = 0.1, 0.5, 0.9$ (solid, dashed, dotted) with $K = 5, \psi = 0.2$. (The true value of the azimuth, 0.2, is marked by a vertical dashed-dot line in (a) and (c).)

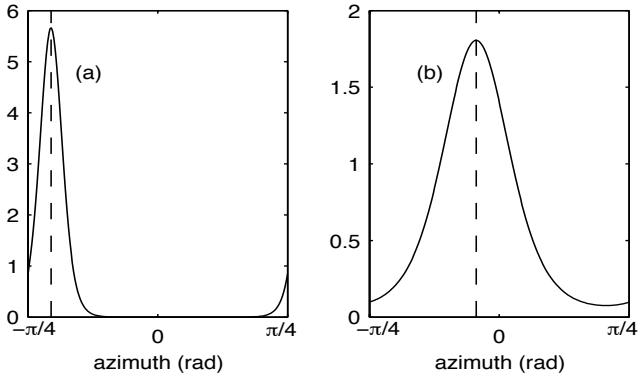

Figure 4. PDFs $f_\Psi(\hat{\psi})$ for (a) $\mathbf{\Sigma}_1$ and (b) $\mathbf{\Sigma}_2$. In both cases $K = 5$. The true value of the azimuth, ψ, is marked by a vertical dashed line.

define the mean and variance for Ψ; in particular, care must be taken with the fact that the distribution is defined on a quarter circle.

We look firstly at the mean direction. Let $A = \mathrm{e}^{\mathrm{i}4\Psi}$. The first trigonometric moment about zero is $\mu' = E\{A\} = E\{\mathrm{e}^{\mathrm{i}4\Psi}\}$. Now write $\mu' = a\mathrm{e}^{\mathrm{i}4\mu_0}$, where μ_0 is the mean direction. The first central trigonometric moment (about the mean

direction) is defined as

$$\mu = E\{e^{i4(\Psi-\mu_0)}\} = E\{\cos 4(\Psi-\mu_0)\} + iE\{\sin 4(\Psi-\mu_0)\} \quad (4.8)$$
$$= e^{-i4\mu_0}E\{e^{i4\Psi}\} = e^{-i4\mu_0}\mu' = e^{-i4\mu_0}ae^{i4\mu_0} = a, \quad (4.9)$$

so that, from (4.8),

$$E\{\sin 4(\Psi-\mu_0)\} = 0, \quad (4.10)$$

and the mean direction is hence given by the value μ_0 satisfying (4.10), agreeing with Mardia (1972, p. 70).

Using (4.2) we can write $f_\Psi(\hat{\psi})$ in (4.4) as

$$f_\Psi(\hat{\psi}) = \frac{C\,\pi^{1/2}}{\Gamma(K)}\sum_{m=0}^{\infty}\frac{\Gamma(K+m)}{\Gamma(m+\frac{1}{2})}\rho^{2m}\,\cos^{2m}[2(\hat{\psi}-\psi)]. \quad (4.11)$$

But $E\{\sin 4(\Psi-\mu_0)\} = E\{2\sin 2(\Psi-\mu_0)\cos 2(\Psi-\mu_0)\}$ and so

$$E\{\sin 4(\Psi-\mu_0)\} = \frac{2C\,\pi^{1/2}}{\Gamma(K)}\sum_{m=0}^{\infty}\frac{\Gamma(K+m)}{\Gamma(m+\frac{1}{2})}\rho^{2m}\,\mathcal{J}_m,$$

where $\mathcal{J}_m$ is given by

$$\int_{-\pi/4}^{\pi/4}\sin 2(\hat{\psi}-\mu_0)\cos 2(\hat{\psi}-\mu_0)\cos^{2m}[2(\hat{\psi}-\psi)]\mathrm{d}\hat{\psi}. \quad (4.12)$$

But $\mathcal{J}_m = 0$ if $\mu_0 = \psi$ since the overall function to be integrated is odd. Hence the solution to (4.10) is $\mu_0 = \psi$, and we see that the mean direction is precisely ψ, so that $\hat{\psi}$ is an unbiased estimator of the exact azimuth ψ.

The circular variance is a measure of circular dispersion for points on the unit circle (Mardia 1972, p. 21); for Ψ it is defined as (Mardia 1972, pp. 45, 71),

$$V_0 = 1 - E\{\cos 4(\Psi-\mu_0)\}. \quad (4.13)$$

After considerable algebra, it is found that V_0 may be written as

$$V_0 = \begin{cases} 1 & \text{if } \rho^2 = 0, \\ (1-\rho^{-2})\log(1-\rho^2) & \text{if } K = 1; 0 < \rho^2 < 1, \\ (1-\rho^{-2})[(1-\rho^2)^{K-1}-1]/(K-1) & \text{if } K \geq 2; 0 < \rho^2 < 1. \end{cases} \quad (4.14)$$

Now (4.13) means that V_0 takes values between 0 and 1, which is very different to the usual variance of a random variable which can take any positive value. To relate V_0 to the standard deviation on the positive real line, the circular standard deviation is defined as (Mardia 1972, p. 74),

$$\sigma_0 = [-2\log(1-V_0)]^{1/2}/4, \quad (4.15)$$

where the divisor of 4 takes into account that $-\pi/4 \leq \hat{\psi} < \pi/4$. σ_0 is plotted as a function of K and ρ^2 in Figure 5; it increases with decreasing K and ρ^2 and is maximized for $\rho^2 = 0$ which corresponds to circular polarization.

