Polarimetric Modeling and Parameter Estimation with Applications to Remote Sensing

Bertrand Hochwald, Student Member, IEEE, and Arye Nehorai, Fellow, IEEE

Abstract—We develop and analyze two parametric models in which electromagnetic plane waves carrying polarimetric information are received. The first model considers estimation of the polarimetric response of a surface by measuring the reflections of actively generated waves. The second considers estimation of the polarization of passively generated waves. Both models have applications to remote sensing. We propose a natural parametrization of the distribution of the received signal. Using the Cramér-Rao bound, we characterize the best possible accuracy of unbiased estimators of these parameters. Simple estimators are given. Both models are fitted into a common framework and compared.

I. INTRODUCTION

ACTIVE methods of determining the characteristics of a scatterer by reflecting an actively generated known signal from its surface have had a long history in the literature. Some employ models that have been in existence since the 1940's (see, for example, [1], [10], and [30]). Since the work of [21] and [29] in the 1950's, the polarization phenomenon has become an integral part of these models. The change in polarization of an electromagnetic plane wave, upon reflection, is represented by a scattering matrix. The relationships between the statistical distribution of the entries of the scattering matrix, or the surface's polarimetric response, and the composition of a surface have been documented and are well known in many cases. We propose and analyze a parametric active model to estimate this distribution.

Passive methods of polarimetry, which are sometimes called radiometry, are usually concerned with determining the polarization of an incoming plane wave. Unlike in active polarimetry, the arrival direction of the wave may not be known. Therefore, to complement the active model, we also develop a passive model that assumes unknown direction. In this model, the processes of estimating the signal's polarization and direction are intertwined. Both the passive and active models have applications to remote sensing. A description of many commonly used polarimetric systems and their uses in remote sensing of the earth's oceans and vegetation may be found, for example, in [12]. The wide range of recent research activities in polarimetry is exemplified in [24], and an introduction to many of them can be found in [4].

The models we use assume that noisy measurements of the complete electromagnetic field are available. To provide the measurements, electromagnetic vector sensors are used. These sensors have only recently come to the attention of the signal processing community; the vector sensor's application to passive direction finding was introduced in [25]. Electromagnetic vector sensors as measuring devices are commercially available and actively researched. EMC Baden Ltd., in Baden, Switzerland, and Flam & Russell, Inc., in Horsham, PA, USA, are companies that manufacture them for signals with a variety of frequencies. Lincoln Laboratory at the Massachusetts Institute of Technology has performed some electromagnetic plane wave localization tests with vector sensors [15]. Some examples of recent research on sensor development are [19] and [20].

Our work uses, but is not limited to, vector sensors. Two- and 1-D sensor analyses can be taken as special cases of the vector-sensor analysis. Because we assume that the 3-D electric and magnetic components of the received wave are measured, the analysis provided herein gives a good idea of what can be achieved, within the limits of the physical theory, when all of the measurable quantities at one point in space are used.

The active and passive models in this paper are intertwined in a way such that results from one are useful for the other. We show their application to terrain identification via remote sensing and provide Cramér-Rao bounds (CRBs) [37] on the variance of unbiased estimators of the polarimetric parameters. The CRB represents the best possible accuracy that can be expected from any unbiased estimator. We propose estimators of the polarimetric parameters and offer a treatment of many aspects of polarimetry that have not been statistically modeled or analyzed before. Some previous analyses include Ulaby et al. [36], where the distributions of some polarimetric quantities are studied, Kong et al. [23], where target classification based on polarimetry is considered, and Novak et al. [26], where detection algorithms are examined.

The active model assumes that a known signal is being reflected from a "distributed" target whose average or large-scale surface normal is known. This assumption is valid in many remote sensing applications. A distributed surface has many randomly placed point scatterers or a randomly varying dielectric composition. Such a surface gives rise to random complex scattering coefficients that affect the polarization of the returned wave. We present a way to decompose the
covariance matrix of the coefficients into parameters that assign the covariance a geometry.

The passive model differs from the active model in that we assume the received signal originates from the surface or object itself, and its direction is unknown. The decomposition of the coefficient covariance matrix used in the active model applies to the signal covariance in this model. We show that the decomposition provides a physically natural parameter set.

A plane wave has electric and magnetic field components varying in a plane, and it is often thought that by using two antennas with orthogonal polarization sensitivities (one for each dimension), all of the information about a signal can be extracted from the combined measurements. However, if we a priori do not know the exact orientation of the plane carrying the fields, as in the passive model, then measurements in all three spatial dimensions are important. The active model does not require the complete 3-D measurement for parameter estimation (as is well known in polarimetry, two dimensions are sufficient), but the common 3-D framework for the passive and active models simplifies the performance comparison. Nevertheless, our analysis of the active model applies almost unchanged to standard 2-D sensors as well.

In Section II, both models are developed. In Section III, we give conditions under which the scattering coefficients have a complex Gaussian distribution. The conditions are of interest because the complex Gaussian distribution, while often used, is rarely justified. The distributions of the observations for both models are derived in Section IV. We compute the Fisher information matrix (FIM) entries (inverse CRB) for the models in Section V. In Section VI, a remote sensing application of the models is presented and analyzed in detail.

We then propose a reparametrization via a decomposition of the scattering coefficient (active) or signal (passive) covariance matrix. The new parameters lend themselves to straightforward physical interpretations and assign a geometry to the covariance matrices. It is shown in the active model that under the new parametrization, the CRB matrix is block diagonal if a certain commonly used diversely polarized signal is transmitted. In the passive model, the parametrization reduces the dependency of the CRB on the coordinate system, thus showing that the variables are intrinsic to the signal and not its direction. In Section VII, we derive simple estimators of the scattering coefficient and signal covariance matrices. The same diversely polarized signal that made the CRB matrix block diagonal in the active model is shown to decouple estimation of a nuisance parameter in the remote sensing application. In Section VIII, we briefly consider applications of the results to models with more traditional (2-D and scalar) sensors and then conclude.

II. THE ACTIVE AND PASSIVE POLARIMETRIC MODELS

A. Assumptions

The electromagnetic vector sensor receiving device measures the complete electromagnetic field corrupted by noise. We assume that the necessary preprocessing is done so that the complex envelopes of the measured signals are available.

Assumptions:

A1) The transmitter-scatterer-sensor distances are large enough to guarantee far-field conditions at the scatterer and sensor. Equivalently, only plane waves are seen at the reflecting surface and sensor.

A2) Additive zero-mean complex Gaussian noise with known covariance influences the complex envelope measurements at the sensor.

Let $e_k(t), e_n(t) \in \mathbb{C}^3, t = 1, 2, \ldots$ be the additive noise seen at, respectively, the electric and magnetic sensors. We suppose that

$$\mathbb{E} \left[ \begin{bmatrix} e_k(t) \\ e_n(t) \end{bmatrix} \right]^* \begin{bmatrix} e_k^*(s) \\ e_n^*(s) \end{bmatrix} = P_e \delta_{t,s},$$

$$P_e \triangleq \begin{bmatrix} \sigma_k^2 I_3 & 0 \\ 0 & \sigma_n^2 I_3 \end{bmatrix}$$

(2.1)

where $I_3$ is the identity matrix in $\mathbb{R}^{3 \times 3}$, that is, the electric and magnetic field measurements are subject to noise having possibly different variances. The diagonal nature of the covariance matrix is due to the independent thermal mechanisms by which the sensor noise is generated. Other possible sources of noise are not considered.

B. Active Model

Consider a reflecting surface scattering a known polarized transmitted signal and a co-located (with the transmitter) vector sensor receiving the so-called "backscatter." Plane waves are incident on the surface and received at the sensor and are parametrized by their direction and polarization. The surface, in general, has a certain polarimetric response. Hence, in addition to experiencing a time delay proportional to the target's range, the signal may have its polarization altered in the reflection process. This change in polarization is now modeled.

We draw on the work of [17] and [21]. The plane wave incident on the target has an electric field vector that moves in the plane perpendicular to the direction of travel of the wave. Let $h^i$ and $v^i$ be two orthonormal vectors that span this plane and $u^i$ be a unit vector pointing from the surface to the transmitter. Then, $\{u^i, h^i, v^i\}$ forms an orthonormal basis for $\mathbb{R}^3$, and we assume that this set is a right-handed triple (see Fig. 1). Designate the frame corresponding to this basis as "T" for "incident."

Since a plane wave is reflected from the target's surface and received at the sensor, we may also define a coordinate system from the set $\{u, h, v\}$, where $u$ points from the sensor to the target, and $h, v$ span the plane traversed by the electric field vector perpendicular to $u$. Because we are receiving backscatter, $u = -u^i$, and letting $h = -h^i, v = v^i$, we have that $\{u, h, v\}$ is a right-handed triple. Let "R" (for reflection) designate the frame corresponding to this basis. Then, vectors in the I and R frames are related by the transformation

$$C_{IR} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$
Let $\omega_c$ be the incident-wave carrier frequency and $\tau$ be the time it takes the transmitted signal, with envelope, say $s(t)$, to go round trip. Plane waves have no electric or magnetic field components along their direction of propagation; therefore, the incident electric field has no component along $\mathbf{w}^t$. The real incident electric field is then given by $\text{Re} E^t(t)e^{j\omega_c t} \in \mathbb{R}^3$, where its complex envelope has the I-frame representation

$$[E^t(t)]_I = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} Q^I \mathbf{w}^s(t - \tau/2)e^{-j\omega_c \tau/2}$$

(2.2)

and

$$Q^I \triangleq \begin{bmatrix} \cos \alpha^I & -\sin \alpha^I \\ \sin \alpha^I & \cos \alpha^I \end{bmatrix}, \quad \mathbf{w}^I \triangleq \begin{bmatrix} \cos \beta^I \\ \sin \beta^I \end{bmatrix}.$$

The matrix $Q^I$ and vector $\mathbf{w}^I$ represent the elliptical polarization of the transmitted wave in the I-frame, where the orientation (with respect to $\mathbf{h}^t$) and ellipticity are given by $\alpha^I$ and $\beta^I$ respectively. See, for example, [6], [11], or [25] for a discussion of polarized waves.

