An Entropy Based Classification Scheme for Land Applications of Polarimetric SAR

Shane Robert Cloude, *Senior Member, IEEE* and Eric Pottier, *Member, IEEE*

***Abstract*—In this paper we outline a new scheme for parameterizing polarimetric scattering problems, which has application in the quantitative analysis of polarimetric SAR data. The method relies on an eigenvalue analysis of the coherency matrix and employs a three-level Bernoulli statistical model to generate estimates of the average target scattering matrix parameters from the data. The scattering entropy is a key parameter is determining the randomness in this model and is seen as a fundamental parameter in assessing the importance of polarimetry in remote sensing problems. We show application of the method to some important classical random media scattering problems and apply it to POLSAR data from the NASA/JPL AIRSAR data base.**

***Index Terms*— Image classification, radar polarimetry, unsupervised classifier.**

I. Introduction

THERE HAVE recently been developed several algorithms for the classification of land features based on their polarimetric microwave signatures [1]–[6]. These methods exploit observed similarities and correlations in feature vectors derived from either complete coherent scattering matrix data or noncoherent multiple channel radar cross section data such as the ratio σ_{HH}/σ_{VV}.

Most such techniques are supervised, in the sense that the feature vector is first derived from measurements over known terrain classes. Unknown terrain is then compared with the training set and a statistical decision made as to class membership. Several unsupervised techniques have also been developed [7]–[11]. They tend to be more physically based and have the advantage that their performance is not data specific, as often arises with supervised methods. In this paper we consider development of just such an unsupervised classification scheme which we suggest has some advantages over those currently employed in the literature.

While classification studies can be used to demonstrate the basic ability of polarimetry to distinguish features in an image, there remains a requirement to extend this basic classification philosophy so that quantitative parameter extraction may be used in radar remote sensing. We can summarize the idea behind such quantitative methods as shown in Fig. 1, where we show how three elements of quantitative microwave scattering (physical modeling, electromagnetic modeling and measurements) interact via observables in the solution of remote sensing problems.

Manuscript received January 8, 1996; revised October 11, 1996.

The authors are with Laboratoire S.E.I. E.P. CNRS 63, IRESTE, La Chantrerie 44087, Nantes, France.

Publisher Item Identifier S 0196-2892(97)00857-7.

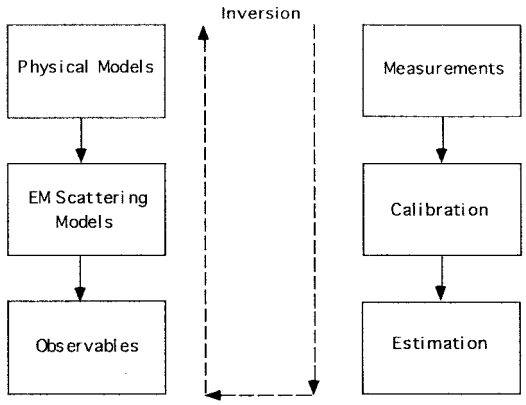

Fig. 1. General structure of quantitative radar remote sensing problems.

In order to achieve reliable inversion of such models, multiparameter measurements must generally be made. For example, multitemporal (both coherent i.e., interferometric and noncoherent), multifrequency (generally at P, L and C bands) and multipolarization data have all been suggested for remote sensing applications. In this paper we consider the use of multipolarization data for quantitative remote sensing applications and the role polarimetric observables play in relating the two key elements of measurement and physical modeling in polarimetric systems.

We propose a method for extracting average parameters from experimental data using a smoothing algorithm based on second order statistics. This method does not rely on the assumption of a particular underlying statistical distribution and so is free of the physical constraints imposed by such multivariate models. For example, the multivariate Gaussian distribution is popular because of the ease of analytical development of moments and marginal distributions and is physically supported on the basis of the central limit theorem, which dictates that within each image cell we have a large number of polarimetric scattering elements, with no one element being a dominant scatterer.

In this paper we employ an alternative statistical model which sets out with the assumption that there is always a dominant "average" scattering mechanism in each cell and then undertakes the task of finding the parameters of this average component. It is these that are then related to the EM scattering and physical model of Fig. 1. We show how this parameterization is able to extract important target information by applying it to some direct random media scattering models. In this way, we suggest that the method is able to make an important contribution to the inversion of data represented by Fig. 1.

II. Scattering Entropy

In previous publications [12]–[14], we have shown that for media with reflection symmetry [15], the coherency matrix may be parameterized in the form shown in (1), at the bottom of the page.

If we make the further assumption that $\lambda_2 = \lambda_3$ (as happens, for example, when we average over rotation about the Radar line of sight), then the scattering by such media can be modeled as a single dominant scattering matrix (1), shown at the bottom of the page.

This dominant matrix is determined from the maximum eigenvector of $\langle[T]\rangle$ and has a degree of randomness or entropy defined from the eigenvalues (in the von Neumann sense) as

$$H = \sum_{i=1}^{n} -P_i \log_n P_i \quad P_i = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j} \tag{2}$$

where $n = 3$ for backscatter problems, so that by construction $0 \leq H \leq 1$. For example, in Bragg surface scatter, the entropy $H = 0$ to first order and the angle α in (1) is then independent of surface roughness and directly related to the angle of incidence and dielectric constant of the surface [14]. In practice, there are two limitations to this simplified model:

1) It is limited to reflection symmetric media, where by definition the co- and cross-linear polarization scattering coefficients are uncorrelated i.e., $\langle S_{hh} S_{hv}^* \rangle = \langle S_{vv} S_{hv}^* \rangle = 0$ [12], [15]. Although common, such media are not the only ones of interest in remote sensing applications.
2) If the minor eigenvalues λ_2 and λ_3 are not equal, then the entropy is not a unique measure of the randomness of the scatterer. In this case the degree of polarization of the scattered wave depends on the polarization of the incident wave. Such behavior has been observed for example in multiple scattering problems involving spherical scatterers [19].

In this paper we propose modifications to the model of (1) which permit its extension into these two important regimes. We then use this extension to develop an unsupervised classification scheme.

III. Probabilistic Model for Random Media Scattering

If the scattering medium does not have reflection symmetry, then there exists a preferred orientation angle β for the scatterer and this we can accommodate by introducing a new parameterization of the eigenvectors of $[T]$ in the form shown in (3)

$$\underline{e} = [\cos\alpha \quad \sin\alpha\cos\beta e^{i\delta} \quad \sin\alpha\sin\beta e^{i\gamma}]^T. \tag{3}$$

This angle arises for example when we consider a tilted Bragg surface scattering model, or when we have a cloud of particles with a preferred orientation or canting angle. With this modification, we obtain a revised parameterization of the coherency matrix in the form in (4), shown at the bottom of the page. The parameterization of a 3×3 unitary matrix in terms of column vectors with different parameters $\alpha_1\beta_1$ etc. is made so as to enable a probabilistic interpretation of the scattering process. In general, the 3×3 matrix has only eight independent parameters (defined by the Gell-Mann matrices [16]) since the columns are not only unitary but mutually orthogonal. This means that in practice $\alpha_1\alpha_2$ and α_3 are not independent. Indeed, for the special case of reflection symmetry (1) we have seen that only the first column is required to determine the whole of the matrix $[U_3]$.

