# Automatic PolSAR Segmentation with the $\mathcal{U}$-distribution and Markov Random Fields.

Anthony P. Doulgeris, Vahid Akbari, and Torbjørn Eltoft
Department of Physics and Technology,
University of Tromsø, 9037 Tromsø, Norway
e-mail: anthony.p.doulgeris@uit.no, Tel: +47 776 45177

## Abstract

A novel unsupervised, non-Gaussian and contextual clustering algorithm is demonstrated for segmentation of Polarimetric SAR images. Previous works have shown the added value of both non-Gaussian modelling and contextual smoothing individually, and goodness-of-fit techniques were introduced to determine the appropriate number of statistically distinct classes. This paper extends our previous work by using the more flexible, two parameter, $\mathcal{U}$-distribution model and includes a Markov Random Field approach for contextual smoothing, without losing the benefits of the goodness-of-fit testing. The proposed, fully automatic, algorithm is demonstrated with both simulated and real data-sets.

## 1 Introduction

We develop and demonstrate an improved automatic clustering algorithm that combines a more flexible non-Gaussian class model, a Markov random field (MRF) for contextual smoothing, and goodness-of-fit testing to optimise the segmentation and determine an appropriate number of classes.

Satellite-borne Polarimetric Synthetic Aperture Radar (PolSAR) systems have many benefits, but analysis is hindered by complicated non-Gaussian statistical methods. PolSAR data models are generally derived from the product model [1], which states that the backscattered signal results from the product between a Gaussian speckle noise component and the textured terrain backscatter.

The scaled Wishart distribution, $\mathcal{W}_d$, is the simplest multi-looked PolSAR model to analyse but contains no texture parameter, describing purely Gaussian speckle. The $\mathcal{K}_d$ (or K-Wishart) distribution [2, 3] and the $\mathcal{G}_d^0$-distribution [4, 5] are more flexible, with one texture parameter, and successful examples for many PolSAR scenes. The two parameter Kummer-U distribution has been used to model PolSAR vector data [6], with promising contiguous segmentation results and demonstrated that the two parameter model is more flexible to fit real data classes. The multivariate extension of the Kummer-U distribution for multi-looked complex (MLC) matrix data, hereafter simply called the $\mathcal{U}_d$-distribution, has not previously been demonstrated, but is expected to yield improved results because of its flexibility to model more varied textures and because it includes the $\mathcal{W}_d$, $\mathcal{K}_d$ and $\mathcal{G}_d^0$ models as asymptotic cases.

As with many of these product models, the probability density functions (PDFs) are complicated and maximum likelihood parameter estimators are not usually available with closed-form solutions. A practical solution is to estimate the model parameters with the method of matrix log-cumulants [7], because they have relatively simple numerical expressions and possess lower bias and variance compared to single channel (marginal) estimates or moment methods for product based distributions. The method of matrix log-cumulants shall be used for parameter estimation within the expectation maximisation algorithm (EM-algorithm) as well as for the goodness-of-fit testing stage that performs the split-and-merge operations to arrive at an appropriate number of classes [8, 9].

Contextual smoothing is desired to improve the accuracy and robustness of the image segmentation. It is achieved in the clustering algorithm with an MRF approach that integrates the $\mathcal{U}_d$-distribution for the PolSAR data statistics conditioned to each image cluster and a Potts model for the spatial context. The parameters of the MRF model are estimated with a mean-field like method [10]. The inclusion of the MRF is not expected to compromise the goodness-of-fit testing stage, because the MRF only affects the local priors and the underlying model remains a mixture of $\mathcal{U}_d$-distributions.

The proposed algorithm combines all the benefits of a flexible, two-parameter, non-Gaussian model for the covariance matrix data classes, an MRF for contextual smoothing, and goodness-of-fit testing to optimise the segmentation and determine an appropriate number of classes.

This state-of-the-art algorithm is described in **Section 2** and shown to give excellent results for both simulated and real data-sets in **Section 3**. Its main drawback seems to be computation time, but this can be partly alleviated with a sub-sampling approach, as in [9], that still finds the major classes of interest but sacrifices smaller sub/side classes for reduced computation time.

## 2 Method

The scope of this algorithm is to analyse multi-look MLC data images, where the data is an image of covariance matrices, $\mathbf{C}$. We assume the scalar product model is valid and that the MLC data is formed by a simple box-car multi-look average from the single-look complex scattering coefficients such that we can assume a global number of looks and "simpler" statistical models. The number of looks, $L$ is in practise substituted with an effective number of looks (ENL) due to pixel correlations. The ENL is optimised during the iterations by a minimum distance method using the log-cumulant expressions given all the current class model parameters simultaneously.