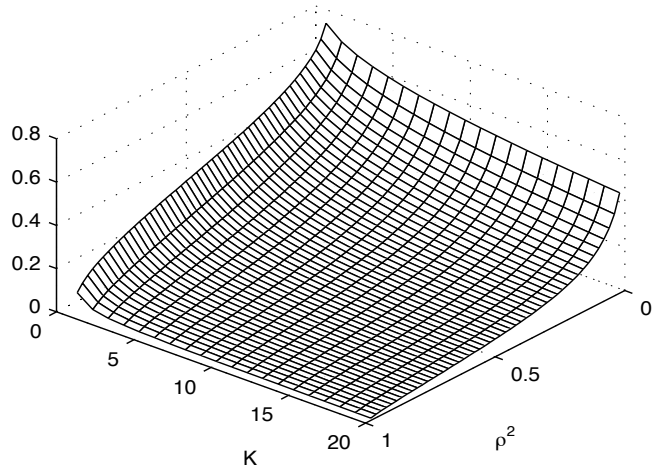

Figure 5. σ_0 as a function of K and ρ^2.

	μ_0	$\hat{\mu}_0$	$\hat{\mu}$	σ_0	$\hat{\sigma}_0$	$\hat{\sigma}$
$\mathbf{\Sigma}_1$	-0.646	-0.647	-0.587	0.078	0.078	0.273
$\mathbf{\Sigma}_2$	-0.139	-0.140	-0.122	0.247	0.244	0.257

Table 1. *Results of the simulation study on azimuth estimation for $K = 5$. See text for notation.*

Consider again the two covariance matrices $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$. The columns labelled μ_0 and σ_0 of Table 1 give the exact values of ψ and the circular standard deviations for these two matrices.

For $\mathbf{\Sigma}_1$ the PDF $f_\Psi(\hat{\psi})$ appears like a periodic Gaussian (Figure 4). The interval $\mu_0 \pm 1.96\sigma_0$ is $[-\pi/4, -0.493) \cup (0.771, \pi/4)$, and numerical integration of the PDF shows that this interval covers probability of 0.95, as in the Gaussian case.

For $\mathbf{\Sigma}_2$ the interval $\mu_0 \pm 1.96\sigma_0$ is $(-0.623, 0.345)$. The corresponding probability coverage is 0.93. It appears then that $\hat{\psi} \pm 1.96\sigma_0$ will be an approximate 95% confidence interval for $\psi = \mu_0$; in practice σ_0 will need to be estimated as $\hat{\sigma}_0$, say, so that

$$\psi \in (\hat{\psi} - 1.96\hat{\sigma}_0, \hat{\psi} + 1.96\hat{\sigma}_0) \tag{4.16}$$

defines a rough 95% confidence interval.

(*c*) *Simulation*

For each of the model covariance matrices $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ we simulated $\hat{\mathbf{\Sigma}}$ in (1.1) a number N of times. (For each of these repetitions $\boldsymbol{Z} = [\boldsymbol{Z}_0, \boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{K-1}]$ was simulated from $\mathcal{N}_2^C(\mathbf{0}, \mathbf{\Sigma})$ with $\mathbf{\Sigma}$ replaced by $\mathbf{\Sigma}_1$ or $\mathbf{\Sigma}_2$, using the technique given in Medkour & Walden (2007, section V). $K = 5$ was used. Each repetition allows the computation of a sample azimuth $\hat{\psi}$. So we obtain $\hat{\psi}_1, \ldots, \hat{\psi}_N$ with $-\pi/4 \le \hat{\psi}_j < \pi/4$. Since $\hat{\psi}_j$ covers only a quarter circle we next create $\tilde{\psi}_j = 4\hat{\psi}_j, j = 1, \ldots, N$

(Mardia 1972, p. 26). Then let $\bar{C} = (1/N)\sum \cos\tilde{\psi}_j$ and $\bar{S} = (1/N)\sum \sin\tilde{\psi}_j$ and calculate $\tilde{\mu} = \tan^{-1}(\bar{S}/\bar{C})$, where $-\pi/2 \le \tilde{\mu} < \pi/2$. The four-quadrant inverse tangent is then obtained from

$$\tilde{\mu}_0 = \begin{cases} \tilde{\mu} & \text{if } \bar{S}, \bar{C} > 0, \\ \tilde{\mu} + \pi & \text{if } \bar{C} < 0, \\ \tilde{\mu} + 2\pi & \text{if } \bar{S} < 0, \bar{C} > 0. \end{cases} \tag{4.17}$$

(The result produced from (4.17) is the same as that obtained with the mathematical function `atan2`$(\bar{S}, \bar{C})$, with $-\pi \le \tilde{\mu}_0 < \pi$.) Finally, the estimated mean direction is given by $\hat{\mu}_0 = \tilde{\mu}_0/4$. For $N = 10\,000$ repetitions, $\hat{\mu}_0$ was obtained as -0.647 for $\mathbf{\Sigma}_1$, and -0.140 for $\mathbf{\Sigma}_2$, very close to the exact azimuths of -0.646 and -0.139 respectively.

The sample circular variance is given by (Mardia 1972, p. 22)

$$\hat{V}_0 = 1 - \frac{1}{N}\sum_{j=1}^{N} \cos(\tilde{\psi}_j - \tilde{\mu}_0),$$

and then this is converted to a measure on the positive real line as in (4.15), $\hat{\sigma}_0 = [-2\log(1 - \hat{V}_0)]^{1/2}/4$. Using the $N = 10\,000$ repetitions, $\hat{\sigma}_0$ was obtained as 0.078 for $\mathbf{\Sigma}_1$,and 0.244 for $\mathbf{\Sigma}_2$, compared to the theoretical values of 0.078 and 0.247, respectively, given by (4.15).