By including the angle β we have allowed for the possibility that the cross and copolar channels may be correlated. The second part of the modification is to allow the possibility of a nondegenerate eigenvalue spectrum ($\lambda_2 \neq \lambda_3$). In this case we can consider the following statistical model of the scatterer. We model the target as a three symbol Bernoulli process i.e., the target is represented by three $[S]$ matrices [the columns of $[U_3]$ in (4)] which occur with probabilities P_i (defined in (2)

$$\begin{aligned}
\langle[T]\rangle &= \frac{1}{2}\left\langle \begin{bmatrix} (S_{HH}+S_{VV})(S_{HH}+S_{VV})^* & (S_{HH}+S_{VV})(S_{HH}-S_{VV})^* & 2(S_{HH}+S_{VV})S_{HV}^* \\ (S_{HH}-S_{VV})(S_{HH}+S_{VV})^* & (S_{HH}-S_{VV})(S_{HH}-S_{VV})^* & 2(S_{HH}-S_{VV})S_{HV}^* \\ 2S_{HV}(S_{HH}+S_{VV})^* & 2S_{HV}(S_{HH}-S_{VV})^* & 4S_{HV}S_{HV}^* \end{bmatrix} \right\rangle \\
&= \begin{bmatrix} e^{i\phi}\cos\alpha & e^{i\delta}\sin\alpha & 0 \\ -e^{-i\delta}\sin\alpha & e^{-i\phi}\cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} e^{-i\phi}\cos\alpha & -e^{i\delta}\sin\alpha & 0 \\ e^{-i\delta}\sin\alpha & e^{i\phi}\cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1\cos^2\alpha + \lambda_2\sin^2\alpha & \cos\alpha\sin\alpha(\lambda_2-\lambda_1)e^{i(\delta+\phi)} & 0 \\ \cos\alpha\sin\alpha(\lambda_2-\lambda_1)e^{-i(\delta+\phi)} & \lambda_2\cos^2\alpha + \lambda_1\sin^2\alpha & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}.
\end{aligned} \tag{1}$$

$$[T] = [U_3] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} [U_3]^{*T} \tag{4}$$

$$[U_3] = \begin{bmatrix} \cos\alpha_1 & \cos\alpha_2 & \cos\alpha_3 \\ \sin\alpha_1\cos\beta_1 e^{i\delta_1} & \sin\alpha_2\cos\beta_2 e^{i\delta_2} & \sin\alpha_3\cos\beta_3 e^{i\delta_3} \\ \sin\alpha_1\sin\beta_1 e^{i\gamma_1} & \sin\alpha_2\sin\beta_2 e^{i\gamma_2} & \sin\alpha_3\sin\beta_3 e^{i\gamma_3} \end{bmatrix}$$

by the normalized eigenvalues so that $P_1 + P_2 + P_3 = 1$). In this way, for example, the parameter α has associated with it a random sequence

$$x(n) = \alpha_1\alpha_2\alpha_1\alpha_3\alpha_2\alpha_1\alpha_1\alpha_3\cdots \tag{5}$$

and our best estimate (in the sense of maximum likelihood) of the parameter is given by the mean of this sequence which is easily evaluated as

$$\bar{\alpha} = P_1\alpha_1 + P_2\alpha_2 + P_3\alpha_3. \tag{6}$$

In this way, we extract the *mean* parameters of the dominant scattering mechanism from the coherency matrix as a vector f such that

$$\underline{f} = [\bar{\alpha} \quad \bar{\beta} \quad \bar{\delta} \quad \bar{\gamma}] \tag{7}$$

where the parameters $\alpha\ \beta\ \delta$ and γ are defined in (4). Equation (7) contains all modifications required for an extension of our model to account for the two limitations of (1).

IV. Application to Random Media Scattering Theory

In this section, we evaluate the vector $\underline{f}$ (7) and the corresponding scattering entropy H (2) for two important direct scattering problems. The first is multiple scattering under a diffusion approximation from a cloud of spherical scatterers. The second is single scattering from a cloud of anisotropic particles. We shall see that the vector $\underline{f}$ contains a great deal of the physical structure embedded in these two problems. This will lend support to our claim that this parametrization is closely related to physical structure in scattering problems, which in turn will justify our use of $\underline{f}$ in an unsupervised classification scheme.

A. Multiple Scattering Entropy

Using a diffusive wave propagation model, we expect random volume scattering to obtain complete wave depolarization very quickly. In this section we examine this assumption a little more closely by using the vector reciprocity theorem to derive the entropy of the polarimetric back scattering coherency matrix as a function of increasing scattering order from a cloud of identical spherical particles. We show that for low orders of multiple scattering, significant polarized structure can be maintained in such random media problems.

The response of a scattering medium is determined by the ensemble averaged field covariance G satisfying the Bethe-Salpeter equation [17], [18] which can be compactly represented as

$$\langle G_{im}G^*_{jn}\rangle = G_{im}G^*_{jn} + G_{im'}G^*_{jn'}W_{ijm'n'}G_{mm'}G^*_{nn'} \tag{8}$$

where W is the intensity operator (the sum of all strongly connected diagrams in the multiple scattering integral equations). In the weak scattering limit, the observed intensity is the incoherent sum of contributions of waves scattered through all possible paths. The incoherent contributions arise from the second term on the right hand side of (8). For vector scattering, employing the ladder approximation to the BS intensity operator, the medium has a diagonal Mueller matrix of the form (see [18] for details) in (9), shown at the bottom of the page, where, under some simplifying assumptions, the correlation terms can be analytically expressed for the nth order of scattering as

$$\begin{aligned}
\langle G_{xx}G^*_{xx}\rangle &= \langle G_{yy}G^*_{yy}\rangle = \frac{1}{2}\left[1 + 2\left(\frac{7}{10}\right)^n\right] \\
\langle G_{xy}G^*_{xy}\rangle &= \frac{1}{2}\left[1 - 2\left(\frac{7}{10}\right)^n\right] \\
\langle G_{xx}G^*_{yy}\rangle &= \frac{3}{4}\left[\left(\frac{7}{10}\right)^n + \left(\frac{1}{2}\right)^n\right] \\
\langle G_{xy}G^*_{yx}\rangle &= \frac{3}{4}\left[\left(\frac{7}{10}\right)^n - \left(\frac{1}{2}\right)^n\right].
\end{aligned} \tag{10}$$

By substitution we then obtain the following general form for the Mueller matrix in the weak scattering limit

$$[M] = \begin{bmatrix} 1 + \frac{1}{2}\left(\frac{7}{10}\right)^n & 0 & 0 & 0 \\ 0 & \frac{3}{2}\left(\frac{7}{10}\right)^n & 0 & 0 \\ 0 & 0 & \frac{3}{2}\left(\frac{7}{10}\right)^n & 0 \\ 0 & 0 & 0 & \frac{3}{2}\left(\frac{1}{2}\right)^n \end{bmatrix}. \tag{11}$$

Now using the relationship between the Mueller matrix $[M]$ and coherency matrix $[T]$ we obtain the following form for the matrix $[T]$, shown in (12), at the bottom of the next page, where $\chi = \frac{3}{2}(\frac{7}{10})^n$. This is the ladder term solution. If we consider the special case of backscatter, then it follows from the vector reciprocity theorem [20] that the cyclic terms have a coherency matrix of the same form as above with a minus sign in the 3, 3 position. Hence, the final backscatter coherency matrix (the sum of ladder and cyclic terms) is given as a function of n, the order of scattering by (13). Since the coherency matrix is diagonal, the nonzero elements in (13) are the required eigenvalues and so we can calculate the entropy and vector $\underline{f}$ of the scattering matrix as a function of order n. Fig. 2 shows the variation of eigenvalues and corresponding entropy. Note the following points:

1) For $n = 0$, the entropy is zero. This indicates that for zero order i.e., single scattering from a random medium, the coherency matrix has only one nonzero eigenvalue (see [14]). In this case, averaging over orientation makes no change to the polarimetric response, due to particle symmetry. Hence, despite having a random scattering problem, we observe a zero entropy system! This has important implications for the use of polarization to observe multiple scattering effects [19]. Since the single

$$[M] = \begin{bmatrix} \langle G_{xx}G^*_{xx}\rangle + \langle G_{xy}G^*_{xy}\rangle & 0 & 0 & 0 \\ 0 & \langle G_{xx}G^*_{xx}\rangle - \langle G_{xy}G^*_{xy}\rangle & 0 & 0 \\ 0 & 0 & \langle G_{xx}G^*_{yy}\rangle + \langle G_{xy}G^*_{yx}\rangle & 0 \\ 0 & 0 & 0 & \langle G_{xx}G^*_{yy}\rangle - \langle G_{xy}G^*_{yx}\rangle \end{bmatrix} \tag{9}$$

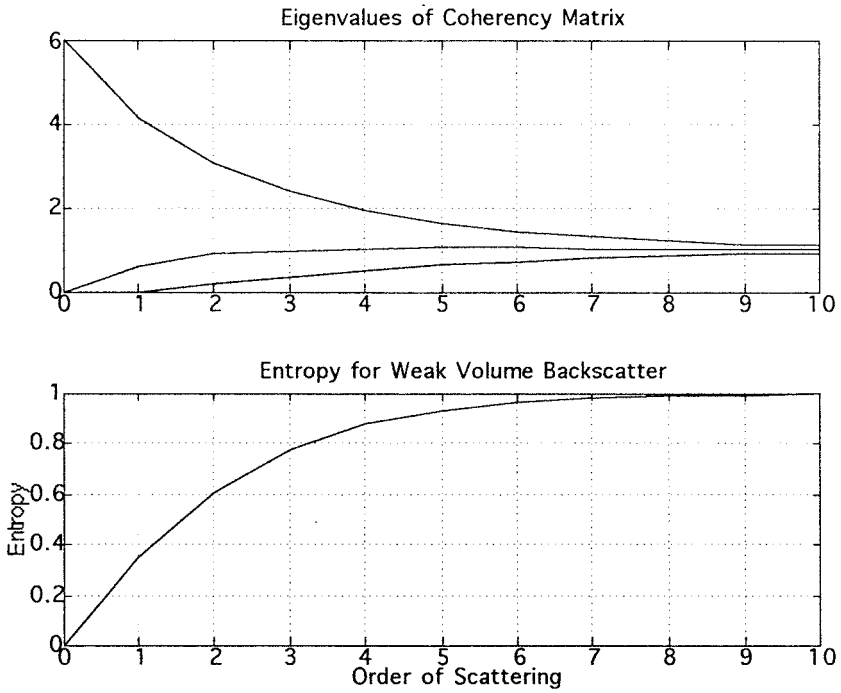

Fig. 2. Coherency matrix eigenvalue parameters as a function of scattering order in volume backscatter.

scattering contributions have $H = 0$, then they can be represented by a single $[S]$ matrix, a scalar multiple of the identity. Hence, choice of sphere reject circular polarization will entirely remove the single scattering contributions from observation. This has been used in experiments to isolate the multiple scattering contributions [19] since the only signal contributions in the sphere reject channel will be those due to scattering from orders $n > 0$.

2) The entropy is a smoothly increasing function of scattering order; as the order increases so the entropy approaches unity (i.e., a random noise process). For large n, the coherency matrix becomes a scalar multiple of the identity matrix and represents a true noise or completely depolarized signal. However, for low order multiple scattering, significant polarized structure survives the averaging. This has important implications for the use of polarization in remote sensing of random media i.e., in the inversion process of Fig. 1, since entropy relates to the inherent reversibility of the scattering process. We shall use this observation in Section V to propose a new classification scheme.
3) The vector $\underline{f}$ can be determined from the eigenvectors of (13), shown at the bottom of the page, as

$$P_1\begin{bmatrix}1\\0\\0\end{bmatrix} + P_2\begin{bmatrix}0\\1\\0\end{bmatrix} + P_3\begin{bmatrix}0\\0\\1\end{bmatrix} \Rightarrow \begin{cases}\bar{\alpha} = P_1 0 + P_2\frac{\pi}{2} + P_3\frac{\pi}{2}\\ \bar{\beta} = P_1\infty + P_2 0 + P_3\frac{\pi}{2} = \infty\\ \bar{\delta} = \bar{\gamma} = \infty\end{cases} \tag{14}$$

where the probabilities P_i are determined from the normalized eigenvalues of (4) (note that the symbol ∞ means that the parameter cannot be uniquely determined). Fig. 3 shows the parameter $\bar{\alpha}$ as a function

$$[T] = \begin{bmatrix} 1+\left(\frac{7}{3}+\left(\frac{5}{7}\right)^n\right)\chi & 0 & 0 & 0\\ 0 & 1+\left(\frac{1}{3}-\left(\frac{5}{7}\right)^n\right)\chi & 0 & 0\\ 0 & 0 & 1+\left(\frac{1}{3}-\left(\frac{5}{7}\right)^n\right)\chi & 0\\ 0 & 0 & 0 & 1+\left(\left(\frac{5}{7}\right)^n-\frac{5}{3}\right)\chi \end{bmatrix} \tag{12}$$

$$[T_3] = \begin{bmatrix} 1+\left(\frac{7}{3}+\left(\frac{5}{7}\right)^n\right)\chi & 0 & 0\\ 0 & 1+\left(\frac{1}{3}-\left(\frac{5}{7}\right)^n\right)\chi & 0\\ 0 & 0 & 1+\left(\left(\frac{5}{7}\right)^n-\frac{5}{3}\right)\chi \end{bmatrix}. \tag{13}$$

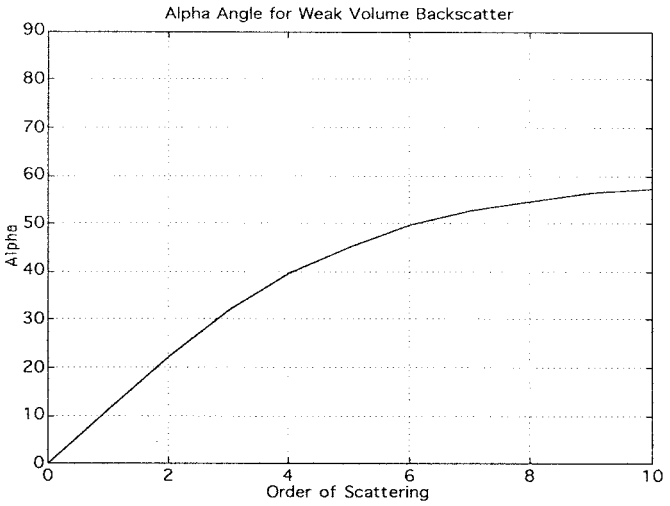

Fig. 3. Alpha parameter variation for volume backscatter as a function of scattering order.

of scattering order. Note that for *single* scattering from spheres we expect $\bar{\alpha} = 0$ and so $\underline{f} = (0, \infty, \infty, \infty)$.

