Shannon Entropy for the $\mathcal{G}_I^0$ Model: A New Segmentation Approach

Jodavid de A. Ferreira and Abraão D. C. Nascimento

***Abstract*—Synthetic aperture radar (SAR) has been successfully used as a remote sensing tool. However, SAR images are contaminated by speckle noise and require specialized postprocessing procedures; e.g, tailored segmenters. The $\mathcal{G}_I^0$ distribution is a flexible model for SAR intensities because of its ability at describing heterogeneous clutters. Furthermore, applying information theory measures (e.g., entropy) to extract features in SAR imagery processing has achieved a prominent position. In this article, we derive both a closed-form expression for the $\mathcal{G}_I^0$ Shannon entropy and some of its mathematical properties. Consequently new entropy-based segmentation procedures for multidimensional SAR intensities—assuming independence or some dependence pattern—are also proposed. Finally, applications to real SAR imagery point out the proposed entropy-based segmenters can be more efficient than other well-defined methods, like the clustering by gamma mixture models.**

***Index Terms*—$\mathcal{G}_I^0$ distribution, clustering model, information theory (IT), stochastic entropy.**

List of Notations

- $\boldsymbol{H}_S$: Parametric matrix of entropies, where its rows indicate the entries of a SAR image and columns represent the different polarization channels.
- $\widehat{\boldsymbol{H}}_S$: Maximum likelihood estimator for $\boldsymbol{H}_S$.
- $\dot{\boldsymbol{H}}_S$: A possible outcome of $\widehat{\boldsymbol{H}}_S$.
- $\boldsymbol{H}_S(i)$: Parametric vector of entropies of the ith entry of a SAR image.
- $\widehat{\boldsymbol{H}}_S(i)$: Maximum likelihood estimator for $\boldsymbol{H}_S(i)$.
- $\dot{\boldsymbol{H}}_S(i)$: A possible outcome of $\widehat{\boldsymbol{H}}_S(i)$.
- $\boldsymbol{\mu}_{S_g}$: Expected value of a vector of stochastic entropies given it belongs to the gth population.
- $\boldsymbol{\Sigma}_{S_g}$: Covariance matrix of a vector of stochastic entropies given it belongs to the gth population.

I. Introduction

SYNTHETIC aperture radar (SAR) has been widely employed to capture and record information of geographic scenes in remote sensing applications. However, by employing coherent illumination in its processing, SAR systems produce images that are affected by a signal-dependent granular noise called "speckle." This noise precludes analyzing images directly.

Supposing probabilistic models to describe SAR features (taking into account the speckle) has become a promising tendency like an imagery processing technique. In particular, the $\mathcal{G}_I^0$ model proposed by Frery *et al.* [1] has been widely used by its capability at describing heterogeneous scenes; i.e., built-up areas.

Many works have provided evidence in favor of the use of information theory (IT) measures for understanding SAR imagery. Morio *et al.* [2] applied the Shannon entropy (SE) for characterizing polarimetric interferometric SAR images, decomposing it into the sum of three terms which have physical meaning. Such a discussion did not involve the statistical properties of entropy measures. In order to outperform this gap, Frery *et al.* [3] have studied the distribution of contrast measures based on scaled complex Wishart entropies into the h-ϕ class. To the best of our knowledge, two interesting open issues are: first, what about closed-form expressions for the $\mathcal{G}_I^0$ entropies and second, how to apply both these quantities and their potential distributions in processing (particularly, in clustering analysis) of SAR returns.

Despite the use of multilook processing imposes a kind of control in the speckle effect, this last becomes SAR imagery segmentation difficult [4]. To outperform this issue, Zaart *et al.* [5] have derived a histogram-based estimation method for thresholds in segmentation issues. Sui *et al.* [6] have tackled a segmentation method by means of a multiscale level set approach. Using multikernel sparse representation, Gu *et al.* [7] furnished new segmenters for SAR imagery. Zhao *et al.* [8] proposed a multilook SAR image segmentation algorithm using gamma mixture model (GaMM) and a Markov random field.

In this article, we derive the $\mathcal{G}_I^0$ SE and some of its mathematical properties. We also provide a new entropy-based segmentation paradigm for SAR imagery, discussing as special cases resulting segmenters when Γ and $\mathcal{G}_I^0$ laws under independence or some dependence structure are used. This approach is formulated from the distribution of the stochastic entropy. Finally, our proposals are applied to multilook multidimensional SAR images of Foulum (Denmark), Munich (Germany), and Flevoland (Denmark) regions. Results give evidence the new segmentation methods can outperform other well-defined segmenters, such as the one based on GaMM.

The rest of this article is organized as follows. Section II approaches a used statistical modeling background. In Section III,

Manuscript received March 18, 2020; revised May 18, 2020; accepted May 20, 2020. Date of publication May 28, 2020; date of current version June 8, 2020. This work was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo a Ciência e Tecnologia de Pernambuco (FACEPE), Brazil. *(Corresponding author: Jodavid de A. Ferreira.)*

The authors are with the Department of Statistics, Universidade Federal de Pernambuco, Recife 50740-540, Brazil (e-mail: jodavid.arts@gmail.com; abraao@de.ufpe.br).

Digital Object Identifier 10.1109/JSTARS.2020.2997666

expressions for SEs are presented for the $\mathcal{G}_I^0$ and Γ distributions. A new entropy-based segmentation paradigm is discussed in Sections IV. Section V displays a comparative study of segmentation methods at real data. Section VI summarizes the main results.

II. Statistical Modeling for SAR Data

The speckled noise—natural from capturing SAR imagery—is neither Gaussian nor additive [1]. The multiplicative model (MM) is a successful approach for describing SAR returns [1], since it emerges from the image formation physics of these systems.

Let Y and X be scalar independent positive random variables such that X models the terrain "backscatter", whereas Y describes the speckle noise. The MM assumes each "intensity" (square norm of a complex polarization channel) behind a SAR image pixel is an outcome of a random variable Z, which is the product of X and Y; i.e., $Z = Y \times X$.

The backscatter carries all the relevant information from the mapped area. In particular, it involves target physical behavior; for instance, moisture, and relief. In this article, we assume a distributed *reciprocal gamma* backscatter, $X \sim \Gamma^{-1}(-\alpha, \gamma)$ [1].

The speckle is exponentially distributed with unitary mean in intensities of single-look images [1]. As one way to extend the previous assumption, the multilook procedure (sample mean) over L independent observations furnishes an intensity speckle variable described by the gamma distribution, $Y \sim \Gamma(L, L)$.