Our main objective is to segment the image pixels into separate clusters based upon the $\mathcal{U}_d$-distribution model. The statistical approach for clustering the images uses the iterative expectation maximisation algorithm with a few modifications, as has been described in detail in [8, 9]. The extension proposed here, is that each class is modelled with the $\mathcal{U}_d$-distribution PDF and that context has been incorporated with an MRF technique based upon the Potts model.

### 2.1 Non-Gaussian modelling: the Kummer-U distribution

**Table 1** lists the probability density functions and the matrix log-cumulant expressions for the matrix variate $\mathcal{W}_d$, $\mathcal{K}_d$, $\mathcal{G}_d^0$, and $\mathcal{U}_d$ distributions.

Bombrun et al. [11] have shown the potential of the $\mathcal{U}_d$ PDF, with texture parameters $\alpha$ and $\lambda$, to model both *extremely heterogeneous*, moderately *heterogeneous* and *homogeneous* clutter. It encompasses the other models as special cases, such that it reverts to the $\mathcal{K}_d$ as $\lambda \to \infty$, the $\mathcal{G}_d^0$ as $\alpha \to \infty$, and the $\mathcal{W}_d^{\mathbb{C}}$ as both $\alpha, \lambda \to \infty$. Therefore, this one model supersedes many previous modelling algorithms.

Parameter estimation is achieved with the method of matrix log-cumulants (MoMLC) because they are fast to compute and achieve the most accurate results [7].

### 2.2 Markov Random Fields

Markov Random Field modelling is a contextual smoothing technique which gives more weight to the class memberships of spatially neighbouring classes. The class label image is modelled as an MRF together with an isotropic second-order neighbourhood system, defining the eight surrounding pixels as the neighbourhood for each site. The class label MRF easily combines with a finite mixture model's spectral clustering, i.e., based on the pixel covariance matrix distributions, by replacing the global class prior probabilities with spatially varying local prior probabilities determined from the local neighbourhoods.

We introduce the MRF for the class labels, $\mathcal{L}$ with sites $\mathcal{S}$, as a Gibbs distribution with the energy function being proportional to the local neighbourhood counts for each class, as in [12, 10, 13].

Therefore, the $k$ class mixture model for the matrix-variate data at the $i^{\text{th}}$ location, $\mathbf{C}^{(i)}$, may be summarised as:

$$P_{\mathbf{C}}(\mathbf{C}^{(i)}) = \sum_{j=1}^{k} \mathcal{U}_d(\mathbf{C}^{(i)}; L, \mathbf{\Sigma}_j, \alpha_j, \lambda_j) \pi_j^{(i)}(\beta, \mathcal{L}) \quad (1)$$

where the local priors for each class, $\pi_j^{(i)}$, are derived from the neighbourhood counts, $m_j^{(i)}$ at each site thus

$$\pi_j^{(i)}(\beta, \mathcal{L}) = \frac{\exp\left(\beta m_j^{(i)}\right)}{\sum_{l=1}^{k} \exp\left(\beta m_l^{(i)}\right)} \quad (2)$$

and the MRF spatial parameter $\beta > 0$, which is a measure of correlation between neighbouring pixels, is found, at each iteration, by maximising

$$\beta = \arg\max_{\beta} \sum_{i \in \mathcal{S}} \sum_{j=1}^{k} P(\mathcal{L}^{(i)} = j | \mathbf{C}^{(i)}, \beta, \mathcal{L}) \log \pi_j^{(i)}(\beta, \mathcal{L}) \quad (3)$$