(*d*) *Comparison with usual mean and standard deviation*

It is interesting to compare our results properly formulated for an angular distribution with what we would obtain using standard methods. We took our simulated values $\hat{\psi}_1, \ldots, \hat{\psi}_N$ and computed the sample mean, $\hat{\mu}$, and standard deviation, $\hat{\sigma}$, in the usual way, obtaining, for $\mathbf{\Sigma}_1$, -0.587 and 0.273, respectively, very different to the values of $\hat{\mu}_0 = -0.647$ and $\hat{\sigma}_0 = 0.078$. The standard method does not recognize that the distribution is angular, finding the mean as the centre of gravity of the PDF exactly as shown in Figure 4(a), and then inflating the standard deviation due to the probability density near $\pi/4$, to the right of the incorrect mean. By way of contrast, the properly formulated approach basically *circularizes* the PDF of Figure 4(a) before determining the centre of gravity and spread.

For $\mathbf{\Sigma}_2$ the sample mean and standard deviation of $\hat{\psi}_1, \ldots, \hat{\psi}_N$ are given by -0.122 and 0.257 compared to the values of $\hat{\mu}_0 = -0.140$ and $\hat{\sigma}_0 = 0.244$. The mean is affected by the increasing tail of the PDF near $\pi/4$ (Figure 4(b)), but the overall effect of this on the standard deviation is relatively small for this particular example.

5. Aspect Ratio

Using the estimators in (3.3), equations (2.3) and (2.4) give the random variable

$$Y = \frac{X_3}{[X_1^2 + X_2^2]^{1/2}} \tag{5.1}$$

$$= \frac{2\mathcal{E}}{1 - \mathcal{E}^2}, \tag{5.2}$$

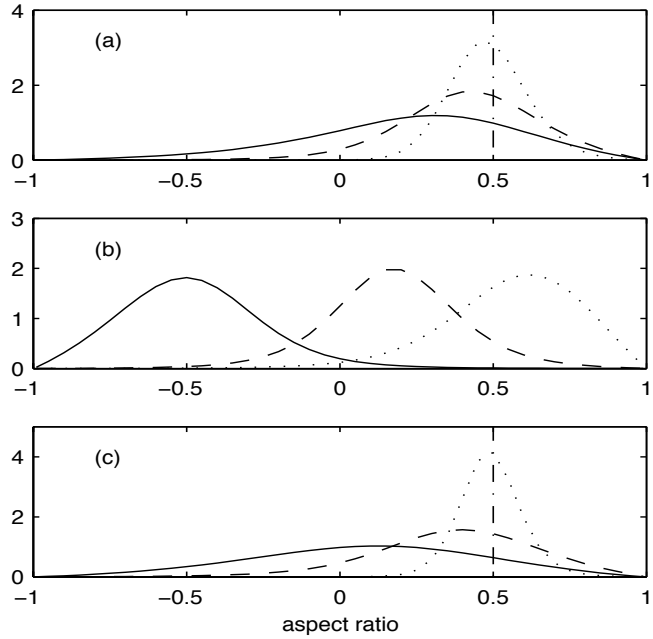

Figure 6. PDFs $f_{\mathcal{E}}(\hat{\varepsilon})$. (a) $K = 1, 5, 20$ (solid, dashed, dotted) with $P = 0.6, \varepsilon = 0.5$. (b) $\varepsilon = -0.6, 0.2, 0.8$ (solid, dashed, dotted) with $P = 0.6, K = 5$. (c) $P = 0.1, 0.5, 0.9$ (solid, dashed, dotted) with $K = 5, \varepsilon = 0.5$. (The true value of the aspect ratio, 0.5, is marked by a vertical dashed-dot line in (a) and (c).)

say, where $\mathcal{E}$ is given by (3.5). Firstly the PDF of Y in (5.1) may be derived, and then the PDF of $\mathcal{E}$ found by the transformation in (5.2). The resulting PDF, $f_{\mathcal{E}}(\hat{\varepsilon})$, is found in Appendix A. For $K = 1$,

$$f_{\mathcal{E}}(\hat{\varepsilon}) = \frac{(1-P^2)(1-\hat{\varepsilon}^2)(\hat{\varepsilon}^2 - 2\delta\hat{\varepsilon} + 1)}{\{(\hat{\varepsilon}^2 - 2\delta\hat{\varepsilon} + 1)^2 - (P^2 - \delta^2)(1-\hat{\varepsilon}^2)^2\}^{3/2}}, \tag{5.3}$$

which agrees with Barakat (1985, eqn. 5.9) as corrected in Eliyahu (1993, p. 2885). For $K \geq 2$,

$$\begin{aligned} f_{\mathcal{E}}(\hat{\varepsilon}) \;=\; & \frac{2^{K+1}\Gamma(K+[1/2])(1-P^2)^K(1+\hat{\varepsilon}^2)^{K-1}(1-\hat{\varepsilon}^2)}{\pi^{1/2}(\hat{\varepsilon}^2 - 2\delta\hat{\varepsilon} + 1)^{K+1}} \\ & \sum_{n=0}^{K-2} \frac{\Gamma(K-1+n)}{\Gamma(K-1-n)\Gamma(K+2+n)} \frac{(-\beta')^n}{n!} \times \\ & {}_4F_3(K, K+[1/2], [3/2], 2; 1, [K+n+2]/2, [K+n+3]/2; \alpha'), \end{aligned} \tag{5.4}$$

with α' and β' given in (A 7). This PDF is plotted in Figures 6(a)-(c) for three values of K, ε and P, respectively. We see, that as K increases, then as expected, the distribution becomes concentrated around the true value of $\varepsilon = 0.5$ in Figure 6(a). The same effect can be seen when the degree of polarization increases for a fixed K in Figure 6(c).