From previous assumptions for X and Y, one can show Z has density [1]

$$f_Z(z; \alpha, \gamma, L) = \frac{L^L \Gamma(L-\alpha)}{\gamma^\alpha \Gamma(-\alpha)\Gamma(L)} \frac{z^{L-1}}{[\gamma + Lz]^{L-\alpha}} \mathbb{I}(0,\infty)(z)$$

where $\mathbb{E}(Z) = \gamma(-\alpha - 1)^{-1}$ for roughness $\alpha < -1$. We denote this situation as $Z \sim \mathcal{G}_I^0(\alpha, \gamma, L)$. In this article, we use $\mathcal{G}_I^0(\alpha, \gamma, L)$ in contrast with $\Gamma(L, L/\mu)$, which is other important MM solution and limit case of the former [for more details, see 1].

III. Stochastic $\mathcal{G}_I^0$ SE and its Asymptotic Distribution

Important tools for statistical inference and image processing [2] are defined from the IT. The concepts of "information" and "entropy" were introduced formally in the context of the communication theory [9]. Thenceforth, the proposition and application of IT measures (in particular, entropy) have become rule in several areas. Zografos and Nadarajah have derived closed-form expressions for Shannon and Rényi entropies for multivariate distributions [10]. In particular, entropy is a fundamental concept, which has been receiving various interpretations, such as the notions of disorder in mechanical statistics [3] and diversity in economics and biology [9]. The (h, ϕ)-entropy class–which generalizes the original entropy concept–and its associated asymptotic distribution are discussed by Pardo [9].

A. The H-ϕ Class of Entropies

Let $f_Z(z; \boldsymbol{\theta})$ be pdf of Z with parameter vector $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^p$. The class (h, ϕ)-entropy of Z is defined by

$$H_\phi^h(\boldsymbol{\theta}) = h\left(\int_{\mathcal{A}} \phi(f_Z(z; \boldsymbol{\theta}))\mathrm{d}z\right)$$

where either $\phi : [0, \infty) \to \mathbb{R}$ is concave and $h : \mathbb{R} \to \mathbb{R}$ is increasing, or ϕ is convex and h is decreasing. In particular, the SE is obtained when $h(x) = x$ and $\phi(x) = -x\log(x)$.

The next result opens a mode of studying entropy-based asymptotic statistical inference methods.

Lemma 1: Let $\widehat{\boldsymbol{\theta}} = [\widehat{\theta_1},\ \widehat{\theta_2},\ \ldots,\ \widehat{\theta_p}]^\top$ be the maximum likelihood (ML) estimator for $\boldsymbol{\theta} = [\theta_1,\ \theta_2,\ \ldots,\ \theta_p]^\top$ based on a N-points random sample, $Z_1, \ldots, Z_N$, under the model $f(z; \boldsymbol{\theta})$. Then,

$$\sqrt{N}\left[H_h^\phi(\widehat{\boldsymbol{\theta}}) - H_h^\phi(\boldsymbol{\theta})\right] \xrightarrow[N\to\infty]{\mathcal{D}} \mathcal{N}\left(0, \sigma_\phi^2(\boldsymbol{\theta})\right) \tag{1}$$

where $\mathcal{N}(\mu, \sigma^2)$ is the Gaussian distribution with mean μ and variance σ^2, "$\xrightarrow{\mathcal{D}}$" denotes *convergence in distribution*

$$\sigma_\phi^2(\boldsymbol{\theta}) = \boldsymbol{\delta}^\top \boldsymbol{\mathcal{K}}(\boldsymbol{\theta})^{-1}\boldsymbol{\delta} \tag{2}$$

$\boldsymbol{\mathcal{K}}(\boldsymbol{\theta}) = \mathbb{E}\{-\partial^2 \log f_Z(Z; \boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top\}$ is the Fisher information matrix, and $\boldsymbol{\delta} = [\delta_1,\ \delta_2,\ \ldots,\ \delta_p]^\top$ such that $\delta_i = \partial H_h^\phi(\boldsymbol{\theta})/\partial\theta_i$ for $i = 1, 2, \ldots, p$.

In what follows, we derive some properties for the $\mathcal{G}_I^0$ model, which are need to construct new segmentation mechanisms.

B. SE-Based Inference for the $\mathcal{G}_I^0$ Model

Now, we look for answering about how to obtain a closed-form expression for $\mathcal{G}_I^0$ distribution entropy and to apply the asymptotic distribution for its stochastic version in processing of SAR intensities. It is important to note Nobre *et al.* [11] have worked with the $\mathcal{G}_A^0$ (counterpart of $\mathcal{G}_I^0$ for modeling amplitude instead of intensity) Rényi entropy, but without using closed-form expressions for the entropy nor working with distributions for entropy. They have obtained good segmentation results, but have indicated the use of numerical integration for entropy as a hard disadvantage of the proposed method. The next result furnishes a solve for the former question we put as well as for the open question in [11].

Theorem 1: Let $Z \sim \mathcal{G}_I^0(\alpha, \gamma, L)$, its SE is given by

$$\begin{aligned} H_{\mathrm{S}}(\mathcal{G}_I^0) = H_{\mathrm{S}}([\alpha, \gamma, L]) &= \mathbb{E}[-\log f_Z(Z; \alpha, \gamma, L)] \\ &= -\log\frac{L^L\Gamma(L-\alpha)}{\Gamma(-\alpha)\gamma^\alpha\Gamma(L)} \\ &\quad + (1-L)\Big\{\psi^{(0)}(L) - \log L + \log\gamma - \psi^{(0)}(-\alpha)\Big\} \\ &\quad + (L-\alpha)[\psi^{(0)}(L-\alpha) + \log\gamma - \psi^{(0)}(-\alpha)]. \end{aligned} \tag{3}$$

The proof of this result is given in Appendix.