**Table 1:** PDFs and MoMLC equations for the covariance matrix distributions under the product model [7].

| Family | $f_{\mathbf{C}}(\mathbf{C})$ of covariance matrix $\mathbf{C}$ | MoMLC equations |
|---|---|---|
| $\mathcal{W}_d^{\mathbb{C}}(L, \mathbf{\Sigma})$ | $\frac{L^{Ld}\lvert\mathbf{C}\rvert^{L-d}}{\Gamma_d(L)\lvert\mathbf{\Sigma}\rvert^{L}} \exp\left(-L\text{tr}(\mathbf{\Sigma}^{-1}\mathbf{C})\right)$ | $\kappa_1\{\mathbf{C}\} = \ln\lvert\mathbf{\Sigma}\rvert + \psi_d^0(L) - d\ln L$<br>$\kappa_{\nu>1}\{\mathbf{C}\} = \psi_d^{\nu-1}(L)$ |
| $\mathcal{K}_d(L, \mathbf{\Sigma}, \alpha)$ | $\frac{2\lvert\mathbf{C}\rvert^{L-d}}{\Gamma_d(L)\Gamma(\alpha)} (L\alpha)^{\frac{\alpha+Ld}{2}} \left(\text{tr}(\mathbf{\Sigma}^{-1}\mathbf{C})\right)^{\frac{\alpha-Ld}{2}}$<br>$\times K_{\alpha-Ld}\left(2\sqrt{L\alpha\text{tr}(\mathbf{\Sigma}^{-1}\mathbf{C})}\right)$ | $\kappa_1\{\mathbf{C}\} = \ln\lvert\mathbf{\Sigma}\rvert + \psi_d^0(L) + d\left(\psi^0(\alpha) - \ln(\alpha L)\right)$<br>$\kappa_{\nu>1}\{\mathbf{C}\} = \psi_d^{\nu-1}(L) + d^{\nu}\psi^{\nu-1}(\alpha)$ |
| $\mathcal{G}_d^0(L, \mathbf{\Sigma}, \lambda)$ | $\frac{L^{Ld}\lvert\mathbf{C}\rvert^{L-d}}{\Gamma_d(L)\lvert\mathbf{\Sigma}\rvert^{L}} \frac{\Gamma(Ld+\lambda)(\lambda-1)^{\lambda}}{\Gamma(\lambda)}$<br>$\left(L\text{tr}(\mathbf{\Sigma}^{-1}\mathbf{C}) + \lambda - 1\right)^{-\lambda-Ld}$ | $\kappa_1\{\mathbf{C}\} = \ln\lvert\mathbf{\Sigma}\rvert + \psi_d^0(L) + d\left(\ln(\frac{\lambda-1}{L}) - \psi^0(\lambda)\right)$<br>$\kappa_{\nu>1}\{\mathbf{C}\} = \psi_d^{\nu-1}(L) + (-d)^{\nu}\psi^{\nu-1}(\lambda)$ |
| $\mathcal{U}_d(L, \mathbf{\Sigma}, \alpha, \lambda)$ | $\frac{L^{Ld}}{\Gamma_d(L)} \frac{\lvert\mathbf{C}\rvert^{L-d}}{\mathbf{\Sigma}^{L}} \frac{\Gamma(\alpha+\lambda)\Gamma(Ld+\lambda)}{\Gamma(\alpha)\Gamma(\lambda)} \left(\frac{\alpha}{\lambda-1}\right)^{Ld}$<br>$\times U\left(Ld+\lambda, Ld-\alpha+1, L\text{tr}(\mathbf{\Sigma}^{-1}\mathbf{C})(\frac{\alpha}{\lambda-1})\right)$ | $\kappa_1\{\mathbf{C}\} = \ln\lvert\mathbf{\Sigma}\rvert + \psi_d^0(L) + d\left(\psi^0(\alpha) - \psi^0(\lambda) + \ln(\frac{\lambda-1}{\alpha L})\right)$<br>$\kappa_{\nu>1}\{\mathbf{C}\} = \psi_d^{\nu-1}(L) + d^{\nu}\left(\psi^{\nu-1}(\alpha) + (-1)^{\nu-1}\psi^{\nu-1}(\lambda)\right)$ |

### 2.3 Goodness-of-fit test stage: Split and Merge

The goodness-of-fit test stage performs a few tasks worthy of mention here. Each cluster's goodness-of-fit to the data is tested at regular intervals and may "split" poorly fitting clusters, and "merge" competing clusters. This process solves two prime problems with general clustering algorithms: initialisation and the number of classes. Always starting as one, undoubtedly "poor", cluster for the whole image, and letting it adapt from that consistent level has proven a robust initialisation method, and the adaptive number of clusters results in the number of statistically distinct classes, given the image data, the model PDF and a chosen confidence level. The current algorithm includes the adaptive sensitivity and sub-sampling ability previously described in [8, 9].