The incident plane wave has two degrees of freedom in its composition. Under assumption A1), in the far-field region relative to the scatterer, the signal received at the sensor is again a plane wave. Let the incident signal be narrowband so that the target may be characterized by its “scattering matrix.” The electric field as measured at the vector sensor is then given by

$$E(t) = \text{Re} E(t)e^{j\omega_c t}$$

(2.3)

where the scattered electric field envelope has the R-frame representation

$$[E(t)]_R = C^R \begin{bmatrix} 0 & 0 \\ 0 & \Gamma \end{bmatrix} [E^s(t - \tau/2)]_I e^{-j\omega_c \tau/2}$$

(2.4)

and the scattering matrix $\Gamma \in \mathbb{C}^{2 \times 2}$ is parametrized by

$$\Gamma = \begin{bmatrix} \gamma_{hh} & \gamma_{hv} \\ \gamma_{vh} & \gamma_{vv} \end{bmatrix}.$$  

(2.5)

Some theoretical and experimental foundations for the above may be found in [17], [21], [23], and [29]. In general, $\Gamma$ is $O(1/\tau)$ as $\tau \to \infty$ because electric fields decay as the reciprocal of the distance between the source and scatterer. Furthermore, $\Gamma$ varies with the aspect angle and frequency of the incident radiation, but these facts are not important to the analysis that follows.

The first column of the scattering matrix contains the polarimetric response of the target to an incident signal polarized along $\mathbf{h}^t$, whereas the second column contains the response to a signal polarized along $\mathbf{v}^t$. The response to an incident signal of any polarization is found by multiplying the matrix by a vector representing the components of the signal in the I-frame (see (2.4)). The variables $\gamma_{hh}$ and $\gamma_{vv}$ are the so-called co-polar scattering coefficients that characterize the scatterer’s ability to reflect radiation along the same axis as the incident $\mathbf{h}^t$ or $\mathbf{v}^t$ radiation, whereas $\gamma_{hv}$ and $\gamma_{vh}$ are the cross-polar coefficients.

We have

$$C^R \begin{bmatrix} 0 & 0 \\ 0 & \Gamma \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & \Gamma \end{bmatrix}$$

where

$$\Gamma = \begin{bmatrix} -\gamma_{hh} & -\gamma_{hv} \\ \gamma_{vh} & \gamma_{vv} \end{bmatrix}.$$  

Using (2.2), we may express (2.4) as

$$[E(t)]_R = \begin{bmatrix} 0 & 0 \\ 0 & \Gamma \end{bmatrix} Q^I \mathbf{w}^s(t - \tau)e^{-j\omega_c \tau}.$$

Then, it is straightforward to show that (2.3) becomes

$$E(t) = \text{Re} [\mathbf{u} \times \mathbf{h}][\begin{bmatrix} 0 & 0 \\ 0 & \Gamma \end{bmatrix} Q^I \mathbf{w}^s(t - \tau)e^{-j\omega_c \tau}] e^{j\omega_c t}$$

$$= \text{Re} V \Gamma Q^I \mathbf{w}^s(t - \tau)e^{j\omega_c (t - \tau)}$$

(2.6)

where

$$V = [\mathbf{h} \times \mathbf{v}].$$

In a plane wave, the electric and magnetic fields are spatially orthogonal. To within a normalizing factor, the component of the electric field along $\mathbf{h}$ is the same as the component of the magnetic field along $\mathbf{u} \times \mathbf{h} = \mathbf{v}$. Similarly, the component of the electric field along $\mathbf{v}$ is along $\mathbf{u} \times \mathbf{v} = -\mathbf{h}$ in the magnetic field [18]. Let $(\mathbf{u} \times)$ be a matrix operator that performs the equivalent of a cross product with the vector $\mathbf{u}$. Then, the complex envelope of the vector-sensor measurement is

$$y(t) \triangleq \begin{bmatrix} y_{h}(t) \\ y_{v}(t) \end{bmatrix} = \begin{bmatrix} I_{S} \\ (\mathbf{u} \times) \end{bmatrix} V \Gamma Q^I \mathbf{w}^s(t - \tau)e^{-j\omega_c \tau} + e(t), \quad t = 1, 2, \cdots$$

(2.7)

The vectors $y_h(t) \in \mathbb{C}^3$ and $y_v(t) \in \mathbb{C}^3$ are the electric and magnetic field measurements, and

$$e(t) = [e_h(t), e_v(t)]^T$$

(2.8)

is the additive noise (see Assumption A2)). We may rewrite (2.7) as

$$y(t) = \begin{bmatrix} I_{S} \\ (\mathbf{u} \times) \end{bmatrix} V \Gamma s(t, \tau) + e(t)$$

(2.9)
where \[
\begin{bmatrix}
s_1(t, \tau) \\
s_2(t, \tau)
\end{bmatrix} = Q^1 u^t s(t - \tau) e^{-i \omega_{0} \tau}.
\] (2.10)

For simplicity of notation, the dependence of \(s(t, \tau)\) on \(\tau\) is omitted in the remainder of the paper.

We have let \(h^1\) and \(v^1\) be any two orthonormal vectors that spanned the electric field's plane. We now orient them to conform to a standard often used in remote sensing models [33]. Let the plane of incidence be defined as the plane spanned by \(u^1\) and \(x\), where \(x\) is the unit normal to the target surface. (In case of a surface that is not smooth, define \(x\) to be perpendicular to the plane about which the average surface height variation is zero [22].) Then, \(h^1 = (x \times u^1)/||x \times u^1||\) is a unit vector in the "horizontal" direction (orthogonal to the plane of incidence and parallel to the surface) and \(v^1 = u^1 \times h^1\) is in the "vertical" direction (in the plane of incidence). Fig. 1 shows this choice of \(h^1\) and \(v^1\). Whenever \(u^1\) and \(x\) are parallel, or the incident wave is normal to surface, then \(h^1\) and \(v^1\) are not unique and may be chosen arbitrarily.

Let the sensor's \(z\) axis be parallel to \(x\). Then, \(u\) and \(V\) may be parametrized as
\[
\begin{bmatrix}
cos \phi \cos \psi \\
\sin \phi \cos \psi \\
\sin \psi
\end{bmatrix} = V = \begin{bmatrix}
-sin \phi & -cos \phi \sin \psi \\
\cos \phi & -sin \phi \sin \psi \\
0 & cos \psi
\end{bmatrix}
\] (2.11)

where \(\phi\) is the azimuth of the scatterer, and \(\psi\) is its elevation relative to the sensor reference frame (Fig. 1). It follows that
\[
(ux) = \begin{bmatrix}
0 & -sin \psi & sin \phi \cos \psi \\
sin \psi & 0 & -cos \phi \cos \psi \\
-sin \phi \cos \psi & cos \phi \cos \psi & 0
\end{bmatrix}.
\] (2.12)

It is assumed that \(\phi, \psi,\) and \(\tau\) are known. We also suppose that the polarization of the transmitted signal, which is represented by \(\alpha^1\) and \(\beta^1\), is known, but the surface's polarimetric response is not.

**C. Passive Model**

There are occasions when an accurate model for the received signal in terms of the transmitted signal cannot realistically be given. Alternatively, the received signal can be generated passively and not under our control. In either case, the passive model is more appropriate than the active model.

The passive model requires a simple modification of the active model. In (2.9), the two components of the received electromagnetic wave are given by \(I(a(t, \tau), \tau)\), and in the passive model, we replace this product by a vector \(\xi(t)\) of two unknown complex scalars representing the time-varying electric field. The equation for the six components of the complex envelope is then (see also [25])
\[
y(t) = \begin{bmatrix}
I(a) \\
\xi(1) \\
\xi(2)
\end{bmatrix} V \xi(t) + \epsilon(t), \quad t = 1, 2, \ldots.
\] (2.13)

As in the active model, let \(u\) be parametrized by an azimuth and elevation. Let \(h = ([0,0,1]^T \times u)/||[0,0,1]^T \times u||\), and \(v = u \times h\). Then, with \(V\) as in (2.6), equation (2.11) applies. Unlike the active model, we assume in (2.13) that \(\phi\) and \(\psi\) are unknown.

**III. DISTRIBUTION OF SCATTERING COEFFICIENTS**

We present the model for the random behavior of the complex scattering coefficients. Models are derived in [1], [2], [27], and [38], where the coefficients are shown sometimes to be Gaussian distributed. We, however, give conditions under which they are complex Gaussian. A central limit theorem argument easily implies their Gaussian nature, but their complex-Gaussian distribution is more difficult to justify. It is possible to have complex random variables that have jointly Gaussian real and imaginary parts but do not have a complex Gaussian distribution in the commonly used sense of [13]. The conditions we give under which this complex Gaussian distribution holds are physically motivated and can be satisfied in practice. Experimental evidence supporting use of the distribution in a remote sensing scenario can be found in [36].

