

# Maximum-Likelihood Parameter Estimation of the Product Model for Multilook Polarimetric SAR Data

Nizar Bouhlef, *Member, IEEE*, and Stéphane Méric, *Member, IEEE*

**Abstract**—The product model is assumed to be an appropriate statistical model for multilook polarimetric synthetic radar data (PolSAR). According to this model, the observed signal is considered as the product of independent random variates of a complex Gaussian speckle and a non-Gaussian texture. With different texture distributions, the product model leads to different expressions for the compound distribution considered as an infinite mixture model. In this paper, the maximum-likelihood (ML) estimator is derived to jointly estimate the speckle and texture parameters in the compound distribution model using the multilook polarimetric radar data. In particular, we estimate: 1) the equivalent number of looks; 2) the covariance matrix of the speckle component; and 3) the texture distribution parameters. The expectation–maximization algorithm is developed to compute the ML estimates of the unknown parameters. The hybrid Cramer–Rao bounds (HCRBs) are also derived for these parameters. First, a general HCRB expression is derived under an arbitrary texture distribution. Then, this expression is simplified for a specific texture distribution. The performance of the ML is compared with the performance of other known estimators using the simulated and real multilook PolSAR data. For real data, a goodness of fit of multilook PolSAR data histograms is used to assess the fitting accuracy of the compound distributions using different estimators.

**Index Terms**—Expectation–maximization (EM) algorithm, maximum likelihood (ML), multilook polarimetric synthetic aperture radar (PolSAR) data, parameter estimation, product model.

## I. INTRODUCTION

**P**OLARIMETRIC synthetic aperture radar (PolSAR) has been widely used to discriminate between different scattering mechanisms. It gives more scattering information than single-polarization-channel SAR data. Radar images are affected by an interference phenomenon known as speckle. This noise can be mitigated by a multilooking processing

step. Multilooking consists in averaging multiple SAR measurements during data formation and postprocessing [1], [2]. The parameter that describes the degree of averaging applied to the SAR measurements is called the number of looks. Due to the random nature of speckles, statistical models have been widely used for PolSAR data. The product model is commonly assumed to be an appropriate statistical model for PolSAR data. Introduced by Yueh *et al.* [3], the multivariate product model considers the observed signal as the product of a complex Gaussian speckle random variate and a non-Gaussian texture random variate. These two components are independent sources of data variability. It has been shown that a single-look PolSAR speckle follows a multivariate zero-mean complex Gaussian distribution [4] and a multi-look PolSAR speckle follows a matrix-variate-scaled complex Wishart distribution [4]. The texture represents the natural spatial variation of the radar cross section and is considered to be a positive real variable. Different distributions are used to model the texture random variable. With different texture distributions, the product model leads to different expressions for the resulting compound distribution. Some of the most important texture distributions proposed in the literature are gamma ( $\gamma$ ), inverse gamma ( $\gamma^{-1}$ ), generalized inverse Gaussian (GIG) ( $\mathcal{N}^{-1}$ ), and Fisher ( $\mathcal{F}$ ), respectively, with the resulting compound distributions  $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ ,  $\mathcal{G}_d$ , and Kummer- $\mathcal{U}_d$  [5]–[10]. A parameter estimation procedure is an important task in the analysis of PolSAR images, and an efficient parameter estimation of the compound distributions is always desired. Three types of parameters need to be estimated: 1) the number of looks; 2) the covariance matrix; and 3) the texture parameters. The first two parameters are from the speckle probability density function (pdf), and the latter concerns the texture pdf and indicates the texture variation level.

In the literature, there are a lot of ways to deal with the parameter estimation. The classical one consists in estimating parameters on each individual single polarization and average to obtain the estimates for the polarimetric distribution [6], [11]. This kind of single-polarization estimator has a performance less than that of polarimetric estimators [12]. This approach is frequently used in the equivalent number of looks (ENL) estimation. Therefore, the traditional approach to an ENL estimation is to identify the homogeneous regions in an image, where the speckle is fully developed and the texture is negligible. These conditions ensure that the distribution of the scattering coefficients can be assumed complex Gaussian. Under this assumption, Anfinsen *et al.* [13] estimated the

N. Bouhlef was with the Laboratoire des Sciences du Numériques (LS2N), École Centrale de Nantes (ECN), 44300 Nantes, France. He is now with the SAR & Hyperspectral multi-modal Imaging and sigNal processing, Electromagnetic modeling team (SHINE), Department of Waves & Signals, Institute of Electronics and Telecommunications of Rennes (IETR), Institut national des sciences appliquées de Rennes (INSA), 35043 Rennes, France (e-mail: nizar.bouhlef@insa-rennes.fr).

S. Méric is with the SAR & Hyperspectral multi-modal Imaging and sigNal processing, Electromagnetic modeling team (SHINE), Institute of Electronics and Telecommunications of Rennes (IETR), Institut national des sciences appliquées de Rennes (INSA), 35043 Rennes, France.

ENL using the coefficient of variation estimator and the fractional moment-based estimator obtained from each available polarization before averaging the results. For full PolSAR, Anfinson *et al.* [13] developed a trace moment-based estimator and a log-determinant moment-based estimator for the ENL estimation via statistics of the polarimetric covariance matrix, which follows the Wishart distribution. These two methods provide a valid estimation when the speckle is fully developed and the texture is considered negligible. Tao *et al.* [14] introduced a novel ENL estimator that could be applied to any distribution of a texture model. The new estimator is the development of trace moments (DTM), which cancels the textural variation using trace moments. This estimator is an extension of the trace moment-based estimator developed by Anfinson *et al.* [13] under the assumption of the Wishart distribution. The most important method developed recently and becoming increasingly popular is the method of matrix log-cumulants (MoMLC) for the multilook polarimetric parameter estimation [12]. The MoMLC is based on the computation of higher derivatives of special functions [12]. With the MoMLC, the ENL is estimated from the first-order MLC equations of the complex Wishart distribution over a textureless area [10], [12].

The estimation of the speckle covariance matrix is important in the statistical modeling of multilook PolSAR data. An important comparative study of covariance matrix estimators for single-look PolSAR data was presented by Tao *et al.* [15]. They compared a few covariance matrix estimators: the maximum-likelihood (ML) estimators under the Gaussian distribution, the  $\mathcal{K}$ -distribution, and Tyler's estimator [16] or the fixed point estimator. For the Gaussian distribution-based ML (G-ML) estimator, the texture variable is considered as constant in the product model. The scattering vector then becomes complex multivariate Gaussian distributed, and the ML estimator reduces to the familiar sample covariance matrix (SCM) estimator. Gini and Greco [17] derived the general ML covariance matrix estimator based on the single-look polarimetric product model. When the texture is gamma-distributed, the scattering vector becomes multivariate  $\mathcal{K}$ -distributed [3]. The  $\mathcal{K}$ -distribution-based ML ( $\mathcal{K}$ -ML) estimator is solved iteratively and it depends on the shape parameter of the texture distribution, which has been preliminarily estimated. The third estimator is Tyler's estimator. It is the approximate ML estimator of the speckle covariance introduced by Gini and Greco [17] and Vasile *et al.* [18], [19]. A particularity of this estimator is that it does not require any prior information about the texture pdf. Tao *et al.* [15] in their comparative study concluded that the  $\mathcal{K}$ -ML estimator was better than the G-ML estimator under a high texture but at increased computational cost and that Tyler's estimator did not provide any distinct advantage for a textureless or textured area. For the multilook PolSAR data, Liu *et al.* [20] extended the spherically invariant random vector (SIRV) model to the multilook case and derived the ML estimator and the approximate ML estimator of the covariance matrix of the speckle.

When the pdf is used to model the texture random variable, additional shape parameters are introduced, which have to be estimated to define the distribution. Several works have been devoted to improve the estimation of the texture

parameters. The moment-based method is a traditional solution for shape parameter estimation. Due to the fact that the variance of a moment-based estimator increases with the moment order, the fractional moments are used instead. Then, the method can be applied to all polarimetric channels before averaging the results. This approach was adopted by Freitas *et al.* [7] and Frery *et al.* [11]. They derived estimators from fractional moments of single polarization by combining half- and quarter-order moments. The final estimate of the shape parameter of the texture is an average of the single-polarization estimates. Alternative estimators were proposed by Frery *et al.* [11] and Doulgeris *et al.* [21], where the former was just a monopoly version of the latter. The estimator was derived from the moments of the Hotelling–Lawley trace. The third method was the MoMLC applied to estimate the texture parameters from the multilook polarimetric  $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ , and  $\mathcal{G}_d$  distributions [10], [12]. When the distribution has only one texture parameter, as  $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ , the parameter will be related to the second-order MLC. If the distribution has two texture parameters, as  $\mathcal{G}_d$  and Kummer- $\mathcal{U}_d$  distributions, the estimation procedure will require the second- and third-order MLC equations, which will involve higher order derivatives of a modified Bessel function of the second kind with respect to its order for  $\mathcal{G}_d$  and higher order derivatives of a multivariate digamma function for the Kummer- $\mathcal{U}_d$  distribution. These derivatives do not have closed-form expressions [10]. Anfinson and Eltoft [12] performed the Monte Carlo simulations to compare the performance of several estimators for texture parameters carried out for simulated multilook PolSAR data. They concluded that polarimetric estimators outperform single-polarization estimators in terms of bias and variance. Recently, Khan and Guida [22] have shown that the method of multivariate fractional moment (MoMFM), which is based on the fractional moments of the multilook polarimetric whitening filter [23], has a lower bias and variance than the MoMLC-based estimators.

The aforementioned methods are used separately in the estimation of parameters and under the condition of homogeneous regions. In fact, when the MLC method is used to estimate parameters, the ENL is estimated under the condition of region homogeneity. Then, it is estimated from the first-order MLC equation of the complex Wishart distribution [12]. In addition, to solve this equation, a prior estimate of the covariance matrix estimated under the Wishart condition is needed. In this paper, a general framework for fully PolSAR data is developed. The objective is a fully automatic estimation procedure that requires no Gaussian assumption or manual intervention, as the selection of homogeneous regions where the speckle is fully developed and the texture is negligible. The ENL, the covariance matrix, and the texture parameters are estimated simultaneously and not separately by the maximum-likelihood estimator. The classical expectation–maximization (EM) algorithm [24] is used to derive the ML parameters. The EM provides an ideal framework for the parameter estimation of the compound distribution, which is the result of the product model, and it guarantees the convergence to a stationary point of the log-likelihood. In addition, the algorithm is easier to implement and computationally efficient,

which requires the computation of special functions and not their higher order derivatives. The aim of this paper is to use the EM algorithm to derive the ML estimates for the multilook polarimetric  $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ , and  $\mathcal{G}_d$ . Then, the polarimetric EM estimator is compared with some known estimators. The rest of this paper is organized as follows. Section II introduces the statistical product model for multilook PolSAR data. The application to parameter estimation for the multilook PolSAR data distribution using the EM algorithm is presented in Section III. The content of Section IV is related to the performance evaluation. The EM estimator is compared with other estimators for multilook polarimetric  $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ , and  $\mathcal{G}_d$  distributions using the simulated PolSAR data. Section V presents the results of the proposed estimators applied to real data. Finally, Section VI concludes this paper.

## II. PRODUCT MODEL FOR MULTILOOK POLARIMETRIC RADAR DATA

The polarimetric scattering vector is defined as

$$\mathbf{s} = [s_{hh}, s_{hv}, s_{vh}, s_{vv}]^T \in \mathbb{C}^d \quad (1)$$

with  $s_{xy}$  the complex scattering coefficients, where  $x$  is the transmit,  $y$  is the receive polarization,  $h$  denotes horizontal,  $v$  denotes vertical [25],  $[\cdot]^T$  means transposition, and  $d = \dim(\mathbf{s})$  is the vector dimension. The single-look polarimetric product model is defined as

$$\mathbf{s} = \sqrt{\tau} \mathbf{x} \quad (2)$$

where  $\tau$  represents a positive scalar texture random variable with a pdf  $f_\tau(\tau)$  and  $\mathbf{x}$  is a  $d$ -dimensional speckle vector, which follows a circular complex Gaussian distribution  $[\mathbf{x} \sim \mathcal{N}_d^{\mathbb{C}}(0, \mathbf{\Sigma})]$ , with a zero-mean vector and a covariance matrix  $\mathbf{\Sigma}$ . The random variable  $\tau$  and the random vector  $\mathbf{x}$  are statistically independent. The multilooking of PolSAR data reduces the speckle effect characteristic of coherent imaging systems. The polarimetric multilooking operation is given by

$$\mathbf{C} = \frac{1}{L} \sum_{l=1}^L \mathbf{s}_l \mathbf{s}_l^H, \quad L \geq d \quad (3)$$

where  $L$  is the number of looks,  $(\cdot)^H$  denotes the Hermitian operator, and  $\mathbf{C} \in \mathbf{\Omega}_+ \subset \mathbb{C}^{d \times d}$  is the multilook polarimetric covariance matrix considered as a random matrix defined on the cone  $\mathbf{\Omega}_+$  of the positive-definite complex Hermitian matrices. The multilook polarimetric product model is then given by

$$\mathbf{C} = \tau \mathbf{X} \quad (4)$$

where  $\mathbf{X}$  is a random speckle matrix defined as  $\mathbf{X} = (1/L) \sum_{l=1}^L \mathbf{x}_l \mathbf{x}_l^H$ . When  $L \geq d$ , the unnormalized SCM defined as  $\mathbf{Z} = L\mathbf{X}$  follows the nonsingular complex Wishart distribution [26] denoted as  $\mathbf{Z} \sim \mathcal{W}_d^{\mathbb{C}}(L, \mathbf{\Sigma})$  and  $\mathbf{X}$  follows a scaled complex Wishart distribution with a pdf given by  $f_{\mathbf{X}}(\mathbf{X}) = f_{\mathbf{Z}}(L\mathbf{X}) |J_{\mathbf{Z} \rightarrow \mathbf{X}}|$ , where  $|J_{\mathbf{Z} \rightarrow \mathbf{X}}| = L^{d^2}$  is the Jacobian determinant of the transformation  $\mathbf{Z} = L\mathbf{X}$  [27]. The pdf of  $\mathbf{X}$  is

$$f_{\mathbf{X}}(\mathbf{X}) = \frac{L^{Ld} |\mathbf{X}|^{L-d}}{\Gamma_d(L) |\mathbf{\Sigma}|^L} \text{etr}(-L\mathbf{\Sigma}^{-1}\mathbf{X}) \quad (5)$$

where  $\text{etr}(\cdot) = \exp(\text{tr}(\cdot))$  is the exponential trace operator,  $|\cdot|$  is the determinant operator, and  $\Gamma_d(L)$  is the multivariate gamma function of the complex kind defined as

$$\Gamma_d(L) = \pi^{d(d-1)/2} \prod_{i=0}^{d-1} \Gamma(L-i) \quad (6)$$

where  $\Gamma(L)$  is the standard Euler gamma function. The pdf of  $\mathbf{C}$  using Bayes' theorem becomes as follows:

$$f_{\mathbf{C}}(\mathbf{C}) = \int_0^{+\infty} f_{\mathbf{C}|\tau}(\mathbf{C}|\tau) f_\tau(\tau) d\tau \quad (7)$$

where  $f_{\mathbf{C}|\tau}(\mathbf{C}|\tau)$  is the pdf of  $\mathbf{C}$  with a specific value of  $\tau$  and is given by  $f_{\mathbf{C}|\tau}(\mathbf{C}|\tau) = f_{\mathbf{X}}(\mathbf{C}/\tau) |J_{\mathbf{X} \rightarrow \mathbf{C}}|$  with  $|J_{\mathbf{X} \rightarrow \mathbf{C}}| = \tau^{-d^2}$ . Then, the pdf is, therefore, given by

$$f_{\mathbf{C}|\tau}(\mathbf{C}|\tau) = \frac{L^{Ld} |\mathbf{C}|^{L-d}}{\Gamma_d(L) |\mathbf{\Sigma}|^L} \frac{1}{\tau^{dL}} \text{etr}\left(-\frac{L}{\tau} \mathbf{\Sigma}^{-1} \mathbf{C}\right). \quad (8)$$