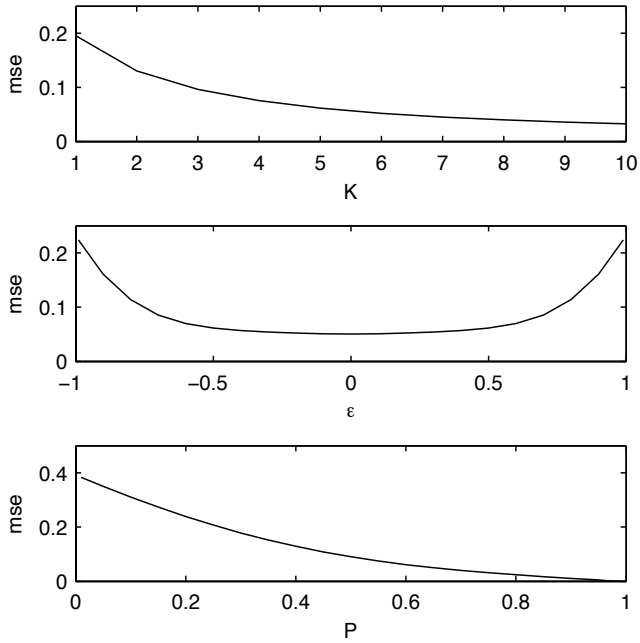

Figure 7. Mean squared error $E\{(\hat{\varepsilon}-\varepsilon)^2\}$ with varying (a) K, for $P=0.6, \varepsilon=0.5$, (b) ε, for $P=0.6, K=5$, and (c) P, for $K=5, \varepsilon=0.5$.

The mean squared error (MSE) of estimation of ε is given by $E\{(\hat{\varepsilon}-\varepsilon)^2\} = E\{\mathcal{E}^2\} - 2\varepsilon E\{\mathcal{E}\} + \varepsilon^2$ and comprises the bias (squared) plus variance; it is a more useful measure of the accuracy of estimation than either the bias or variance singly. Due to the complexity of the PDF the rth moments, $E\{\mathcal{E}^r\} = \int_{-1}^{1} \hat{\varepsilon}^r f_{\mathcal{E}}(\hat{\varepsilon}) \mathrm{d}\hat{\varepsilon}$, are most easily computed by numerical integration. The MSE is shown in Figure 7; it falls with increasing K or with increasing degree of polarization P. We also see that the MSE increases with the absolute value of the aspect ratio, so that, as would be expected, the estimate worsens the closer the polarization approaches circular.

6. ULF wave example

The two-component narrowband series for the ULF wave in the solar magnetic field recorded by each of the the four Cluster mission spacecraft were shown in Figure 1. The measurement unit is nanoTeslas, each series has 70 values, the sample interval is 2s, and the bandpass interval 0.02-0.05Hz. Archer *et al* (2005) found that the wave could be assumed uniform and planar over the separations of the craft. Figure 8 shows plots of the instantaneous estimated values of (a) ρ^2 defined in (4.5), estimated using $\hat{\rho}^2 = (X_1^2 + X_2^2)/(X_0^2 - X_3^2)$, (b) the azimuth ψ estimated using (3.4), and (c) the standard deviation of the azimuth estimator σ_0 given in (4.15) estimated using (4.14) with ρ^2 replaced by its estimate $\hat{\rho}^2$, (d) the degree of polarization P, defined in (2.7), estimated using $\hat{P} = (X_1^2 + X_2^2 + X_3^2)^{1/2}/X_0$, (e) the aspect ratio ε estimated using (3.5). The degree of polarization is just a little less than unity at all times.

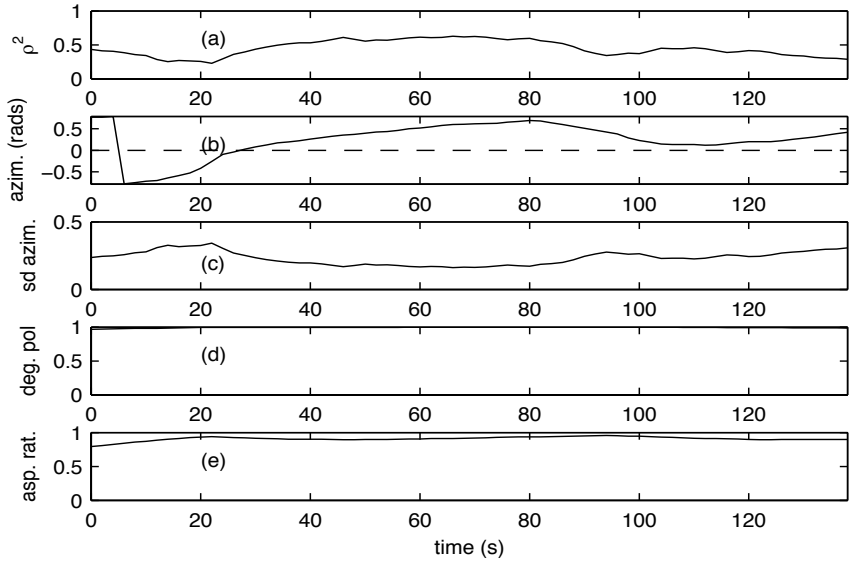

Figure 8. Analysis of ULF series. Instantaneous estimated values of (a) ρ^2, (b) azimuth ψ, (c) the standard deviation of the azimuth estimator, (d) the degree of polarization, P, and (e) the aspect ratio ε.

No assumption of temporal stationarity was necessary to produce the estimates of Figure 8. However, the PDFs for the azimuth and aspect ratio estimators, which assume Gaussianity, could change with time if any of the governing parameters in (4.4) and (5.4) are time-varying.