**Assumption:**

A3) The surface we are observing has many closely spaced point scatterers, making the returned signal at a point far from the surface a superposition of the reflections of the incident signal from each of these scatterers.

Assumption A3) is valid, for example, if the target has a rough surface with height variations of the same order as the incident signal wavelength or the target has a randomly varying dielectric composition. Theoretical and experimental arguments for the validity of A3) under various circumstances may be found in, for example, [1], [2], [27], and [28].

For simplicity, in this section, we allow the scattering coefficients \(\gamma_{1h}, \gamma_{2h}, \gamma_{3h}, \gamma_{4h}\) to be labeled \(\gamma_1, \ldots, \gamma_4\). Assume there are \(n\) scatterers on the surface. It follows from A3) (see also [2, p. 120]) that at a point far from the target the \(k\)th scattering coefficient may be expressed as
\[
\gamma_{kn} = \sum_{l=1}^{n} c_{kl} e^{i\varphi_{kl}}, \quad k = 1, \ldots, 4.
\] (3.1)

where \(c_{kl} \in \mathbb{R}\) and \(\varphi_{kl} \in (-\pi, \pi]\) are the amplitude and phase of the contribution of scatterer \(l\) to the coefficient. Clearly, the values of \(c_{kl}\) and \(\varphi_{kl}\) depend on the unknown and possibly complicated structure and composition of the target surface. We therefore model \(c_{kl}\) and \(\varphi_{kl}\) as random variables. Their precise distributions will not be important since, under mild conditions, the central limit theorem will apply.

By definition, every complex Gaussian distribution in the sense of [13] has a real Gaussian distribution counterpart with a specific covariance structure. To prove that the vector of scattering coefficients has a complex Gaussian distribution, we must compose a real vector from the real and imaginary components of the complex vector and show that the covariance of the real vector has this structure. Define \(\gamma_{n} \in \mathbb{C}^4\) and
\[ \tilde{\gamma}_n \in \mathbb{R}^n \] by
\[ \gamma_n \triangleq [\gamma_{n1}, \ldots, \gamma_{n4}]^T = \sum_{i=1}^{n} g_i, \]
\[ \tilde{\gamma}_n \triangleq [\text{Re} \gamma_{n1}, \text{Im} \gamma_{n1}, \ldots, \text{Re} \gamma_{n4}, \text{Im} \gamma_{n4}]^T = \sum_{i=1}^{n} \tilde{g}_i, \]
where we have used (3.1) and
\[ g_i \triangleq [c_{l1} e^{i\varphi_{l1}}, \ldots, c_{l4} e^{i\varphi_{l4}}]^T, \]
\[ \tilde{g}_i \triangleq [g_{i1}, \ldots, g_{i4}]^T, \quad \varphi_{kl} \triangleq \frac{c_{kl} \cos \varphi_{kl}}{c_{kl} \sin \varphi_{kl}}. \]
Furthermore, let \( P_n = \sum_{i=1}^{n} \mathbb{E} g_i g_i^T \) and \( \tilde{P}_n = \sum_{i=1}^{n} \mathbb{E} \tilde{g}_i \tilde{g}_i^T \).
The following give \( \tilde{P}_n \) the necessary structure.

Assumptions:

T1) For all \( k, k' = 1, \ldots, 4 \) and \( l, l' = 1, \ldots, n \), \( k \neq k' \) and \( l \neq l' \), \( \varphi_{kl} \) and \( \varphi_{k'l'} \) are independent; similarly, \( c_{kl} \) and \( c_{k'l'} \) are independent. Moreover, for all \( k, k' \), \( l, l' \), \( c_{kl} \) and \( c_{k'l'} \) are independent.

T2) For any \( k, \{ \varphi_{kl} \}_{l=1, \ldots, n} \) is a sequence of (independent) identically distributed random variables uniformly distributed over \((0, \pi]\). 

T3) For any \( l \), the joint distribution of \( \varphi_{1l}, \ldots, \varphi_{4l} \) is an arbitrary function of only the pairwise phase differences, \( \varphi_{kl} \). 

T4) \(|c_{kl}| \leq C \) for some \( 0 < C < \infty \) not dependent on \( k \) or \( l \).

T5) \( \tilde{P}_n^{-1/2} \tilde{\gamma}_n \) tends, in distribution, to a standard Gaussian (zero-mean, unit-covariance) random variable as \( n \to \infty \). Furthermore, \( \tilde{P}_n^{-1/2} \gamma_n \) tends, in distribution, to a standard complex Gaussian random variable as \( n \to \infty \), where \( P_n \) is congruent to \( 2 \tilde{P}_n \) in the sense of [13, Section II].

Proof: See the Appendix.

Remarks: Assumption T1) requires that the scatterers act independently of each other and the amplitude and phase fluctuations be generated by independent processes. That the scatterers give no particular phase preference is the content of T2). Assumption T3) demands that for any given scatterer, the joint distribution of the phases of its contributions to the coefficients be a function of only the pairwise phase differences. In combination with T2), assumption T3) says that only phase differences between coefficients affect the limiting distribution. Assumption T4) requires each scatterer to give a bounded reflection. Assumption T5) is a technicality related to the central limit theorem ensuring nondegeneracy of the limiting distribution. The only assumption that may appear a little unusual is T3). However, some experimental evidence making T3) reasonable is given by Kong et al. [23] and Ulaby et al. [34], [36] (in the context of remote sensing). These authors show that many scattering surfaces give preferences to pairwise phase differences rather than individual phase values.

Theorem 1 does not require the full strength of T3). As can be seen in the Appendix, the theorem’s proofs use only the joint distribution between any two phases rather than all four. However, T3), as stated, is more physically meaningful because it is invariant to a rotational basis change for the scattering coefficients. The proof of this is tedious and is omitted.

According to Theorem 1, a reflecting surface with a sufficiently large number of scatterers satisfying (T1)–(T5) has scattering coefficients that are zero-mean complex Gaussian random variables. The behavior of the coefficients \( \gamma \triangleq [\gamma_1, \ldots, \gamma_4]^T \) (dependence on \( n \) is now dropped) is then completely determined by the covariance matrix \( P_n \triangleq \mathbb{E} \gamma \gamma^T \).

It has been found by many authors—among the more recent [23], [28], and [36]—that the elements of \( P_n \) depend strongly on the surface’s characteristics. Consequently, \( P_n \) may be used to identify the surface. In practice, an estimate of \( P_n \) is compared with a table of known values for various surfaces. It is therefore of interest to establish lower bounds on estimation error, and this is done in Section V after the distributions of the observations are obtained.

IV. DISTRIBUTIONS OF THE OBSERVATIONS

We derive the distributions of the observations for the polarimetric models (2.9) and (2.13). Let
\[ \theta = [\phi, \psi]^T, \]
\[ A = \begin{bmatrix} I_4 \\ \nu \times \end{bmatrix} V. \]

A. Active Model

Equation (2.9) may be written
\[ y(t) = AS(t)\gamma + e(t), \quad t = 1, 2, \ldots, \]
where
\[ S(t) = \begin{bmatrix} -s_1(t) & 0 \\ s_2(t) & 0 \end{bmatrix}. \]
The expressions for \( s_1(t) \) and \( s_2(t) \) are given in (2.10). Define
\[ p = [[P_1]_{1,1}, \text{Re}[P_1]_{1,2}, \text{Im}[P_1]_{1,2}, \ldots, [P_4]_{4,4}] \in \mathbb{R}^{16} \]
(4.5) as the vector of unknown parameters in this model.

We hypothesize that the transmitted signal consists of a series of pulses or “snapshots,” and during each, the scattering coefficients are constant. If the pulse duration is short compared with the time constants in the target’s and observer’s dynamics, then this is a reasonable assumption. Let the number of snapshots (sometimes also called “looks”) be denoted \( N_s \), and let the number of samples within each snapshot be denoted \( N \). Designate observation \( t \) within snapshot \( s \) of (4.3) by \( y_s(t) \).

Then
\[ Y(s) \triangleq [y_s^T(1), \ldots, y_s^T(N)]^T, \quad s = 1, \ldots, N_s \]
is a record of snapshot \( s \).

It is assumed that the scattering coefficients are realized independently with each snapshot. This holds if, for example, each snapshot illuminates a different part of the surface. Implicitly, it is also supposed that \( \phi, \psi \), and the Gaussian distribution governing the behavior of the scattering coefficients remain approximately constant while the target is
being observed. This supposition is often reasonable under far-field conditions. Let \( \{ \gamma_s \}_{s=1, \ldots, N_s} \) be the collection of realized coefficients. It follows that \( E \gamma_s \gamma_s^* = P_\gamma \), and \( \gamma_s \) is independent of \( \gamma_{s'} \) when \( s' \neq s \).

Many targets obey these conditions, yet Yueh et al. [38] show how some targets yield coefficients with a \( K \)-distribution (the \( K \) indicates a modified Bessel function).

The \( K \)-distribution is indicated when a target has spatially varying statistics, making the Gaussian distribution a poor approximation. Some classes of scatterers for which this happens are detailed in [38], but we rule them out in the analysis that follows. This is not very restrictive, as [23], [36], and [38] give several classes of scatterers where the Gaussian distribution fits experimental data very well (see also [28] for use of the Gaussian in favor of the \( K \)-distribution).

