UNIVERSITY OF TROMSØ UIT

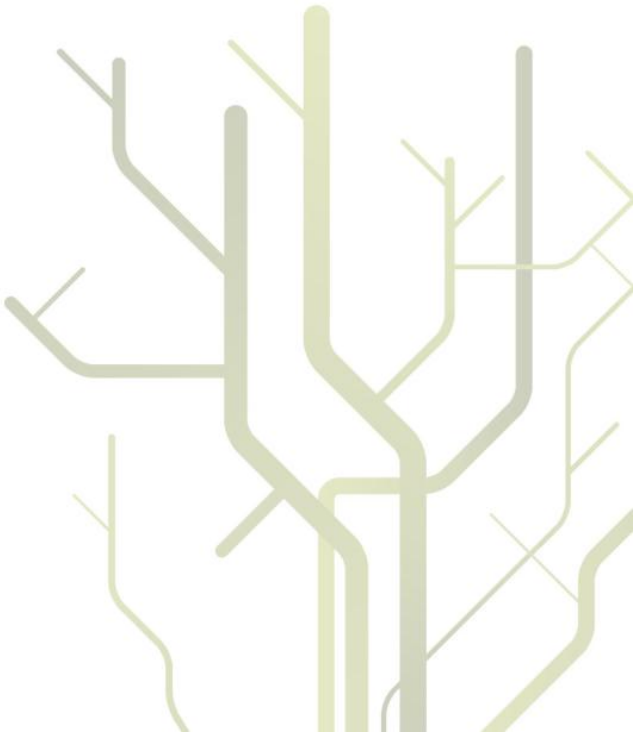
FACULTY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF PHYSICS AND TECHNOLOGY

Multitemporal Analysis of Multipolarization Synthetic Aperture Radar Images for Robust Surface Change Detection

Vahid Akbari

A dissertation for the degree of Philosophiae Doctor

June 2013



Dedicated to my love Roya

«If you knew what you are doing,
it wouldn't be called research.»
Albert Einstein

«If you judge a fish by its ability to climb a tree,
it will spend its entire life believing its stupid.»
Albert Einstein

The hands of fate play our game
We the players are given a name
Some are tame, others gain fame
Yet in the end, we're all the same.

ما لعبتکانیم و فلک لعبت باز
از روی حقیقتی نه از روی مجاز
یک چند درین بساط بازی کردیم
رفتیم به صندوق عدم یک یک باز

Khayyam Neishabouri

Abstract

This thesis addresses change detection from multipolarization, multilooked SAR images. Change detection can be viewed as a particular case of the multitemporal image classification problem. Change detection from SAR images is particular challenging due to several intrinsic properties related to the way these images are formed. SAR images are affected by physical properties such as target permittivity, imaging geometry, and surface topography, which call for accurate pre-processing and co-registration. In addition, SAR images of distributed targets are disturbed by speckle, a noise-like phenomenon, which complicates analysis and interpretation. On the other hand, radar signals in the micro frequency bands commonly used in satellite borne SAR systems, are almost insensitive to light and atmospheric conditions. This makes the possibility of performing change detection from SAR interesting, especially in the Northern and Arctic regions.

We discuss two main approaches: a post-classification comparison algorithm and a direct change detection algorithm.

We consider the complete processing chain associated with performing postclassification change detection from time series of multi-polarimetric SAR (PolSAR) images acquired with different imaging geometries and polarimetric configurations. The application is connected to monitoring of changes in Arctic glaciers. The images are corrected for terrain effects by thoroughly reducing topographic effects on both geolocation, radiometry and polarization signature, and subsequently stacked into proper time series for further analysis. Multilooked, multi-polarimetric SAR images are generally represented as matrix-variates. The matrix-variate $\mathcal{U}_d$-distribution is found to enable proper statistical representation of the variable texture observed in our multitemporal PolSAR data. An unsupervised Markov random field (MRF) based contextual non-Gaussian clustering algorithm, named the $\mathcal{U}$-MRF segmentation algorithm, is employed for segmentation of the terrain corrected images. This algorithm has built in contextual smoothing by MRF modeling, and yields homogeneous segmentation, leading to robust change detection results. The clustered PolSAR data is subsequently labeled into glacier zones with the aid of ground truth data. The consistency of the segmentation algorithm is also demonstrated by characterizing the expected random error level for SAR images under different imaging conditions. This allows us to determine whether an observed variation is statistically significant and therefore can be used for glaciers change detection. Finally, the classified images of succeeding years are compared and analyzed, and changes are identified as the detected temporal changes in the location of boundaries between glacier distinct zones.