The pdf of  $\mathbf{C}$  is obtained by substituting (8) in (7)

$$f_{\mathbf{C}}(\mathbf{C}) = \frac{L^{Ld} |\mathbf{C}|^{L-d}}{\Gamma_d(L) |\mathbf{\Sigma}|^L} \int_0^{+\infty} \frac{1}{\tau^{dL}} \text{etr}\left(-\frac{L}{\tau} \mathbf{\Sigma}^{-1} \mathbf{C}\right) f_\tau(\tau) d\tau. \quad (9)$$

This equation is the extension of the SIRV distribution model to the multilook case [9], [20]. The same pdf of the multilook polarimetric covariance matrix  $\mathbf{C}$  can be defined by different couples  $(\mathbf{\Sigma}, \theta_\tau)$ , where  $\theta_\tau$  are the texture distribution parameters. In order to obtain a unique correspondence between  $f_{\mathbf{C}}(\mathbf{C})$  and the parameters, some constraint on the mean of the texture distribution or on the scale of  $\mathbf{\Sigma}$  is considered. The first constraint is to normalize the mean of the texture distribution to 1, and then,  $E\{\tau\} = 1$ . This normalization was adopted in [12], [13], and [28]. The second constraint is to normalize the determinant of the covariance matrix of the speckle to 1 by dividing  $\mathbf{\Sigma}$  by its determinant  $|\mathbf{\Sigma}|^{1/d}$  [10], [21] or to normalize the trace of the covariance matrix of the speckle to  $d$  by dividing  $\mathbf{\Sigma}$  by  $\text{tr}(\mathbf{\Sigma})/d$  [17]. In fact, normalization is required, but any normalization can be adopted. Table I provides the analytical expression of the covariance matrix pdfs of possible choices of  $f_\tau(\tau)$ : gamma ( $\bar{\gamma}$ ), inverse gamma ( $\bar{\gamma}^{-1}$ ), and GIG ( $\mathcal{N}^{-1}$ ) distributed textures. They are, respectively, the matrix-variate  $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ , and  $\mathcal{G}_d$  distributions. The overbar in the symbol of gamma ( $\bar{\gamma}$ ) and inverse gamma distributions ( $\bar{\gamma}^{-1}$ ) means that these distributions have been normalized to a unit mean. The determinant of the covariance matrix of the speckle of the matrix-variate  $\mathcal{G}_d$  distribution is normalized to 1. Substituting the pdf of  $\bar{\gamma}(\alpha)$  and  $\mathcal{N}^{-1}(\alpha, w, \eta)$  into (9), where  $(\alpha, w)$  are two shape parameters and  $\eta$  is a scale parameter, and using [29, eq. (3.471.9)] lead to the pdf of  $\mathcal{K}_d(\mathbf{\Sigma}, L, \alpha)$  and  $\mathcal{G}_d(\mathbf{\Sigma}, L, \alpha, w, \eta)$  depending on  $K_{\alpha-dL}(\cdot)$ , which is the modified Bessel function of the second kind of order  $\alpha - dL$ .

## III. EM FOR COMPOUND DISTRIBUTION

Let us consider  $N$  independent and identically distributed (i.i.d.) realizations  $\mathbf{C}_i, i = \{1, \dots, N\}$  of a random matrix  $\mathbf{C}$  distributed according to the density function  $f_{\mathbf{C}}(\mathbf{C})$ . One can suppose that each observed covariance matrix  $\mathbf{C}_i$  has a corresponding unobserved and hidden texture  $\tau_i$ . The sequence

TABLE I  
TEXTURE AND COVARIANCE MATRIX DISTRIBUTION UNDER THE MULTILOOK POLARIMETRIC PRODUCT MODEL

Texture pdf		Compound pdf		Ref.
Symbol	$f_{\tau}(\tau)$	Symbol	$f_{\mathbf{C}}(\mathbf{C})$	
$\bar{\gamma}(\alpha)$	$\frac{\alpha^{\alpha} \tau^{\alpha-1}}{\Gamma(\alpha)} \exp(-\alpha\tau)$ $\tau, \alpha \in \mathbb{R}_+$	$\mathcal{K}_d(\boldsymbol{\Sigma}, L, \alpha)$	$\frac{L^{Ld}  \mathbf{C} ^{L-d}}{\Gamma_d(L)  \boldsymbol{\Sigma} ^L} \frac{2\alpha^{\frac{\alpha+Ld}{2}}}{\Gamma(\alpha)} (\text{Ltr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}))^{\frac{\alpha-Ld}{2}} \times$ $K_{\alpha-Ld} \left( 2\sqrt{L\alpha\tau}(\boldsymbol{\Sigma}^{-1}\mathbf{C}) \right)$	[12]
$\bar{\gamma}^{-1}(\lambda)$	$\frac{(\lambda-1)^{\lambda}}{\Gamma(\lambda)} \frac{1}{\tau^{1+\lambda}} \exp(-\frac{\lambda-1}{\tau})$ $\tau \in \mathbb{R}_+, \lambda > 1$	$\mathcal{G}_d^0(\boldsymbol{\Sigma}, L, \lambda)$	$\frac{L^{Ld}  \mathbf{C} ^{L-d}}{\Gamma_d(L)  \boldsymbol{\Sigma} ^L} \frac{(\lambda-1)^{\lambda} \Gamma(dL+\lambda)}{\Gamma(\lambda)} (\text{Ltr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}) + \lambda - 1)^{-(dL+\lambda)}$	[7]
$\mathcal{N}^{-1}(\alpha, w, \eta)$	$\frac{1}{\eta^{\alpha}} 2K_{\alpha}(w) \tau^{\alpha-1} \exp(-\frac{w}{2}(\frac{\eta}{\tau} + \frac{\tau}{\eta}))$ $\eta, \tau, w \in \mathbb{R}^+, \alpha \in \mathbb{R}$	$\mathcal{G}_d(\boldsymbol{\Sigma}, L, \alpha, w, \eta)$	$\frac{L^{Ld}  \mathbf{C} ^{L-d}}{\eta^{\alpha} \Gamma_d(L)  \boldsymbol{\Sigma} ^L} \frac{1}{K_{\alpha}(w)} \left( \frac{2\text{Ltr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}) + w\eta}{w/\eta} \right)^{\frac{\alpha-Ld}{2}} \times$ $K_{\alpha-Ld} \left( \sqrt{\frac{w}{\eta}} (2\text{Ltr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}) + w\eta) \right)$	[10]

$\{\tau_i, i = 1, \dots, N\}$  is also assumed to be i.i.d. The pdf of  $\mathbf{C}$  depends on a set of unknown parameters  $\theta = (L, \boldsymbol{\Sigma}, \theta_{\tau})$ , where  $\theta_{\tau}$  are the parameters of the texture distribution. The ML estimator  $\hat{\theta}$  of the set parameter  $\theta$  of the compound distribution is given by

$$\hat{\theta} = \arg \max_{\theta} f_{\mathbf{C}}(\mathbf{C}|\theta) = \arg \max_{\theta} \prod_{i=1}^N f_{\mathbf{C}}(\mathbf{C}_i|\theta). \quad (10)$$

In the absence of a closed-form solution to (10), the EM algorithm is used to find the estimation  $\hat{\theta}$  given a current estimate  $\theta'$ . The new equation is then given by

$$\begin{aligned} \hat{\theta} &= \arg \max_{\theta} E_{\tau|\mathbf{C}}\{\ln f_{\tau, \mathbf{C}}(\mathbf{C}, \tau|\theta) | \mathbf{C}, \theta'\} \\ &= \arg \max_{\theta} E_{\tau|\mathbf{C}} \left\{ \sum_{i=1}^N \ln f_{\tau, \mathbf{C}}(\mathbf{C}_i, \tau_i|\theta) | \mathbf{C}_i, \theta' \right\} \\ &= \arg \max_{\theta} \sum_{i=1}^N E_{\tau|\mathbf{C}}\{\ln f_{\tau, \mathbf{C}}(\mathbf{C}_i, \tau_i|\theta) | \mathbf{C}_i, \theta'\}. \end{aligned} \quad (11)$$

Let  $\theta_{\mathbf{X}} = (\boldsymbol{\Sigma}, L)$  and using the fact that  $f_{\tau, \mathbf{C}}(\mathbf{C}_i, \tau_i|\theta) = f_{\mathbf{C}|\tau}(\mathbf{C}_i|\tau_i, \theta_{\mathbf{X}}) f_{\tau}(\tau_i|\theta_{\tau})$ , then (11) can be maximized separately according to the speckle parameter  $\theta_{\mathbf{X}}$  and  $\theta_{\tau}$

$$\hat{\theta}_{\mathbf{X}} = \arg \max_{\theta_{\mathbf{X}}} \sum_{i=1}^N E_{\tau|\mathbf{C}}\{\ln f_{\mathbf{C}|\tau}(\mathbf{C}_i|\tau_i, \theta_{\mathbf{X}}) | \mathbf{C}_i, \theta'\} \quad (12)$$

$$\hat{\theta}_{\tau} = \arg \max_{\theta_{\tau}} \sum_{i=1}^N E_{\tau|\mathbf{C}}\{\ln f_{\tau}(\tau_i|\theta_{\tau}) | \mathbf{C}_i, \theta'\}. \quad (13)$$

#### A. Estimation of Speckle Parameters

Substituting (8) in (12), differentiating with respect to  $L$  and  $\boldsymbol{\Sigma}$  by taking into account the properties [30]:  $\partial \ln |\boldsymbol{\Sigma}| / \partial \boldsymbol{\Sigma} = (\boldsymbol{\Sigma}^{-1})^T$  and  $\partial \text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{W}) / \partial \boldsymbol{\Sigma} = -(\boldsymbol{\Sigma}^{-1}\mathbf{W}\boldsymbol{\Sigma}^{-1})^T$ , and setting the result to zero yield

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{N} \sum_{i=1}^N E_{\tau|\mathbf{C}} \left\{ \frac{1}{\tau_i} | \mathbf{C}_i, \theta' \right\} \mathbf{C}_i \quad (14)$$

$$\begin{aligned} d \ln \hat{L} + d - \psi_d(\hat{L}) \\ &= \frac{1}{N} \sum_{i=1}^N \left( E_{\tau|\mathbf{C}} \left\{ \frac{1}{\tau_i} | \mathbf{C}_i, \theta' \right\} \text{tr}(\hat{\boldsymbol{\Sigma}}^{-1} \mathbf{C}_i) - \ln |\hat{\boldsymbol{\Sigma}}^{-1} \mathbf{C}_i| \right) \\ &+ \frac{d}{N} \sum_{i=1}^N E_{\tau|\mathbf{C}}\{\ln \tau_i | \mathbf{C}_i, \theta'\} \end{aligned} \quad (15)$$

where  $E_{\tau|\mathbf{C}}\{\ln \tau_i | \mathbf{C}_i, \theta'\}$  and  $E_{\tau|\mathbf{C}}\{(1/\tau_i) | \mathbf{C}_i, \theta'\}$  are posterior expectations depending on the texture distributions, and their closed-form expressions are given in Section III-B. The function  $\psi_d(L)$  is the multivariate digamma function defined as  $\psi_d(L) = \partial \ln \Gamma_d(L) / \partial L = \sum_{j=0}^{d-1} \psi(L-j)$ , where  $\psi(\cdot)$  is the digamma function. Equation (14) is the ML estimator of the covariance matrix of the speckle. Equation (14) is equivalent to the ML estimator of the covariance matrix of the speckle derived by Liu *et al.* [20], which is given as

$$\hat{\boldsymbol{\Sigma}}_{\text{ML}} = \frac{1}{N} \sum_{i=1}^N \frac{h_{dL+1}(\text{Ltr}(\hat{\boldsymbol{\Sigma}}_{\text{ML}}^{-1} \mathbf{C}_i))}{h_{dL}(\text{Ltr}(\hat{\boldsymbol{\Sigma}}_{\text{ML}}^{-1} \mathbf{C}_i))} \mathbf{C}_i \quad (16)$$

where the function  $h_{dL}(q)$  is defined as  $h_{dL}(q) = \int_0^{+\infty} \tau^{-dL} \exp(-q/\tau) f_{\tau}(\tau) d\tau$ . A complete proof is given in Appendix A. For single-look PolSAR data, we get  $L = 1$  and  $\mathbf{C} = \mathbf{s}\mathbf{s}^H$ . By applying the identity  $\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{s}\mathbf{s}^H) = \mathbf{s}^H \boldsymbol{\Sigma}^{-1} \mathbf{s}$ , (16) is reduced to the ML covariance matrix estimator derived by Gini and Greco [17] given by  $\hat{\boldsymbol{\Sigma}}_{\text{ML}} = (1/N) \sum_{i=1}^N (h_{d+1}(\mathbf{s}_i^H \hat{\boldsymbol{\Sigma}}_{\text{ML}}^{-1} \mathbf{s}_i) / h_d(\mathbf{s}_i^H \hat{\boldsymbol{\Sigma}}_{\text{ML}}^{-1} \mathbf{s}_i)) \mathbf{s}_i \mathbf{s}_i^H$ . It is important to point out that to guarantee the unicity of the covariance matrix for the case of GIG texture distribution, the determinant of the estimated covariance matrix of the speckle is normalized to 1 by dividing  $\hat{\boldsymbol{\Sigma}}$  by its determinant  $|\hat{\boldsymbol{\Sigma}}|^{1/d}$ . Equation (15) is needed to solve the ML estimator of the ENL  $L$ . An explicit solution for  $\hat{L}$  is not obtainable directly from (15). The Newton–Raphson method is proposed to solve it iteratively. As regards the estimation of  $L$ , (15) is equivalent to  $E_{\mathbf{Y}}\{\text{tr}(\mathbf{Y}_i) - \ln |\mathbf{Y}_i|\} = E_{\mathbf{C}}\{E_{\mathbf{Y}|\mathbf{C}}\{\text{tr}(\mathbf{Y}_i) - \ln |\mathbf{Y}_i| | \mathbf{C}_i\}\}$ , where  $\mathbf{Y}_i = (1/\tau_i) \boldsymbol{\Sigma}^{-1/2} \mathbf{C}_i \boldsymbol{\Sigma}^{-1/2}$  is distributed as a scaled Wishart distribution with the parameters  $(L, \mathbf{I}_d)$ , where  $\mathbf{I}_d$  is the  $d \times d$  identity matrix,  $E_{\mathbf{Y}}\{\text{tr}(\mathbf{Y}_i)\} = d$ , and  $E_{\mathbf{Y}}\{\ln |\mathbf{Y}_i|\} = \psi_d(L) - d \ln L$ . A complete description is provided in Appendix B.

#### B. Estimation of Texture Parameters

The texture distribution parameters  $\hat{\theta}_{\tau}$  are estimated using (13), which needs the texture pdf  $f_{\tau}(\tau_i|\theta_{\tau})$  and the posterior distribution  $f_{\tau|\mathbf{C}}(\tau_i|\mathbf{C}_i, \theta')$ . In what follows, four texture pdfs are used.