The probability density function of each snapshot is then

\[
f_y(s) = \frac{1}{\pi \sigma^N \det R} \exp \left\{ -Y^*(s) R^{-1} Y(s) \right\}
\]

where

\[
R = \bar{A} \bar{A}^* + \bar{P}_e,
\]

\[
\bar{A} = I_N \otimes A,
\]

\[
\bar{P}_e = I_N \otimes P_e,
\]

\[
S = \begin{bmatrix} S(1) \\ \vdots \\ S(N) \end{bmatrix}
\]

where \( \bar{P}_e \) is given in (2.1), and \( \otimes \) is the Kronecker matrix product.

**B. Passive Model**

Equation (2.13) may be rewritten as

\[
y(t) = A \xi(t) + e(t), \quad t = 1, 2, \ldots
\]

We assume the signal \( \xi(t) \) is stochastic and independent of \( e(t) \). Moreover, \( \xi(t) \) has a Gaussian distribution, and

\[
E \xi(t) = 0, \quad E \xi(s) \xi^*(t) = P_\xi \delta_{s,t}, \quad s, t = 1, 2, \ldots
\]

Let \( P_\xi \) be parametrized as

\[
P_\xi = \begin{bmatrix} q_1 & q_2 & i q_3 \\ q_2 & q_3 & q_4 \end{bmatrix}
\]

and define \( q = [q_1, \ldots, q_4]^T \). Then, the unknown parameters are \( \theta \) and \( q \). In view of the independent Gaussian distributions of the signal and noise vectors, (2.1) and (4.10), we immediately obtain

\[
f_y(t) = \frac{1}{\pi \sigma^N \det R} \exp \left\{ -y^*(t) R^{-1} y(t) \right\}
\]

where \( R = A P_\xi A^* + P_e \). We assume \( t = 1, \ldots, N_t \), that is, each sample is counted as one snapshot or look.

Equation (4.11) can be viewed as a special case of (4.7), calculated by substituting \( P_\xi \) for \( P_\gamma, 1 \) for \( N \), and \( J_2 \) for \( S(1) \).

**V. FISHER INFORMATION AND CRB**

The Cramér-Rao bound for a parameter vector \( \eta \), which is denoted CRB(\( \eta \)), is a lower bound on the covariance of any unbiased estimator of \( \eta \). If \( \eta \) is such an estimator, then \( \text{CRB}(\eta) = \text{FIM}(\eta) \) is positive semidefinite. The CRB and Fisher information matrices (FIM’s) are related through \( \text{CRB}^{-1}(\eta) = \text{FIM}(\eta) \), and the latter has entries given by the negative expectation of the second derivative of the log-density (see [37] for more details). The following theorem computes the FIM entries for both polarimetric models.

**Theorem 2:** Let \( f_y(1, \ldots, N_t) \) be the joint density of \( N_t \) independent observations from (4.7). Then, for \( i, j = 1, 2, k, l = 1, \ldots, 16, \)

\[
E \frac{\partial^2 \log f_y(1, \ldots, N_t)}{\partial \theta^i \partial \theta^j} = -2 N_t \text{Re} \text{tr} \left\{ K \bar{S}^* \left[ A^* P_e^{-1} \frac{\partial A}{\partial \theta} \bar{S} K \bar{S}^* A^* P_e \right] \frac{\partial A}{\partial \theta} \right\}
\]

\[
+ \frac{\partial A}{\partial \theta} \text{Re} \text{tr} \left\{ \bar{S}^* \bar{S} P_e \left( A^* P_e^{-1} \frac{\partial A_e}{\partial \theta} \right) \right\}
\]

\[
- \frac{\partial A}{\partial \theta} \text{Re} \text{tr} \left\{ \bar{S}^* \bar{S} P_e \left( A^* P_e^{-1} \frac{\partial A_e}{\partial \theta} \right) \right\}
\]

(5.1)

\[
E \frac{\partial^2 \log f_y(1, \ldots, N_t)}{\partial \theta^k \partial \theta^l} = -2 N_t \text{Re} \text{tr} \left\{ K \bar{S}^* \frac{\partial P_e}{\partial \theta} L \bar{S}^* A^* P_e^{-1} \frac{\partial A}{\partial \theta} \right\}
\]

(5.2)

\[
E \frac{\partial^2 \log f_y(1, \ldots, N_t)}{\partial \theta^k \partial \theta^l} = -2 N_t \text{Re} \text{tr} \left\{ S^* S L \frac{\partial P_e}{\partial \theta} S L \frac{\partial P_e}{\partial \theta} \right\}
\]

(5.3)

where \( \theta^i \) is the \( i \)th component of \( \theta \), etc., and

\[
K = \sigma_\eta^2 (\sigma_\eta^2 I + P_e \bar{S} \bar{S}^*)^{-1} P_e, \quad L = (\sigma_\eta^2 I + P_e \bar{S} \bar{S}^*)^{-1},
\]

\[
\sigma_\eta^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_a^2}.
\]

(5.4)

**Proof:** See [16].

In the active model, only (5.3) is needed since only \( \theta \) is unknown. In the passive model, all of the expressions are needed with the substitutions \( P_\xi \) for \( P_\gamma \), \( 1 \) for \( N \), \( J_2 \) for \( S(1) \), and \( q \) for \( p \).

**VI. REMOTE SENSING APPLICATION**

In this section, we apply the models and Theorem 2 to an active and passive problem in remote sensing. In both problems, Cramér-Rao bounds for estimating the polarimetric parameters will be derived and analyzed. Throughout, intermediate algebraic details are left out since most of the manipulations were performed using a commercial symbol-manipulating software package.

**A. Active Model**

Suppose that an airborne transceiver is transmitting a known polarized wave towards the earth and receiving the backscatter. It is assumed that the reflections are all due to an area of
rough terrain with time-invariant scattering properties. The polarimetric responses $P_{\gamma}$ of various land and water masses have been computed and tabulated by many authors, among them [23] and [33]. We wish to find lower bounds, given by the CRB, on the error with which $P_{\gamma}$ can be estimated. The results should have practical applications to the geosciences [12], [33], [36].

We assume that the reflecting surface is composed of a reciprocal medium. For such a medium, $\gamma_{hm} = \gamma_{sh}$; see, for example, [22, Section 5.5] for proofs of the reciprocity of various media. Thus, there are effectively only three unknown scattering coefficients. It has been shown theoretically and experimentally in [5], [23], and [36] that to a second-order approximation

$$P_{\gamma} = \begin{bmatrix} P_{\gamma}^\prime & 0 & 0 \\ 0 & P_{\gamma}^x & P_{\gamma}^x \\ 0 & P_{\gamma}^x & P_{\gamma}^x \end{bmatrix}, \quad P_{\gamma}^x \in \mathbb{C}^{2 \times 2}, \quad p_x \geq 0 \quad (6.1)$$

for azimuthally symmetric terrains, where the surface identifying information is contained in $P_{\gamma}^x$ only. In other words, the information-carrying co-polar coefficients are uncorrelated with the cross-polar coefficients, and $p_x$ is a nuisance parameter. Let $P_{\gamma}^x$ have the parameterization

$$P_{\gamma}^x = \begin{bmatrix} p_1 & p_2 + ip_3 \\ p_2 - ip_3 & p_4 \end{bmatrix}. \quad (6.2)$$

Then, the unknown polarimetric parameters are $p = [p_1, \cdots, p_4, p_x]^T$.

As shown in [16], for a transmitted polarized signal of the form (2.10), where $Q^1$ and $w^1$ are time-invariant, FIM ($p$) is singular, and therefore, CRB ($p$) does not exist. This is because a time-invariant polarized signal allows us to observe only two linearly independent combinations of the scattering coefficients, and more than two combinations are needed to uniquely identify $\gamma$. This is a well-known problem that is corrected by using polarization diversity [14]. Polarization diversity involves transmitting a signal whose polarization varies. For example, the polarization of the signal may be altered at $t = N/2$ ($N$ is assumed even for simplicity) by varying $Q^1$ and $w^1$.

There are many possible ways to vary $Q^1$ and $w^1$, and we will assume a method is chosen satisfying

$$\sum_{t \leq N} s_1(t)s_2^*(t) = 0, \quad (6.3a)$$

$$\sum_{t \leq N} |s_1(t)|^2 = \sum_{t \leq N} |s_2(t)|^2. \quad (6.3b)$$

For example, to achieve (6.3), let a signal envelope $s(t), t = 1, \cdots, N/2$, be transmitted twice in immediate succession with $\beta = 0$, once with $\alpha^1 = -\pi/4$, and once with $\alpha^2 = \pi/4$. This corresponds to a signal with spatially orthogonal linear polarizations since $s_2(t) = -s_1(t) = -s(t)/\sqrt{2}$ for $t = 1, \cdots, N/2$, and $s_2(t) = s_1(t) = s(t)/\sqrt{2}$ for $t = N/2+1, \cdots, N$. It is easy to verify that identifiability of $p$ is now assured. A possible variation is sending, in succession, two orthogonal circularly polarized signals [14].