The thesis also proposes a novel method for direct unsupervised change detection from PolSAR data. We assume that the matrix variates follow the complex Wishart distribution, and the complex Hotelling-Lawley (HL) trace statistic is applied as a new test statistic for measuring the similarity between two complex covariance matrices. In the null hypothesis (no change), the statistical moments of the HL statistic are functions of the number of looks and the number of polarimetric channels. The sampling distribu-

tion of the test statistic is then approximated by a Fisher-Snedecor (FS) distribution. The proposed method is to match the empirical population moments of the FS distribution with those of the HL statistic. The no change hypothesis of equal covariance matrices may then be rejected at a predefined false alarm rate. The performance of the proposed method is demonstrated with good results on simulated and real PolSAR data sets.

To ease the reading for non-expert readers, the thesis also provides several chapter of appropriate background material.

Acknowledgments

I would like to express my gratitude to all those who has helped and inspired me during my doctoral study.

I would like to thank the research council of Norway for giving me the scholarship for being a research fellow at the Department of Physics and Technology of the University of Tromsø.

I am deeply indebted to my supervisor Professor Torbjøn Eltoft whose help, stimulating suggestions and encouragement helped me in all the research time for completion of this thesis. A huge thank to my co-advisors, Anthony Doulgeris and Stian Normann Anfinsen, they spent all the time as much as possible helping me, for being always available and for their technical supports that they have given me during my PhD. I also thank them for their detailed and precise comments on my manuscripts and dissertation. I want to thank the members of our Earth Observation Laboratory, Camilla Brekke, saïd Faozi, Ane Schwenke Fors, Mari-Ann Moen, Tao Ding, Stine Skrunes, Thomas Kræmer, and Necip Gökhan Kasapoglu. I have thoroughly enjoyed being part of this research group and look forward to future collaborations.

I would like to take this opportunity to thank Professor Sebastian Bruno Serpico for allowing me to spend a great time at the Signal Processing and Telecommunications (SPT) research group of the Department of Telecommunications, Electronic, Electrical, and Naval Engineering, University of Genoa. I am very grateful for having this opportunity. In particular, I would like to thank Gabriele Moser for his technical supports and valuable comments on my manuscripts . I also thank support of all other members in the research group, Giulia Troglio, Michaela De Martino. I also want to thank the PhD students from this university, Mauricio Soto Alvarez, Pantea Nadimi, Tewodros Atanaw, Henry Andrade, and Jose Francisco for sharing their funny times with me.

A special thank goes to Yngvar Larsen from Norut for sharing his knowledge of SAR geocoding with me and spending time on answering my questions patiently and kindly.

I am also very grateful to all friends from my home county living in Tromsø, for making all good times. My warm thanks go to my dear friend, Sajjad, who shared his time and gave his supports all the time.

Also, I thank my office roommates, Pal Løvhaugen, Balpreet Ahluwalia, Sanat Wagle, Thomas Kræmer, and Tao Ding. Without you, there have been lonely days in the office.

I take this opportunity to record my sincere thanks to my fiends at the department for providing me some funny times in ping-pong playing.

Nothing is more important than family. To my wife Roya whose smiles, encouragements, and endless patience enabled me to complete this work. And to my family who their loving supports and encouragements from the beginning of my life was with me. And to my parents-in-law. Without their help I may never complete writing the thesis.

I dedicate this thesis to you all.

Vahid Akbari
Tabriz, March 2013

Contents

List of Tables

List of Figures

Nomenclature

List of Notation

$\mathbf{C}$	sample covariance matrix
d	number of polarimetric channels
D	length of the radar antenna
M	width of the radar antenna
H_{sat}	height of the satellite orbit above the Earth
$\mathbf{S}$	scattering matrix
$\underline{\mathbf{k}}$	Pauli scattering vector
$\underline{\mathbf{\Omega}}$	Lexicographic target vector
$\mathbf{W}$	complex Wishart distributed matrix
$\mathcal{W}_d^{\mathbb{C}}$	complex Wishart distribution
$s\mathcal{W}_d^{\mathbb{C}}$	complex scaled Wishart distribution
$\mathcal{KW}$	$\mathcal{K}$-Wishart distribution
$\mathcal{K}$	$\mathcal{K}$ distribution
μ	location parameter
L	nominal number of looks
$\mathbf{\Sigma}$	scale matrix of matrix distributions
$\Gamma_d(L)$	multivariate gamma function of the complex kind
$\Gamma(\cdot)$	standard Euler gamma function
α	shape parameter
η	shape parameter
$K_\rho(\cdot)$	modified Bessel function of the second kind with order ρ
Z	texture variable
κ_ν	νth-order matrix-variate matrix log-cumulant
U	energy function
$\psi_d^\nu(\cdot)$	multivariate polygamma function
$\psi_d^0(\cdot)$	multivariate digamma function
$\psi^\nu(\cdot)$	ordinary polygamma function
c	a clique
$\mathcal{C}$	collection of all cliques
V_c	potential associated with clique c
S_{rt}	scattering coefficient subscripted with associated receive and transmit polarization