Let $Z \sim \Gamma^*(L, \mu) := \Gamma(L, L/\mu)$, its SE is given by

$$\begin{aligned} H_{\mathrm{S}}(\Gamma^*) = H_{\mathrm{S}}([L, \mu]) &= L\left(\log L - \log\mu\right) \\ &\quad + \log\Gamma(L) + (1-L)\psi^{(0)}(L). \end{aligned} \tag{4}$$

(i) For the $\mathcal{G}_I^0$ FIM, $\mathcal{K}^{\mathcal{G}}_{\bullet\bullet}$ [beyond some first- and second-order derivatives in terms of L, elements extracted from 12]: $\mathcal{K}^{\mathcal{G}}_{\alpha\alpha} = N[\psi^{(1)}(-\alpha) - \psi^{(1)}(L-\alpha)]$; $\mathcal{K}^{\mathcal{G}}_{\alpha\gamma} = N\left[\frac{1}{\gamma} - \frac{\alpha}{\gamma(\alpha-L)}\right]$; $\mathcal{K}^{\mathcal{G}}_{\alpha L} = N\left[\psi^{(1)}(L-\alpha) - (L-\alpha)^{(-1)}\right]$; $\mathcal{K}^{\mathcal{G}}_{\gamma\gamma} = N\left[-\frac{\alpha}{\gamma^2} - \frac{\alpha(\alpha-1)}{(L-\alpha+1)\gamma^2}\right]$; $\mathcal{K}^{\mathcal{G}}_{\gamma L} = N\left[\frac{\alpha}{\gamma(\alpha-L)} + \frac{\alpha}{\gamma(L-\alpha+1)}\right]$; and $\mathcal{K}^{\mathcal{G}}_{LL} = N\left[\psi^{(1)}(L) - \frac{1}{L} - \psi^{(1)}(L-\alpha) + \frac{2}{(L-\alpha)} - \left(\frac{L+1}{L}\right)\frac{1}{(L-\alpha+1)}\right]$.

(ii) For the Γ^* FIM, $\mathcal{K}^{\Gamma}_{\bullet\bullet}$: $\mathcal{K}^{\Gamma}_{LL} = N\left[-L^{-1} + \psi^{(1)}(L)\right]$; $\mathcal{K}^{\Gamma}_{\mu\mu} = \frac{NL}{\mu^2}$; and $\mathcal{K}^{\Gamma}_{L\mu} = 0$.

Fig. 1. Fisher information matrix elements for the Γ^* and $\mathcal{G}^0$ models.

In this article, we assume the term "stochastic entropy" as the expression of entropy after replacing the model (on which it is obtained) parameters by their ML estimators. In order to study the asymptotic distribution of the $\mathcal{G}_I^0$ stochastic SE, say $\widehat{H}_{\mathrm{S}}(\mathcal{G}_I^0)$, the next result gives theoretical quantities to define the variance of $\widehat{H}_{\mathrm{S}}(\mathcal{G}_I^0)$.

Proposition 1: Let $Z \sim \mathcal{G}_I^0(\alpha,\gamma,L)$, $\boldsymbol{\delta}^{\top}_{\mathcal{G}_I^0} = (\delta^{\mathcal{G}}_{\alpha}, \delta^{\mathcal{G}}_{\gamma}, \delta^{\mathcal{G}}_{L}) := (\frac{\partial H(\mathcal{G}_I^0)}{\partial\alpha}, \frac{\partial H(\mathcal{G}_I^0)}{\partial\gamma}, \frac{\partial H(\mathcal{G}_I^0)}{\partial L})$ in (2), where

$$\delta^{\mathcal{G}}_{\alpha} = (1-\alpha)\psi^{(1)}(-\alpha) - (L-\alpha)\psi^{(1)}(L-\alpha), \quad \delta^{\mathcal{G}}_{\gamma} = \frac{1}{\gamma}$$

$$\text{and } \delta^{\mathcal{G}}_{L} = (L-\alpha)\psi^{(1)}(L-\alpha) - (L-1)\psi^{(1)}(L) - \frac{1}{L}.$$

The analog expressions for the speckle noise with nonunitary mean are given as follows.

Proposition 2: Let $Y \sim \Gamma^*(L,\mu)$

$$\boldsymbol{\delta}^{\top}_{\Gamma} = \left(\delta^{\Gamma}_{L}, \delta^{\Gamma}_{\mu}\right) := \left(\frac{\partial H(\Gamma^*)}{\partial L}, \frac{\partial H(\Gamma^*)}{\partial \mu}\right) \text{ in (2)}$$

where $\delta^{\Gamma}_{L} = (L-1)/L + (1-L)\psi^{(1)}(L)$ and $\delta^{\Gamma}_{\mu} = \mu^{-1}$.

Expressions in Propositions 1 and 2 are important because both Γ^* [in (4)] and $\mathcal{G}_I^0$ [in (3)] SEs as well as their distributions will be used as core tools to construct new segmentation procedures. To determine the variances of $\widehat{H}(\Gamma^*)$ and $\widehat{H}(\mathcal{G}_I^0)$, we need to find elements of the Fisher information matrices due to the Γ^* and $\mathcal{G}_I^0$, respectively. Fig. 1 displays these elements.

IV. New Entropy-Based Segmentation Algorithm

In what follows, we develop a new segmentation paradigm for multidimensional SAR images, taking into account the stochastic Γ^* and $\mathcal{G}_I^0$ SEs.

A. Segmentation Model

Polarimetric SAR (PolSAR) returns can be understood as outcomes from a sample either of complex random vectors (called "single look") or of Hermitian positive definite random matrices (called "multilook") [3]. In this article, we address the multilook case, proposing new segmentation methods for elements of the main diagonal of resulting matrices from multilook PolSAR images. Backscattered PolSAR signals are recorded in terms of four possible combinations of linear reception and transmission polarizations: HH, HV, VH, and VV (H for horizontal and V for vertical polarization). In practice, satisfied the reciprocity theorem conditions, HV and VH polarizations are identical. Formally, each multilook PolSAR entry, say $\boldsymbol{Z}$, is represented by a 3×3 Hermitian positive definite matrix

$$\boldsymbol{Z} = \begin{bmatrix} |Z_{\text{HH}}|^2 & Z_{\text{HH-HV}} & Z_{\text{HH-VV}} \\ Z^*_{\text{HH-HV}} & |Z_{\text{HV}}|^2 & Z_{\text{HV-VV}} \\ Z^*_{\text{HH-VV}} & Z^*_{\text{HV-VV}} & |Z_{\text{VV}}|^2 \end{bmatrix},$$

where $\{|Z_{\text{HH}}|^2, |Z_{\text{HV}}|^2, |Z_{\text{VV}}|^2\}$ represents the set of intensities from Z_{HH}, Z_{HV}, and Z_{VV} polarization channels values (complex random elements) and $\{Z_{\text{HH-HV}}, Z_{\text{HH-VV}}, Z_{\text{HV-VV}}\}$ indicates the set of possible inner products between two different polarization channels such that $Z_{\text{A-B}} = \langle Z_{\text{A}}, Z_{\text{B}}\rangle = Z_{\text{A}}Z^*_{\text{B}}$, for A, B $\in \{\text{HH}, \text{HV}, \text{VV}\}$, and * denotes the conjugate of a complex number. The intensities of the echoed signal polarization channels play an important role, since they depend on the physical properties of the target surface.