The result of using a signal that obeys (6.3), with $N_x = 1$, is shown in the expression at the bottom of the page, where we have not shown the result for the nuisance parameter $p_x$, and

$$E_s = \sum_{t \leq N} |s_1(t)|^2 = \sum_{t \leq N} |s_2(t)|^2.$$

Because the CRB matrix is symmetric, the upper triangular entries are displayed explicitly, whereas redundant entries are marked with a "*". When $N_x > 1$, the matrix entries must be multiplied by $1/N_x$.

The expressions for CRB ($p$) are not very revealing, primarily because the entries of $p$ do not have clear physical or geometrical interpretations. We remedy this with the following lemma on the decomposition of a covariance matrix that we will apply to $P_{\gamma}^x$ and later to $P_{\gamma}$.

**Lemma 1:** Let $P \in \mathbb{C}^{2 \times 2}$ be Hermitian positive semidefinite. Then, $P$ can be decomposed as

$$P = \sigma_u^2 I + \sigma_p^2 Q w w^* Q^* \quad (6.4)$$

where either $\sigma_u^2$ or $\sigma_p^2$ can be zero, and

$$Q = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}, \quad w = \begin{bmatrix} \cos \beta \\ i \sin \beta \end{bmatrix}, \quad \alpha \in (-\pi/2, \pi/2], \quad \beta \in [-\pi/4, \pi/4]. \quad (6.5)$$

Furthermore, assuming $\sigma_p^2 > 0$, this decomposition is unique if and only if $|\beta| \neq \pi/4$. If $|\beta| = \pi/4$, then $\alpha$ is arbitrary. If $\sigma_p^2 = 0$, then both $\alpha$ and $\beta$ are arbitrary.

**Proof:** In [6, p. 551], there is a similar decomposition that may be extended to (6.4) by employing the standard elliptical decomposition of a polarized wave. The details are given in [16].

There is a physical interpretation: The first term in (6.4) corresponds to the unpolarized component of $P$ with power $\sigma_u^2$, whereas the second term corresponds to the elliptically polarized component with power $\sigma_p^2$ and orientation and ellipticity given, respectively, by $\alpha$ and $\beta$. Hence, this lemma effectively assigns the covariance matrix a geometry.
Applying the lemma to \( P'_\gamma \), we effectively decompose the covariance \((6.1)\). Another decomposition of \((6.1)\) is proposed in \([39]\), where an outer product decomposition of Cloude
\([8]\), \([9]\) is employed. However, Lemma 1 is more similar
to the decomposition of a Stokes vector into polarized and
unpolarized components \([7]\).

We study the implications. As \( \sigma_p^2 \) tends to zero, so does
\[
\det P'_\gamma, \text{ the limit corresponding to fully correlated co-polar
scattering coefficients (}\gamma_{hh} \text{ and } \gamma_{ee}\text{). The parameters } \alpha, \beta, \text{ and } \sigma_p^2 \text{ become, respectively, the orientation, ellipticity, and size
of an ellipse defined by the scatterer. These parameters are similar
to the electric field ellipse parameters of a polarized wave
\([11], [25]\), but their values here assign a shape to the target's
co-polar coefficients. When } |\beta| = \pi/4, \text{ the ellipse becomes a
circle, and it is easy to verify that the orientation } \alpha \text{ is arbitrary.}

When \( \sigma_p^2 \) > 0 we may regard } \alpha, \beta, \text{ and } \sigma_p^2 \text{ as generalized
to the case in which there is an uncorrelated component in
\( \gamma_{hh} \text{ and } \gamma_{ee}\). As } \sigma_p^2 \text{ tends to zero, } P'_\gamma \to (\sigma_p^2/2)I \text{ and,}
clearly, } \alpha \text{ and } \beta \text{ become arbitrary. This is because the co-
polar coefficients become completely uncorrelated with the
same power (variance) and therefore cannot be assigned any shape
whatsoever.}

From \((6.2)\), it is easy to show that
\[
\begin{align*}
\alpha &= \frac{1}{2} \tan^{-1} \left( \frac{2p_2}{p_1 - p_4} \right), \\
\beta &= \tan^{-1} \left( \frac{d - 2p_3}{\sqrt{d + 2p_3}} \right) - \frac{\pi}{4}, \\
\sigma_p^2 &= d - \sigma_a^2 = \text{tr } P'_\gamma - d
\end{align*}
\]
\((6.6a)\)
\[(6.6b)\]
where
\[(6.7)\]
is the square root of the discriminant of the characteristic
polynomial of \( P'_\gamma \). Define \( p' = [\alpha, \beta, \sigma_p^2, \sigma_a^2]^T \). Then CRB \((p')\)
and CRB \((p)\) are related by \([37, p. 83]\)
\[
\text{CRB } (p') = \frac{\partial p'}{\partial p}^{\top} \text{CRB } (p) \frac{\partial p'}{\partial p}. \tag{6.6a}
\]
We get
\[
\text{CRB } (p') = \left[ \begin{array}{cccc}
\delta_1 & 0 & 0 & 0 \\
0 & \frac{1}{2d^2} & 0 & 0 \\
0 & 0 & -d^2 + 2d_1 & -d^2 + 2 \delta_1 + 2 [\delta_1 - \delta_a^2/E_a] \\
0 & 0 & 0 & 2(d^2 - d \text{ tr } P'_\gamma + 2[\delta_1 - \delta_a^2/E_a])
\end{array} \right]
\]
where
\[
\delta_1 = \det (\sigma_a^2/E_a I + P'_\gamma), \quad d' = \sqrt{(p_1 - p_4)^2 + 4p_3^2}.
\]
Observe that CRB \((p')\) is nearly diagonal or, equivalently,
the entries of \( p' \) are information-decoupled. It follows that knowledge of some of the entries does not affect our ability
to estimate the remaining entries. The elements of \( p' \) may be
thought of as components of the polarimetric model that can
be estimated in statistically uncorrelated ways. The elements
of \( p \) have no such interpretation, as we could not find a
signal that diagonalized CRB \((p)\). The information-decoupling
property of \( p' \), however, depends strongly on heeding \((6.3)\).
If \((6.3a)\) were to hold but not \((6.3b)\), then CRB \((p')\) would
contain many nonzero off-diagonal entries having the factor
\( \Sigma_{l \leq N} |s_l(t)|^2 - \Sigma_{l \leq N} |s_l(t)|^2 \). Additionally, CRB \((\alpha)\)
and CRB \((\beta)\) would become strictly larger with additive terms
proportional to \( \Sigma_{l \leq N} |s_l(t)|^2 - \Sigma_{l \leq N} |s_l(t)|^2 \). The signal
obeying \((6.3)\) therefore minimizes the variance bound on \( \alpha \)
and \( \beta \).

There are some singularities in CRB \((p')\) that are natural and
not problematic. As \( d = \sigma_p^2 \to 0 \), CRB \((\alpha)\) and CRB \((\beta)\)
tend to infinity. This is because \( \alpha \) and \( \beta \) become undefined (and
physically meaningless) as \( P \to (\sigma_a^2/2)I \). Now, let \( d' \to 0 \)
and \( d \to 2|p_3| > 0 \). Then, CRB \((\alpha)\) tends to infinity, whereas
CRB \((\beta)\) remains bounded. This is because \( |\beta| \to \pi/4 \) (ellipse
becomes a circle) and the orientation \( \alpha \) becomes arbitrary.

When there is a high SNR, or \( E_a/\sigma_a^2 \) is large, then
\[
\text{CRB } (p') = \frac{\sigma_a^2}{d' E_a} \text{CRB } (p) \frac{\sigma_a^2}{d' E_a} \tag{6.6a}
\]
\[
\left[ \begin{array}{cccc}
\frac{\delta_1}{2d^2} & 0 & 0 & 0 \\
0 & \frac{1}{2d^2} & 0 & 0 \\
0 & 0 & -d^{2} + 2d_1 & -d^{2} + 2 \delta_1 + 2 \delta_1 [\delta_1 - \delta_a^2/E_a] \\
0 & 0 & 0 & 2(d^{2} - d \text{ tr } P'_\gamma + 2 \delta_1 [\delta_1 - \delta_a^2/E_a])
\end{array} \right]
\]
\[
+ O(\sigma_a^2/E_a).
\]
Hence, when a strong signal is used and there is little unpolarized
reflection (such that \( \text{det } P'_\gamma \) is small), we can estimate the
shape parameters \( \alpha \) and \( \beta \) with great accuracy. It is expected
that \( \alpha \) and \( \beta \) will be the two most useful parameters for
identifying the scatterer in this case.