From Fig. 3 we see that the alpha parameter varies as a function of increasing entropy, from 0° at $n = 0$ and tending to 60° as $n \to \infty$. Hence, in order to use this parameter for target classification, it is necessary to always associate the $\underline{f}$ vector parameters with an entropy value, in order to avoid ambiguities as to the nature of the scattering process. In fact, we shall show in Section V that variation of the parameters of $\underline{f}$ with increasing entropy is generic to the classification scheme and can be explained physically on the basis of increasing uncertainty about the scattering process.

B. Single Scattering Particle Anisotropy

As a second important example of the use of this eigenvector formalism in random media scattering, consider the case where we have backscatter from a cloud of identical anisotropic particles with a scattering matrix of the form

$$[S] = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} \tag{15}$$

where a and d are complex scattering coefficients in the particle characteristic coordinate system. In this case we can generate the effect of rotation about the line of sight on the coherency matrix as

$$\begin{aligned}[T_3(\theta)] &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta \\ 0 & \sin 2\theta & \cos 2\theta \end{bmatrix} \begin{bmatrix} \varepsilon & \mu & 0 \\ \mu^* & \nu & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &\quad \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\theta & \sin 2\theta \\ 0 & -\sin 2\theta & \cos 2\theta \end{bmatrix} \\ &= \begin{bmatrix} \varepsilon & \mu \cos 2\theta & \mu \sin 2\theta \\ \mu^* \cos 2\theta & \nu \cos^2 2\theta & \nu \sin 2\theta \cos 2\theta \\ \mu^* \sin 2\theta & \nu \sin 2\theta \cos 2\theta & \nu \sin^2 2\theta \end{bmatrix} \end{aligned} \tag{16}$$

where $\varepsilon = |(a+d)|^2/2, \nu = |(a-d)|^2/2$ and $\mu = (a+d)(a^* - d^*)/2$. If we now average over all angles θ then we obtain a form of the coherency matrix as shown in (17). This matrix represents the two basic observables (ε and ν) for single scattering systems of this type.

$$\langle [T_3(\theta)] \rangle = \frac{1}{2} \begin{bmatrix} 2\varepsilon & 0 & 0 \\ 0 & \nu & 0 \\ 0 & 0 & \nu \end{bmatrix}. \tag{17}$$

We note that the coherency matrix is diagonal and that we have all the conditions required for (1) to apply i.e., degenerate minor eigenvalues and reflection symmetry. However, we can use it to illustrate the power of the new probabilistic model.

For diagonal coherency matrices, the vector $\underline{f}$ has only one uniquely defined parameter (α) and is given by

$$f = \left[\frac{\pi}{2}(P_2 + P_3) \quad \infty \quad \infty \quad \infty \right]. \tag{18}$$

In the case of (17), $P_2 = P_3$ so the vector has the simplified form $f = [\pi P_2 \infty \infty \infty]$. It follows that for problems with the symmetry of (17) we need only consider one target parameter, namely α. However, we still need to ascertain whether this parameter is closely related to the physical structure of the problem or not.

There are three interesting special cases to be considered.

- $a = d$
  In this case we have $\nu = 0$ in (17), the entropy is zero and the probability for the first eigenvector $P_1 = 1$. Thus we have a completely deterministic problem even though we have averaged over all angles θ. In this case the vector becomes $\underline{f} = (0 \quad \infty \quad \infty \quad \infty)$ where the values of β, δ and γ are all indeterminate. The dominant scattering mechanism thus corresponds to an eigenvector of the form $\underline{e} = (1 \quad 0 \quad 0)^T$. Such a situation arises in the single scattering from a random cloud of spherical objects, as we saw in Figs. 2 and 3 for $n = 0$.
- $a = -d$
  In this case $\varepsilon = 0$ in (17) and the coherency matrix has the form

$$[T] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \nu & 0 \\ 0 & 0 & \nu \end{bmatrix} \tag{19}$$

  and the vector $\underline{f} = [\pi/2 \quad \infty \quad \infty \quad \infty]$ i.e., the average scattering mechanism is correctly identified as being due to an eigenvector $\underline{e} = (0 \quad 1 \quad 0)^T$ i.e., dihedral scattering, but with a uniform distribution of rotation angle (β is indeterminate). This uncertainty as to the angle of orientation is manifest as an increase in entropy (to $H = 0.62$).
- $a \gg d$
  In this case we assume that the particles are highly anisotropic (dipole scatterers, for example, when $d = 0$). The coherency matrix then has the limiting form

$$[T] \propto \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{20}$$

  In this case, using (18), we obtain $\underline{f} = [\pi/4 \quad \infty \quad \infty \quad \infty \quad \infty]$. Again we see that the dominant scattering mechanism has been correctly identified as an eigenvector

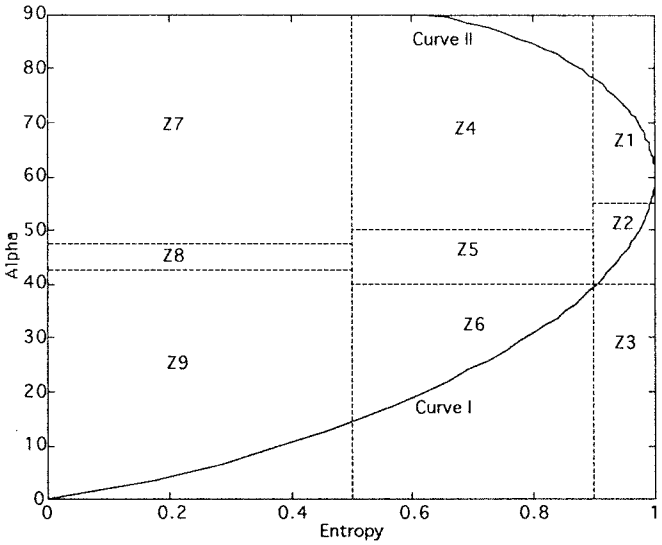

Fig. 4. Feasible region in α-H plane for random media scattering problems.

$\underline{e} = (0.707 \quad 0.707 \quad 0)^T$ but that the target has been averaged over all angles (again β is indeterminate).

We can see from these examples that the averaging suggested in (7) leads to parameter estimates which relate directly to underlying physical scattering mechanisms and hence may be used to associate observables with physical properties of the medium. In the next section we consider application of this idea to an unsupervised classification scheme.

V. SAR Classification Scheme

Although the vector $\underline{f}$ has four components, from the above analysis it is clear that for random media problems, the main parameter for identifying the dominant scattering mechanism is α. This angle has a useful range of 90° and corresponds to a continuous change from surface scattering in the geometrical optics (GO) limit $(\alpha = 0°)$ through surface scattering under physical optics (PO) to the Bragg surface model, encompassing dipole scattering and moving into double bounce scattering mechanisms between two dielectric surfaces, finally reaching dihedral scatter from metallic surfaces at $\alpha = 90°$.

We can therefore consider a two-dimensional (2-D) H-α classification space as shown in Fig. 4. All random scattering mechanisms can be represented in this space. However, because of the averaging inherent in (6), not all regions are equally populated. For example, when $H = 1$ there is only one possible value for alpha $(\alpha = 60°)$. This reflects our increasing inability to distinguish between scattering mechanisms as the underlying entropy increases. On the other hand at $H = 0$ we have access to the full range of possible alpha values. We can quantify the variation in this feasible region of points by considering the curves I and II, as shown in Fig. 4.