## 3 Results

### 3.1 Evidence for Kummer-U: Log-Cumulant Diagrams

Recent investigations have found that a single texture parameter model is sometimes insufficient to capture the range of data distributions observed in PolSAR images. **Figure 1** shows a log-cumulant diagram (discussed in [7]) for an ENVISAT ASAR dual-pol scene of an arctic glacier on Svalbard. The blue, green and yellow sample data clusters are plotted for known regions of glacier ice, superimposed ice and firn, respectively. The black circle represents the (non-textured) $\mathcal{W}_d$ model, the red and blue lines represent the $\mathcal{K}_d$ and $\mathcal{G}_d^0$ models, respectively, and the yellow region depicts the parametric coverage of the $\mathcal{U}_d$-distribution. (The more extreme Wittaker W and M models are not considered here.) The data clusters clearly fall within the region of the $\mathcal{U}_d$-distribution, which should, therefore, achieve a better fit to the data classes, and improve the clustering results.

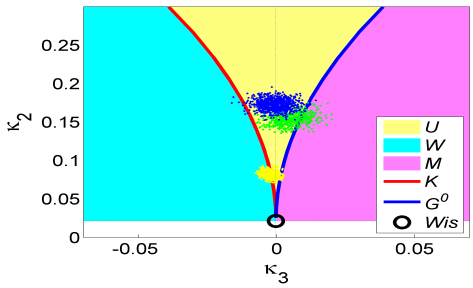


**Figure 1:** Log-cumulant diagram for real dual-polarisation ASAR scene of an Arctic glacier, showing clusters for firn, superimposed and glacier ice regions. Note that the clusters are clearly within the yellow $\mathcal{U}_d$-distribution region in the diagram.

### 3.2 Simulated Data-set

We generated a six-class $\mathcal{U}_d$-distribution test image to validate the clustering algorithm. The simulated data was 5-look, dual-pol, with a range of texture, brightness and polarimetry values taken from real images. The clustering algorithm was given only the MLC covariance matrix image as input, with no prior knowledge about the number of clusters, and resulted in the perfect clustering into six classes. **Figure 2** shows the clustering result, which perfectly matches the original simulated image. **Figure 3** shows the resulting log-cumulant space with an ellipse for each model cluster found, and indicates the spread in texture space covered by the simulated $\mathcal{U}_d$-distribution data. These figures clearly demonstrate a successful clustering of the simulated data image.

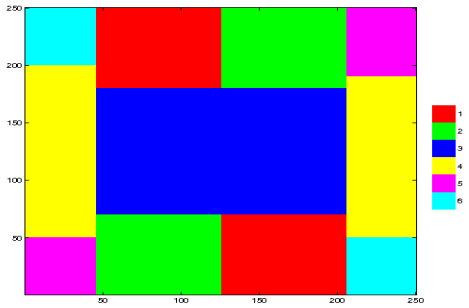


**Figure 2:** Perfect clustering result of a six-class simulated test image.

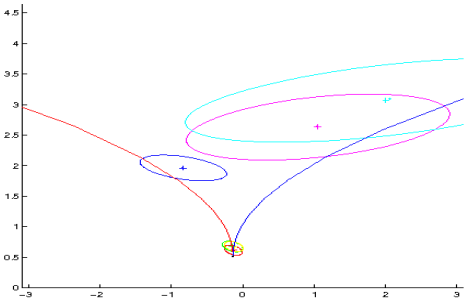


**Figure 3:** Log-cumulant diagram of simulated clustering result. The coloured ellipses indicate the six different classes and their expected sample scatter, and the red and blue curved lines indicate the $\mathcal{K}_d$ and $\mathcal{G}_d^0$ models, respectively. Note that several classes are nearly $\mathcal{W}_d$ distributed where the curves meet, but others vary in texture.

### 3.3 Real Data Example

Our real data example is the same Arctic glacier, Kongsvegen, on the Svalbard archipelago, as the log-cumulant diagram in **Figure 1** and is an ENVISAT, dual-pol VV/VH, image from May 2005, MLC processed with 24-looks.

**Figure 4** shows the pseudo-Pauli RGB image (top) and the automatically clustered image (below) which found six clusters. A visual inspection of the found clusters supports that it has found real, distinct classes in the image, although a rigorous investigation with ground truth has not yet been performed. This segmentation compares favourably to our earlier work [8], using the $\mathcal{K}_d$ model without contextual smoothing, which showed a much more fragmented cluster image.