1) *Gamma Distribution*: Substituting  $\bar{\gamma}(\alpha)$  in (13), differentiating with respect to  $\alpha$ , and setting the result to zero yield

$$\ln(\hat{\alpha}) - \psi(\hat{\alpha}) + 1 = \frac{1}{N} \sum_{i=1}^N E_{\tau|\mathbf{C}}\{\tau_i - \ln \tau_i | \mathbf{C}_i, \theta'\}. \quad (17)$$

TABLE II  
POSTERIOR PDF OF  $\tau$  GIVEN THE MULTILOOK POLARIMETRIC COVARIANCE MATRIX  $\mathbf{C}$

Texture pdf	Posterior pdf	
Symbol	Symbol	$f_{\tau \mathbf{C}}(\tau_i \mathbf{C}_i)$
$\bar{\gamma}(\alpha)$	$\mathcal{N}^{-1}(\alpha - dL, 2\sqrt{L\alpha\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i)}, \sqrt{\frac{L}{\alpha}\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i)})$	$\frac{\tau_i^{\alpha-dL-1} \exp\left(-\alpha\tau_i - \frac{L}{\tau_i}\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i)\right)}{\left(\frac{L}{\alpha}\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i)\right)^{\frac{\alpha-dL}{2}} 2K_{\alpha-dL}\left(2\sqrt{L\alpha\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i)}\right)}$
$\bar{\gamma}^{-1}(\lambda)$	$\gamma^{-1}(dL + \lambda, L\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i) + \lambda - 1)$	$\frac{(L\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i) + \lambda - 1)^{dL+\lambda}}{\Gamma(dL + \lambda)} \frac{1}{\tau_i^{dL+\lambda+1}} \exp\left(-\frac{L\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i) + \lambda - 1}{\tau_i}\right)$
$\mathcal{N}^{-1}(\alpha, w, \eta)$	$\mathcal{N}^{-1}\left(\alpha - dL, \sqrt{\frac{w}{\eta}}(2L\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i) + w\eta), \sqrt{\frac{2L\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i) + w\eta}{w/\eta}}\right)$	$\frac{\tau_i^{\alpha-dL-1} \exp\left(-\frac{L}{\tau_i}\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i) - \frac{w}{2}\left(\frac{\eta}{\tau_i} + \frac{\tau_i}{\eta}\right)\right)}{\left(\frac{2L\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i) + w\eta}{w/\eta}\right)^{\frac{\alpha-dL}{2}} 2K_{\alpha-dL}\left(\sqrt{\frac{w}{\eta}}(2L\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i) + w\eta)\right)}$

TABLE III  
POSTERIOR EXPECTATION OF  $\tau$  GIVEN THE MULTILOOK POLARIMETRIC COVARIANCE MATRIX  $\mathbf{C}$

Texture pdf	Compound pdf	Posterior expectation
$f_{\tau}(\tau)$	$f_{\mathbf{C}}(\mathbf{C})$	$E_{\tau \mathbf{C}}\{\cdot \mathbf{C}_i, \theta'\}$
$\bar{\gamma}(\alpha)$	$\mathcal{K}_d(\boldsymbol{\Sigma}, L, \alpha)$	$E_{\tau \mathbf{C}}\{\tau_i \mathbf{C}_i, \theta'\} = \sqrt{\frac{L'}{\alpha'}\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i)} \frac{K_{\alpha'-dL'+1}\left(2\sqrt{L'\alpha'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i)}\right)}{K_{\alpha'-dL'}\left(2\sqrt{L'\alpha'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i)}\right)}$ $E_{\tau \mathbf{C}}\left\{\frac{1}{\tau_i} \mathbf{C}_i, \theta'\right\} = \frac{1}{\sqrt{\frac{L'}{\alpha'}\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i)} \frac{K_{\alpha'-dL'-1}\left(2\sqrt{L'\alpha'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i)}\right)}{K_{\alpha'-dL'}\left(2\sqrt{L'\alpha'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i)}\right)}$ $E_{\tau \mathbf{C}}\{\ln \tau_i \mathbf{C}_i, \theta'\} = \ln \sqrt{\frac{L'}{\alpha'}\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i)} + \frac{\partial}{\partial \alpha} K_{\alpha'-dL'+\alpha}\left(2\sqrt{L'\alpha'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i)}\right) \Big _{\alpha=0}$ $K_{\alpha'-dL'}\left(2\sqrt{L'\alpha'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i)}\right)$
$\bar{\gamma}^{-1}(\lambda)$	$\mathcal{G}_d^0(\boldsymbol{\Sigma}, L, \lambda)$	$E_{\tau \mathbf{C}}\{\tau_i \mathbf{C}_i, \theta'\} = \frac{L'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i) + \lambda' - 1}{dL' + \lambda' - 1}$ $E_{\tau \mathbf{C}}\left\{\frac{1}{\tau_i} \mathbf{C}_i, \theta'\right\} = \frac{dL' + \lambda' - 1}{L'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i) + \lambda' - 1}$ $E_{\tau \mathbf{C}}\{\ln \tau_i \mathbf{C}_i, \theta'\} = -\psi(dL' + \lambda') + \ln(L'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i) + \lambda' - 1)$
$\mathcal{N}^{-1}(\alpha, w, \eta)$	$\mathcal{G}_d(\boldsymbol{\Sigma}, L, \alpha, w, \eta)$	$E_{\tau \mathbf{C}}\{\tau_i \mathbf{C}_i, \theta'\} = \sqrt{\frac{2L'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i) + w'\eta'}{w'/\eta'}} \frac{K_{\alpha'-dL'+1}\left(\sqrt{\frac{w'}{\eta'}}(2L'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i) + w'\eta')\right)}{K_{\alpha'-dL'}\left(\sqrt{\frac{w'}{\eta'}}(2L'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i) + w'\eta')\right)}$ $E_{\tau \mathbf{C}}\left\{\frac{1}{\tau_i} \mathbf{C}_i, \theta'\right\} = \sqrt{\frac{w'/\eta'}{2L'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i) + w'\eta'}} \frac{K_{\alpha'-dL'-1}\left(\sqrt{\frac{w'}{\eta'}}(2L'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i) + w'\eta')\right)}{K_{\alpha'-dL'}\left(\sqrt{\frac{w'}{\eta'}}(2L'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i) + w'\eta')\right)}$ $E_{\tau \mathbf{C}}\{\ln \tau_i \mathbf{C}_i, \theta'\} = \ln \sqrt{\frac{2L'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i) + w'\eta'}{w'/\eta'}} + \frac{\partial}{\partial \alpha} K_{\alpha'-dL'+\alpha}\left(\sqrt{\frac{w'}{\eta'}}(2L'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i) + w'\eta')\right) \Big _{\alpha=0}$ $K_{\alpha'-dL'}\left(\sqrt{\frac{w'}{\eta'}}(2L'\text{tr}(\boldsymbol{\Sigma}'^{-1}\mathbf{C}_i) + w'\eta')\right)$

The ML-based  $\alpha$  parameter estimator is the solution to (17) involving  $f_{\tau|\mathbf{C}}(\tau_i|\mathbf{C}_i)$ . Using Bayes' rule, the posterior pdf of  $\tau_i$  given  $\mathbf{C}_i$  is provided as follows:

$$f_{\tau|\mathbf{C}}(\tau_i|\mathbf{C}_i) = \frac{f_{\mathbf{C}}(\mathbf{C}_i|\tau_i)f_{\tau}(\tau_i)}{f_{\mathbf{C}}(\mathbf{C}_i)}. \quad (18)$$

The expression of  $f_{\tau|\mathbf{C}}(\tau_i|\mathbf{C}_i)$  is presented in Table II in the case where the texture is gamma-distributed. This posterior distribution corresponds to the GIG distribution (see Appendix C), with the parameters ( $\alpha_1 = \alpha - dL, w_1 = 2(L\alpha\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i))^{1/2}, \eta_1 = ((L/\alpha)\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i))^{1/2}$ ). The posterior expectation expressions in terms of these parameters are shown in Table III. It is worth noticing that (17) can be interpreted as the difference between  $E_{\tau}\{\tau_i\}$  and  $E_{\tau}\{\ln \tau_i\}$ , with  $E_{\tau}\{\tau_i\} = 1 = E_{\mathbf{C}}\{E_{\tau|\mathbf{C}}\{\tau_i|\mathbf{C}_i\}\}$  and  $E_{\tau}\{\ln \tau_i\} = \psi(\alpha) - \ln(\alpha) = E_{\mathbf{C}}\{E_{\tau|\mathbf{C}}\{\ln \tau_i|\mathbf{C}_i\}\}$ .

2) *Inverse Gamma Distribution*: In a similar manner, substituting  $\bar{\gamma}^{-1}(\lambda)$  in (13), differentiating with respect to  $\lambda$ , and setting the result to zero yield

$$\ln(\hat{\lambda} - 1) - \psi(\hat{\lambda}) + \frac{\hat{\lambda}}{\hat{\lambda} - 1} = \frac{1}{N} \sum_{i=1}^N E_{\tau|\mathbf{C}} \left\{ \frac{1}{\tau_i} + \ln \tau_i | \mathbf{C}_i, \theta' \right\}. \quad (19)$$

The posterior pdf of  $\tau_i$  given  $\mathbf{C}_i$  is provided in a closed form in Table II. It corresponds to an inverse gamma distribution with parameters ( $\alpha_1 = dL + \lambda, \beta_1 = L\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}_i) + \lambda - 1$ ). The posterior expectation expression in terms of these parameters is defined in Table III (see Appendix D). Equation (19) represents the sum of  $E_{\tau}\{(1/\tau_i)\}$  and  $E_{\tau}\{\ln \tau_i\}$ , where  $E_{\tau}\{(1/\tau_i)\} = \lambda/(\lambda - 1) = E_{\mathbf{C}}\{E_{\tau|\mathbf{C}}\{(1/\tau_i)|\mathbf{C}_i\}\}$  and  $E_{\tau}\{\ln \tau_i\} = \ln(\lambda - 1) - \psi(\lambda) = E_{\mathbf{C}}\{E_{\tau|\mathbf{C}}\{\ln \tau_i|\mathbf{C}_i\}\}$ .

3) *Generalized Inverse Gaussian Distribution*: Substituting  $\mathcal{N}^{-1}(\alpha, w, \eta)$  in (13), differentiating with respect to  $\alpha, w$ , and  $\eta$ , and setting the result to zero yield

$$\frac{K_{\hat{\alpha}-1}(\hat{w}) + K_{\hat{\alpha}+1}(\hat{w})}{K_{\hat{\alpha}}(\hat{w})} = \frac{1}{N} \sum_{i=1}^N E_{\tau|\mathbf{C}} \left\{ \frac{\hat{\eta}}{\tau_i} + \frac{\tau_i}{\hat{\eta}} | \mathbf{C}_i, \theta' \right\} \quad (20)$$

$$\frac{1}{K_{\hat{\alpha}}(\hat{w})} \frac{\partial K_{\hat{\alpha}}(\hat{w})}{\partial \alpha} = \frac{1}{N} \sum_{i=1}^N E_{\tau|\mathbf{C}} \left\{ \ln \frac{\tau_i}{\hat{\eta}} | \mathbf{C}_i, \theta' \right\} \quad (21)$$

$$2 \frac{\hat{\alpha}}{\hat{w}} = \frac{1}{N} \sum_{i=1}^N E_{\tau|\mathbf{C}} \left\{ \frac{\tau_i}{\hat{\eta}} - \frac{\hat{\eta}}{\tau_i} | \mathbf{C}_i, \theta' \right\}. \quad (22)$$

The posterior pdf of  $\tau_i$  given  $\mathbf{C}_i$  is given in a closed form in Table II. This is a  $\mathcal{N}^{-1}$  distribution with the parameters

$[\alpha_1 = \alpha - dL, w_1 = ((w/\eta)(2L\text{tr}(\mathbf{\Sigma}^{-1}\mathbf{C}_i) + w\eta))^{1/2}$ , and  $\eta_1 = ((2L\text{tr}(\mathbf{\Sigma}^{-1}\mathbf{C}_i) + w\eta)/w\eta)^{1/2}E_{\tau}\{\tau_i + (1/\tau_i)\} = E_{\mathbf{C}}\{E_{\tau|\mathbf{C}}\{\tau_i + (1/\tau_i)|\mathbf{C}_i\}\}$  and  $E_{\tau}\{\ln \tau_i\} = \ln \eta + (1/K_{\alpha}(w))(\partial K_{\alpha}(w)/\partial \alpha) = E_{\mathbf{C}}\{E_{\tau|\mathbf{C}}\{\ln \tau_i|\mathbf{C}_i\}\}$ , respectively, with  $E_{\tau}\{\tau_i\} = \eta(K_{\alpha+1}(w)/K_{\alpha}(w))$  and  $E_{\tau}\{(1/\tau_i)\} = (1/\eta)(K_{\alpha-1}(w)/K_{\alpha}(w))$ .

### C. Implementation of EM Algorithm

Equations (14), (15), and (17), (14), (15), and (19), and (14), (15), and (20)–(22) represent the iterative EM algorithm for the estimation of the matrix-variate  $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ , and  $\mathcal{G}_d$  distribution parameters, respectively. The algorithm can be initialized with any suitable values of  $\mathbf{\Sigma}'$ ,  $L'$ , and  $\theta'_{\tau}$ . The initial parameter values for the EM algorithm can be produced by the MFM method or the MLC method. We have to note that the choice of initialization value too far from the final solution can cause the divergence of the algorithm. The algorithm is stopped when the convergence criterion is satisfied. The partial derivative  $(\partial/\partial \alpha)K_{\alpha}(\cdot)$  is approximated using a numerical differentiation given as follows:

$$\frac{\partial}{\partial \alpha} K_{\alpha}(x) \approx \frac{K_{\alpha+h}(x) - K_{\alpha-h}(x)}{2h} \quad (23)$$

with  $h = 10^{-3}$ . The approximation accuracy of the numerical derivatives depends on the smoothness of the functions. The algorithm is stopped once the difference between the successive parameter estimates is less than 0.5% or if the maximum iteration number 300 is reached. Since  $\ln(z) - \psi(z) + 1$  and  $d \ln(z) + d - \psi_d(z)$  are monotonically decreasing with respect to  $z$ , the value of  $\hat{\alpha}$  ( $\mathcal{K}_d$ ) and  $\hat{L}$  is uniquely determined if a solution exists. The function  $\ln(z-1) - \psi(z) + z/(z-1)$  is monotonically increasing with respect to  $z$ , and  $\hat{\lambda}$  is uniquely determined. The trust-region method [31] is used to solve (17) for  $\mathcal{K}_d$ , (19) for  $\mathcal{G}_d^0$ , and the system of equations (20)–(22) for  $\mathcal{G}_d$ . Specifically, the *fsolve* function in MATLAB is utilized to obtain the parameter estimates. The estimation approach of the parameters is summarized in Algorithm 1. In our case, the Frobenius norm is used in step 11.

## IV. ESTIMATORS PERFORMANCE COMPARISON

In this section, the performance of the proposed method is compared with the performance of different estimators for ENL and texture parameters by comparing their bias, variance, and mean square error (mse) using Monte Carlo methods. This is computed for the matrix-variate  $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ , and  $\mathcal{G}_d$  distributions. In Section IV-A, a lower bound on the variance of these parameters is established and the hybrid Cramer–Rao bound (HCRB) introduced by Gini and Reggiannini [32] is used. The HCRB is easy to calculate at the cost of bound tightness [33].