For each value of entropy H, we can identify a possible variation of alpha lying between curves I and II. These curves then represent bounds on the maximum and minimum observable alpha values as a function of entropy. These curves are determined by the H-α variation for a coherency matrix with degenerate minor eigenvalues with amplitude $m(0 \leq m \leq 1)$ of the form shown in (21)

$$[T]_{\mathrm{I}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{bmatrix} \quad 0 \leq m \leq 1$$

$$\begin{cases} [T]_{\mathrm{II}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2m \end{bmatrix} & 0 \leq m \leq 0.5 \\ [T]_{\mathrm{II}} = \begin{bmatrix} 2m-1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & 0.5 \leq m \leq 1. \end{cases} \tag{21}$$

By calculating the entropy, $H(m)$ for T_{I} and T_{II} the two curves shown in Fig. 4 result. We shall see that all experimental results lie inside this feasible region and so classification must take place inside this limited zone of H-α space.

In Fig. 4 we also show, with dotted lines, nine basic zones characteristic of classes of different scattering behavior. Note that we can extend this idea into a hierarchical system by further sub-dividing each zone in H and α, but this top level classification will generally be applied first in order to separate the data into basic scattering mechanisms. It would be interesting to try and relate the actual value of the parameters inside these zones to physical characteristics, such as surface roughness or biomass, in order to complete the inversion process of Fig. 1. Such studies will form the basis for future extensions of this work.

Note that the bounds in Fig. 4 are chosen generically i.e., based on the general properties of the scattering mechanisms and are not dependent on a particular data set. In this way we will achieve an unsupervised classification scheme. We now present a summary of the physical scattering characteristics of each of the nine zones outlined in Fig. 4:

- **Zone 9: Low Entropy Surface Scatter**
  In this zone occur low entropy scattering processes with alpha values less than 42.5°. These include GO and PO surface scattering, Bragg surface scattering and specular scattering phenomena which do not involve 180° phase inversions between hh and vv. Physical surfaces such as water at L and P-Bands, sea-ice at L-Band, as well as very smooth land surfaces, all fall into this category.
- **Zone 8: Low Entropy Dipole Scattering**
  In this zone occur strongly correlated mechanisms which have a large imbalance between hh and vv in amplitude. An isolated dipole scatterer would appear here, as would scattering from vegetation with strongly correlated orientation of anisotropic scattering elements. The mean angle of orientation would then be given by the β parameter in $\underline{f}$. The width of this zone is determined by the ability of the Radar to measure the hh/vv ratio i.e., on the quality of the calibration.
- **Zone 7: Low Entropy Multiple Scattering Events**
  This zone corresponds to low entropy double or 'even' bounce scattering events, such as provided by isolated dielectric and metallic dihedral scatterers. These are characterized by $\alpha > 47.5°$. The lower bound chosen for this zone is dictated by the expected dielectric constant of the dihedrals and by the measurement accuracy of the Radar. For $e_r \geq 2$, for example, and using a Bragg surface model for each surface, it follows that $\alpha \geq 50°$.

The upper entropy boundary for these first three zones is chosen on the basis of tolerance to perturbations of first order scattering theories (which generally yield zero entropy for all scattering processes). By estimating the level of entropy change due to second and higher order events, tolerance can be built into the classifier so that the important first order process can still be correctly identified. Note also that system measurement noise will act to increase the entropy and so the system noise floor should also be used to set the boundary. $H = 0.5$ is chosen as a typical value accounting for these two effects.

- **Zone 6: Medium Entropy Surface Scatter**
  This zone reflects the increase in entropy due to changes in surface roughness and due to canopy propagation effects. In surface scattering theory the entropy of low frequency theories like Bragg scatter is zero. Likewise, the entropy of high frequency theories like Geometrical Optics is also zero. However, in between these two extremes, there is an increase in entropy due to the physics of secondary wave propagation and scattering mechanisms. Thus as the roughness/correlation length of a surface changes, its entropy will increase. Further, a surface cover comprising oblate spheroidal scatterers (leafs or discs for example) will generate an entropy betwee 0.6 and 0.7 (see [21]). In Fig. 4 we set a bound of $H = 0.9$ as an upper limit for these changes.
- **Zone 5: Medium Entropy Vegetation Scattering**
  Here again we have moderate entropy but with a dominant dipole type scattering mechanism. The increased entropy is due to a central statistical distribution of orientation angle. Such a zone would include scattering from vegetated surfaces with anisotropic scatterers and moderate correlation of scatterer orientations.
- **Zone 4: Medium Entropy Multiple Scattering**
  This zone accounts for dihedral scattering with moderate entropy. This occurs for example in forestry applications, where double bounce mechanisms occur at P and L bands following propagation through a canopy. The effect of the canopy is to increase the entropy of the scattering process. A second important process in this category is urban areas, where dense packing of localized scattering centres can generate moderate entropy with low order multiple scattering dominant.
  The boundary between zones 4, 5, 6 and 1, 2, 3 is set as 0.9. This is chosen on the basis of the upper limit for surface, volume and dihedral scattering before random distributions apply. The two classes which lie above this entropy value follow directly from (14) and (20).
- **Zone 3: High Entropy Surface Scatter**
  This class is not part of the feasible region in H-α space i.e., we cannot distinguish surface scattering with entropy $H > 0.9$. This is a direct consequence of our increasing inability to classify scattering types with increasing entropy. It is included to reinforce the idea that increasing entropy really does limit our ability to use polarimetric behavior to classify targets. Radar Polarimetry will then, in our view, be most successfully applied to low entropy problems.
- **Zone 2: High Entropy Vegetation Scattering**
  High entropy volume scattering arises when $\alpha = 45°$ and $H = 0.95$. This can arise for single scattering from a cloud of anisotropic needle like particles or from multiple scattering from a cloud of low loss symmetric particles. In both cases however, the entropy lies above 0.9, where the feasible region of H-α space is rapidly shrinking. Scattering from forest canopies lies in this region, as does the scattering from some types of vegetated surfaces with random highly anisotropic scattering elements. The extreme behavior in this class is random noise i.e., no polarization dependence, a point which lies to the extreme right of Fig. 4.
- **Zone 1: High Entropy Multiple Scattering**
  From T_{II} in (21), we see that in the $H > 0.9$ region we can still distinguish double bounce mechanisms in a high entropy environment. Again such mechanisms can be observed in forestry applications or in scattering from vegetation which has a well developed branch and crown structure.

There is of course some degree of arbitrariness over where to locate the boundaries in Fig. 4, based for example on knowledge of Radar calibration, measurement noise floor, variance of parameter estimates etc. We offer this segmentation of the H-α space merely to illustrate our classification strategy and to emphasise the geometrical segmentation of physical scattering processes. It is this key feature which makes this an unsupervised, data independent approach to the target classification problem. Nonetheless, we shall see in Section VI that these boundaries, although rather simply chosen, do offer sensible segmentation of experimental SAR data for well known test sites. This again lends support to our idea that the scheme is closely linked with physical scattering mechanisms.