Let $\boldsymbol{I}(i) = [I_1(i), I_2(i), I_3(i)]^{\top} \in \mathbb{R}^3_+$ for $i = 1, \ldots, T$ (number of pixels of the under-study image) be the vector of intensities of HH, HV, and VV channels at the ith entry of a multilook PolSAR image and $\boldsymbol{H}_S(i) = [H_1(i), H_2(i), H_3(i)]^{\top}$ the vector of entropies at the ith entry. Assume also the random representations for entropies at entries of one image (for a specified channel) are independent.

Consider that the under-study image has c regions and there is a membership matrix $\boldsymbol{U} = [U_{ig}]_{T\times c}$ such that $U_{ig} \in \{0,1\}$ is the index which indicates if the ith pixel belongs to the gth group ($U_{ig} = 1$) or not ($U_{ig} = 0$) and $\sum_{g=1}^{c} U_{ig} = 1$. In the segmentation practice, an estimate for $\boldsymbol{H}_S = [\boldsymbol{H}_S(i)|\cdots|\boldsymbol{H}_S(T)]^{\top}$, say $\dot{\boldsymbol{H}}_S$, is an observable database, $\boldsymbol{U}$ is unobservable one, and $\{\dot{\boldsymbol{H}}_S, \boldsymbol{U}\}$ is the complete data set. A segmenter can be understood as an algorithm for estimating the unobservable matrix $\boldsymbol{U}$ given $\dot{\boldsymbol{H}}_S$. In what follows, we detail these quantities.

Sample entropies are obtained by plugging ML estimates based on the neighborhood of $\boldsymbol{I}(i)$ in entropies (3) and (4). When the result (1) holds, the variable $\widehat{\boldsymbol{H}_S}(i)$ conditioned the

gth region follows the multivariate normal distribution having density

$$p\left(\dot{\boldsymbol{H}}_S(i)|U_{ig}=1;\boldsymbol{\theta}_g\right)=f_{\mathcal{N}}(\dot{\boldsymbol{H}}_S(i);\boldsymbol{\theta}_g)=|2\pi\boldsymbol{\Sigma}_g|^{-\frac{1}{2}} \times\exp\left\{-\frac{1}{2}\left[\dot{\boldsymbol{H}}_S(i)-\boldsymbol{\mu}_g\right]^{\top}\boldsymbol{\Sigma}_g^{-1}\left[\dot{\boldsymbol{H}}_S(i)-\boldsymbol{\mu}_g\right]\right\} \quad (5)$$

where $\dot{\boldsymbol{H}}_S(i)=[\dot{H}_1(i),\dot{H}_2(i),\dot{H}_3(i)]^{\top}$ is an outcome of $\widehat{\boldsymbol{H}_S}(i)$ with (at the gth region) mean $\boldsymbol{\mu}_g=[\mu_{g:1},\mu_{g:2},\mu_{g:3}]^{\top}=\mathbb{E}[\widehat{\boldsymbol{H}_S}(i)|U_{ig}=1]$ and covariance matrix $\boldsymbol{\Sigma}_g=\{\boldsymbol{\sigma}_{g:l_1,l_2}\}_{l_1,l_2=1,2,3}=\mathrm{Cov}[\widehat{\boldsymbol{H}_S}(i)|U_{ig}=1]$, $\boldsymbol{\theta}_g=[\boldsymbol{\mu}_g^{\top},\mathrm{vec}(\boldsymbol{\Sigma}_g)^{\top}]^{\top}$ is the parametric configuration in the gth region, and $\mathrm{vec}(\cdot)$ is the vectorization operator.

According to the Bayesian theory [13], the joint density of $\widehat{\boldsymbol{H}}_S(i)$ and U_{ig} is

$$p(\dot{\boldsymbol{H}}_S(i),U_{ig};\boldsymbol{\theta}_g)=p(U_{ig}=1)p\left(\dot{\boldsymbol{H}}_S(i)|U_{ig}=1;\boldsymbol{\theta}_g\right) \quad (6)$$

where $\boldsymbol{\theta}=[\boldsymbol{\theta}_1^{\top},\ldots,\boldsymbol{\theta}_c^{\top}]^{\top}$. Assuming $\widehat{\boldsymbol{H}}_S(i)$ and $\widehat{\boldsymbol{H}}_S(j)$ are independent and adopting the prior probability of the ith pixel belongs to gth group as $p(U_{ig}=1)=\pi_g\in[0,1]$ such that $\sum_{g=1}^{c}\pi_g=1$, then the whole image joint density, $\widehat{\boldsymbol{H}}_S$, is

$$p\left(\dot{\boldsymbol{H}}_S;\boldsymbol{\theta}\right)=\prod_{i=1}^{T}\left\{\sum_{g=1}^{c}\pi_g\left[f_{\mathcal{N}}(\dot{\boldsymbol{H}}_S(i);\boldsymbol{\theta}_g)\right]\right\} \quad (7)$$

which is the segmentation model.

B. Estimation of Model Parameters Based on the EM Algorithm

Based on (7), the conditional probability of the ith pixel belonging to the gth region in the kth iteration given an image is [14]

$$p\left(U_{ig}=1|\dot{\boldsymbol{H}}_S(i);\hat{\boldsymbol{\theta}}^{(k)}\right)=\hat{\tau}_g\left(\hat{\boldsymbol{\theta}}^{(k)}|\dot{\boldsymbol{H}}_S(i)\right)=\frac{\hat{\pi}_g^{(k)}(i)\left[f_{\mathcal{N}}\left(\dot{\boldsymbol{H}}_S(i);\hat{\boldsymbol{\theta}}_g^{(k)}\right)\right]}{\sum_{g=1}^{c}\hat{\pi}_g^{(k)}(i)\left[f_{\mathcal{N}}\left(\dot{\boldsymbol{H}}_S(i);\hat{\boldsymbol{\theta}}_g^{(k)}\right)\right]}. \quad (8)$$

The updating equation of $\hat{\pi}_g^{(k+1)}(i)$ is

$$\hat{\pi}_g^{(k+1)}(i)=\frac{1}{T}\sum_{i=1}^{T}\hat{\tau}_g(\hat{\boldsymbol{\theta}}^{(k)}|\dot{\boldsymbol{H}}_S(i)). \quad (9)$$