### A. Hybrid Cramer–Rao Bound

The performance of an unbiased estimator is judged by the Cramer–Rao lower bound (CRLB), which is known to be a

### Algorithm 1 Parameter estimation

- 
- 1: **Input:**  $N, \mathbf{C}_i, d, \epsilon$
  - 2: **Output:**  $\hat{\theta} = (\hat{\mathbf{\Sigma}}, \hat{L}, \hat{\theta}_{\tau})$
  - 3: **Initialization:**
  - 4:     Set initial parameters  $\theta' = (\mathbf{\Sigma}' = \mathbf{I}_d, L_0, \theta_{\tau_0})$
  - 5: **Main Loop:**
  - 6:     **Repeat**
  - 7:         Calculate  $E_{\tau|\mathbf{C}}\{\cdot|\mathbf{C}_i, \theta'\}$  according to the compound pdf using Table III
  - 8:         Estimate  $\hat{\mathbf{\Sigma}}$  using (14). For the case of  $\mathcal{G}_d$  distribution use:  $\hat{\mathbf{\Sigma}} \leftarrow \frac{\hat{\mathbf{\Sigma}}}{|\hat{\mathbf{\Sigma}}|^{1/d}}$
  - 9:         Estimate  $\hat{L}$  using  $\hat{\mathbf{\Sigma}}$  and (15)
  - 10:         Estimate  $\hat{\theta}_{\tau}$  using (17) for  $\mathcal{K}_d$ , or (19) for  $\mathcal{G}_d^0$ , or (20, 21, 22) for  $\mathcal{G}_d$ .
  - 11:         Calculate stop criterion:  $D \leftarrow \|\hat{\theta} - \theta'\|$
  - 12:         Set inputs for next iteration:  $\mathbf{\Sigma}', L', \theta'_{\tau} \leftarrow \hat{\mathbf{\Sigma}}, \hat{L}, \hat{\theta}_{\tau}$
  - 13:         **Until**  $D < \epsilon$
  - 14:     **Return**  $(\hat{\mathbf{\Sigma}}, \hat{L}, \hat{\theta}_{\tau})$
- 

tight bound for a wide class of estimators [32]. However, in many practical situations, the computation of this bound is infeasible. The computation of a less optimal bound is considered in this paper. It is the HCRB which is defined by Gini and Reggiannini [32]. As it is mentioned in their paper, there are other alternatives to the Cramer–Rao bound, such as the HCRB, the modified Cramer–Rao bound (MCRB) [34], the Miller–Chang bound (MCB) [35], and the extended MCB (EMCB) [32]. The HCRB is tighter than the MCRB, since the latter is generally lower than the former. In addition, both the bounds apply to a wide class of estimators that are unbiased on the average [32]. Contrariwise, the MCB and the EMCB apply to a more restricted class of estimators than the conventional CRB [32]. In this paper, the definition of the HCRB is adapted for complex parameters following the methodology of Anfinson *et al.* [13] for the computation of a CRLB for complex parameters [36]. The information matrix of the HCRB is defined as

$$\begin{aligned} \mathbf{J}(\Theta) &= E_{\mathbf{C}, \tau} \left\{ \left( \frac{\partial \ln f_{\mathbf{C}, \tau}(\mathbf{C}, \tau|\Theta)}{\partial \Theta^T} \right)^H \frac{\partial \ln f_{\mathbf{C}, \tau}(\mathbf{C}, \tau|\Theta)}{\partial \Theta^T} \right\} \\ &= -E_{\mathbf{C}, \tau} \left\{ \frac{\partial}{\partial \Theta^*} \left( \frac{\partial \ln f_{\mathbf{C}, \tau}(\mathbf{C}, \tau|\Theta)}{\partial \Theta} \right)^T \right\} \quad (24) \end{aligned}$$

where  $\Theta = [\theta_{\tau}^T, \Theta_{\Sigma}^T]^T$  and  $\Theta_{\Sigma} = [L, \Sigma_{11}, \Sigma_{21}, \dots, \Sigma_{dd}]^T = [L, \text{vec}(\mathbf{\Sigma})^T]^T$  is the complex-valued parameter vector of the speckle, with the vectorization operator denoted as  $\text{vec}(\cdot)$ . The superscript  $*$  denotes complex conjugation.  $\Sigma_{ij}, i, j \in \{1, \dots, d\}$  are the entries of the covariance matrix. As seen in (24), the expectation is carried out with respect to both  $\mathbf{C}$  and  $\tau$ . If the estimator  $\hat{\Theta}$  is unbiased, then the covariance matrix of the parameters will satisfy the following relation:

$$\text{cov}(\hat{\Theta}) \geq \text{HCRB}(\Theta) = \mathbf{J}^{-1}(\Theta). \quad (25)$$

After simplification, the information matrix of the HCRB is a block diagonal matrix with two blocks given

TABLE IV  
HCRB FOR TEXTURE DISTRIBUTIONS

Symbol	$\mathbf{J}(\theta_\tau)$	$\mathbf{J}^{-1}(\theta_\tau)$
$\bar{\gamma}(\alpha)$	$N(-\frac{1}{\alpha} + \psi^{(1)}(\alpha))$	$\text{var}(\hat{\alpha}) > \frac{1}{N} \frac{\alpha}{\alpha \psi^{(1)}(\alpha) - 1}$
$\bar{\gamma}^{-1}(\lambda)$	$N(\frac{1}{(\lambda-1)^2} + \psi^{(1)}(\alpha) - \frac{1}{\lambda-1})$	$\text{var}(\hat{\lambda}) > \frac{1}{N} \frac{(\lambda-1)^2}{(\lambda-1)^2 \psi^{(1)}(\lambda) - \lambda + 2}$
$\mathcal{N}^{-1}(\alpha, w, \eta)$	$N \begin{bmatrix} \frac{\partial^2}{\partial \alpha^2} \ln K_\alpha(w) & \frac{\partial^2}{\partial \alpha \partial w} \ln K_\alpha(w) & \frac{1}{\eta} \\ \frac{\partial^2}{\partial w \partial \alpha} \ln K_\alpha(w) & \frac{\partial^2}{\partial w^2} \ln K_\alpha(w) & \frac{1}{2\eta} \frac{K_{\alpha-1}(w) - K_{\alpha+1}(w)}{K_\alpha(w)} \\ \frac{1}{\eta} & \frac{1}{2\eta} \frac{K_{\alpha-1}(w) - K_{\alpha+1}(w)}{K_\alpha(w)} & \frac{1}{\eta^2} (-\alpha + w \frac{K_{\alpha+1}(w)}{K_\alpha(w)}) \end{bmatrix}$	$\text{var}(\hat{\alpha}) > [\mathbf{J}^{-1}(\theta_\tau)]_{11} = \frac{1}{[\mathbf{J}(\theta_\tau)]} \left[ \frac{\partial^2 \ln K_\alpha(w)}{\partial w^2} \times \frac{1}{\eta^2} (-\alpha + w \frac{K_{\alpha+1}(w)}{K_\alpha(w)}) - \frac{1}{4\eta^2} (\frac{K_{\alpha-1}(w) - K_{\alpha+1}(w)}{K_\alpha(w)})^2 \right]$ $\text{var}(\hat{w}) > [\mathbf{J}^{-1}(\theta_\tau)]_{22} = \frac{1}{[\mathbf{J}(\theta_\tau)]} \left[ \frac{\partial^2 \ln K_\alpha(w)}{\partial \alpha^2} \times \frac{1}{\eta^2} (-\alpha + w \frac{K_{\alpha+1}(w)}{K_\alpha(w)}) - \frac{1}{\eta^2} \right]$

as

$$\begin{aligned} \mathbf{J}(\Theta) &= -E_{\mathbf{C}, \tau} \left\{ \frac{\partial}{\partial \Theta^*} \left( \frac{\partial \ln f_{\mathbf{C}|\tau}(\mathbf{C}|\tau, \Theta_\Sigma)}{\partial \Theta} \right)^T \right. \\ &\quad \left. + \frac{\partial}{\partial \Theta^*} \left( \frac{\partial \ln f_\tau(\tau|\theta_\tau)}{\partial \Theta} \right)^T \right\} \\ &= \begin{bmatrix} \mathbf{J}(\theta_\tau) & 0 \\ 0 & \mathbf{J}(\Theta_\Sigma) \end{bmatrix} \end{aligned} \quad (26)$$

with

$$\mathbf{J}(\theta_\tau) = -E_\tau \left\{ \frac{\partial}{\partial \theta_\tau} \left( \frac{\partial \ln f_\tau(\tau|\theta_\tau)}{\partial \theta_\tau} \right)^T \right\} \quad (27)$$

$$\mathbf{J}(\Theta_\Sigma) = -E_{\mathbf{C}, \tau} \left\{ \frac{\partial}{\partial \Theta_\Sigma^*} \left( \frac{\partial \ln f_{\mathbf{C}|\tau}(\mathbf{C}|\tau, \Theta_\Sigma)}{\partial \Theta_\Sigma} \right)^T \right\}. \quad (28)$$

The HCRB is defined as the inverse of  $\mathbf{J}(\Theta)$

$$\mathbf{J}^{-1}(\Theta) = \frac{1}{N} \begin{bmatrix} \frac{1}{N} \mathbf{J}(\theta_\tau) & 0 & 0 \\ 0 & -\frac{d}{L} + \psi_d^{(1)}(L) & 0 \\ 0 & 0 & L(\Sigma^{-1} \otimes \Sigma^{-1}) \end{bmatrix}^{-1} \quad (29)$$

where  $\otimes$  denotes the Kronecker product. The derivation is shown in Appendix E.  $\mathbf{J}(\theta_\tau)$  is given in Table IV for each texture distribution.

The bound on the variance of the ENL estimator thus becomes

$$\text{var}(\hat{L}) > [\mathbf{J}^{-1}(\Theta_\Sigma)]_{11} = \frac{1}{N} \frac{L}{L \psi_d^{(1)}(L) - d} \quad (30)$$

where  $[\mathbf{J}^{-1}(\cdot)]_{ij}$  denotes the element  $(i, j)$  of  $\mathbf{J}^{-1}(\cdot)$ . The other lower bounds on the variance of the texture parameters are given in Table IV.

### B. Estimation of Equivalent Number of Looks

We present, in this section, a brief description of the MoMLC and the DTM for the estimation of the ENL.

1) *Method of Matrix Log-Cumulant*: The first and the  $\nu$ th-order MLCs of the multilook polarimetric covariance matrix  $\mathbf{C}$ , denoted by  $\kappa_1\{\mathbf{C}\}$  and  $\kappa_\nu\{\mathbf{C}\}$ , respectively, evaluated under the product model for general texture variable, are given by [12]

$$\kappa_1\{\mathbf{C}\} = \psi_d(L) + \ln |\Sigma| - d(\ln L - \kappa_1\{\tau\}) \quad (31)$$

$$\kappa_\nu\{\mathbf{C}\} = \psi_d^{(\nu)}(L) + d^\nu \kappa_\nu(\tau), \quad \nu > 1 \quad (32)$$

TABLE V  
LOG-CUMULANTS OF SOME TEXTURE DISTRIBUTIONS

Symbol	Log-cumulants $\kappa_\nu\{\tau\}$
$\bar{\gamma}(\alpha)$	$\kappa_1\{\tau\} = \psi(\alpha) - \ln(\alpha),$ $\kappa_\nu\{\tau\} = \psi^{(\nu-1)}(\alpha), \quad \nu > 1$
$\bar{\gamma}^{-1}(\lambda)$	$\kappa_1\{\tau\} = \ln(\lambda - 1) - \psi(\lambda),$ $\kappa_\nu\{\tau\} = (-1)^\nu \psi^{(\nu-1)}(\lambda), \quad \nu > 1$
$\mathcal{N}^{-1}(\alpha, w, \eta)$	$\kappa_1\{\tau\} = \ln \eta + \ln K_\alpha^{(1)}(w),$ $\kappa_\nu\{\tau\} = \ln K_\alpha^{(\nu)}(w), \quad \nu > 1$

where  $\kappa_\nu\{\tau\}$  is  $\nu$ th-order log-cumulant of the texture variable,  $\psi_d^{(\nu)}(L)$  is the  $\nu$ th-order multivariate polygamma function defined as  $\psi_d^{(\nu)}(L) = \sum_{i=0}^{d-1} \psi^{(\nu)}(L - i)$ , and  $\psi^{(\nu)}(L)$  is the polygamma function of order  $\nu$ . The MLCs are related to the matrix log-moments (MLMs) denoted by  $\mu_\nu\{\mathbf{C}\}$ . The first MLCs are expressed by [12]

$$\kappa_1\{\mathbf{C}\} = \mu_1\{\mathbf{C}\} \quad (33)$$

$$\kappa_2\{\mathbf{C}\} = \mu_2\{\mathbf{C}\} - \mu_1^2\{\mathbf{C}\} \quad (34)$$

$$\kappa_3\{\mathbf{C}\} = \mu_3\{\mathbf{C}\} - 3\mu_1\{\mathbf{C}\}\mu_2\{\mathbf{C}\} + 2\mu_1^3\{\mathbf{C}\}. \quad (35)$$

To evaluate the MLMs, the sample MLMs are used, which are calculated with the sample mean estimator  $\mu_\nu\{\mathbf{C}\} \triangleq (1/N) \sum_{i=1}^N (\ln |\mathbf{C}_i|)^\nu$ . When the specific texture log-cumulant is inserted in (31) and (32), the MLCs of  $\mathbf{C}$  will then be determined. The log-cumulants of the texture distribution are given in Table V.  $\ln K_\alpha^{(\nu)}(w) = (d^\nu/d\alpha^\nu) \ln K_\alpha(w)$  is the  $\nu$ th derivative with respect to order of the logarithm of the modified Bessel function of the second kind. The log-cumulants of the GIG distribution are only computable using numerical differentiation.

To estimate the ENL, Anfinson *et al.* [12] proposed the first MLC-based estimator, denoted by  $\hat{L}_{A1}$ , estimated from the first-order MLC equation of the complex Wishart distribution under a textureless area condition. This means that the log-cumulant of the texture variable is omitted from (31). The equation to solve for  $\hat{L}_{A1}$  is then as follows:

$$\psi_d(\hat{L}_{A1}) - d \ln \hat{L}_{A1} = \kappa_1\{\mathbf{C}\} - \ln |\Sigma| \quad (36)$$

where the covariance matrix  $\Sigma$  is estimated before solving  $L$  by using the SCM.

The second MLC-based estimator, denoted by  $\hat{L}_{A2}$ , is based on a system of equations (31) and (32). These MLC equations contain texture parameters from the first and the higher order.  $\hat{L}_{A2}$  and the texture parameters need to be estimated jointly from the system of equations. Section IV-C describes the estimation of  $\hat{L}_{A2}$  for different stochastic models.

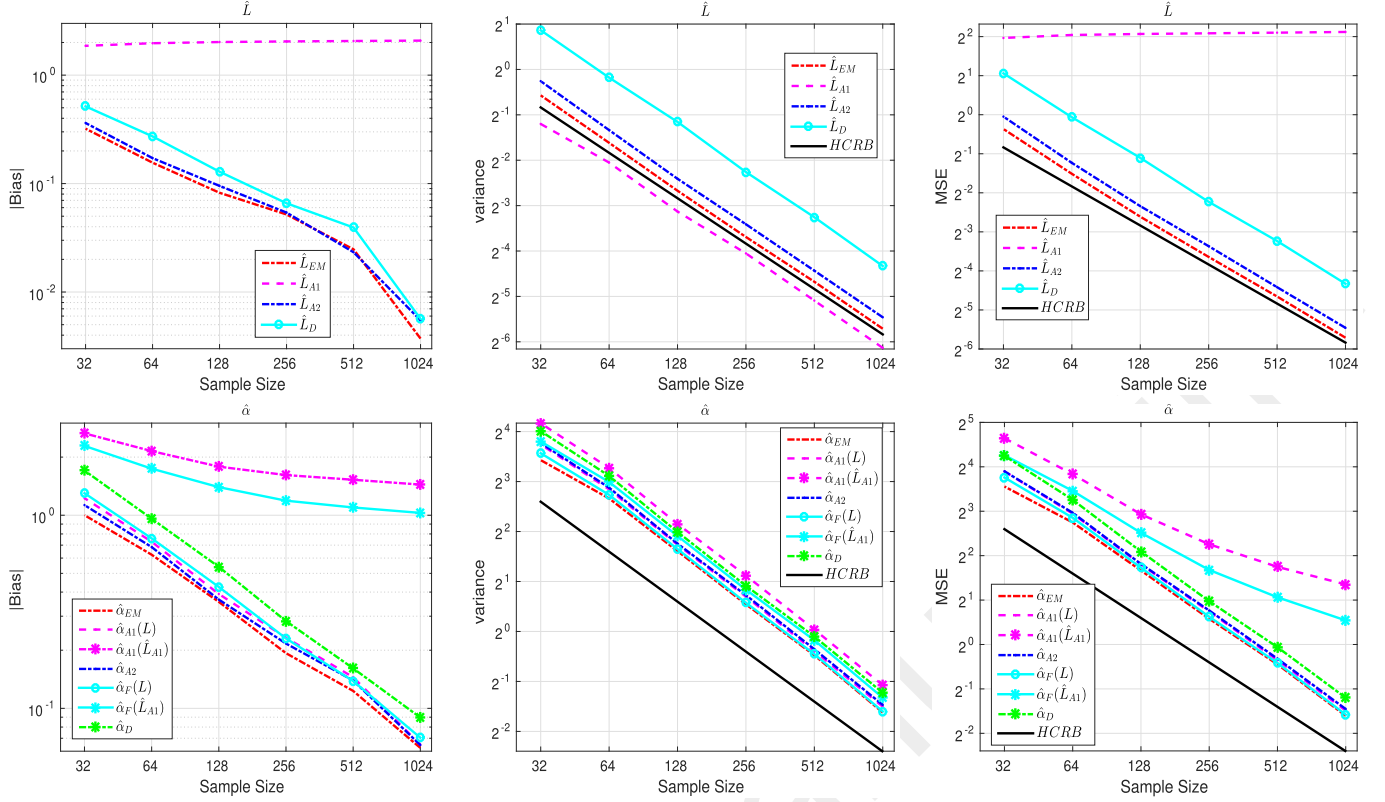


Fig. 1. (Left) Bias, (Middle) variance, and (Right) mse of estimators for (Top) ENL  $L$  and (Bottom)  $\mathcal{K}_d$  distribution texture parameter  $\alpha$  as a function of sample size  $N$ .