But, if a temporally stationary vector process is assumed, then the governing parameters will be time-invariant. Suppose we treat the estimated azimuth values in Figure 8(b) as a random sample of azimuth estimates from the distribution in (4.4) where we take $K = 4$, the now time-invariant ψ to be the estimated mean direction equal to 0.4, calculated as in §4c and the now time-invariant ρ^2 to be the median of the $\hat{\rho}^2$ values in Figure 8(a). A histogram of the sample values is compared to the resulting theoretical PDF in Figure 9(a). The fit appears quite good, despite the fact that the estimated azimuth values are of course correlated in time, rather than independent as for a random sample. Our distributional results thus appear consistent with temporal stationarity holding here.

Again, under the assumption of temporal stationarity, suppose we treat the estimated aspect ratio values in Figure 8(e) as a random sample of azimuth estimates from the distribution in (5.4) where we take $K = 4$, the now time-invariant ε to be the median of the $\hat{\varepsilon}$ values in Figure 8(e), and the now time-invariant P to be the median of the $\hat{P}$ values in Figure 8(d). A histogram of the sample values is compared to the theoretical PDF in Figure 9(b); again the fit is reasonable considering the necessarily crude approximations, and appears consistent with temporal stationarity.

The standard deviation of the azimuth estimator in Figure 8(c) defines, via (4.16), an approximate 95% confidence interval of the form $\pm 1.96\hat{\sigma}_0$ about the instantaneous estimate. With $\hat{\sigma}_0$ averaging about 0.3, we can see that the value of

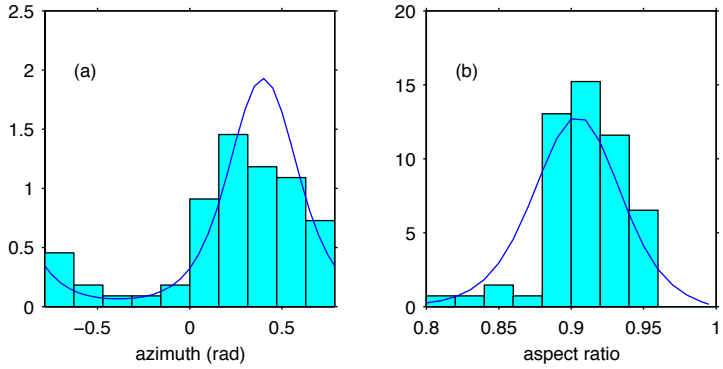

Figure 9. Sample histograms versus theoretical PDFs for (a) estimated azimuth, $\hat{\psi}$ and (b) estimated aspect ratio, $\hat{\varepsilon}$.

ψ is barely constrained within its range. The difficulty of estimating the azimuth is consistent with the fact that here polarization is near circular — the aspect ratio has a median of 0.91.

7. Summary

We have derived the statistical distributions for instantaneous estimators of the azimuth and aspect ratio of the polarization ellipse derived from a sample covariance (coherence) matrix formed from a number K of independent samples under the Gaussian assumption. We showed that the estimation accuracy improves as (i) K increases, as expected, or (ii) as ρ^2 increases (azimuth) or P increases (aspect ratio) for a fixed K. The distribution of the azimuth is angular and special techniques were used to derive its mean and standard deviation. An assumption of temporal stationarity of the vector process is not necessary for the validity of the theoretical results.

For the ULF wave data analysis, the empirical temporal distributions of azimuth and aspect ratio matched quite well the theoretical distributions with fixed parameters, suggesting that in this example at least temporal stationarity might hold.

Tarek Medkour thanks the government of the People's Democratic Republic of Algeria for financial support. The authors are very grateful to Dr. Tim Horbury for supplying the Cluster mission time series.

Appendix A.

We seek first the PDF of Y given by (5.1). For $K \geq 2$, the joint PDF of the four Stokes variables $(X_0,\, X_1,\, X_2,\, X_3)$ is (Touzi & Lopes 1996, eqn. 59)

$$f(x_0,\, x_1,\, x_2,\, x_3) \;=\; \frac{C_0}{H^K}(x_0^2 - x_1^2 - x_2^2 - x_3^2)^{K-2}\mathrm{e}^{-\frac{2K}{H}(s_0x_0 - s_1x_1 - s_2x_2 - s_3x_3)}, \quad \text{(A 1)}$$

where $C_0 = 2K^{2K}/[\pi\Gamma(K)\Gamma(K-1)]$ and $H = s_0^2 - s_1^2 - s_2^2 - s_3^2 = 4\det\{\mathbf{\Sigma}\} > 0$.

Starting with (A 1) we first transform to variables X_1, X_2, X_3 and $Z = X_0/(X_1^2 + X_2^2 + X_3^2)^{1/2}$ and integrating with respect to Z over $[1, \infty)$ using (Gradshteyn & Ryzhik 1980, 3.387(3)), we get,

$$f(x_1,x_2,x_3) \;=\; B_0(x_1^2+x_2^2+x_3^2)^{[K-(3/2)]/2} \exp\left[\frac{2K}{H}(s_1x_1+s_2x_2+s_3x_3)\right] \times \mathcal{K}_{K-(3/2)}\left[\frac{2K}{H}s_0(x_1^2+x_2^2+x_3^2)^{1/2}\right],$$

where $B_0 = 2K^{K+(3/2)}s_0^{(3/2)-K}/[(\pi H)^{3/2}\Gamma(K)]$ and $\mathcal{K}_n(\cdot)$ denotes the modified Bessel function of the second kind and order n.