The update equations for $\widehat{\boldsymbol{\mu}}_g^{(k+1)}=[\widehat{\mu}_{g:1}^{(k+1)},\widehat{\mu}_{g:2}^{(k+1)},\widehat{\mu}_{g:3}^{(k+1)}]^{\top}$ and $\widehat{\boldsymbol{\Sigma}}_g^{(k+1)}=\{\hat{\sigma}_{g:l_1,l_2}^{(k+1)}\}_{l_1,l_2=1,2,3}$ are given by, respectively

$$\widehat{\boldsymbol{\mu}}_g^{(k+1)}=\frac{\sum_{i=1}^{T_g}\left[\hat{\pi}_g^{(k)}(i)f_{\mathcal{N}}\left(\dot{\boldsymbol{H}}_S(i);\hat{\boldsymbol{\theta}}_g^{(k)}\right)\right]\dot{\boldsymbol{H}}_S(i)}{\sum_{i=1}^{T_g}\left[\hat{\pi}_g^{(k)}(i)f_{\mathcal{N}}\left(\dot{\boldsymbol{H}}_S(i);\hat{\boldsymbol{\theta}}_g^{(k)}\right)\right]} \quad (10)$$

and

$$\widehat{\boldsymbol{\Sigma}}_g^{(k+1)}=\frac{\sum_{i=1}^{T_g}\left[\hat{\pi}_g^{(k)}(i)f_{\mathcal{N}}\left(\dot{\boldsymbol{H}}_S(i);\hat{\boldsymbol{\theta}}_g^{(k)}\right)\right]\boldsymbol{\Delta}_g^{(i)}}{\sum_{i=1}^{T_g}\left[\hat{\pi}_g^{(k)}(i)f_{\mathcal{N}}\left(\dot{\boldsymbol{H}}_S(i);\hat{\boldsymbol{\theta}}_g^{(k)}\right)\right]} \quad (11)$$

where $\boldsymbol{\Delta}_g^{(i)}=(\dot{\boldsymbol{H}}_S(i)-\widehat{\boldsymbol{\mu}}_g^{(k+1)})(\dot{\boldsymbol{H}}_S(i)-\widehat{\boldsymbol{\mu}}_g^{(k+1)})^{\top}$ and T_g is the pixels numbers at the gth region for an under-study image. Once there are $p(p+1)/2$ parameters to be estimated in (11), it is common to impose more parsimony structures in Gaussian mixture models.

Celeux and Govaert [15] have introduced fourteen ones by means of different eigen-decompositions of the covariance matrix, as

$$\boldsymbol{\Sigma}_g=\lambda_g\mathbf{D}_g\mathbf{A}_g\mathbf{D}_g^{\top}$$

where $\mathbf{D}_g$ is the matrix of eigenvectors, $\mathbf{A}_g$ is a diagonal matrix with entries proportional to the eigenvalues, and λ_g is a constant of proportionality. Model names are associated with volume, shape, and orientation. The model we used in this article is called VVV, meaning that the volume, shape, and orientation are variable. This model has showed the best results for SAR images according to the proposed methodology [for more details, see 16].

C. Stopping Criterion

At each iteration it is necessary to verify the convergence of the algorithm. There are several ways of measuring convergence of an algorithm. One of them uses criteria based on the absolute error between steps. In this case, we use the c-sample entropy-based statistics to calculate the absolute error [9]

$$S^{(k)}=\sum_{l=1}^{3}S_l^{(k)}\equiv\sum_{l=1}^{3}\left[\sum_{g=1}^{c}\frac{T_g\left(\widehat{\mu}_{g:l}^{(k+1)}-\overline{D}_l^{(k+1)}\right)^2}{\hat{\sigma}_{g:l,l}^{(k+1)}}\right] \quad (12)$$

where

$$\overline{D}_l^{(k+1)}=\frac{1}{\sum_{g=1}^{c}\frac{T_g}{\hat{\sigma}_{g:l,l}^{(k+1)}}}\sum_{g=1}^{c}\frac{T_g}{\hat{\sigma}_{g:l,l}^{(k+1)}}\widehat{\mu}_{g:l}^{(k+1)}.$$

If the absolute error between steps k and $(k+1)$ is lower than a certain specified threshold (i.e., $|S^{(k+1)}-S^{(k)}|<\epsilon$), then the algorithm converges. In this study was used a threshold of $\epsilon=1\times10^{-5}$. A pseudocode of the proposed segmentation method is presented in the Algorithm 1.

V. Results

This section addresses three real experiments on multidimensional intensities of multilook PolSAR images. We assess the performance of eight segmenters resulting from the method discussed in Section IV—on these images:

1) independent SE-based segmenter for Γ^* with L known (IEL$_{\Gamma^*}$);
2) independent SE-based segmenter for Γ^* with L estimated (IELE$_{\Gamma^*}$);

TABLE I
VALUES OF PERFORMANCE MEASURES FOR THE FOULUM, MUNICH, AND FLEVOLAND IMAGES SEGMENTATION

	Foulum			Munich			Flevoland		
Methods	Accuracy	Kappa	var. kappa	Accuracy	Kappa	var. kappa	Accuracy	Kappa	var. kappa
k-means	0.391	0.250	3.1×10^{-5}	0.345	0.248	1.1×10^{-5}	0.199	0.109	3.5×10^{-5}
GaMM-L	0.484	0.374	2.5×10^{-5}	0.527	0.291	1.5×10^{-5}	0.171	0.088	2.2×10^{-5}
GaMM-LE	0.374	0.238	3.9×10^{-5}	0.561	0.342	1.5×10^{-5}	0.121	0.061	6.7×10^{-5}
IEL_{Γ^*}	0.692	0.624	2.1×10^{-5}	0.548	0.322	1.4×10^{-5}	0.331	0.269	2.9×10^{-5}
IELE_{Γ^*}	0.869	0.839	1.2×10^{-5}	0.711	0.567	1.1×10^{-5}	0.506	0.453	3.7×10^{-5}
$\text{IEL}_{\mathcal{G}_I^0}$	0.696	0.69	2.1×10^{-5}	**0.764**	**0.647**	$\mathbf{1.0}\times10^{-5}$	0.411	0.351	3.0×10^{-5}
$\text{IELE}_{\mathcal{G}_I^0}$	0.674	0.603	2.2×10^{-5}	0.733	0.601	1.1×10^{-5}	0.496	0.447	3.3×10^{-5}
DEL_{Γ^*}	**0.916**	**0.897**	$\mathbf{8.3\times10^{-6}}$	0.743	0.615	1.0×10^{-5}	**0.685**	**0.646**	$\mathbf{3.0\times10^{-5}}$
DELE_{Γ^*}	0.775	0.724	1.8×10^{-5}	0.724	0.587	1.1×10^{-5}	0.665	0.628	3.0×10^{-5}
$\text{DEL}_{\mathcal{G}_I^0}$	0.914	0.893	8.5×10^{-6}	0.760	0.640	1.1×10^{-5}	0.673	0.635	3.1×10^{-5}
$\text{DELE}_{\mathcal{G}_I^0}$	0.845	0.809	1.4×10^{-5}	0.732	0.599	1.0×10^{-5}	0.599	0.554	3.1×10^{-5}