B. Passive Model

Both the signal direction \( \theta \) and the polarimetric parameters
\( q \) are now unknown. The CRB matrix in this case has 21
different entries, all of which are generally nonzero. Due to
space limitations, we present only the diagonal ones:

\[
\begin{align*}
\text{CRB } (\theta)_{1,1} &= \frac{u}{2 \Delta \cos^2 \psi}, \\
\text{CRB } (\theta)_{2,2} &= \frac{u'}{2 \Delta}, \\
\text{CRB } (q)_{1,1} &= (q_1 + \sigma_q^2)^2 + \frac{2 \nu q_2 \tan^2 \psi}{\Delta}, \\
\text{CRB } (q)_{2,2} &= \frac{1}{2} (q_1 q_4 - q_2^2 + q_3^2 + \sigma_q^2 \text{ tr } P_t + \sigma_t^2) \\
&\quad + \frac{u (q_1 - q_4)^2 \tan^2 \psi}{2 \Delta}, \\
\text{CRB } (q)_{3,3} &= \frac{1}{2} (q_1 q_4 - q_2^2 + q_3^2 + \sigma_q^2 \text{ tr } P_t + \sigma_t^2), \\
\text{CRB } (q)_{4,4} &= (q_4 + \sigma_q^2)^2 + \frac{2 \nu q_2 \tan^2 \psi}{\Delta},
\end{align*}
\]
where
\[ U = \text{Re}(P_k + \sigma_n^2 I)^{-1} P_k^2, \]
\[ \Delta = \text{det}(U), \]
\[ u = (1/\sigma_n^2)U_{11}, \]
\[ u' = (1/\sigma_n^2)U_{12}. \]

Observe that unlike \( \text{CRB}(p) \), \( \text{CRB}(q) \) depends on \( \psi \). That is, we are leading to the disturbing conclusion that the chosen coordinate system, while arbitrary, seems to affect our ability to estimate the polarimetric parameters accurately. This raises doubts about the usefulness of the parameters. Only the bound on \( g_2 \) is independent of \( \psi \), suggesting that it is the only parameter that may be estimated reliably, no matter what the signal’s direction. It is interesting to note that the expressions for \( \text{CRB}(q) \) and \( \text{CRB}(p) \) are virtually identical, except for the terms involving \( \psi \) (simply set \( E_s = 1 \)). The presence of \( \psi \) shows \( q' \)’s sensitivity to lack of knowledge of the source direction.

However, applying Lemma 1 to decompose \( P_k \) and letting \( q' = [\alpha, \sigma_n^2, \sigma_n^2]^T \), we see that \( \text{CRB}(q') \) is block diagonal and almost free of the coordinate system’s influence, as shown in the expression at the bottom of this page, where
\[ \delta_2 = \text{det}(P_k + \sigma_n^2 I), \quad d = \sqrt{(q_1 - q_4)^2 + 4q_2^2 + 4q_3^2}, \]
\[ d' = \sqrt{(q_1 - q_4)^2 + 4q_2^2}, \quad \tilde{u} = (1/\sigma_n^2 - 1/\sigma_n^2)U_{12}. \]

We note that except for a single term involving \( \psi \), \( \text{CRB}(q') \) and \( \text{CRB}(p') \) are virtually identical. Of the new polarimetric parameters, only \( \alpha \) has a bound depending on \( \psi \), but this dependence turns out to be easy to explain. Applying the geometrical interpretation of \( P_k' \)’s decomposition in the active model to the decomposition of \( P_k \), we see that \( \alpha \) is the orientation of the polarized portion of the plane wave’s ellipse. The orientation is a rotation about \( u \) relative to \( h \), but \( h \) is not uniquely defined when \( \psi = \pm \pi/2 \) (see Section II-C for the definitions of \( h \) and \( v \), and note that \( u = [0, 0, 0]^T \)).

Hence, the rotation is also not uniquely defined when \( \psi = \pm \pi/2 \). \( \text{CRB}(\alpha) \) therefore increases to infinity as \( |\psi| \to \pi/2 \). This connection between \( \alpha \) and \( \psi \) suggests that \( \alpha \) is a not a very useful polarimetric parameter in the passive model.

However, the remaining parameters \( \beta, \sigma_n^2, \) and \( \sigma_n^2 \) are useful since they do not suffer from directional sensitivity and are therefore intrinsic to the signal. Indeed, they represent the ellipticity and power of the polarized component and the power of the unpolarized component of the received signal. Clearly, \( q' \) has natural physical interpretations not available to \( q \).

Consider also estimation of the degree of polarization defined as \( d_p = \sigma_n^2/(\sigma_n^2 + \sigma_n^2) \) (see also [6, p. 552]). The degree of polarization is the ratio of polarized power to total signal power and is another parameter commonly used in polarimetry [36]. We may easily calculate its Cramér-Rao bound by noting that
\[ \frac{\partial d_p}{\partial q^T} = \left[ 0, 0, \frac{\text{tr} P_k - d}{(\text{tr} P_k)^2}, -d \right]. \]

Thus
\[ \text{CRB}(d_p) = \frac{\partial d_p}{\partial q^T} \text{CRB}(q') \frac{\partial d_p}{\partial q} = \frac{4}{(\text{tr} P_k)^4} \left( 2 \text{det} P_k^2 + 2 \sigma_n^2 (\text{tr} P_k) \text{det} P_k \right) \]
\[ + \sigma_n^4 \left( \text{tr} P_k^2 - 2 \text{det} P_k \right) \]
which is not coordinate system dependent.

The results of Sections VI-A and B can be summarized as follows. The standard covariance matrix parametrizations of \( P_k' \) and \( P_k \) did not lend themselves to simple analysis: In the active model, we could not easily interpret the parameters; in the passive model, most of the CRB expressions had a disturbing dependence on \( \psi \). On the other hand, the reparametrization that used Lemma 1 exhibited several nice properties: The new parameters assigned a geometry to the covariance matrices; in the active model, the parameters defined an ellipse whose size and shape were functions of only the target’s co-polar scattering characteristics. With the proper transmitted signal, the CRB matrix was nearly diagonal or information decoupled, and two of the diagonal entries were minimized; in the passive model, only \( \alpha \)’s CRB depended on \( \psi \). It was argued that \( \alpha \) was tied to the coordinate system, explaining this dependence.

VII. ESTIMATION OF \( P_k \) AND \( P_k' \)

Given \( N_s \) snapshots of data from an active polarimetric vector-sensor system, we are interested in computing an estimate of \( P_k' \). We propose the maximum likelihood estimate. We will also propose an estimate of \( P_k \) for the passive model that is adapted from the estimate of \( P_k' \) but is not maximum likelihood.
A. The Maximum Likelihood Estimate of $P_\gamma$

In the active model, $\tau$, $\phi$, and $\psi$ are known, and it is simple to compute the maximum likelihood estimate of $P_\gamma$ from (4.7). Given the observations $Y(s), s = 1, \cdots, N_s$, the maximum likelihood estimate satisfies

$$
\hat{P}_\gamma = \arg \max_{P_\gamma} \prod_{s=1}^{N_s} \frac{1}{\pi^{N_s} N_s} \exp \left\{ -Y^*(s)R^{-1}Y(s) \right\}.
$$

By dropping terms not dependent on $P_\gamma$, we obtain

$$
g(R, \hat{R}) = \log \det R + \text{tr} \{ R^{-1} \hat{R} \}.
$$

(7.1)

Defining $B = \hat{P}_e^{-1/2}A\tilde{S} \in C^{N_x \times 4}$ and $M = BP_eB^* + I$, it follows from (4.8) that $R = \hat{P}_e^{-1/2}MP_e^{1/2}$, whence $g(R, \hat{R}) = g(M, \hat{M}) + \log \det P_e$, where $\hat{M} = \hat{P}_e^{-1/2} \hat{R} \hat{P}_e^{1/2}$. Consequently, $\hat{P}_\gamma = \arg \min_{P_\gamma} g(M, \hat{M})$. The result of this minimization, when $P_\gamma$ is Hermitian but otherwise unstructured, is well known (see, for example [32])

$$
\hat{P}_\gamma = (B^*B)^{-1}B^*\hat{M}B(B^*B)^{-1} - (B^*B)^{-1}
$$

(7.2)

assuming that $B^*B$ is nonsingular. Because $B^*B = S^*A\hat{P}_e^{-1}A\tilde{S} = \sigma_0^{-2}S^*S$, we are left with

$$
\hat{P}_\gamma = \sigma_0^2(S^*S)^{-1}S^*A\hat{P}_e^{-1}\hat{R}\hat{P}_e^{-1}A\tilde{S}(S^*S)^{-1}
\quad - \sigma_0^2(S^*S)^{-1}.
$$

(7.3)

B. Specialization to Remote Sensing Application

1) The Maximum Likelihood Estimate of $P_\gamma$: Equation (7.2) may be applied to estimate the covariance of the scattering coefficients in the remote sensing application detailed in Section VI. However, as seen in (6.1), $P_\gamma$ has a special structure, the knowledge of which is not used to construct (7.2). Therefore, in our application, (7.2) is not the maximum likelihood estimate of $P_\gamma$. This becomes evident if we consider that (7.2) attempts to estimate all of the entries of $P_\gamma$: even the entries in (6.1) known to be zero and the four lower right entries known to be $p_x$.

Nevertheless, a few simple modifications to $\hat{P}_\gamma$ yield the maximum likelihood estimate, provided the transmitted signal obeys (6.3). Let $\hat{P}_\gamma^0$ be the matrix $\hat{P}_\gamma$ modified so that zeros replace the estimates of the known zero entries of $P_\gamma$ and the average of the four lower right estimates of $p_x$ replaces the individual estimates. We show that $\hat{P}_\gamma^0$ is the maximum likelihood estimate of the structured $P_\gamma$.