Next we transform from X_1, X_2, X_3 to R, φ, Y where $X_1 = R\cos\varphi, X_2 = R\sin\varphi, X_3 = RY$. (We note that Y is thus of the form in (5.1).) Integrating with respect to φ over $(0, 2\pi)$ using Gradshteyn & Ryzhik (1980, 3.937(2)), we get,

$$f(r,y) \;=\; 2\pi B_0 r^{K+(1/2)}(1+y^2)^{[K-(3/2)]/2}\exp\left[\frac{2K}{H}s_3yr\right] \times \mathcal{K}_{K-(3/2)}\left[\frac{2K}{H}s_0r(1+y^2)^{1/2}\right]\mathcal{I}_0\left[\frac{2K}{H}r(s_1^2+s_2^2)^{1/2}\right].$$

$\mathcal{I}_0(\cdot)$ denotes the modified Bessel function of the first kind and order 0. But since

$$I_0(a) = \sum_{l=0}^{\infty}(a/2)^{2l}/(l!)^2, \tag{A 2}$$

(Gradshteyn & Ryzhik 1980, 8.447(1)), we can write $f(r, y)$ as

$$f(r,y) \;=\; 2\pi B_0(1+y^2)^{[K-(3/2)]/2}\exp\left[\frac{2K}{H}s_3yr\right]\mathcal{K}_{K-(3/2)}\left[\frac{2K}{H}s_0r(1+y^2)^{1/2}\right] \times \sum_{l=0}^{\infty}\left[\frac{K^2}{H^2}(s_1^2+s_2^2)\right]^l \frac{r^{K+2l+(1/2)}}{(l!)^2}.$$

Using Gradshteyn & Ryzhik (1980, 6.621(3)) to integrate with respect to r over $(0, \infty)$, followed by considerable manipulation yields,

$$f_Y(y) \;=\; B_1\frac{(1+y^2)^{K-(3/2)}}{[s_0(1+y^2)^{1/2}-s_3y]^{2K}} \times \sum_{l=0}^{\infty}\left[\frac{s_1^2+s_2^2}{4[s_0(1+y^2)^{1/2}-s_3y]^2}\right]^l \frac{\Gamma(2K+2l)\Gamma(2l+3)}{(l!)^2\Gamma(K+2l+2)} \times {}_2F_1\left(2K+2l, K-1; K+2l+2; \frac{s_3y+s_0(1+y^2)^{1/2}}{s_3y-s_0(1+y^2)^{1/2}}\right),$$

where $B_1 = H^K/[2\Gamma(K)]$. This is valid if $(2K/H)[s_0(1+y^2)^{1/2} - s_3y] > 0$. Now

$$s_3 \;=\; [s_1^2+s_2^2+s_3^2]^{1/2}\sin 2\chi = \left[\frac{[s_1^2+s_2^2+s_3^2]^{1/2}}{s_0}\right]s_0\sin 2\chi = Ps_0\sin 2\chi,$$

where P is the degree of polarization. But $|P| < 1$, since we have assumed that $\det\{\mathbf{\Sigma}\} > 0$. So, since $H, K, s_0 > 0$ also, the condition reduces to showing $g(y;q) = (1+y^2)^{1/2} - qy \ > 0$, where $q = P\sin 2\chi$ and $-1 < q < 1$. Differentiating $g(y,q)$ we find that there is a single minimum at $y = q/(1-q)^{1/2}$ at which $g(y,q)$ takes the positive value $(1-q^2)^{1/2}$. Hence $g(y,q) > 0$, as required.

The form of the PDF, $f(y)$, just found, is not convenient for computational purposes as it involves an infinite sum of hypergeometric functions. Using Abramowitz & Stegun (1965, 15.3.5) we can rewrite the hypergeometric function in the sum as

$$\left[\frac{2s_0(1+y^2)^{1/2}}{s_0(1+y^2)^{1/2} - s_3 y}\right]^{1-K} {}_2F_1\left(2-K, K-1; K+2l+2; \frac{s_3 y}{2s_0(1+y^2)^{1/2}} + \tfrac{1}{2}\right).$$

Hence,

$$f_Y(y) \ = \ B_1 \frac{(2s_0)^{1-K}(1+y^2)^{(K-2)/2}}{[s_0(1+y^2)^{1/2} - s_3 y]^{K+1}} \times$$
$$\sum_{l=0}^{\infty} \frac{\alpha^l \Gamma(2K+2l)\Gamma(2l+3)}{4^l (l!)^2 \Gamma(K+2l+2)} {}_2F_1(2-K, K-1; K+2l+2; \beta)\,,$$

where we have defined

$$\alpha = \frac{s_1^2 + s_2^2}{[s_0(1+y^2)^{1/2} - s_3 y]^2}; \quad \beta = \frac{s_3 y}{2s_0(1+y^2)^{1/2}} + \tfrac{1}{2}.$$

Using Abramowitz & Stegun (1965, 15.4.1), ${}_2F_1(2-K, K-1; K+2l+2; \beta)$ can be written as

$$\sum_{n=0}^{K-2} \frac{\Gamma(2-K+n)\Gamma(K-1+n)\Gamma(K+2l+2)}{\Gamma(2-K)\Gamma(K-1)\Gamma(K+2l+2+n)} \frac{\beta^n}{n!},$$

so

$$f_Y(y) \ = \ B_1 \frac{(2s_0)^{1-K}(1+y^2)^{(K-2)/2}}{[s_0(1+y^2)^{1/2} - s_3 y]^{K+1}} \sum_{n=0}^{K-2} \frac{\Gamma(2-K+n)\Gamma(K-1+n)}{\Gamma(2-K)\Gamma(K-1)} \frac{\beta^n}{n!} \times$$
$$\sum_{l=0}^{\infty} \frac{\alpha^l \Gamma(2K+2l)\Gamma(2l+3)}{4^l (l!)^2 \Gamma(K+2l+2+n)}.$$