2) *Method of DTM*: An alternative estimator was proposed by Tao *et al.* [14]. They applied the trace moments of the covariance matrices to estimate the ENL for both the Wishart model and the product model. The DTM estimator is given as follows [14]:

$$\hat{L}_D = \frac{E\{\text{tr}^2(\mathbf{C})\}\text{tr}^2(E\{\mathbf{C}\}) - E\{\text{tr}(\mathbf{C}^2)\}\text{tr}(E^2\{\mathbf{C}\})}{E\{\text{tr}(\mathbf{C}^2)\}\text{tr}^2(E\{\mathbf{C}\}) - E\{\text{tr}^2(\mathbf{C})\}\text{tr}(E^2\{\mathbf{C}\})}. \quad (37)$$

### C. Texture Parameter Estimation

The matrix-variate  $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ , and  $\mathcal{G}_d$  are the most widely used stochastic models to describe the product model. The performance of the EM and some other estimators, such as the MoMLC, the MoMFM, and the Doulgeris method for texture parameter estimation, are evaluated using these distribution models.

1) *Matrix-Variate  $\mathcal{K}$  Distribution*: Under this distribution, the MLC-based estimator, denoted by  $\hat{\alpha}_{A1}$ , is related to the second-order MLC and proposed by Anfinsen *et al.* [12]. This estimator is obtained by solving the following equation:

$$\psi^{(1)}(\hat{\alpha}_{A1}) = \frac{\kappa_2(\mathbf{C}) - \psi_d^{(1)}(L)}{d^2} \quad (38)$$

where  $L$  is supposed to be a known constant. Two choices are made for the value of  $L$ . The first choice is to use the true value of  $L$  and the estimator is then denoted by  $\hat{\alpha}_{A1}(L)$ . The second choice is to use the first MLC-based estimator  $\hat{L}_{A1}$  estimated from the first-order MLC equation of the complex Wishart distribution given by (36), and then, the estimator is denoted by  $\hat{\alpha}_{A1}(\hat{L}_{A1})$ .

The second MLC-based estimator, denoted by  $\hat{\alpha}_{A2}$ , is based on the resolution of a system of equations dependent on a multiple MLCs and can be defined as

$$\begin{aligned} \kappa_1\{\mathbf{C}\} &= \psi_d(\hat{L}_{A2}) + \ln|\boldsymbol{\Sigma}| - d \ln \hat{L}_{A2} + d(\psi(\hat{\alpha}_{A2}) - \ln \hat{\alpha}_{A2}) \\ \kappa_2\{\mathbf{C}\} &= \psi_d^{(1)}(\hat{L}_{A2}) + d^2 \psi^{(1)}(\hat{\alpha}_{A2}) \end{aligned} \quad (39)$$

where  $\hat{L}_{A2}$  also needs to be solved and  $\boldsymbol{\Sigma}$  is estimated using the SCM, since the texture is normalized to a unit mean resulting in  $E\{\mathbf{C}\} = \boldsymbol{\Sigma}$ .

The MoMFM-based estimator proposed by Khan and Guida [22] for the texture parameters is obtained by solving the following equation:

$$\frac{E\{M^\nu\}}{E^\nu\{M\}} = \frac{E\{\tau^\nu\}}{E^\nu\{\tau\}} \left(\frac{1}{Ld}\right)^\nu \frac{\Gamma(Ld + \nu)}{\Gamma(Ld)} \quad (40)$$

where  $M = \text{tr}(\boldsymbol{\Gamma}^{-1}\mathbf{C})$ ,  $\boldsymbol{\Gamma} = E\{\mathbf{C}\}$ , and  $\nu \in \mathbb{R}^+$ . The new variable  $M$  can be expressed as a function of the variable texture:  $M = \tau/E\{\tau\}\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{X})$ , where  $\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{X})$  follows the gamma distribution with the parameters  $(Ld, L)$ . The  $\nu$ th moment of  $M$  is calculated with the sample mean estimator  $1/N \sum_{i=1}^N M_i^\nu$  given a sample of  $\{M_i = \text{tr}(\boldsymbol{\Gamma}^{-1}\mathbf{C}_i)\}_{i=1}^N$ . For the case of the matrix-variate  $\mathcal{K}_d$  distribution, the MFM estimator, denoted by  $\hat{\alpha}_F$ , can be obtained as a solution for the following equation:

$$\frac{E\{M^\nu\}}{E^\nu\{M\}} = \frac{\Gamma(\hat{\alpha}_F + \nu)}{\hat{\alpha}_F^\nu \Gamma(\hat{\alpha}_F)} \left(\frac{1}{Ld}\right)^\nu \frac{\Gamma(Ld + \nu)}{\Gamma(Ld)}. \quad (41)$$

The value of  $\nu$  is chosen to be equal to 1/8 in the experiments. The choice is arbitrary, but generally, lower  $\nu$  makes lower



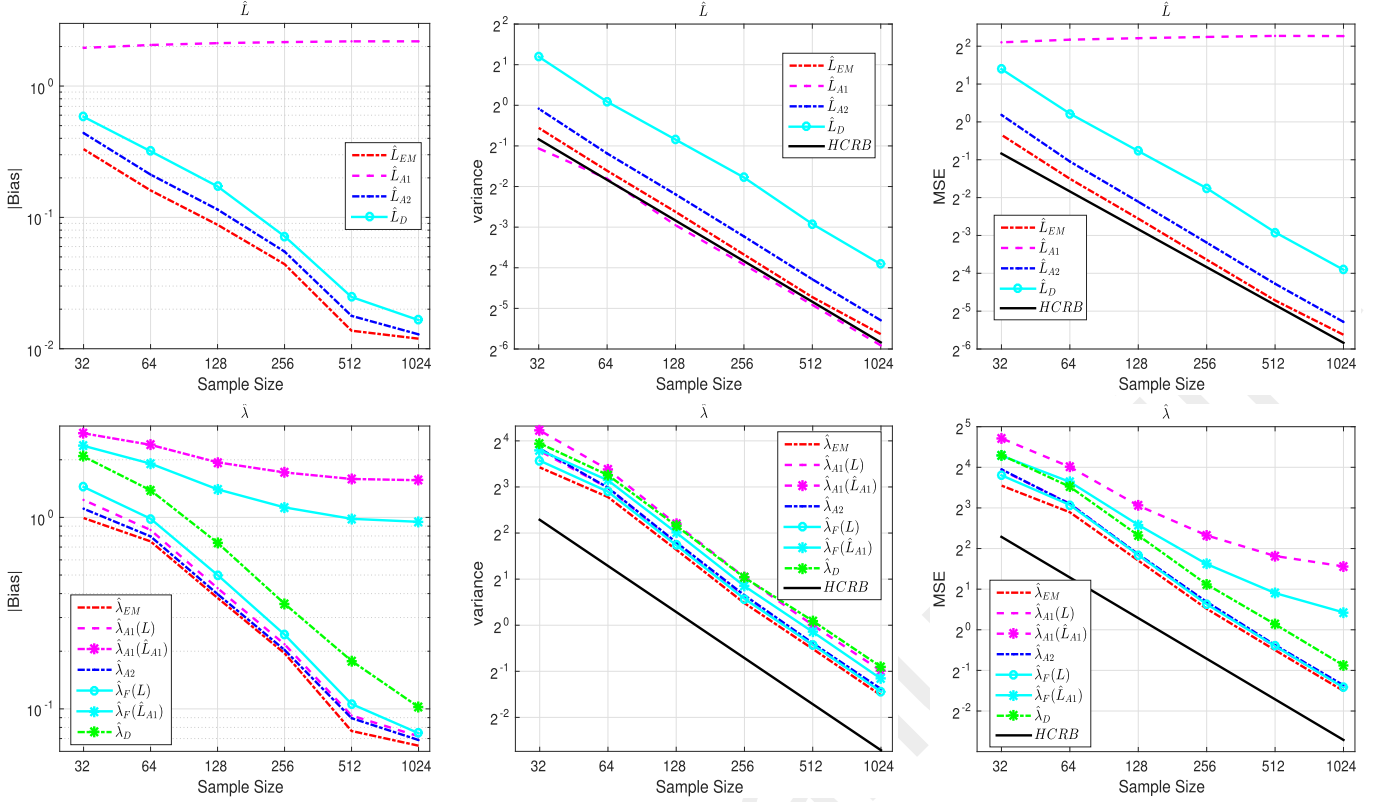


Fig. 2. (Left) Bias, (Middle) variance, and (Right) mse of estimators for (Top) ENL  $L$  and (Bottom)  $\mathcal{G}_d^0$  distribution texture parameter  $\lambda$  as a function of sample size  $N$ .

variance [22]. The MFM estimator depends on the value of  $L$ .  $\hat{\alpha}_F(L)$  and  $\hat{\alpha}_F(\hat{L}_{A1})$  are the estimators dependent on the true value and an estimate of  $L$ , respectively.

Another known estimator is Doulgeris' estimator [21] which is given for the matrix-variate  $\mathcal{K}_d$  distribution by

$$\hat{\alpha}_D = \frac{d(Ld + 1)}{L \cdot \text{var}(M) - d}. \quad (42)$$

The variance of  $M$ ,  $\text{var}(M)$ , is estimated with a standard variance estimator from a sample of Hotelling–Lawley traces  $\{M_i = \text{tr}(\mathbf{\Gamma}^{-1}\mathbf{C}_i)\}_{i=1}^N$  [12], [21] and considering that  $\tau \sim \bar{\gamma}(\alpha)$  with  $E\{\tau\} = 1$  and  $E\{\tau^2\} = (\alpha + 1)/\alpha$ .

Fig. 1 shows the value of bias, variance, and mse of all the estimators after 2000 Monte Carlo simulations with  $\alpha = 10$  and  $L = 10$ . It shows that the EM estimator outperforms the other estimators. The estimator  $\hat{L}_{EM}$  is better than  $\hat{L}_{A2}$ . The worst estimator is  $\hat{L}_{A1}$ . The DTM estimator  $\hat{L}_D$  has a less performance than  $\hat{L}_{A2}$ . It is clearly seen that the variance of  $\hat{L}_{A1}$  is less than the HCRB which is normal, given that the HCRB is calculated assuming a textureless area condition. It is observable that the MoMFM-based estimator  $\hat{\alpha}_F(\hat{L}_{A1})$  and the first MLC-based estimator  $\hat{\alpha}_{A1}(\hat{L}_{A1})$  perform the worst, where the former exhibits the lower bias, variance, and mse than the latter. The EM- and MLC2-based estimators exhibit very close performances. The better overall performance of the  $\hat{\alpha}_{EM}$  estimator is shown through the lower mse compared with the other estimators. When using the true value of  $L$ , the performance of the MoMFM-based estimator  $\hat{\alpha}_F(L)$  and  $\hat{\alpha}_{A1}(L)$  is very close to the performance of the EM estimator

$\hat{\alpha}_{EM}$  and the second MLC estimator  $\hat{\alpha}_{A2}$ . It is clear that the underestimation of the ENL  $L$  affects the quality of the estimation of the other parameters, especially when the ENL is estimated using the first MLC-based estimator. The performance of Doulgeris' estimator  $\hat{\alpha}_D(L)$  is between the performances of  $\hat{\alpha}_{EM}$  and  $\hat{\alpha}_F(\hat{L}_{A1})$ . In terms of the computational time comparison between the different algorithms, the EM algorithm takes more time than the others due to the iterative nature of the algorithm. The ratio of the EM to MFM computation time for  $N = 256$  and for the texture parameter  $\alpha = 10$  is 6.5. In addition, the EM algorithm converges quickly at low  $\alpha$  values and slowly at higher values.

In addition to plotting the evolution of the bias and the mse as a function of the sample size, we have evaluated them in Fig. 3 as a function of the shape parameter  $\alpha$  that describes the degree of homogeneous or heterogeneous of the texture. The goal is to assess the influence of the homogeneity of the texture on the quality of the estimators. The sample size is fixed to 512, and 2000 Monte Carlo simulations are realized for each value of  $\alpha \in [5, 80]$ . We can clearly observe that the estimators  $\hat{L}_{EM}$  and  $\hat{L}_{A2}$  maintain their performances regardless of the value of  $\alpha$ . The mse of  $\hat{L}_{A1}$  decreases, as the value of  $\alpha$  increases. This behavior is different from that of Fig. 1, where the mse of  $\hat{L}_{A1}$  is invariant as a function of the number of samples. In fact, with homogeneous texture (large value of  $\alpha$ ), the matrix-variate  $\mathcal{K}$  distribution converges toward the complex scaled Wishart distribution and the first MLC-based estimator  $\hat{L}_{A1}$  is then more justified. This is also valid for the matrix-variate  $\mathcal{G}^0$  and  $\mathcal{G}$  distribution. As we increase the

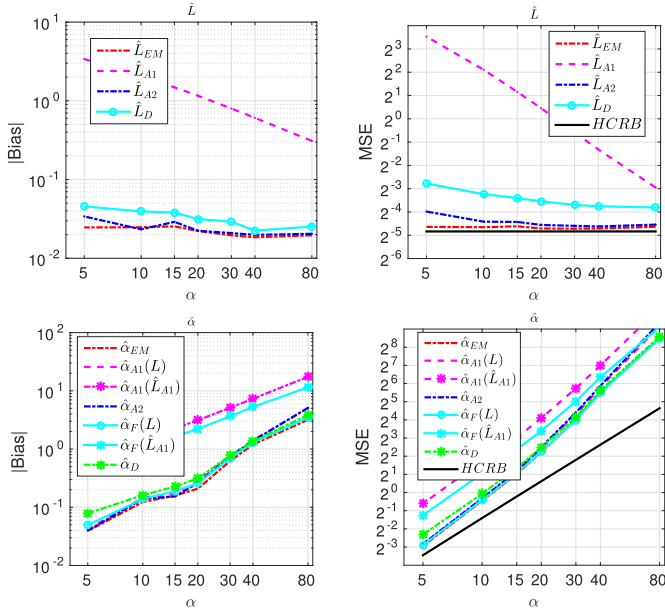


Fig. 3. (Left) Bias and (Right) mse of estimators for (Top) ENL  $L$  and (Bottom)  $\mathcal{K}_d$  distribution texture parameter  $\alpha$  as a function of shape parameter  $\alpha$ .

value of  $\alpha$ , the mses of the estimators, such as  $\hat{\alpha}_{EM}$ ,  $\hat{\alpha}_F(\cdot)$ ,  $\hat{\alpha}_{A1}(\cdot)$ ,  $\hat{\alpha}_{A2}$ , and  $\hat{\alpha}_D$ , increase. Thus, the mses move away from the HCRB and the performances deteriorate gradually.

2) *Matrix-Variate  $\mathcal{G}^0$  Distribution*: For this distribution, the MLC-based estimator, denoted by  $\hat{\lambda}_{A1}$ , is estimated by solving the following second-order MLC equation [12]:

$$\psi^{(1)}(\hat{\lambda}_{A1}) = \frac{\kappa_2(\mathbf{C}) - \psi_d^{(1)}(L)}{d^2}. \quad (43)$$

$\hat{\lambda}_{A1}(L)$  and  $\hat{\lambda}_{A1}(\hat{L}_{A1})$  are the estimators dependent on the true value and an estimate of  $L$ , respectively.

The second MLC-based estimator, denoted by  $\hat{\lambda}_{A2}$ , is obtained in the same way as done in the previous section for the matrix-variate  $\mathcal{K}_d$  distribution. However, the MLCs for the  $\mathcal{G}^0$  distribution are used, instead, as follows:

$$\kappa_1\{\mathbf{C}\} = \psi_d(\hat{L}_{A2}) + \ln|\Sigma| - d \ln \hat{L}_{A2} \quad (44)$$

$$+ d(\ln(\hat{\lambda}_{A2} - 1) - \psi(\hat{\lambda}_{A2}))$$

$$\kappa_2\{\mathbf{C}\} = \psi_d^{(1)}(\hat{L}_{A2}) + d^2 \psi^{(1)}(\hat{\lambda}_{A2}). \quad (45)$$

The MoMFM-based estimator, which is denoted by  $\hat{\lambda}_F$ , is the solution for the following equation [22]:

$$\frac{E\{M^\nu\}}{E^\nu\{M\}} = (\hat{\lambda}_F - 1)^\nu \frac{\Gamma(\hat{\lambda}_F - \nu)}{\Gamma(\hat{\lambda}_F)} \left(\frac{1}{Ld}\right)^\nu \frac{\Gamma(Ld + \nu)}{\Gamma(Ld)}. \quad (46)$$

The method of Douglgeris developed for the  $\mathcal{G}_d^0$  distribution to estimate  $\hat{\lambda}_D$  is given by [21]

$$\hat{\lambda}_D = \frac{2L \cdot \text{var}(M) + d(Ld - 1)}{L \cdot \text{var}(M) - d}. \quad (47)$$

Monte Carlo simulations for the estimation of  $\lambda$  are performed with the parameters  $\lambda = 10$  and  $L = 10$  of the matrix-variate  $\mathcal{G}_d^0$  distribution. The behavior of bias, variance, and mse shown in Fig. 2 is similar to those reported for the estimators of  $\alpha$ .