VI. Application to Polarimetric SAR Data

In the analysis of experimental POLSAR data, we generally have access either to complete coherent scattering matrix data or to multi-look averaged Stokes matrix data. In either case, we can form local estimates of the coherency matrix using pixel averaging as

$$\langle [T] \rangle = \frac{1}{N} \sum_{i=1}^{N} \underline{k}_i \underline{k}_i^{*T}$$

$$\underline{k} = \frac{1}{\sqrt{2}} [S_{HH} + S_{VV} \quad S_{HH} - S_{VV} \quad 2S_{HV}] \tag{22}$$

$$\langle [T] \rangle = F\left(\frac{1}{N} \sum_{i=1}^{N} [M_i] \right)$$

where $[M_i]$ is the 4×4 real symmetric Stokes reflection matrix and the function $F(\cdots)$ relating M and T is detailed in [14]. From this estimate, we can then calculate the eigenvalues and eigenvectors of $[T]$ and use these to calculate the secondary parameters such as entropy H, the probabilities P_1 P_2 and P_3 and the average feature vector $\underline{f}$.

Here we consider three very different SAR data sets, to demonstrate the generic properties of the classifier and to illustrate that the feature vector $\underline{f}$ does indeed contain the physics of the scattering processes. All data has been taken from the NASA/JPL AIRSAR data base and we shall consider

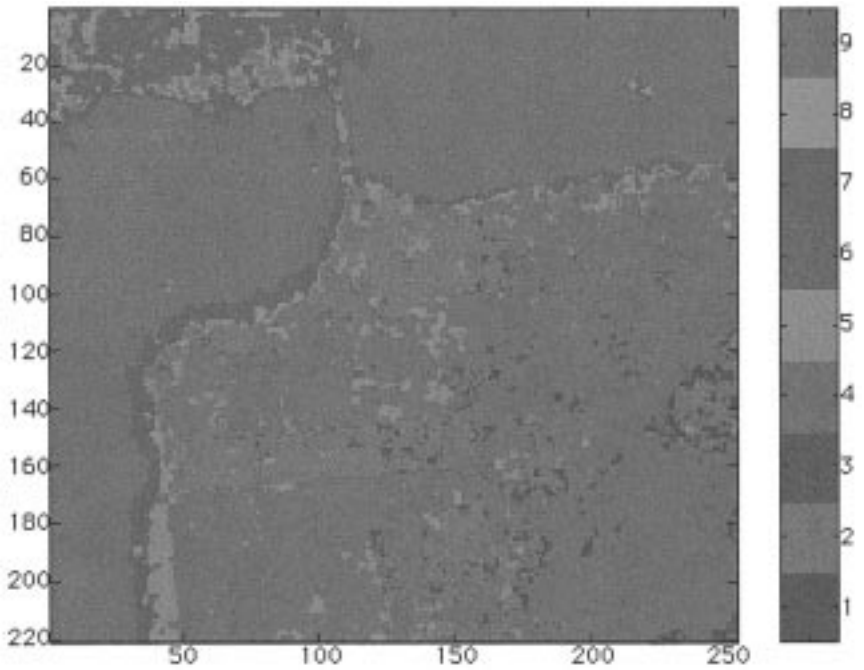

Fig. 5. L-band AIRSAR classification map: San Francisco region (red = zones 1-4-7, green = zones 2-5-8, blue = zones 3-6-9 of Fig. 4).

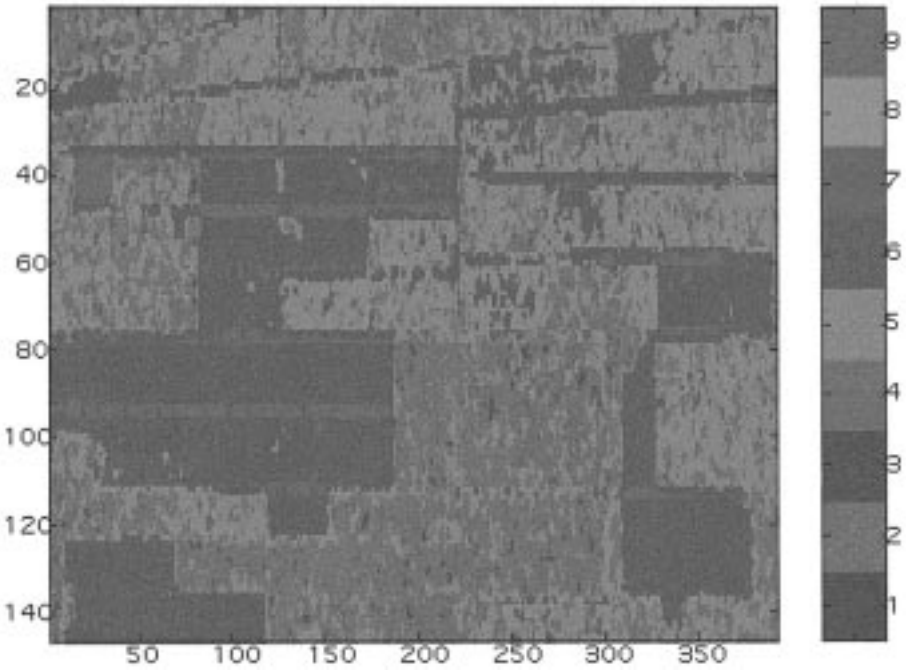

Fig. 6. P-band AIRSAR classification map: Landes forest(red = zones 1-4-7, green = zones 2-5-8, blue = zones 3-6-9 of Fig. 4).

the following examples; L-Band data for the San Francisco mixed urban/vegetation/ocean scene and two frequency (P and C-Band) data for the Landes forest in south-west France. Figs. 5 and 8 show results for the San Francisco data set. Fig. 5 shows the nine-zone classification results in a color coded format. We can see that the classifier has correctly identified the key physical features of this scene such as the sea surface, the coastal zone, the urban areas (with moderate entropy dihedral scattering) and finally the vegetated areas which are clearly identified on the basis of their H-α behavior. Note that both H and α are required to correctly classify these zones. Also note that no absolute reflectivity levels have been used in the classification. The entropy and alpha parameters are both relative eigenvalue measures and this alleviates problems associated with the need for absolute radiometric calibration of the radar. It is important to realize that no training data was required to implement this classifier: the data is classified according to its location in the H-α plane and to our physical understanding of different scattering mechanisms in this plane.

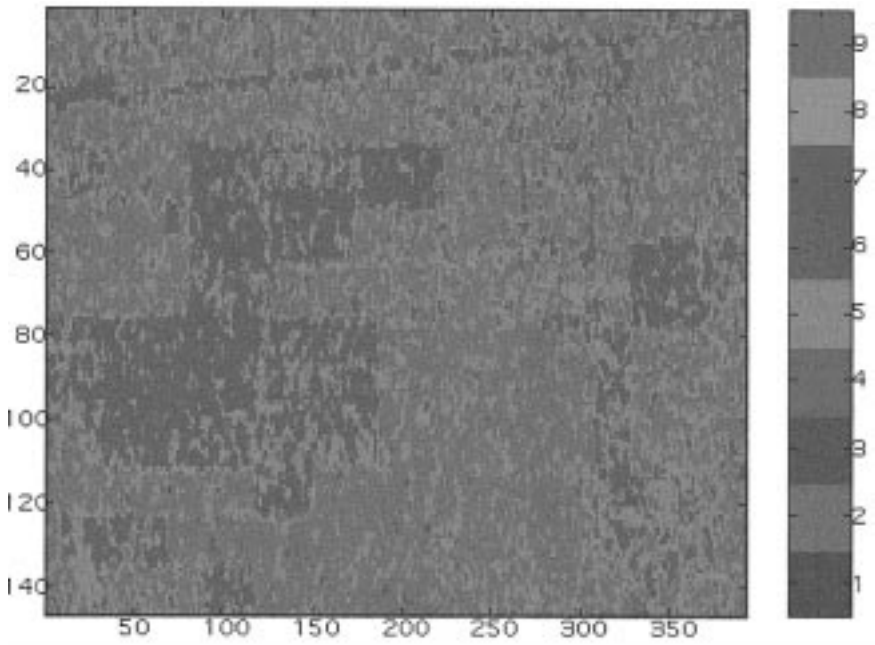

Fig. 7. C-band AIRSAR classification map: Landes forest (red = zones 1-4-7, green = zones 2-5-8, blue = zones 3-6-9 of Fig. 4).