Next we use the duplication formula for gamma functions (Abramowitz & Stegun 1965, 6.1.18), namely $\Gamma(2z) = \pi^{-1/2} 2^{2z-1} \Gamma(z)\Gamma(z+[1/2])$. Applying this to the gamma terms involving $2l$ in the infinite summation we get

$$f_Y(y) \ = \ \frac{B_1}{\pi^{1/2}} \frac{(2s_0)^{1-K}(1+y^2)^{(K-2)/2}}{[s_0(1+y^2)^{1/2} - s_3 y]^{K+1}} \sum_{n=0}^{K-2} 2^{K-n} \frac{\Gamma(2-K+n)\Gamma(K-1+n)}{\Gamma(2-K)\Gamma(K-1)} \frac{\beta^n}{n!} \times$$
$$\sum_{l=0}^{\infty} \frac{\Gamma(K+l)\Gamma(K+l+[1/2])\Gamma(l+[3/2])\Gamma(l+2)\alpha^l}{\Gamma([K+n]/2+1+l)\Gamma([K+n]/2+[3/2]+l)\Gamma(l+1)l!}.$$

But from (4.3), the infinite sum is equal to

$${}_4F_3(K, K+[1/2], [3/2], 2; 1, 1+[K+n]/2, [K+n+3]/2; \alpha)$$
$$\times \frac{\Gamma(K)\Gamma(K+[1/2])\Gamma(3/2)\Gamma(2)}{\Gamma([K+n+2]/2)\Gamma([K+n+3]/2)}.$$

Application of the duplication formula to the ratio of gamma functions shows that this term can be written as $\pi^{1/2}2^{-K+n}\Gamma(2K)\Gamma(3)/\Gamma(K+2+n)$. So

$$f_Y(y) = B_1 \frac{(2s_0)^{1-K}(1+y^2)^{(K-2)/2}}{[s_0(1+y^2)^{1/2}-s_3y]^{K+1}} \sum_{n=0}^{K-2} \frac{\Gamma(2-K+n)\Gamma(K-1+n)\Gamma(2K)\Gamma(3)}{\Gamma(2-K)\Gamma(K-1)\Gamma(K+2+n)} \frac{\beta^n}{n!}$$
$$\times {}_4F_3(K, K+[1/2], [3/2], 2; 1, 1+[K+n]/2, [K+n+3]/2; \alpha).$$

For computational purposes we need to evaluate $\Gamma(2-K+n)/\Gamma(2-K)$. We note that $\Gamma(z)\Gamma(1-z) = \pi/\sin\pi z$, (Abramowitz & Stegun 1965, 6.1.17). Hence,

$$\begin{aligned} \Gamma(2-K+n)\Gamma(K-1-n) &= \pi\sin\pi(2-K+n) = (-1)^n\pi\sin\pi(2-K) \\ \Gamma(2-K)\Gamma(K-1) &= \pi\sin\pi(2-K), \end{aligned}$$

so $[\Gamma(2-K+n)/\Gamma(2-K)] = (-1)^n[\Gamma(K-1)/\Gamma(K-1-n)]$. Then $f_Y(y)$ becomes

$$f_Y(y) = B_2 \frac{s_0^{1-K}(1+y^2)^{(K-2)/2}}{[s_0(1+y^2)^{1/2}-s_3y]^{K+1}} \sum_{n=0}^{K-2} \frac{\Gamma(K-1+n)}{\Gamma(K-1-n)\Gamma(K+2+n)} \frac{(-\beta)^n}{n!} \times$$
$${}_4F_3(K, K+[1/2], [3/2], 2; 1, 1+[K+n]/2, [K+n+3]/2; \alpha),$$

where $B_2 = B_1\Gamma(2K)\Gamma(3)$. Use of the duplication formula for gamma functions gives $B_2 = (2H)^K\pi^{-1/2}\Gamma(K+[1/2])$. This form for the PDF of Y consists of a finite sum of hypergeometric series.

Now from (5.1), $Y = 2\mathcal{E}/[1-\mathcal{E}^2]$, and the transformation is one-to-one, and its derivative is continuous on $(-1,1)$, so

$$f_{\mathcal{E}}(\hat{\varepsilon}) = 2\frac{(1+\hat{\varepsilon}^2)}{(1-\hat{\varepsilon}^2)^2} f_Y\left(\frac{2\hat{\varepsilon}}{1-\hat{\varepsilon}^2}\right). \tag{A 3}$$

The terms α and β now become

$$\alpha = \frac{(s_1^2+s_2^2)(1-\hat{\varepsilon}^2)^2}{[s_0(1+\hat{\varepsilon}^2)-2s_3\hat{\varepsilon}]^2}, \quad \beta = \frac{s_3\hat{\varepsilon}}{s_0(1+\hat{\varepsilon}^2)} + \tfrac{1}{2}, \tag{A 4}$$

and finally

$$f_{\mathcal{E}}(\hat{\varepsilon}) = 2B_2 \frac{s_0^{1-K}(1-\hat{\varepsilon}^2)(1+\hat{\varepsilon}^2)^{K-1}}{[s_0(1+\hat{\varepsilon}^2)-2s_3\hat{\varepsilon}]^{K+1}} \sum_{n=0}^{K-2} \frac{\Gamma(K-1+n)}{\Gamma(K-1-n)\Gamma(K+2+n)} \frac{(-\beta)^n}{n!} \times$$
$${}_4F_3(K, K+[1/2], [3/2], 2; 1, 1+[K+n]/2, [K+n+3]/2; \alpha), \tag{A 5}$$

the desired PDF for $\mathcal{E}$, valid for $K \geq 2$ and $-1 \leq \hat{\varepsilon} \leq 1$.