3) *Matrix-Variate  $\mathcal{G}$  Distribution*: The texture parameter estimation is more complicated, since there are two shape parameters to estimate instead of only one texture parameter. The first MLC-based estimator requires the second- and third-order MLC equations

$$\kappa_2\{\mathbf{C}\} = \psi_d^{(1)}(L) + d^2 \ln K_{\hat{\alpha}_{A1}}^{(2)}(\hat{w}_{A1}) \quad (48)$$

$$\kappa_3\{\mathbf{C}\} = \psi_d^{(2)}(L) + d^3 \ln K_{\hat{\alpha}_{A1}}^{(3)}(\hat{w}_{A1}). \quad (49)$$

The ENL  $L$  is supposed to be a known constant. In a similar way,  $\hat{\alpha}_{A1}(L)$  [resp.  $\hat{w}_{A1}(L)$ ] and  $\hat{\alpha}_{A1}(\hat{L}_{A1})$  [resp.  $\hat{w}_{A1}(\hat{L}_{A1})$ ] are the estimators dependent on the true and the estimate of  $L$ , respectively. The second MLC-based estimators, denoted by  $\hat{\alpha}_{A2}$  and  $\hat{w}_{A2}$ , are obtained by using the first-, second-, and third-order MLC equations. The ENL  $\hat{L}_{A2}$  needs to be estimated simultaneously with the texture parameters. The determinant of the covariance matrix  $\Sigma$  is normalized to 1. Then,  $\ln|\Sigma|$  is omitted from the first MLC equation. The first-, second-, and third-order MLC equations are given by

$$\kappa_1\{\mathbf{C}\} = \psi_d(\hat{L}_{A2}) - d \ln \frac{\hat{L}_{A2}}{\hat{\eta}} + d \ln K_{\hat{\alpha}_{A2}}^{(1)}(\hat{w}_{A2}) \quad (50)$$

$$\kappa_2\{\mathbf{C}\} = \psi_d^{(1)}(\hat{L}_{A2}) + d^2 \ln K_{\hat{\alpha}_{A2}}^{(2)}(\hat{w}_{A2}) \quad (51)$$

$$\kappa_3\{\mathbf{C}\} = \psi_d^{(2)}(\hat{L}_{A2}) + d^3 \ln K_{\hat{\alpha}_{A2}}^{(3)}(\hat{w}_{A2}). \quad (52)$$

After replacing  $\hat{\eta} = |\Sigma_{SCM}|^{1/d} K_{\hat{\alpha}}(\hat{w}) / K_{\hat{\alpha}+1}(\hat{w})$ , where  $\Sigma_{SCM} = E\{\mathbf{C}\}$ , in the first MLC equation, the parameters  $\hat{L}_{A2}$ ,  $\hat{\alpha}_{A2}$ , and  $\hat{w}_{A2}$  can be estimated. The disadvantage of the first and second MLC-based estimators is that the estimation procedure requires the second- and third-order MLC equations, which involves higher order derivatives of a modified Bessel function of the second kind with respect to its order. These derivatives are only computable using numerical differentiation.

The equation for the MoMFM-based estimator for the matrix-variate  $\mathcal{G}_d$  distribution, denoted by  $\hat{\alpha}_F$  and  $\hat{w}_F$ , is given by

$$\frac{E\{M^\nu\}}{E^\nu\{M\}} = \frac{K_{\hat{\alpha}_F + \nu}(\hat{w}_F) K_{\hat{\alpha}_F}^{\nu-1}(\hat{w}_F)}{K_{\hat{\alpha}_F + 1}^\nu(\hat{w}_F)} \frac{\Gamma(Ld + \nu)}{(Ld)^\nu \Gamma(Ld)}. \quad (53)$$

We note here that similar to the previous estimator, (53) depends on  $L$ , and then,  $\hat{\alpha}_F(L)$  [resp.  $\hat{w}_F(L)$ ] and  $\hat{\alpha}_F(\hat{L}_{A1})$  [resp.  $\hat{w}_F(\hat{L}_{A1})$ ] are possible estimators. Since two parameters have to be estimated, two equations derived from (53), for  $\nu = 1/4$  and  $1/8$ , are simultaneously solved. In fact, these two values are taken by Khan and Guida [22].

Fig. 4 shows the value of bias, variance, and mse of all the estimators after 3000 Monte Carlo simulations with  $\alpha = 10$ ,  $w = 10$ ,  $\eta = 1$ , and  $L = 10$ . The best estimator for  $L$  is the EM estimator, followed by the  $\hat{L}_D$  estimator. Although the estimator  $\hat{L}_{A2}$  has a good performance when we deal with only one texture parameter such as the matrix-variate  $\mathcal{K}$  and  $\mathcal{G}^0$  distributions, it has registered the worst performance when we deal with the  $\mathcal{G}$  distribution with two shape parameters. This is due to numerical instabilities during the computation of higher order (second- and third-order) derivative of the modified Bessel function of the second kind.

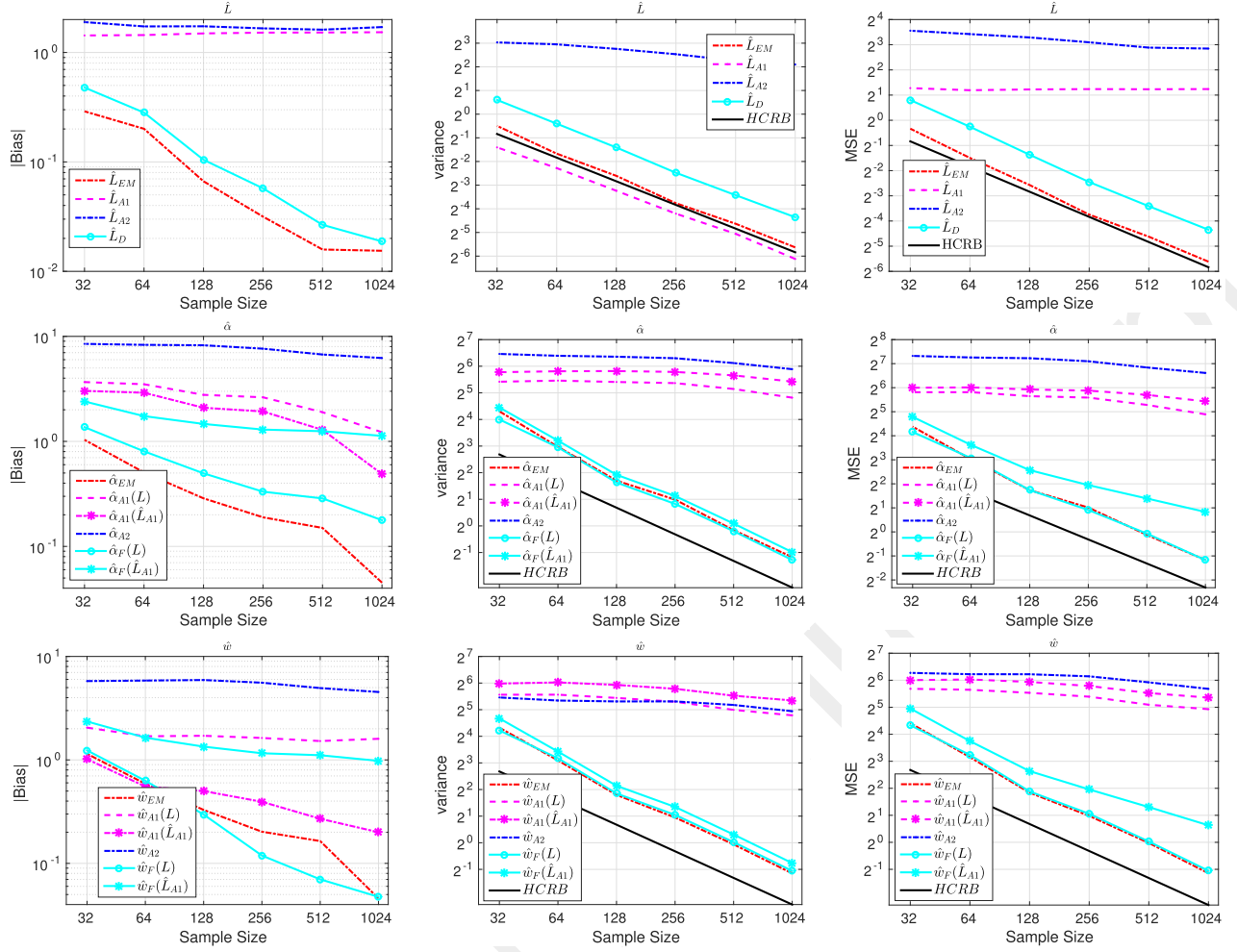


Fig. 4. (Left) Bias, (Middle) variance, and (Right) mse of estimators for (Top) ENL  $L$  and the  $\mathcal{G}_d$  distribution texture parameters (Middle row)  $\alpha$  and (Bottom)  $w$  as a function of sample size  $N$ .

The preferred estimators for texture parameters are  $\hat{\alpha}_{EM}$  and  $\hat{w}_{EM}$  due to their lowest bias, variance, and mse. The  $\hat{\alpha}_F(L)$  and  $\hat{w}_F(L)$  estimators have a very close performance to that given by the EM. We recall that they use the true value of  $L$ . The  $\hat{\alpha}_F(\hat{L}_{A1})$  and  $\hat{w}_F(\hat{L}_{A1})$  estimators are less performant than  $\hat{\alpha}_F(L)$  and  $\hat{w}_F(L)$  estimators. The worst performance is registered for the first and second MLC-based estimators. They are far from the performance of the EM and the MoMFM.

## V. APPLICATION TO REAL POLSAR DATA

To evaluate the performance of the EM estimator on real data and to prove the outperformance of the proposed estimator versus some known estimators, two multilook PolSAR data sets are used for this purpose. The former was acquired over San Francisco Bay by the NASA/Jet Propulsion Laboratory (JPL) Airborne Synthetic Aperture Radar (AIRSAR) airborne platform. The latter data set was acquired over Oberpfaffenhofen in Germany by the DLR's ESAR airborne platform. Three different areas with varying degrees of texture are selected from each PolSAR data set and indicated in the panes in Fig. 5. The selected areas are airside, forest, and urban for the Oberpfaffenhofen data set, and sea, vegetation, and urban

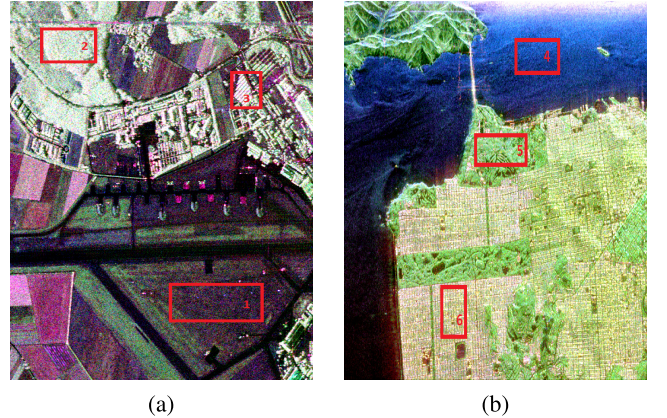


Fig. 5. (a) Color-coded image in Pauli basis in Oberpfaffenhofen. Area 1 indicates airside, area 2 indicates vegetation, and area 3 indicates urban. (b) Color-coded image in Pauli basis in San Francisco. Area 4 indicates sea, area 5 indicates vegetation, and area 6 indicates urban.

for the San Francisco Bay data set. The data have been 6-look and 12-looked processed images. The dimension of the multi-look polarimetric covariance matrix ( $d = 3$ ) makes it difficult to verify the pdf with a histogram from data. The variable  $\text{tr}(\Sigma^{-1}\mathbf{C})$  transforms the matrix-variate data into the scalar,

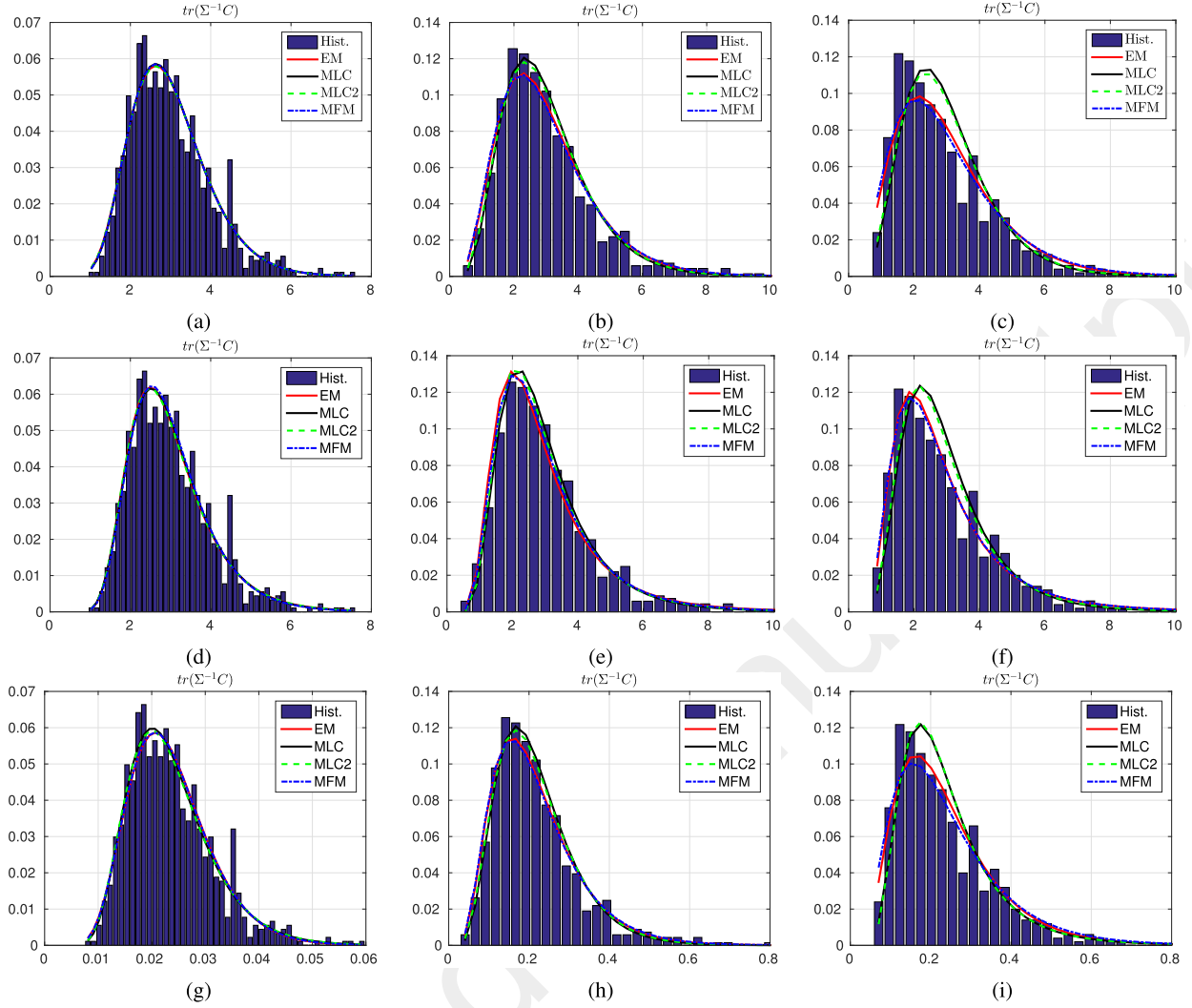


Fig. 6.  $\text{tr}(\Sigma^{-1}\mathbf{C})$  normalized histograms and the estimated pdfs when  $\mathbf{C}$  follows  $\mathcal{K}_d$  (first row),  $\mathcal{G}_d^0$  (second row), and  $\mathcal{G}_d$  (third row) distributions. Data are from the San Francisco data set. (a), (d), and (g) Area 4. (b), (e), and (h) Area 5. (c), (f), and (i) Area 6.