Algorithm 1: Proposed Segmentation Method for SAR Images.

1: First allocate randomly each pixel to one of among c groups.
2: Start the weights $\hat{\pi}_g(i)$ as $\hat{\pi}_g^{(0)}(i) = \frac{T_g}{T}$, such that $\sum_{g=1}^{c} \hat{\pi}_g^{(0)}(i) = 1$;
3: Estimate $\widehat{\boldsymbol{\mu}}_{g:l}^{(k+1)}$ and $\hat{\sigma}_{g:l,l}^{(k+1)}$, by Eqs. (10) and (11).
4: The conditional probability of observation $\dot{\boldsymbol{H}}_S(i)$ belong to gth class is calculated by $\hat{\tau}_g(\hat{\boldsymbol{\theta}}|\dot{\boldsymbol{H}}_S(i))$ in Eq. (8).
5: The observations $\dot{\boldsymbol{H}}_S(i)$ are reallocated in the groups according to $\hat{c} = g \in c\text{arg max}\hat{\tau}_g(i)$;
6: Compute $S^{(k)}$ in Eq. (12) and, through absolute error, verify the convergence.
7: If there is convergence, stop the process; otherwise, compute the weights by Eq. (9) and come back to the third step.

3) independent SE-based segmenter for $\mathcal{G}_I^0$ with L known ($\text{IEL}_{\mathcal{G}_I^0}$);
4) independent SE-based segmenter for $\mathcal{G}_I^0$ with L estimated ($\text{IELE}_{\mathcal{G}_I^0}$);
5) dependent SE-based segmenter for Γ^* with L known (DEL_{Γ^*});
6) dependent SE-based segmenter for Γ^* with L estimated (DELE_{Γ^*});
7) dependent SE-based segmenter for $\mathcal{G}_I^0$ with L known ($\text{DEL}_{\mathcal{G}_I^0}$);
8) dependent SE-based segmenter for $\mathcal{G}_I^0$ with L estimated ($\text{DELE}_{\mathcal{G}_I^0}$).

The proposed methods are compared with other three well-defined segmenters: k-means, GaMM with L known (GaMM-L), and with L estimated (GaMM-LE). These two last methods can be understood as particular cases of Wishart-based segmenter in [17] or resulting of Zhao *et al.* [8]. In order to compare the previous methods, we use two figures of merit: accuracy percentage and kappa coefficient.

As exhibited in Fig. 2(c), we first consider a clipping of the Foulum (Denmark) image, which has been obtained by the EMISAR sensor. Its capturing has been made in L-band and quad-pol under a number of looks eight and represented five agricultural areas according to its reference mapping in Fig. 2(a) and (b): wheat (in white), rapeseed (in stronger intermediate gray), oats (in light gray) rye (in gray), coniferous (in lower intermediate gray), and the background (in black). Additionally, Fig. 2(c) highlights the pure areas, on which we quantified assessment criteria.

Now we interpret by visual inspection Fig. 2(d)–(n). It is noticeable the worst results have been obtained by the *k*-means and GaMM-LE. With respect to the methods which assume independence among stochastic entropies, $\text{IEL}_\bullet$ and $\text{IELE}_\bullet$, procedures equipped with known number of looks and the $\mathcal{G}_I^0$ entropy have produced the best results. However, all previous methods have failed at recognizing and splitting wheat and rapeseed areas. The last problem has been outperformed by procedures which adopting the hypothesis of dependence among stochastic entropies, $\text{DEL}_\bullet$ and $\text{DELE}_\bullet$. For these methods, the fewer estimators were used, the better results were obtained. Table I displays the values of assessment criteria. Bold results represent the best accuracy and confirm the previous discussion. The DEL_{Γ^*} has obtained the best performance. Moreover, IEL_{Γ^*} and $\text{IEL}_{\mathcal{G}_I^0}$ performed equivalently.

Second, we took one scene of the E-SAR image with 3.2 (equivalent number of) looks of the surroundings of Munich, Germany. According to the reference map in Fig. 2(o), the studied image has three regions. Consider a qualitative discussion of Fig. 2(q)–(ab); k-means has presented the worst result. In general, as in the previous experiment, the group of procedures $\{\text{DEL}_\bullet, \text{DELE}_\bullet\}$ tended to outperform $\{\text{IEL}_\bullet, \text{IELE}_\bullet\}$. Table I exhibits values of used performance criteria at this experiment, confirming the previous analysis. The best segmentations were achieved when using the SE for $\mathcal{G}_I^0$ with L known; i.e., $\text{IEL}_{\mathcal{G}_I^0}$ and $\text{DEL}_{\mathcal{G}_I^0}$.