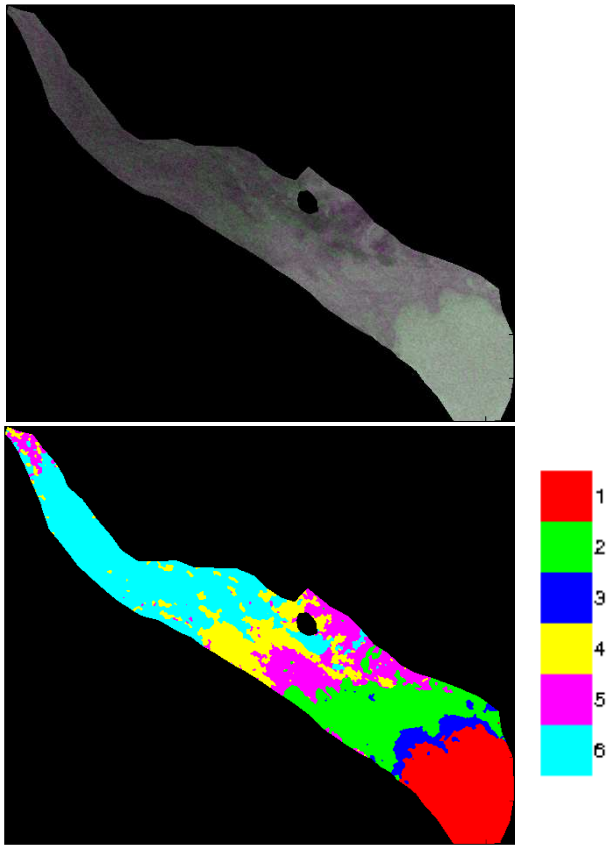


**Figure 4:** Real ENVISAT, dual-pol, 24-look example of an Arctic glacier. Pseudo-Pauli RGB image (top) and the automatically clustered image (below) which found six clusters in good (visual) correspondence to real glacier classes.

# 4 Conclusions

We have proposed an improved non-Gaussian clustering algorithm which incorporates the $\mathcal{U}_d$-distribution for multi-look covariance matrix data classes, Markov random fields for contextual smoothing, and goodness-of-fit testing to optimise the number of clusters. Visual inspection indicates that it achieves good results that appear valid for real data images.

Further real examples using this algorithm for post classification change detection may be seen in another EUSAR 2012 submission by the authors.

# References


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

[2] J. S. Lee, D. L. Schuler, R. H. Lang, and K. J. Ranson. K-Distribution for multi-look processed polarimetric SAR imagery. In *IEEE Int. Geosci. Remote Sensing Symp.*, vol. 4, pp. 2179–2181, Pasedena, USA, 1994.

[3] A. Doulgeris and T. Eltoft. Scale mixture of Gaussians modelling of polarimetric SAR data. In *International POLinSAR Workshop (POLinSAR2007)*, Frascati, Italy, January 22-26 2007.

[4] C.C. Freitas, A.C. Frery, and A.H. Correia. The polarimetric G distribution for SAR data analysis. *Environmetrics*, 16(1):13–31, 2005.

[5] M.M. Horta, N. Mascarenhas, A.C. Frery, and A. Levada. Clustering of fully polarimetric sar data using finite gp0 mixture model and sem algorithm. In *Systems, Signals and Image Processing, 2008. IWSSIP 2008. 15th International Conference on*, pp. 81 –84, June 2008.

[6] L. Bombrun, G. Vasile, M. Gay, and F. Totir. Hierarchical segmentation of polarimetric SAR images using heterogeneous clutter models. *Geoscience and Remote Sensing, IEEE Transactions on*, pp. 1–12, 2010.

[7] S.N. Anfinsen and T. Eltoft. Application of the Matrix-Variate Mellin Tranform to Analysis of Polarimetric Radar Images. *IEEE Trans. Geoscience and Remote Sensing*, 49(7), July 2011. 15 pp.

[8] A.P. Doulgeris and T. Eltoft. Automated Non-Gaussian clustering of polarimetric SAR. In *8th European Conference on Synthetic Aperture Radar (EUSAR2010)*, Aachen, Germany, June 7-10 2010.

[9] A.P. Doulgeris, S.N. Anfinsen, and T. Eltoft. Automated non-Gaussian clustering of polarimetric synthetic aperture radar images. *IEEE Trans. Geoscience and Remote Sensing*, 49(10), 2011.

[10] G. Celeux, F. Forbes, and N. Peyrand. Em procedures using mean field-like approximations for markov model-based image segmentation. *Pattern Recognition*, 36(1):131–144, 2003.

[11] L. Bombrun and J.-M. Beaulieu. Fisher distribution for texture modeling of polarimetric sar data. *Geoscience and Remote Sensing Letters, IEEE*, 5(3):512 –516, July 2008.

[12] S. Li. *Markov random field modeling in image analysis*. Springer, 2009.

[13] J. Besag. Efficiency of pseudo-likelihood estimation for simple gaussian fields. *Biometrika*, 64:616– 618, 1977.