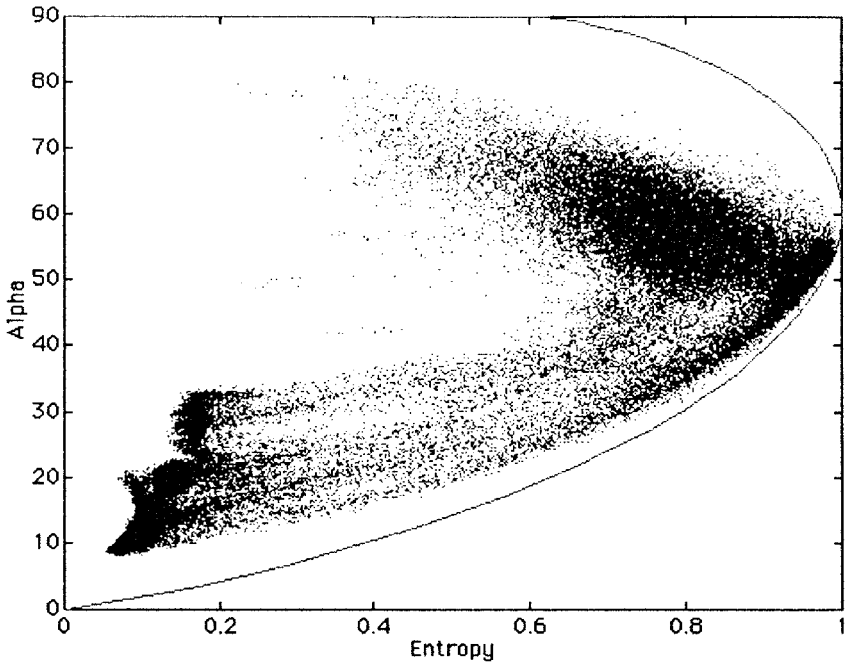

Fig. 8. L-band AIRSAR α-H distribution: San Francisco region (data from Fig. 5).

Fig. 8 shows the distribution of pixels for this scene, mapped point-wise into the parameter space of Fig. 4. We can clearly see the sea surface data in zone 9, while the urban zones are seen as a group of pixels in the upper right hand portion of the map. Note that all data points are bounded by the curves I and II, as explained in Section V.

Fig. 6 shows the classification map for the Landes forest data at P-Band. We can see a clear distinction between clear cut and forested regions (the former having low entropy) as well as seeing zones (in red) where significant moderate to high entropy dihedral scattering occurs. This is due to penetration of the canopy by the radar signal at the longer P-Band wavelength. For comparison, Fig. 7 shows the same scene imaged at C-Band. Here the scattering is dominated by the top layer of the forest and so the classification behavior is more uniform. The dihedral scattering zones are no longer so apparent and the clear cut zones are less well resolved, due to the similarity of scattering behavior in entropy and alpha. In

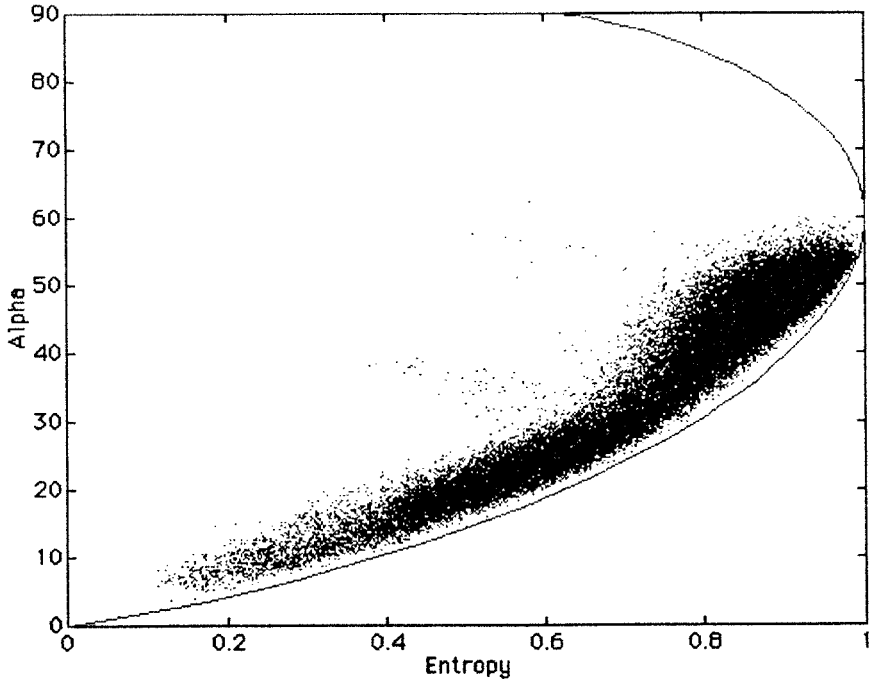

Fig. 9. P-band AIRSAR α-H distribution: Landes forest (data from Fig. 6).

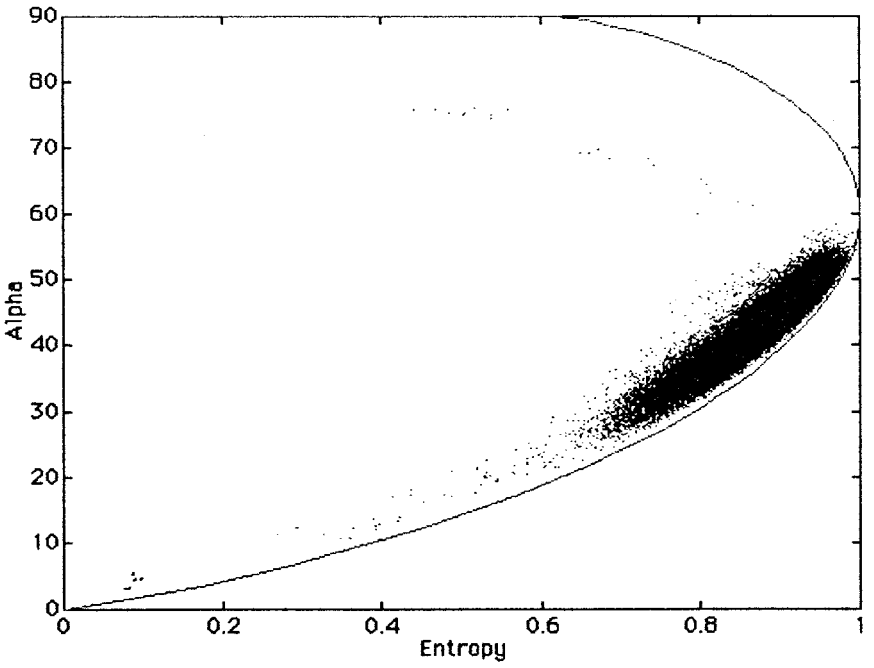

Fig. 10. C-band AIRSAR α-H distribution: Landes forest (data from Fig. 7).

this way, the H-α classification procedure reflects changes in the physical scattering mechanisms with radar wavelength.