Third experiment was made over an AIRSAR image of Flevoland, the Netherlands, took on August 1989 with four nominal looks. From Fig. 2(ac) and (ad), this scene has eleven

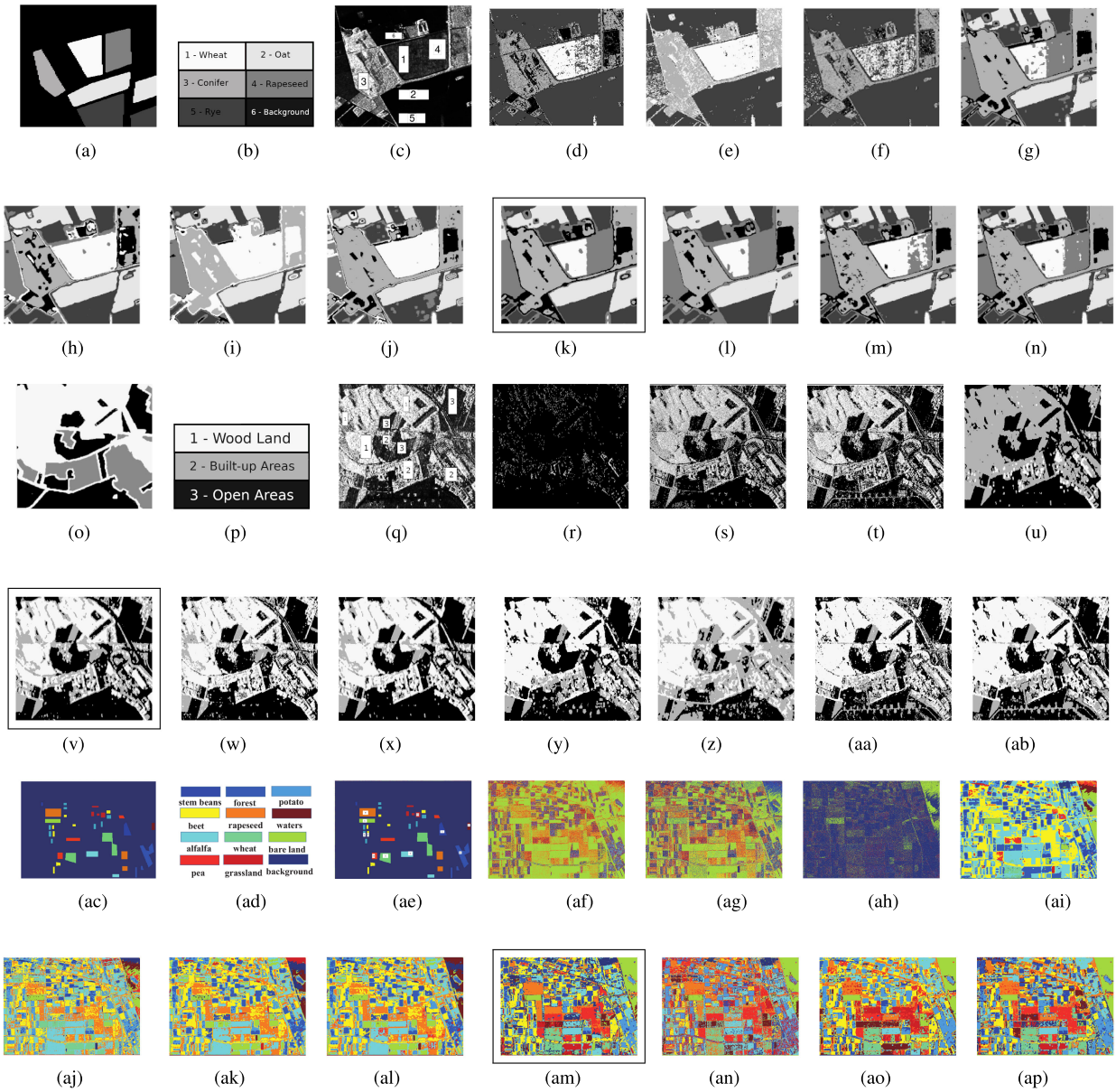


Fig. 2. Segmentation study over scenes of an EMISAR image of Foulum (Denmark), an E-SAR image of Munich (Germany), and a AIRSAR image of Flevoland (Netherlands). The best segmentation maps are highlighted with frames. (a) Reference map. (b) Areas names. (c) Selected areas. (d) k-means. (e) GaMM-L. (f) GaMM-LE. (g) IEL_{Γ^*}. (h) $\text{IEL}_{\mathcal{G}^0}$. (i) IELE_{Γ^*}. (j) $\text{IELE}_{\mathcal{G}^0}$. (k) DEL_{Γ^*}. (l) $\text{DEL}_{\mathcal{G}^0}$. (m) DELE_{Γ^*}. (n) $\text{DELE}_{\mathcal{G}^0}$. (o) Reference map. (p) Areas names. (q) Selected areas. (r) emphk-means. (s) GaMM-L. (t) GaMM-LE. (u) IEL_{Γ^*}. (v) $\text{IEL}_{\mathcal{G}^0}$. (w) IELE_{Γ^*}. (x) $\text{IELE}_{\mathcal{G}^0}$. (y) DEL_{Γ^*}. (z) $\text{DEL}_{\mathcal{G}^0}$. (aa) DELE_{Γ^*}. (ab) $\text{DELE}_{\mathcal{G}^0}$. (ac) Reference map. (ad) Areas names. (ae) Selected areas. (af) k-means. (ag) GaMM-L. (ah) GaMM-LE. (ai) IEL_{Γ^*}. (aj) $\text{IEL}_{\mathcal{G}^0}$. (aj) IELE_{Γ^*}. (ak) $\text{IELE}_{\mathcal{G}^0}$. (al) DEL_{Γ^*}. (am) $\text{DEL}_{\mathcal{G}^0}$. (an) DELE_{Γ^*}. (ao) $\text{DELE}_{\mathcal{G}^0}$. (ap) $\text{DELE}_{\mathcal{G}^0}$.

areas: steam beans, forest, potato, beet, rapeseed, waters, alfalfa, wheat, bare land, pea, and grassland. First, from a qualitative analysis of Fig. 2(af)–(ad), the GaMM-L presented the worst segmentation, followed by GaMM-LE and k-means. In general, the use of a dependence pattern in $\widehat{\boldsymbol{H}}_S(i)$ improved the segmentation results. Table I displays values of adopted performance measures. The best segmentation has been made by DEL_{Γ^*}.

In terms of dependence, the proposed segmenters based on Γ^* worked better under the dependence condition. The ones equipped with $\mathcal{G}_I^0$ obtained the best performance according to the rule: in homogeneous regions, the dependence condition was required; while the independence assumption was most recommended in heterogeneous scenarios. This seems to be motivated by fact that the Γ distribution is used for smooth intensity returns (and, therefore, the use of dependence structure yields more classification flexibility) and the $\mathcal{G}_I^0$ law assumes often high return values in heterogeneous scenarios (like in Munich) and gives low values of intensity for homogeneous ones (like in Foulum and Flevoland).