TABLE VI  
PDF OF  $Y = \text{tr}(\Sigma^{-1}\mathbf{C})$  FOR DIFFERENT TEXTURE DISTRIBUTIONS

$\tau$	$f_Y(y)$
$\bar{\gamma}(\alpha)$	$\frac{2(L\alpha)^{\frac{\alpha+dL}{2}}}{\Gamma(\alpha)\Gamma(dL)} y^{\frac{\alpha+dL}{2}-1} K_{\alpha-dL}(2\sqrt{L\alpha y})$
$\bar{\gamma}^{-1}(\lambda)$	$\frac{(\lambda-1)^{\lambda} L^{dL}}{\Gamma(\lambda)} \frac{\Gamma(dL+\lambda)}{\Gamma(dL)} y^{dL-1} (Ly + \lambda - 1)^{-(dL+\lambda)}$
$\mathcal{N}^{-1}(\alpha, w, \eta)$	$\frac{L^{dL} y^{dL-1}}{\eta^{\alpha} \Gamma(dL)} \frac{1}{K_{\alpha}(w)} \left( \frac{2Ly + w\eta}{w/\eta} \right)^{\frac{\alpha-dL}{2}} \times K_{\alpha-dL} \left( \sqrt{\frac{w}{\eta}} (2Ly + w\eta) \right)$

such that it can be used to fit the histogram of each area. The theoretical statistical model for  $\text{tr}(\Sigma^{-1}\mathbf{C}) = \tau \text{tr}(\Sigma^{-1}\mathbf{X})$  is derived using the distribution of the texture  $\tau$  and the distribution of  $\text{tr}(\Sigma^{-1}\mathbf{X})$ , which follows a gamma distribution with parameters  $\gamma(Ld, L)$ , and its expectation is given by  $E\{\text{tr}(\Sigma^{-1}\mathbf{X})\} = d$ . The pdfs of the scalar variable  $Y = \text{tr}(\Sigma^{-1}\mathbf{C})$  are given in Table VI for each texture pdf. The fit ability is evaluated qualitatively by visual comparison between the estimated pdfs and the normalized histograms of  $\text{tr}(\Sigma^{-1}\mathbf{C})$  and quantitatively using the Kolmogorov–Smirnov (KS) test.

The smaller value of KS indicates the better, and the hypothesized model fits with the empirical distribution. A small value of  $p$ -value of the test indicates strong incompatibilities of the data with the employed distribution hypothesis. Table VII lists the values of KS and the  $p$ -values of the test (in percentage) obtained for the three selected areas from the Oberpfaffenhofen data set when  $\mathbf{C}$  follows  $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ , and  $\mathcal{G}_d$  distributions. The parameters of these distributions are estimated using the proposed EM and other estimators, such as MFM, MLC, and MLC2. For the area 1, EM-based estimators  $\mathcal{G}_d^0$  and  $\mathcal{G}_d$  show the highest  $p$ -values of 81.02% and 84.43% followed by the MFM-based estimators at 59.79% and 80.98%, respectively. Contrariwise, the MFM-based estimator shows the best  $p$ -value of 73.21% for the  $\mathcal{K}_d$  distribution. The remaining estimators show the  $p$ -values close to 0. For the area 2, the highest  $p$ -values of 90.66% and 25.84% are obtained for  $\mathcal{K}_d$  and  $\mathcal{G}_d^0$  with EM-based estimators, respectively. For the  $\mathcal{G}_d$ , it is the MFM-based estimator that shows the highest  $p$ -value of 79.66% followed closely by a value of 78.43% for the EM-based estimator. Finally, for urban area 3, the highest  $p$ -values of 71.41%, 19.55%, and 86.77% are obtained with the EM-based

TABLE VII  
VALUES OF KS AND  $p$ -VALUES (%) OF THE OBERPFAFFENHOFEN  
DATA OBTAINED BY USING  $\text{tr}(\Sigma^{-1}\mathbf{C})$  WHEN  $\mathbf{C}$  FOLLOWS  
 $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ , AND  $\mathcal{G}_d$  DISTRIBUTIONS

$\mathcal{K}_d(\Sigma, L, \alpha)$						
Estimator	Area 1		Area 2		Area 3	
	KS	$p$ -value	KS	$p$ -value	KS	$p$ -value
EM	0.0151	69.41	<b>0.018</b>	<b>90.66</b>	<b>0.0359</b>	<b>71.41</b>
MFM	<b>0.0146</b>	<b>73.21</b>	0.0186	88.41	0.0367	68.60
MLC	0.0521	0	0.0357	16.59	0.1141	0
MLC2	0.059	0	0.0321	26.71	0.0836	1.06
$\mathcal{G}_d^0(\Sigma, L, \lambda)$						
EM	<b>0.0136</b>	<b>81.02</b>	<b>0.0324</b>	<b>25.84</b>	<b>0.0388</b>	<b>19.55</b>
MFM	0.0163	59.79	0.0350	18.27	0.0392	17.24
MLC	0.0516	0	0.0415	7.02	0.0876	0
MLC2	0.0591	0	0.0423	6.14	0.0563	1.53
$\mathcal{G}_d(\Sigma, L, \alpha, w, \eta)$						
EM	<b>0.0131</b>	<b>84.43</b>	0.0209	78.43	<b>0.0306</b>	<b>86.77</b>
MFM	0.0136	80.98	<b>0.0206</b>	<b>79.66</b>	0.0408	55.60
MLC	0.0579	0	0.0269	48.09	0.1059	0
MLC2	0.0619	0	0.0250	57.29	0.0586	15.12

TABLE VIII

VALUES OF KS AND  $p$ -VALUES (%) OF THE SAN FRANCISCO DATA  
OBTAINED BY USING  $\text{tr}(\Sigma^{-1}\mathbf{C})$  WHEN  $\mathbf{C}$  FOLLOWS  $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ , AND  $\mathcal{G}_d$   
DISTRIBUTIONS

$\mathcal{K}_d(\Sigma, L, \alpha)$						
Estimator	Area 4		Area 5		Area 6	
	KS	$p$ -value	KS	$p$ -value	KS	$p$ -value
EM	<b>0.0256</b>	<b>58.56</b>	0.0362	32.32	0.0604	4.94
MFM	0.0270	51.84	<b>0.0345</b>	<b>38.17</b>	<b>0.0533</b>	<b>11.24</b>
MLC	0.0275	49.31	0.0463	10.34	0.1075	0
MLC2	0.0262	55.60	0.0443	13.21	0.0998	0
$\mathcal{G}_d^0(\Sigma, L, \lambda)$						
EM	<b>0.0199</b>	<b>85.97</b>	<b>0.0256</b>	<b>75.07</b>	<b>0.0258</b>	<b>88.34</b>
MFM	0.0208	82.19	0.0269	69.57	0.0309	71.19
MLC	0.0217	78.20	0.0391	23.82	0.1025	0
MLC2	0.0214	79.34	0.0289	60.78	0.0902	0
$\mathcal{G}_d(\Sigma, L, \alpha, w, \eta)$						
EM	<b>0.0149</b>	<b>98.63</b>	<b>0.0284</b>	<b>62.74</b>	<b>0.0436</b>	<b>28.96</b>
MFM	0.0184	91.36	0.0300	55.61	0.0443	27.04
MLC	0.0263	55.16	0.0441	13.47	0.0880	0
MLC2	0.0258	57.47	0.0397	22.37	0.0902	0

estimators for  $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ , and  $\mathcal{G}_d$ . Table VIII lists the values of KS and the  $p$ -values of the test (in percentage) obtained for the three selected areas from the San Francisco data set when  $\mathbf{C}$  follows the polarimetric  $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ , and  $\mathcal{G}_d$  distributions, respectively. For area 4 which corresponds to the sea, the EM-based estimators  $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ , and  $\mathcal{G}_d$  give the highest  $p$ -values of 58.56%, 85.97%, and 98.63%, respectively, whereas among the remaining estimators, the MFM-based estimators show the higher  $p$ -values of 51.84%, 82.19%, and 91.36% than the MLC and MLC2-based estimators. For the vegetation area,  $\mathcal{G}_d^0$  shows the best fitting for all estimators with the  $p$ -value of 75.07%. Again, the EM-based estimators  $\mathcal{G}_d^0$  and  $\mathcal{G}_d$  show the highest  $p$ -values of 75.07% and 62.74%, respectively, except for the  $\mathcal{K}_d$  distribution where the MFM-based estimator shows the best  $p$ -value of 38.17%. Finally, the MLC and MLC2-based estimators show very low  $p$ -values for urban area and, then, the worst fitting for all distributions. Fig. 6 shows the comparison between the normalized histograms of  $\text{tr}(\Sigma^{-1}\mathbf{C})$  and the estimated pdfs for the selected three areas in San Francisco Bay. It is clear that the estimated pdfs curves fit well the normalized histograms, especially when it comes to

the EM-based estimator which is consistent with the results of Table VIII.

According to Tables VII and VIII, the  $p$ -values and the KS test for the EM-based estimators are generally higher than those of the other estimators. The  $p$ -values for the MFM-based estimators are close to the EM-based estimators. It is clear to conclude that the estimated parameters by the EM method give a better estimation performance and higher precision by acquiring higher  $p$ -values than the other estimators.

## VI. CONCLUSION

In this paper, the ML estimates of some important distribution parameters, such as the matrix-variate  $\mathcal{K}_d$ ,  $\mathcal{G}_d^0$ , and  $\mathcal{G}_d$  distributions, are analytically derived based on the EM algorithm. The simulation results reveal the outperformance of the proposed EM-based algorithm versus some known estimators. The former leads to the lowest mse of the parameter estimator. Furthermore, it gives a better estimation of the parameters for small and moderate numbers of samples. The results indicate that the accuracy of the estimation method improves when the sample size is increased. The EM estimates of these distribution parameters are asymptotically unbiased and efficient. The proposed estimation method has a good performance and more easily implementable. Results on the real multilook polarimetric data show that the ML estimator outperforms the MFM and MLC polarimetric estimators.

## APPENDIX A

### ML ESTIMATOR OF COVARIANCE MATRIX OF SPECKLE

For  $N$  i.i.d. data, the likelihood function is given by

$$\begin{aligned}
 L_{\mathbf{C}}(\mathbf{C}_1, \dots, \mathbf{C}_N, \Sigma, \tau_1, \dots, \tau_N) & \\
 &= \prod_{i=1}^N \frac{L^{Ld} |\mathbf{C}_i|^{L-d}}{\Gamma_d(L) |\Sigma|^L} \int_0^\infty \frac{1}{\tau_i^{dL}} \text{etr} \left( -\frac{L}{\tau_i} \Sigma^{-1} \mathbf{C}_i \right) f_{\tau_i}(\tau_i) d\tau_i \\
 &= \left( \frac{L^{Ld}}{\Gamma_d(L) |\Sigma|^L} \right)^N \prod_{i=1}^N |\mathbf{C}_i|^{L-d} h_{dL}(L \text{tr}(\Sigma^{-1} \mathbf{C}_i)). \quad (55)
 \end{aligned}$$

The ML estimator of  $\Sigma$  is obtained by maximizing the likelihood  $L_{\mathbf{C}}(\cdot)$  with respect to  $\Sigma$ . Differentiating the log-likelihood with respect to the latter and setting the result to zero yield

$$-N_L \frac{\partial \ln |\Sigma|}{\partial \Sigma} + \sum_{i=1}^N \frac{g_{dL}(L \text{tr}(\Sigma^{-1} \mathbf{C}_i))}{h_{dL}(L \text{tr}(\Sigma^{-1} \mathbf{C}_i))} \frac{\partial L \text{tr}(\Sigma^{-1} \mathbf{C}_i)}{\partial \Sigma} = 0 \quad (56)$$

where  $g_{dL}(x) = \partial h_{dL}(x) / \partial x$ . Knowing that  $\partial \ln |\Sigma| / \partial \Sigma = (\Sigma^{-1})^T$  and  $\partial \text{tr}(\Sigma^{-1} \mathbf{W}) / \partial \Sigma = -(\Sigma^{-1} \mathbf{W} \Sigma^{-1})^T$ , and observing that  $g_{dL}(x) = -h_{dL+1}(x)$ , we reach the expression of the ML estimator of the covariance

$$\hat{\Sigma} = \frac{1}{N} \sum_{i=1}^N \frac{h_{dL+1}(L \text{tr}(\hat{\Sigma}^{-1} \mathbf{C}_i))}{h_{dL}(L \text{tr}(\hat{\Sigma}^{-1} \mathbf{C}_i))} \mathbf{C}_i. \quad (57)$$

The function  $h_{dL}(L\text{tr}(\Sigma^{-1}\mathbf{C}_i))$  is developed as follows:

$$\begin{aligned}
& h_{dL}(L\text{tr}(\Sigma^{-1}\mathbf{C}_i)) \\
&= \int_0^{+\infty} \tau_i^{-dL} \exp\left(-\frac{L\text{tr}(\Sigma^{-1}\mathbf{C}_i)}{\tau_i}\right) f_\tau(\tau_i) d\tau_i \\
&= \frac{|\Sigma|^L \Gamma_d(L)}{L^{Ld} |\mathbf{C}_i|^{L-d}} \int_0^{+\infty} f_{\mathbf{C}|\tau}(\mathbf{C}_i|\tau_i) f_\tau(\tau_i) d\tau_i \\
&= \frac{|\Sigma|^L \Gamma_d(L)}{L^{Ld} |\mathbf{C}_i|^{L-d}} \int_0^{+\infty} f_{\mathbf{C},\tau}(\mathbf{C}_i, \tau_i) d\tau_i \\
&= \frac{|\Sigma|^L \Gamma_d(L)}{L^{Ld} |\mathbf{C}_i|^{L-d}} f_{\mathbf{C}}(\mathbf{C}_i) \tag{58}
\end{aligned}$$

and also the expression of  $h_{dL+1}(L\text{tr}(\Sigma^{-1}\mathbf{C}_i))$

$$\begin{aligned}
& h_{dL+1}(L\text{tr}(\Sigma^{-1}\mathbf{C}_i)) \\
&= \frac{|\Sigma|^L \Gamma_d(L)}{L^{Ld} |\mathbf{C}_i|^{L-d}} \int_0^{+\infty} \frac{1}{\tau_i} f_{\mathbf{C}|\tau}(\mathbf{C}_i|\tau_i) f_\tau(\tau_i) d\tau_i \\
&= \frac{|\Sigma|^L \Gamma_d(L)}{L^{Ld} |\mathbf{C}_i|^{L-d}} \int_0^{+\infty} \frac{1}{\tau_i} f_{\tau|\mathbf{C}}(\tau_i|\mathbf{C}_i) f_{\mathbf{C}}(\mathbf{C}_i) d\tau_i \\
&= \frac{|\Sigma|^L \Gamma_d(L)}{L^{Ld} |\mathbf{C}_i|^{L-d}} f_{\mathbf{C}}(\mathbf{C}_i) E_{\tau|\mathbf{C}} \left\{ \frac{1}{\tau} \middle| \mathbf{C}_i \right\}. \tag{59}
\end{aligned}$$

Dividing the expression of  $h_{dL+1}(L\text{tr}(\Sigma^{-1}\mathbf{C}_i))$  by  $h_{dL}(L\text{tr}(\Sigma^{-1}\mathbf{C}_i))$ , we get the result (14).