For $K = 1$ the calculations are different but similar in terms of the steps involved; the joint distribution of X_1, X_2, X_3 is given by (Eliyahu 1994, eqn. 10)

$$f(x_1,x_2,x_3) = \frac{1}{\pi H}(x_1^2+x_2^2+x_3^2)^{-(1/2)}\mathrm{e}^{-\frac{2}{H}(s_0[x_1^2+x_2^2+x_3^2]^{1/2}-s_1x_1-s_2x_2-s_3x_3)}.$$

Next we transform from X_1, X_2, X_3 to R, φ, Y where $X_1 = R\cos\varphi, X_2 = R\sin\varphi$, $X_3 = RY$. Integrating with respect to φ over $(0, 2\pi)$ using Gradshteyn & Ryzhik (1980, 3.937(2)), we get,

$$f(r,y) = \frac{2r}{H(1+y^2)^{1/2}} \exp\left[-\frac{2}{H}[s_0(1+y^2)^{1/2}-s_3y]r\right] \mathcal{I}_0\left[\frac{2r}{H}(s_1^2+s_2^2)^{1/2}\right].$$

Expanding I_0 again as in (A 2), and integrating with respect to r over $(0,\infty)$ using Gradshteyn & Ryzhik (1980, 8.310(1)) we obtain

$$f_Y(y)=\frac{H[s_0(1+y^2)^{1/2}-s_3y]}{2(1+y^2)^{1/2}\{[s_0(1+y^2)^{1/2}-s_3y]^2-(s_1^2+s_2^2)\}^{3/2}}.$$

Transforming from Y to $\mathcal{E}$ using (A 3), gives

$$f_{\mathcal{E}}(\hat{\varepsilon}) = \frac{H(1-\varepsilon^2)[s_0(1+\varepsilon^2)-2s_3\varepsilon]}{\{[s_0(1+\varepsilon^2)-2s_3\varepsilon]^2-(s_1^2+s_2^2)(1-\varepsilon^2)^2\}^{3/2}}, \tag{A 6}$$

the PDF for $\mathcal{E}$ when $K=1$, with $-1\le\hat{\varepsilon}\le 1$.

Taking δ as in (4.6), and noting that α and β in (A 4) can be written as

$$\alpha' = \frac{(P^2-\delta^2)(1-\hat{\varepsilon}^2)^2}{(\hat{\varepsilon}^2-2\delta\hat{\varepsilon}+1)^2}, \quad \beta' = \frac{\delta\hat{\varepsilon}}{(1+\hat{\varepsilon}^2)}+\tfrac{1}{2}, \tag{A 7}$$

then the PDF of the estimator of the aspect ratio can be written in terms of the degree of polarization, P, the true aspect ratio ε, and the number of samples, K. For $K=1$ by rewriting (A 6) we get (5.3) and for $K\ge 2$, by rewriting (A 5) we get (5.4).

References

Abramowitz, M & Stegun, I. 1965 *Handbook of Mathematical Functions*. New York: Dover.

Archer, M. Horbury, T. S., Lucek, E. A., Mazelle, C., Balogh, A. & Dandouras, I. 2005 Size and shape of ULF waves in the terrestrial foreshock. *J. Geophys. Res.* **A110**, A05208.

Barakat, R. 1985 The statistical properties of partially polarized light. *Opt. Acta* **32**, 295–312.

Bendat, J. S. & Piersol, A. G. 1986 *Random Data*, 2nd Ed. New York: John Wiley.

Born, M. & Wolf, E. 1970 *Principles of Optics*, 4th Ed. Cambridge University Press.

Brosseau, C. 1998 *Fundamentals of Polarized Light: A Statistical Optics Approach*. Wiley.

Eliyahu, D. 1993 Vector statistics of correlated Gaussian fields. *Phys. Rev. E* **47**, 2881–92.

Eliyahu, D. 1994 Statistics of Stokes variables for correlated Gaussian fields. *Phys. Rev. E* **50**, 2381–4.

Emery, W. J. & Thomson, R. E. 1998 *Data Analysis Methods in Physical Oceanography*. Pergamon.

Goodman, N. R. 1963 Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). *Ann. Math. Stat.* **34**, 152–77.

Gradshteyn, I. S. & Ryzhik, I. M. 1980 *Table of Integrals, Series, and Products*, Corrected and Enlarged Edition. NY: Academic Press.

Hayashi, Y. 1979. Space-time spectral analysis of rotary vector series. *J. Atmosph. Sci.* **36** 757-66.

Mardia, K. V. 1972 *Statistics of Directional Data*. NY: Academic Press.

Medkour, T. & Walden, A. T. 2007 A variance equality test for two correlated complex Gaussian variables with application to spectral power comparison. *IEEE Trans. Signal Process.* **55**, 881-8.

Miller, K. S. 1980 *Hypothesis Testing with Complex Distributions*. NY: Robert E. Krieger.

Park, J., Vernon III, F. L. & Lindberg, C. R. 1987 Frequency dependent polarization analysis of high-frequency seismograms. *J. Geophys. Res* **B92**, 12664–74.

Picinbono, B. 1993 *Random Signals and Systems*. Englewood Cliffs, NJ: Prentice-Hall.

Schreier, P. J. 2007 Polarization analysis of nonstationary random signals. Preprint.

Simmons J. F. L. & Stewart, B. G. 1985 Point and interval estimation of the true unbiased degree of linear polarization in the presence of low signal-to-noise ratios. *Astron. Astrophys.* **142**, 100–6.

Touzi, R. & Lopes, A. 1996 Statistics of the Stokes parameters and of the complex coherence parameters in one-look and multilook speckle fields. *IEEE Trans. Geosci. Remote Sens.* **34**, 519–31.