$\langle \mu_\nu \rangle$	sample matrix log-moments
W	partition function
β	spatial interaction parameter
$\mathcal{G}^0$	$\mathcal{G}^0$ distribution
$\mathcal{U}$	$\mathcal{U}$ distribution
$\boldsymbol{\theta}$	vector of parameters of the matrix-variate distribution
δ_r	slant range resolution
δ_a	azimuth resolution
δ_g	ground range resolution
ω	azimuth slope angle
γ	ground range slope angle
χ^2	χ^2 distribution
β^0	normalized radar brightness
$\underline{\mathbf{\Omega}}_{\beta^0}$	area normalized scattering vector in the slant range plane
$\mathbf{C}_{\text{RTC}}$	radiometrically terrain corrected covariance matrix
$\mathbf{C}_{\text{OAC}}$	covariance matrix after orientation angle compensation
$\delta(z-1)$	Dirac delta function
$\bar{\gamma}$	gamma distribution
$\bar{\gamma}^{-1}$	gamma inverse distribution
$\overline{\mathcal{FS}}$	Fisher-Snedecor distribution
$U(\cdot,\cdot,\cdot)$	confluent hypergeometric function of the second kind (KummerU)
Ψ	projection angle
θ	radar look angle
θ_E	ellipsoidal incidence angle
θ_ℓ	local incidence angle
$\mathbf{R}(\vartheta)$	rotation matrix
R_E	Earth's radius
R_e	equatorial radius of the Earth model
R_p	polar radius of the Earth model
$\underline{\mathbf{r}}_{\text{tar}}$	target position vector in the ECEF reference frame
$\underline{\mathbf{r}}_{\text{sat}}$	satellite position vector in the ECEF reference frame
$\underline{\mathbf{v}}_{\text{sat}}$	satellite velocity vector in the ECEF reference frame
$\mathbb{C}$	complex plane
ω_h	antenna beam width in the azimuth direction
ω_v	antenna beam width in the range direction
ν	order parameter
B	bandwidth of the radar signal
k	wave number
$\mathcal{N}_s$	a set of neighbors of the site s
N	number of scatterers in the resolution cell
A_β	reference area for the β^0 backscatter
A_σ	reference area for the σ^0 backscatter
A_γ	reference area for the γ^0 backscatter
ϕ	phase of scattering coefficient
τ	pulse length
W_G	radar swath
λ	wavelength

R	slant range distance
ϑ	polarization orientation angle shift
τ_{CR}	contrast ratio
τ_{E}	ellipticity
τ_{ν}	Minkowski distance of order ν
τ_{2}	Euclidean distance
τ_{1}	Manhattan distance
τ_{CH}	Chebyshev distance
τ_{CAD}	Canberra distance
τ_{BH}	Bhattacharyya distance
τ_{KL}	Kullback-Libeler distance
τ_{W}	Wishart distance
τ_{RW}	revised Wishart distance
τ_{SRW}	symmetrized revised Wishart distance
τ_{B}	Bartlett distance
τ_{H}	Hellinger distance
τ_{WC}	Wishart-Chernoff distance
τ_{M}	M-distance
τ_{HL}	complex Hotelling-Lawley trace statistic
$m_{\nu}^{(\mathrm{HL})}$	ν^{th} order moment of the Hotelling-Lawley trace statistic
$m_{\nu}^{(\mathrm{FS})}$	ν^{th} order moment of Fisher-Snedecor distribution
α_c	significance level
z_{α_c}	threshold at the significance level α_c