The problem is to find

$$
\hat{P}_\gamma = \arg \min_{P_\gamma} g(M, \hat{M})
$$

constraining $P_\gamma$ to have the structure given in (6.1). Let the column of $B$ be denoted by $b_i, i = 1, \cdots, 4$ and define $B' = [b_1, b_2]$. Then

$$
g(M, \hat{M}) = \log \det (B'P_eB'^* + p_x(b_3 + b_4)(b_3 + b_4)^* + I) + \text{tr} \{(B'P_eB'^*)^{-1}
\quad + p_x(b_3 + b_4)(b_3 + b_4)^* + I}^{-1}\hat{M}
\quad = g(T^{-1/2}B'P_eB'^*T^{-1/2} + I, T^{-1/2}\hat{M}T^{-1/2})
\quad + \log \det T
$$

(7.4)

where $T = p_x(b_3 + b_4)(b_3 + b_4)^* + I$. Since $T$ does not depend on $P_\gamma$, we may minimize the first term on the right-hand side of (7.4) to obtain the maximum likelihood estimate of $P_\gamma$.

The result is

$$
\hat{P}_\gamma' = (B^*T^{-1}B')^{-1}B^*T^{-1}\hat{M}T^{-1}B'(B^*T^{-1}B')^{-1}
\quad - (B^*T^{-1}B')^{-1}.
$$

(7.5)

An easy manipulation yields

$$
T^{-1} = I - (p_x^{-1} + ||b_3 + b_4||^2)^{-1}(b_3 + b_4)(b_3 + b_4)^*.
$$

Condition (6.3) gives $B^*B = \sigma_0^{-2}S^*S = \sigma_0^{-2}\xi_4I$. Therefore, the columns of $B$ are orthogonal, implying that $B^*T^{-1} = B^*$. Equation (7.5) then simplifies to

$$
\hat{P}_\gamma' = (B^*B')^{-1}B^*\hat{M}B'(B^*B')^{-1} - (B^*B')^{-1}.
$$

(7.6)

Since $B^*B'$ is the upper left block of $B^*B$ and is therefore also a multiple of the identity matrix, we conclude that $\hat{P}_\gamma'$ is simply the upper left block of $\hat{P}_\gamma$, given by (7.2). Note that no estimate of $p_x$ was needed to calculate the maximum likelihood estimate of $P_\gamma$. If the transmitted signal were to disobey (6.3), this would not necessarily be true.

Although $p_x$ is considered a nuisance parameter, for completeness, its maximum likelihood estimate is also given. Substituting $\hat{P}_\gamma'$ for $P_\gamma'$ in $g(M, \hat{M})$ yields the concentrated minimization function

$$
\log \det (B'\hat{P}_\gamma'B'^* + p_x(b_3 + b_4)(b_3 + b_4)^* + I) + \text{tr} \{(B'\hat{P}_\gamma'B'^*)^{-1}
\quad + p_x(b_3 + b_4)(b_3 + b_4)^* + I}^{-1}\hat{M}
\quad = \log \det T
$$

The $p_x$ that minimizes this function is

$$
p_x = \|b_3 + b_4\|^{-4}(b_3 + b_4)^*\hat{M}(b_3 + b_4) - \|b_3 + b_4\|^{-2}
\quad - \sigma_0^2(b_3 + b_4)^*\hat{M}(b_3 + b_4) - \sigma_0^2/2\xi_2.
$$

It is easily verified that this expression coincides with the average of the lower right four entries of $\hat{P}_\gamma$ given by (7.2). Thus, when, and generally only when, a signal obeying (6.3) is transmitted, $\hat{P}_\gamma^0$ is the maximum likelihood estimate of the structured $P_\gamma$.

The maximum likelihood estimate of the parameterization suggested by Lemma 1 is obtained by applying (6.6) to $\hat{P}_\gamma^0$ because maximum likelihood estimates are invariant to reparametrizations.
2) An Estimate of $P_\xi$: The passive model's distribution (4.11) can be computed by substituting $P_\xi$ for $P_\gamma$ in (4,1) for $N$, and $I_\xi$ for $S(1)$ into (4.7). The equivalent of (7.3) is therefore

$$\hat{P}_\xi = \sigma^2 \tilde{A}^* P_\xi^{-1} \tilde{R} P_\xi^{-1} A - \sigma^2 I,$$

(7.7)

where $\tilde{R} = (1/N_N) \sum_{N=1}^{N_N} y(t) y^*(t)$. However, $u$ and, hence, $A = [I_\xi, (u \times T)^T] V$ are unknown and must be estimated. A simple estimate of $u$, which is denoted $\hat{u}$, is the unit vector parallel to $(1/N_R) \sum_{N=1}^{N_R} \text{Re} \{y(t) \times y^*(t)\}$. Then, $\hat{h} = [(0, 0, 1)^T \times \hat{u}] / ||(0, 0, 1)^T \times \hat{u}||$ and $\hat{v} = \hat{u} \times \hat{h}$ provide an estimate of $V$. Combining the results, we get the estimate $\hat{A}$ that is substituted for $A$ into (7.7). Clearly, once modified this way, (7.7) is not necessarily maximum likelihood. However, the angular error between $\hat{u}$ and $u$ has been shown in [25] to nearly attain its lower bound in a variety of scenarios as $N_0 \to \infty$. Therefore, this method of first obtaining $\hat{A}$ and then using it to get $\hat{P}_\xi$ should perform reasonably well. Once $\hat{P}_\xi$ is computed, it may be decomposed as in Lemma 1 to obtain estimates of $\alpha, \beta, \sigma^2_\alpha$, and $\sigma^2_\beta$.

VIII. DISCUSSION AND CONCLUSION

We have presented active and passive models for estimating polarimetric parameters associated with electromagnetic plane waves. The models and their applications to problems in remote sensing were comprehensively analyzed. We proposed a parametrization of the scattering-coefficient (active) and signal (passive) covariance matrices and compared it with a standard parametrization. The parametrization we proposed was argued to be preferable since it had many desirable interpretive and analytical features the standard one did not. Experimental confirmation of the theoretical results and tests of the reliability of the new parameters for surface identification and classification would clearly be desirable.

Also of interest would be a comparison of the proposed parametrization with those of Stokes and Mueller [33], which are often used in remote sensing models. A possible application of the results of this paper would be the design of signals (in the active model) that minimize the parameterizations' estimation errors.

Estimators of the scattering coefficient covariance $P_\gamma$ and the signal covariance $P_\xi$ were derived. The estimate of $P_\gamma$ was maximum likelihood, whereas the estimate of $P_\xi$ was not. Maximum likelihood estimators are generally asymptotically normally distributed with covariance given by the CRB, and therefore, the active model's CRB should be a good indication of the performance of $\hat{P}_\gamma$. We argued that the estimate of $P_\xi$ should also be accurate; a performance analysis would be of interest.

Analysis with vector sensor measurements allows one to examine what can be accomplished if all of the information available to a sensor at a single point in space is used. The models provide a clear portrait of the physics that underlie the measured processes. More traditional receiving antennas or scalar sensors act as projectors and filters of the electromagnetic field. Nevertheless, the vector-sensor framework is easily adapted to cases where only part of the electromagnetic field is measured. For example, the theoretical results we derived for the active model apply directly to a system in which two orthogonally polarized antennas are used. A vector-sensor model with $\phi = \psi = 0$ (where this fact is known) is equivalent to a model with two co-located orthogonal linear-polarization sensitive receiving antennas in the $y-z$ plane because the received plane wave has no electric or magnetic field components along the $x$ axis. Therefore, all of the analysis carries over to the polarization sensitive two-colocated-antenna system. Moreover, if the magnetic field measurements are not available, our expressions may still be used by simply letting $\sigma^2_\beta \to \infty$.

Most of our discussion, in the active model, centered around transmitting a narrowband signal of fixed carrier frequency. This scenario was chosen to facilitate the scattering coefficient analysis, but it should be noted that our models do not otherwise carry a narrowband restriction. As an aid to reflector identification, it has been suggested by some authors, for example [31] and [33], that frequency diversity is important. Signals of different frequencies tend to excite different types of scatterers on the reflector surface, and if the reflector is a complex object, it may have scatterers contributing at selective frequencies. The analysis we have given would then apply to each frequency of interest, provided that the return signal is Gaussian in each case; that is, a large number of each type of scatterer is present. Each transmitted frequency would have a corresponding reflection ellipse whose size and shape could be used to determine geometrical features of the reflector.

Extensions of the models to unknown $P_\xi$ and spatially correlated measurement noise would also be important. Such models would be realistic when there is external electric or magnetic field noise and the noise behavior has to be estimated along with the polarimetric parameters.

APPENDIX

PROOF OF THEOREM 1

Before we can prove the theorem, we need a lemma concerning the asymptotic normality of sums of independent random variables. Since the scatterers on the rough surface are not assumed to behave exactly like one another, the random variables will not, in general, have identical distributions.

Assumptions:

L1) $\{Z_i\}_{i=1,2,\ldots} \in \mathbb{R}^p$ is a sequence of independent vectors such that $E Z_i = 0$ and $||Z_i|| \leq C < \infty$ are (uniformly) bounded.

L2) $\Lambda_n \triangleq \Sigma_{i \leq n} E Z_i Z_i^T$ obeys $||\Lambda_n^{-1/2}|| \to 0$ as $n \to \infty$.

Lemma A: Assumptions L1)–L2) imply $\Lambda_n^{-1/2} \Sigma_{i \leq n} Z_i$ tends, in distribution, to a standard Gaussian random variable as $n \to \infty$.