VI. Conclusion

In this article, we have introduced a new SAR imagery segmentation paradigm which is based on asymptotic properties of stochastic SEs in both independence and dependence assumptions, termed $\{\text{IEL}_\bullet, \text{IELE}_\bullet\}$ and $\{\text{DEL}_\bullet, \text{DELE}_\bullet\}$, respectively. Two particular cases in this approach have been discussed assuming the Γ (base for understanding speckle) and $\mathcal{G}_I^0$ (for pronounced texture intensities) models, which are successful members in the multiplication modeling family.

Three applications to real images have been realized. For two of them–the EMISAR image of Foulum (homogeneous with five areas) and for the Flevoland AIRSAR image (homogeneous with eleven areas), the Γ entropy-based segmentation procedure under dependence DEL_{Γ^*} has obtained the best result. For the Munich E-SAR image (more heterogeneous), the $\mathcal{G}_I^0$ entropy-based segmentation procedure under independence $\text{IEL}_{\mathcal{G}^0}$ has provided the best result.

We will work with other (Pol)SAR distributions in future researches.

Appendix
Derivation for the SE for $\mathcal{G}_I^0$

Let $Z \sim \mathcal{G}_I^0(\alpha, \gamma, L)$. After algebraic manipulation,

$$\mathbb{E}\log Z = \psi^{(0)}(L) - \psi^{(0)}(-\alpha) + \log\frac{\gamma}{L} \tag{13}$$

and

$$\mathbb{E}\log(\gamma + LZ) = \psi^{(0)}(L-\alpha) + \log\gamma - \psi^{(0)}(-\alpha). \tag{14}$$

From these identities, the $\mathcal{G}_I^0$ SE is given by

$$\begin{aligned} H_S(Z) = \mathbb{E}\left[-\log f(Z,\alpha,\gamma,L)\right] &= -L\log L + \alpha\log\gamma \\ &- \{\log\Gamma(L-\alpha) - \log\Gamma(L), -\log\Gamma(-\alpha)\} \\ &- (L-1)\mathbb{E}\log Z + (L-\alpha)\mathbb{E}\log(\gamma + LZ). \end{aligned} \tag{15}$$

The expression follows from applying (13) and (14) in (15).

References

[1] A. C. Frery, H.-J. Muller, C. D. C. F. Yanasse, and S. J. S. Sant'Anna, "A model for extremely heterogeneous clutter," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 3, pp. 648–659, May 1997.

[2] J. Morio, P. Réfrégier, F. Goudail, P. C. Dubois-Fernandez, and X. Dupuis, "Information theory-based approach for contrast analysis in polarimetric and/or interferometric SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2185–2196, Aug. 2008.

[3] A. C. Frery, R. J. Cintra, and A. D. Nascimento, "Entropy-based statistical analysis of PolSAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 6, pp. 3733–3743, Jun. 2013.

[4] R. Shang *et al.*, "A spatial fuzzy clustering algorithm with kernel metric based on immune clone for SAR image segmentation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 4, pp. 1640–1652, Apr. 2016.

[5] A. El Zaart, D. Ziou, S. Wang, and Q. Jiang, "Segmentation of SAR images," *Pattern Recognit.*, vol. 35, no. 3, pp. 713–724, 2002.

[6] H. Sui, C. Xu, J. Liu, K. Sun, and C. Wen, "A novel multi-scale level set method for SAR image segmentation based on a statistical model," *Int. J. Remote Sens.*, vol. 33, pp. 5600–5614, 2012.

[7] J. Gu, L. Jiao, S. Yang, F. Liu, B. Hou, and Z. Zhao, "A multi-kernel joint sparse graph for SAR image segmentation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 3, pp. 1265–1285, Mar. 2016.

[8] Q. Zhao, X. Li, and Y. Li, "Multilook SAR image segmentation with an unknown number of clusters using a gamma mixture model and hierarchical clustering," *Sensors*, vol. 17, no. 5, pp. 1114–1131, 2017.

[9] L. Pardo, *Statistical Inference Based on Divergence Measures*. Boca Raton, FL, USA: CRC Press, 2005.

[10] K. Zografos and S. Nadarajah, "Expressions for Rényi and Shannon entropies for multivariate distributions," *Statist. Probability Lett.*, vol. 71, pp. 71–84, 2005.

[11] R. H. Nobre *et al.*, "SAR image segmentation with Rényi's entropy," *IEEE Signal Process. Lett.*, vol. 23, no. 11, pp. 1551–1555, Nov. 2016.

[12] K. L. P. Vasconcellos, A. C. Frery, and L. B. Silva, "Improving estimation in speckled imagery," *Comput. Statist.*, vol. 20, pp. 503–519, 2005.

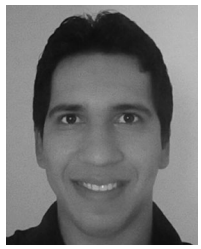
[13] C. Nikou, A. C. Likas, and N. P. Galatsanos, "A Bayesian framework for image segmentation with spatially varying mixtures," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2278–2289, Sep. 2010.

[14] D. P. Geoffrey McLachlan, *Finite Mixture Models* (Series Wiley Series in Probability and Statistics), 1st ed. New York, NY, USA: Wiley, 2000.

[15] G. Celeux and G. Govaert, "Gaussian parsimonious clustering models," *Pattern Recognit.*, vol. 28, pp. 781–793, 1995.

[16] R. P. Browne and P. D. Mcnicholas, "Estimating common principal components in high dimensions," *Adv. Data Anal. Classification*, vol. 8, pp. 217–226, 2014.

[17] M. M. Horta, N. D. Mascarenhas, and A. C. Frery, "Analyzing polarimetric imagery with G0p mixture models and SEM algorithm," in *Proc. Brazilian Symp. Comput. Graph. Image Process.*, vol. 20, 2007, pp. 1–2.



Jodavid de A. Ferreira received the B.Sc. degree in statistics from the Universidade Federal da Paraiba, João Pessoa, Brazil, in 2015 and the M.Sc. degree in statistics from the Universidade Federal de Pernambuco, Recife, Brazil, in 2017.

His research interests include computational statistics, multivariate analysis, statistical IT, and inference on random matrices (with emphasis for applications on PolSAR imagery).

Abraão D. C. Nascimento received the B.Sc., M.Sc., and D.Sc. degrees in statistics from Universidade Federal de Pernambuco (UFPE), Recife, Brazil, in 2005, 2007, and 2012, respectively.

In 2014, he joined the Department of Statistics with UFPE as an Adjoint Professor. His research interests include statistical IT, inference on random matrices (with emphasis for applications on PolSAR imagery), statistical theory of shape, spatio-temporal processes, survival analysis, and asymptotic theory.