#### APPENDIX B

Let  $\mathbf{Y} = (1/\tau)\Sigma^{-1/2}\mathbf{C}\Sigma^{-1/2} = \Sigma^{-1/2}\mathbf{X}\Sigma^{-1/2}$ . The Jacobian determinant of  $\mathbf{X}$  written as  $\mathbf{Y}$  is given as follows [27]:  $|J_{\mathbf{X}\rightarrow\mathbf{Y}}| = |\Sigma|^d$ . The pdf of  $\mathbf{Y}$  is then given as

$$f_{\mathbf{Y}}(\mathbf{Y}) = f_{\mathbf{X}}(\Sigma^{1/2}\mathbf{Y}\Sigma^{1/2})|J_{\mathbf{X}\rightarrow\mathbf{Y}}|. \tag{60}$$

$\mathbf{X}$  follows a scaled complex Wishart distribution with a pdf given by (5). Equation (60) can be written as follows:

$$\begin{aligned}
f_{\mathbf{Y}}(\mathbf{Y}) &= \frac{L^{Ld} |\Sigma^{1/2}\mathbf{Y}\Sigma^{1/2}|^{L-d}}{\Gamma_d(L) |\Sigma|^{L-d}} \text{etr}(-L\Sigma^{-1}\Sigma^{1/2}\mathbf{Y}\Sigma^{1/2}) \\
&= \frac{L^{Ld} |\mathbf{Y}|^{L-d}}{\Gamma_d(L)} \text{etr}(-L\mathbf{Y}). \tag{61}
\end{aligned}$$

It is clear that  $E\{\mathbf{Y}\} = \mathbf{I}_d$ . Therefore,  $\text{tr}(E\{\mathbf{Y}\}) = E\{\text{tr}(\mathbf{Y})\} = d$ . The Mellin kind characteristic function of the complex random matrix  $\mathbf{Y}$  is

$$\phi_{\mathbf{Y}}(s) = E\{|\mathbf{Y}|^{s-d}\} = \int_{\Omega_+} |\mathbf{Y}|^{s-d} f_{\mathbf{Y}}(\mathbf{Y}) d\mathbf{Y} \tag{62}$$

$$= \frac{\Gamma_d(L+s-d)}{\Gamma_d(L)} L^{-d(s-d)}. \tag{63}$$

The  $\nu$ th-order Mellin kind matrix moment of  $\mathbf{Y}$  is defined as

$$\mu_\nu(\mathbf{Y}) = \frac{d^\nu}{ds^\nu} \phi_{\mathbf{Y}}(s) \Big|_{s=d} = E\{(\ln |\mathbf{Y}|)^\nu\}. \tag{64}$$

$$\begin{aligned}
\text{For } \nu = 1, \quad \frac{d}{ds} \phi_{\mathbf{Y}}(s) \Big|_{s=d} &= \psi_d(L) - d \ln L = E\{\ln |\mathbf{Y}|\}. \\
& \tag{65}
\end{aligned}$$

#### APPENDIX C

##### GENERALIZED INVERSE GAUSSIAN DISTRIBUTION

The pdf of the GIG distribution is given by

$$f(z; \alpha_1, w_1, \eta_1) = \frac{1}{\eta_1^{\alpha_1} 2K_{\alpha_1}(w_1)} z^{\alpha_1-1} \exp\left(-\frac{w_1}{2} \left(\frac{\eta_1}{z} + \frac{z}{\eta_1}\right)\right). \tag{66}$$

The following moments exist and are finite:

$$E\{Z^k\} = \eta_1^k \frac{K_{\alpha_1+k}(w_1)}{K_{\alpha_1}(w_1)} \tag{67}$$

$$E\{\ln Z\} = \ln \eta_1 + \frac{\frac{\partial}{\partial a} K_{\alpha_1+a}(w_1) \Big|_{a=0}}{K_{\alpha_1}(w_1)}. \tag{68}$$

#### APPENDIX D

##### INVERSE GAMMA DISTRIBUTION

The pdf of the inverse gamma distribution is given by

$$f(z; \alpha_1, \beta_1) = \frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)} \frac{1}{z^{1+\alpha_1}} \exp\left(\frac{-\beta_1}{z}\right). \tag{69}$$

The following moments exist and are finite:

$$E\{Z\} = \frac{\beta_1}{\alpha_1 - 1} \text{ for } \alpha_1 > 1, \quad E\{Z^{-1}\} = \frac{\alpha_1}{\beta_1} \tag{70}$$

$$E\{\ln Z\} = -\psi(\alpha_1) + \ln \beta_1. \tag{71}$$

#### APPENDIX E

##### DERIVATION OF INFORMATION MATRIX OF HCRB

The log-likelihood function can be expanded to

$$\begin{aligned}
& \ln f_{\mathbf{C}|\tau}(\mathbf{C}|\tau, \Theta) \\
&= \ln \prod_{i=1}^N f_{\mathbf{C}|\tau}(\mathbf{C}_i|\tau_i, \Theta) \\
&= \sum_{i=1}^N Ld \ln L + (L-d) \ln |\mathbf{C}_i| - \frac{L}{\tau_i} \text{tr}(\Sigma^{-1}\mathbf{C}_i) \\
&\quad - \ln \Gamma_d(L) - L \ln |\Sigma| - dL \ln \tau_i. \tag{72}
\end{aligned}$$

The first and the second partial derivative with respect to  $L$  as follows:

$$\begin{aligned}
& \frac{\partial}{\partial L} \ln f_{\mathbf{C}|\tau}(\mathbf{C}|\tau, \Theta) \\
&= \sum_{i=1}^N d \ln L + d + \ln |\mathbf{C}_i| - \frac{1}{\tau_i} \text{tr}(\Sigma^{-1}\mathbf{C}_i) \\
&\quad - \psi_d(L) - \ln |\Sigma| - d \ln \tau_i \tag{73}
\end{aligned}$$

$$\begin{aligned}
& \frac{\partial^2}{\partial L^2} \ln f_{\mathbf{C}|\tau}(\mathbf{C}|\tau, \Theta) \\
&= N \left( \frac{d}{L} - \psi_d^{(1)}(L) \right). \tag{74}
\end{aligned}$$

The partial derivative with respect to  $\Sigma$  by using a complex matrix differentiation [30] is given by

$$\frac{\partial}{\partial \Sigma} \ln f_{\mathbf{C}|\tau}(\mathbf{C}|\tau, \Theta) = -NL\Sigma^{-1} + \sum_{i=1}^N \frac{L}{\tau_i} \Sigma^{-1}\mathbf{C}_i\Sigma^{-1}. \tag{75}$$

To calculate the second partial derivative, the following relations [30] are used:

$$\frac{\partial(\mathbf{AB})}{\partial\boldsymbol{\Sigma}} = \frac{\partial\mathbf{A}}{\partial\boldsymbol{\Sigma}}(\mathbf{I}_d \otimes \mathbf{B}) + (\mathbf{I}_d \otimes \mathbf{A})\frac{\partial\mathbf{B}}{\partial\boldsymbol{\Sigma}} \quad (76)$$

$$\frac{\partial\boldsymbol{\Sigma}^{-1}}{\partial\boldsymbol{\Sigma}} = -\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1} \quad (77)$$

where  $\mathbf{A}$  and  $\mathbf{B}$  are two arbitrary complex matrices with the same dimensions. The second partial derivative is then given by (see [13] for more calculation details)

$$\begin{aligned} \frac{\partial^2}{\partial\boldsymbol{\Sigma}^2} \ln f_{\mathbf{C}|\tau}(\mathbf{C}|\tau, \Theta) &= N\mathbf{L}\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1} \\ &- \sum_{i=1}^N \frac{L}{\tau_i} [(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1} \mathbf{C}_i \boldsymbol{\Sigma}^{-1}) + (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \mathbf{C}_i)], \end{aligned} \quad (78)$$

Finally,

$$\begin{aligned} \frac{\partial}{\partial L} \left( \frac{\partial \ln f_{\mathbf{C}|\tau}(\mathbf{C}|\tau, \Theta)}{\partial\boldsymbol{\Sigma}} \right) &= \frac{\partial}{\partial\boldsymbol{\Sigma}} \left( \frac{\partial \ln f_{\mathbf{C}|\tau}(\mathbf{C}|\tau, \Theta)}{\partial L} \right) \\ &= -N\boldsymbol{\Sigma}^{-1} + \sum_{i=1}^N \frac{1}{\tau_i} (\boldsymbol{\Sigma}^{-1} \mathbf{C}_i \boldsymbol{\Sigma}^{-1}). \end{aligned} \quad (79)$$

Using the relation  $\partial/\partial\text{vec}(\mathbf{A}) = \text{vec}(\partial/\partial\mathbf{A})$  [30],  $\mathbf{J}(\Theta_{\boldsymbol{\Sigma}})$  can be expressed as

$$\begin{aligned} [\mathbf{J}(\Theta_{\boldsymbol{\Sigma}})]_{11} &= -E_{\mathbf{C},\tau} \left\{ \frac{\partial^2}{\partial L^2} \ln f_{\mathbf{C}|\tau}(\mathbf{C}|\tau, \Theta) \right\} \\ &= N \left( -\frac{d}{L} + \psi_d^{(1)}(L) \right) \end{aligned} \quad (80)$$

$$\begin{aligned} [\mathbf{J}(\Theta_{\boldsymbol{\Sigma}})]_{22} &= -E_{\mathbf{C},\tau} \left\{ \frac{\partial}{\partial\text{vec}(\boldsymbol{\Sigma})^*} \left( \frac{\partial \ln f_{\mathbf{C}|\tau}(\mathbf{C}|\tau, \Theta)}{\partial\text{vec}(\boldsymbol{\Sigma})} \right)^T \right\} \\ &= N\mathbf{L}(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \end{aligned} \quad (81)$$

$$[\mathbf{J}(\Theta_{\boldsymbol{\Sigma}})]_{12} = -E_{\mathbf{C},\tau} \left\{ \frac{\partial}{\partial\text{vec}(\boldsymbol{\Sigma})^*} \left( \frac{\partial \ln f_{\mathbf{C}|\tau}(\mathbf{C}|\tau, \Theta)}{\partial L} \right) \right\} = 0 \quad (82)$$

$$[\mathbf{J}(\Theta_{\boldsymbol{\Sigma}})]_{21} = -E_{\mathbf{C},\tau} \left\{ \frac{\partial}{\partial L} \left( \frac{\partial \ln f_{\mathbf{C}|\tau}(\mathbf{C}|\tau, \Theta)}{\partial\text{vec}(\boldsymbol{\Sigma})} \right)^T \right\} = 0. \quad (83)$$

#### ACKNOWLEDGMENT

The authors would like to thank J. Idier from the Laboratoire des Sciences du Numérique de Nantes (LS2N) for the constructive comments. They would also like to thank the anonymous reviewers for their suggestions that have helped improve this paper.

#### REFERENCES

[1] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*, 2nd ed. Raleigh, NC, USA: SciTech Publishing, 2004.

[2] G. Cumming and F. H. Wong, *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation*, Norwood, MA, USA: Artech House, 2005.

[3] S. H. Yueh, J. A. Kong, J. K. Jao, R. T. Shin, and L. M. Novak, "K-distribution and polarimetric terrain radar clutter," *J. Electromagn. Waves Appl.*, vol. 3, no. 8, pp. 747–768, Jan. 1989.

[4] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution: An introduction," *Ann. Math. Statist.*, vol. 34, no. 81, pp. 152–177, Mar. 1963.

[5] G. Gao, "Statistical modeling of SAR images: A survey," *Sensors*, vol. 10, no. 1, pp. 775–795, 2010.

[6] J. S. Lee, D. L. Schuler, R. H. Lang, and K. J. Ranson, "K-distribution for multi-look processed polarimetric SAR imagery," in *Proc. Int. Geosci. Remote Sens. Symp. (IGARSS), Surf. Atmos. Remote Sens., Technol., Data Anal. Interpretation*, vol. 4, Aug. 1994, pp. 2179–2181.

[7] C. C. Freitas, A. C. Frery, and A. H. Correia, "The polarimetric  $\mathcal{G}$  distribution for SAR data analysis," *Environmetrics*, vol. 16, no. 1, pp. 13–31, 2005.

[8] L. Bombrun and J. M. Beaulieu, "Fisher distribution for texture modeling of polarimetric SAR data," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 3, pp. 512–516, Jul. 2008.

[9] L. Bombrun, S. N. Anfinsen, and O. Harant, "A complete coverage of log-cumulant space in terms of distributions for polarimetric SAR data," in *Proc. PolInSAR*, Frascati, Italy, 2011, p. 136.

[10] S. Khan and R. Guida, "Application of Mellin-kind statistics to polarimetric  $\mathcal{G}$  distribution for SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3513–3528, Jun. 2014.

[11] A. C. Frery, A. H. Correia, and C. D. C. Freitas, "Classifying multifrequency fully polarimetric imagery with multiple sources of statistical evidence and contextual information," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3098–3109, Oct. 2007.

[12] S. N. Anfinsen and T. Eltoft, "Application of the matrix-variate Mellin transform to analysis of polarimetric radar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2281–2295, Jun. 2011.

[13] S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, "Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3795–3809, Nov. 2009.

[14] L. Tao, C. Hao-Gui, X. Ze-Min, and G. Jun, "Texture-invariant estimation of equivalent number of looks based on trace moments in polarimetric radar imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 6, pp. 1129–1133, Jun. 2014.

[15] D. Tao, S. N. Anfinsen, and C. Brekke, "A comparative study of sea clutter covariance matrix estimators," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 5, pp. 1010–1014, May 2014.

[16] D. E. Tyler, "A distribution-free M-estimator of multivariate scatter," *Ann. Statist.*, vol. 15, no. 1, pp. 234–251, 1987.

[17] F. Gini and M. Greco, "Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter," *Signal Process.*, vol. 82, no. 12, pp. 1847–1859, Dec. 2002.

[18] G. Vasile, J. P. Ovarlez, F. Pascal, and C. Tison, "Coherency matrix estimation of heterogeneous clutter in high-resolution polarimetric SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 1809–1826, Apr. 2010.

[19] G. Vasile, F. Pascal, J. P. Ovarlez, P. Formont, and M. Gay, "Optimal parameter estimation in heterogeneous clutter for high-resolution polarimetric SAR data," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 6, pp. 1046–1050, Nov. 2011.

[20] M. Liu, H. Zhang, C. Wang, and F. Wu, "Change detection of multilook polarimetric SAR images using heterogeneous clutter models," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 12, pp. 7483–7494, Dec. 2014.

[21] A. P. Doulgeris, S. N. Anfinsen, and T. Eltoft, "Classification with a non-Gaussian model for PolSAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2999–3009, Oct. 2008.

[22] S. Khan and R. Guida, "On fractional moments of multilook polarimetric whitening filter for polarimetric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3502–3512, Jun. 2014.

[23] A. Lopès and F. Séry, "Optimal speckle reduction for the product model in multilook polarimetric SAR imagery and the Wishart distribution," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 3, pp. 632–647, May 1997.

- [24] A. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc., B (Methodol.)*, vol. 39, no. 1, pp. 1–38, 1977.
- [25] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton, FL, USA: CRC Press, 2009.
- [26] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)," *Ann. Math. Statist.*, vol. 34, no. 1, pp. 152–177, 1963.
- [27] A. M. Mathai, *Jacobians of Matrix Transformations and Functions of Matrix Argument*. Singapore: World Scientific, 1997.
- [28] A. P. Doulgeris, S. N. Anfinsen, and T. Eltoft, "Automated non-Gaussian clustering of polarimetric synthetic aperture radar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3665–3676, Oct. 2011.
- [29] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. New York, NY, USA: Academic, 2007.
- [30] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2740–2746, Jun. 2007.
- [31] Y. X. Yuan, "A review of trust region algorithms for optimization," in *Proc. 4th Int. Congr. Ind. Appl. Math. (ICIAM)*, Edinburgh, U.K., vol. 99, 1999, pp. 271–282.
- [32] F. Gini and R. Reggiannini, "On the use of Cramer-Rao-like bounds in the presence of random nuisance parameters," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2120–2126, Dec. 2000.
- [33] J. Wang, A. Dogandzic, and A. Nehorai, "Maximum likelihood estimation of compound-Gaussian clutter and target parameters," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3884–3898, Oct. 2006.
- [34] F. Gini, "A radar application of a modified Cramer-Rao bound: Parameter estimation in non-Gaussian clutter," *IEEE Trans. Signal Process.*, vol. 46, no. 7, pp. 1945–1953, Jul. 1998.
- [35] R. Miller and C. Chang, "A modified Cramér-Rao bound and its applications (Corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 398–400, May 1978.
- [36] A. Van Den Bos, "A Cramer-Rao lower bound for complex parameters," *IEEE Trans. Signal Process.*, vol. 42, no. 10, p. 2859, Oct. 1994.