Proof of Lemma: A standard method to prove convergence results about vectors is to make a vector proof from a series of scalar proofs. We use the Cramér-Wold device [3, Theorem 29.4]. Let $t$ be a deterministic vector in $\mathbb{R}^p$ different from zero, and define $x_n = t^T \Lambda_n^{-1/2} Z_i$ and $z_n = \Sigma_{i \leq n} x_n = t^T \Lambda_n^{-1/2} \Sigma_{i \leq n} Z_i$. Let $z$ be a $p$-dimensional standard Gaussian random variable. We have to show that $x_n$ converges in distribution to $t^T z$ for all $t$. Observe that $E x_n = 0$ and
\[ E z_n^2 = t^T \Lambda_n^{-1/2} E \zeta \zeta^T \Lambda_n^{-1/2} t. \] Letting \( s_n^2 = \Sigma_{i \leq n} E z_n^2 \), we see that \( s_n^2 = ||t||^2. \)

We have
\[
\frac{1}{||t||^2} \sum_{i=1}^{n} E |z_i|^2 \leq \frac{1}{||t||^2} \left( C ||t|| \cdot ||\Lambda_n^{-1/2}|| \right) \sum_{i=1}^{n} E z_i^2 \\
= C ||\Lambda_n^{-1/2}||
\]
and, using the assumption \( ||\Lambda_n^{-1/2}|| \to 0, \lim_{n \to \infty} (1/||t||^2) \sum_{i=1}^{n} E |z_i|^2 = 0. \) Because \( ||t||^2 = s_n^2, \) we have verified Lyapunov's condition \( \lim_{n \to \infty} (1/\kappa_n s_n^2) \sum_{i=1}^{n} E |z_i|^2 = 0 \) for \( \delta = 1 \) \( \text{[3, p. 371].} \) Hence, by the Lindeberg-Feller central limit theorem for scalar random variables, \( z_n/||t|| \) converges in distribution to a standard Gaussian random variable, or \( z_n \) to \( t \) in \( Z \). This is true for all \( t \neq 0 \) (for \( t = 0 \), the result is trivial), and the proof is concluded.

**Proof of Theorem 1:** We apply Lemma A to show that \( \hat{P}_n^{-1/2} \hat{\tau}_n \) has a multivariate Gaussian distribution asymptotically, and hence, we must first verify L1) and L2) for the random vectors \( \{g_{il}\}_{i=1,2,\ldots} \).

L1): It follows from assumptions T1) and T4) that \( g_{il} \) and \( g_{il'} \) are independent for all \( i \neq i' \) and are uniformly bounded. Note that \( E c_{ik} \cos \varphi_{kl} = E c_{ik} \cos \varphi_{kl} = 0 \) and similarly \( E c_{ik} \sin \varphi_{kl} = 0 \), and therefore \( E g_{il} = 0 \).

L2): Directly implied by T5).

Hence, we immediately have that \( \hat{P}_n^{-1/2} \hat{\tau}_n \) tends in distribution to a standard Gaussian random variable as \( n \to \infty \).

The vector \( P_n^{-1/2} \tau_n \) is asymptotically complex Gaussian if and only if \( P_n \) has the special structure detailed in Sections II and III of [13]. We now show that this is so. By assumptions T1) and T2)
\[
E \hat{g}_{kl} \hat{g}_{kl}^T = E \hat{c}_{kl} \begin{bmatrix}
E \cos^2 \varphi_{kl} & E \cos \varphi_{kl} \sin \varphi_{kl} \\
E \cos \varphi_{kl} \sin \varphi_{kl} & E \sin^2 \varphi_{kl}
\end{bmatrix}
= \begin{bmatrix}
b_{kl} & 0 \\
0 & b_{kl}
\end{bmatrix}
\]
for some \( b_{kl} > 0 \). Let the joint distribution between \( \varphi_{kl} \) and \( \varphi_{kl'} (k \neq k'), \) which is a function of only \( \varphi_{kl} - \varphi_{kl'} \) (assumption T3)), be denoted by \( f_{\varphi_{kl} \varphi_{kl'}} \). We make the following observation:
\[
E \sin (\varphi_{kl} + \varphi_{kl'}) = \int_{-\pi}^{\pi} \frac{d\varphi_{kl}}{2\pi} \int_{-\pi}^{\pi} \sin (\varphi_{kl} + \varphi_{kl'}) \\
= \frac{1}{2} \int_{-\pi}^{\pi} \frac{dt}{2\pi} \int_{-\pi}^{\pi} \sin s ds = 0 \quad (A.1)
\]
using the change of variables \( s = \varphi_{kl} + \varphi_{kl'} \) and \( t = \varphi_{kl} - \varphi_{kl'}. \)

Equation (A.1) and the familiar trigonometric identity for \( \sin (\varphi_{kl} + \varphi_{kl'}) \) then imply
\[
E \sin \varphi_{kl} \cos \varphi_{kl'} = -E \cos \varphi_{kl} \sin \varphi_{kl'}. \quad (A.2)
\]
Using similar reasoning
\[
E \cos (\varphi_{kl} + \varphi_{kl'}) = \frac{1}{2} \int_{-\pi}^{\pi} \frac{dt}{2\pi} \int_{-\pi}^{\pi} \cos t ds \\
= \frac{1}{2} \int_{-\pi}^{\pi} f_{\varphi_{kl} \varphi_{kl'}} (t) \sin (2\pi - |t|) dt \\
= \int_{-\pi}^{\pi} f_{\varphi_{kl} \varphi_{kl'}} (t) \sin |t| dt \\
= \int_{-\pi}^{\pi} [f_{\varphi_{kl} \varphi_{kl'}} (\pi + t) + f_{\varphi_{kl} \varphi_{kl'}} (\pi - t)] \sin t dt \\
= \int_{-\pi}^{\pi} [f_{\varphi_{kl} \varphi_{kl'}} (\pi + t) + f_{\varphi_{kl} \varphi_{kl'}} (\pi - t)] \sin t dt
\]
since \( f_{\varphi_{kl} \varphi_{kl'}} (t) \) has period \( 2\pi \). Plainly, \( f_{\varphi_{kl} \varphi_{kl'}} (\pi + t) + f_{\varphi_{kl} \varphi_{kl'}} (\pi - t) \) is an even function of \( t \); therefore, the final integral is zero.

A trigonometric identity for \( \cos (\varphi_{kl} + \varphi_{kl'}) \) yields
\[
E \cos \varphi_{kl} \cos \varphi_{kl'} = E \sin \varphi_{kl} \sin \varphi_{kl'} \quad (A.3)
\]

Hence, for \( k \neq k' \)
\[
E \hat{g}_{kl} \hat{g}_{kl'}^T = E \hat{c}_{kl} c_{kl'} \begin{bmatrix}
E \cos \varphi_{kl} \cos \varphi_{kl'} & E \cos \varphi_{kl} \sin \varphi_{kl'} \\
E \cos \varphi_{kl} \sin \varphi_{kl'} & E \sin \varphi_{kl} \sin \varphi_{kl'}
\end{bmatrix}
= \begin{bmatrix}
b_{kl} & -b_{kl} \\
b_{kl} & b_{kl}
\end{bmatrix}
\]
for some \( b_{kl} \in R \). These structures for \( E \hat{g}_{kl} \hat{g}_{kl}^T \) and \( E \hat{g}_{kl} \hat{g}_{kl'}^T \) imply that \( E \hat{g}_{kl} \hat{g}_{kl'}^T \) is congruent to \( 2E \hat{g}_{kl} \hat{g}_{kl}^T \) in the sense of [13, Section II]. Since \( P_n = \Sigma_{i<\infty} \hat{g}_{kl} \hat{g}_{kl}^T \) and \( P_n = \Sigma_{i<\infty} \hat{g}_{kl} \hat{g}_{kl}^T \), it is immediate that \( P_n \) and \( 2P_n \) are congruent, and from the first part of this proof, \( P_n^{-1/2} \tau_n \) converges in distribution to a standard complex-Gaussian random variable.

**ACKNOWLEDGMENT**

The authors gratefully acknowledge the helpful comments of Prof. W.-M. Boerner at the University of Illinois.

**REFERENCES**


Bertrand Hochwald (S'90) was born in New York, NY, He received the bachelor's degree from Swarthmore College, Swarthmore, PA, and the M.S. degree in electrical engineering from Duke University, Durham, NC.

From 1986 to 1989, he worked for the Department of Defense at Fort Meade, MD. In 1989, he enrolled at Yale University, New Haven, CT, where he received the M.A. degree in statistics. He has completed the Ph.D. degree in electrical engineering at Yale and is now with the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign. His interests include statistical signal processing, probability theory, and information theory.

Mr. Hochwald received several achievement awards while employed at the Department of Defense and received the Prize Teaching Fellowship at Yale. He is a member ofEta Kappa Nu.

Aryeh Nehorai (S'80-M'83-SM'90-F'94) received the B.Sc. and M.Sc. degrees in electrical engineering from the Technion—Israel Institute of Technology, in 1976 and 1979, respectively, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1983.

From 1983 to 1984, he was a Research Associate at Stanford University. From 1984 to 1985, he was a Research Engineer at Systems Control Technology, Inc., Palo Alto, CA. Since 1985, he has been with the Department of Electrical Engineering at Yale University, New Haven, CT, where he is an Associate Professor.

Dr. Nehorai is an Associate Editor of the journals Circuits, Systems, and Signal Processing and The Journal of the Franklin Institute and was an Associate Editor of the IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. He serves as the Chairman of the Connecticut IEEE Signal Processing Society's Senior Award for an outstanding paper. He is a member of Sigma Xi.