Figs. 9 and 10 show the corresponding distributions of pixel α-H values for the P and C Band data sets, respectively. We can see that the distribution is more concentrated at high entropy for the C-Band data, resulting from the fact that canopy volume scattering mechanisms are predominate at the shorter wavelengths.

VII. Conclusion

In this paper we have extended the eigenvector parameterization of the coherency matrix in two ways: to account for both lack of target reflection symmetry and for the case of nondegenerate minor eigenvalues. By maintaining the idea that each region has a dominant scattering mechanism, we employ a three-symbol Bernoulli model to obtain estimates of

the average parameters for the dominant target. In this way we can maintain the simple idea that each zone in a SAR image has a single average scattering mechanism. Further, these parameters have been shown to be directly related to physical structure in the object and so of potential use in the inversion of experimental data.

We have applied the formalism to two direct scattering problems, namely multiple scattering by spherical particles in the diffusive wave approximation and single scattering by a cloud of anisotropic particles. We have shown from the first problem that entropy arises as a natural measure of the inherent reversibility of the scattering data and have further shown from the second that the alpha angle (derived from the eigenvectors of $[T]$) can be used to identify the underlying average scattering mechanism. For example, for a cloud of dipoles, the dominant scattering eigenvector is of the form $\underline{e} = (1, 0, 0)^T$ which is associated with scattering from an object with spherical symmetry. However, by adopting the averaging procedure outlined in this paper, the dominant scattering eigenvector becomes $\underline{e} = (0.707, 0.707, 0)^T$ i.e., dipole scattering, which more closely matches the true average scattering behavior of the cloud.

We suggest that this scheme may be used for unsupervised classification of multi-frequency polarimetric SAR data as a first stage process before attempting inversion of the data, based on particular physical models. It is our aim to next extend this classification scheme into inversion, based on physical and scattering models for surface and volume features.

REFERENCES

[1] J. A. Kong, S. H. Yueh, R. T. Shin, and J. J. van Zyl, "Classification of earth terrain using polarimetric synthetic aperture radar images," *PIER*, J. A. Kong, Ed. New York: Elsevier, vol. 3, ch. 6, 1990.

[2] H. A. Zebker, J. J. van Zyl, S. L. Durden, and L. Norikane, "Calibrated imaging radar polarimetry: Techniques examples and applications," *IEEE Trans. Geosci. Remote Sensing*, vol. 29, pp. 942–961, 1991.

[3] Y. Hara, R. G. Atkins, S. H. Yueh, R. T. Shin, and J. A. Kong, "Application of neural networks to radar image classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 100–110, Jan. 1994.

[4] E. Pottier, "On full polarimetric target decomposition theorems with application to classification and identification of real target cross section," in *Proc. Int. Radar Conf.*, Paris, France, May 1994, pp. 330–335.

[5] J. S. Lee and M. R. Grunes, "Classification of multi-look polarimetric SAR imagery based on the complex Wishart distribution," *Int. J. Remote Sensing*, vol. 15, no. 11, pp. 2299–2311, 1994.

[6] J. J. van Zyl and C. F. Burnette, "Bayesian classification of polarimetric SAR images using adaptive a-priori probabilities," *Int. J. Remote Sensing*, vol. 13, pp. 835–840, 1992.

[7] J. J. van Zyl, "Unsupervised classification of scattering behavior using radar polarimetry data," *IEEE Trans. Geosci. Remote Sensing*, vol. 27, pp. 36–45, 1989.

[8] A. Freeman and S. Durden, "A three component scattering model to describe polarimetric SAR data," *SPIE*, Radar Polarimetry, vol. 1748, pp. 213–225, 1992.

[9] A. Freeman, S. Durden, and R. Zimmerman, "Mapping sub-tropical vegetation using multi-frequency multi-polarization SAR data," in *Proc. IGARSS*, Houston, TX, June 1992, pp. 1686–1689.

[10] J. J. van Zyl, "Application of Cloude's target decomposition theorem to polarimetric imaging radar data," *SPIE*, Radar Polarimetry, vol. 1748, pp. 184–212, 1992.

[11] M. Ferri, L. Castellano, A. Siciliano, R. Vigliotti, and P. Murino, "Using polarimetric SAR data in morphological analyses: The island of Ischia, Southern Italy," in *Proc. IGARSS'94*, CalTech, Pasadena, CA, July 1994.

[12] S. R. Cloude, "Symmetry, zero correlations and target decomposition theorems," in *Proc. 3rd Int. Workshop on Radar Polarimetry (JIPR'95)*, IRESTE, University of Nantes, Mar. 1995, pp. 58–68.

[13] ——, "An entropy based classification scheme for polarimetric SAR data," in *Proc. IGARSS'95*, Florence, Italy, July 1995, pp. 2000–2002.

[14] S. R. Cloude and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 498–518, Mar. 1996.

[15] Ngheim, S. H. Yueh, R. Kwok, and F. K. Li, "Symmetry properties in polarimetric remote sensing," *Radio Sci.*, vol. 27. no. 5, pp. 693–711, Oct. 1992.

[16] S. R. Cloude, "Lie groups in electromagnetic wave propagation and scattering," *J. Electromag. Waves Applicat.*, vol. 6, no. 8, pp. 947–974, 1992.

[17] L. Tsang, J. A. Kong, and R. T. Shin, *Theory of Microwave Remote Sensing*. New York: Wiley Interscience, 1985.

[18] D. Bicout and C. Brosseau, "Multiply scattered waves through a spatially random medium: Entropy production and depolarization," *J. Phys. I*, France, vol. 2, pp. 2047–2063, Nov. 1992.

[19] D. S. Wiersma, M. P. van Albada, and A. Lagendijk, "Precise weak localization experiments reveal recurrent light scattering in random structures," in *Proc. PIERS'94*, Noordwijk, The Netherlands, July 1994, p. 192.

[20] S. R. Cloude and E. Pottier, "The concept of polarization entropy in optical scattering," *Opt. Eng.*, vol. 34, no. 6, ISSN 0091-3286, pp. 1599–1610.

[21] S. R. Cloude, "Uniqueness of target decomposition theorems in radar polarimetry," *Direct and Inverse Methods in Radar Polarimetry*, Part 1, NATO-ARW, W. M. Boerner *et al.*, Eds. New York: Kluwer Academic, 1992, pp. 267–296.

Shane Robert Cloude (M'87–SM'96), for a photograph and biography, see p. 518 of the March 1996 issue of this TRANSACTIONS.

Eric Pottier (M'95) was born in 1962. He received the "Signal Processing and Telecommunications" D.E.A. degree in 1987 and the Ph.D. degree in 1990, both from the University of Rennes, France.

In 1988, he joined the laboratory "Systemes et Signaux Hautes Frequences" at IRESTE, University of Nantes, France, where he was Research and Teaching Assistant and involved in analog electronic and microwave systems. Since 1991 he has held an Assistant Professor position. His current activities of research and education are centered in the topics of analog electronic, microwave theory, and radar imaging, with emphasis in radar polarimetry. Since November 1994, he has been the Head of the Radar Polarimetry Group of the S.E.I. (Systemes Electroniques et Informatiques) Laboratory, which is associated with C.N.R.S. (EP 0063). He has more than 50 publications in the area of radar polarimetry and has supervised 15 research students to graduation.