List of Acronyms

ALOS	Advanced Land Observing Satellite
AP	alternating polarization
ASAR	Advanced Synthetic Aperture Radar
BD	Bhattacharyya distance
BSA	backscattered alignment
CAD	Canberra distance
CDF	cumulative distribution function
CFAR	constant false alarm rate
CR	contrast ratio
DA	detection accuracy
DCD	direct change detection
DEM	digital elevation model
ECEF	Earth Centered Earth Fixed
ELD	equivalent linear displacement
EM	expectation maximization
ENL	equivalent (or effective) number of looks
ENVISAT	environmental satellite
EO	Earth Observation
FAR	false alarm rate
FS	Fisher-Snedecor
GI	glacier ice
GPR	ground penetrating radar
GRF	Gibbs random field
GTC	geometric terrain correction
HL	Hotelling-Lawley

IEEE	Institute of Electrical and Electronics Engineers
KL	Kullback-Leibler
KS	Kolmogorov-Smirnov
LRT	likelihood ratio test
LUT	look up table
MAL	maximum asymptotic likelihood
MI	mutual information
MoMLC	method of matrix log-cumulant
MLC	multi-look complex
MRF	Markov random field
MAP	maximum a posteriori
ML	maximum likelihood
NPI	Norwegian Polar Institute
OAC	orientation angle compensation
OER	overall error rate
PA	projection angle
PALSAR	Phased Array type L-band Synthetic Aperture Radar
PC	projection cosine
PCCD	post-classification change detection
PD	probability of detection
PDF	probability density function
PFA	probability of false alarm
PMF	probability mass function
POA	polarization orientation angle
PolSAR	polarimetric synthetic aperture radar
PRF	pulse repetition frequency

RCS	radar cross section
RDE	range-Doppler-Earth model
RLOS	radar line of sight
RTC	radiometric terrain correction
SEM	stochastic expectation maximization
SAR	synthetic aperture radar
SI	superimposed ice
SL	significance level
SLC	single-look complex
SNR	signal-to-noise ratio
WC	Wishart-Chernoff

Chapter 1

Introduction

This chapter is intended to give an overall overview of the whole thesis. It introduces the motivation for the research, summarizes and discusses the three research articles that make up the main content, and lists associated conference works.

1.1 Motivation

Advances in Earth Observation (EO) technologies have improved our abilities to monitor and study processes on the Earth's surface. One of the main applications of remote sensing data is to detect and monitor changes in surface cover in multitemporal images. A number of different applications relies on robust and accurate change detection algorithms. Examples of such applications are; detection and monitoring of volcanic activity, surveillance of disasters (e.g. landslides, floods, forest fires), monitoring and tracking of sea-ice motion, monitoring glaciers, mapping of snow cover areas, and surveillance of coastal areas and growth of urban areas.

Imaging synthetic aperture radar (SAR) is presently a key instrument onboard EO satellites. Images acquired by SARs have to a lesser extent (contrary to optical sensors) been exploited in the context of change detection. This is due to the fact that SAR images suffer from the presence of *speckle* phenomenon, which largely complicates the analysis. Also, it is difficult to directly compare SAR images acquired with different polarizations, different incidence angles, different satellite fight paths and different look directions. However, the use of SAR sensors is attractive in temporal studies from an operational viewpoint, because microwave SAR systems offer the advantage of being insensitive to atmospheric and light conditions. Research studies demonstrate that the use of SAR images have great potential in change detection and time series analysis.

Radar polarimetry use the complete information which can be gained with microwave remote sensing about targets of any kind. This technology has advanced rapidly the last 10-18 years, and is now at a stage where full-polarimetric radar systems operating on many different frequencies are deployed on EO satellites (ALOS PALSAR, RADARSAT-

2, TerraSAR-X, COSMO-SkyMed). The polarimetric SAR (PolSAR) systems are expected to improve the ability to infer change information from multitemporal data. This is due to the fact that PolSAR data also allow for classification of scattering mechanisms, which potentially may provide an increased discrimination capability.

The overall objective of the thesis "Multitemporal Analysis of Multipolarization Synthetic Aperture Radar Images for Robust Surface Change Detection" is to develop statistical models, divergence measures and analysis strategies for temporal analysis of SAR and PolSAR data, which can result in robust change detection algorithms. Our goal is to adapt and improve existing, and develop new methods for analyzing these data. The usefulness of remote sensing data in monitoring various phenomena and detecting change is highly dependent on the ability to create reliable time series of images, which can reveal dynamics at high temporal and spatial resolution. Through the thesis, we want to develop algorithms that can do this, by enhancing the information retrieval from multidimensional and multitemporal SAR data. We note that a difference in radar backscattering between multitemporal data may be caused by several factors such as actual change in land cover, differences in viewing geometry, differences in satellite path, differences in sensor calibration, differences in atmospheric moisture conditions, differences in meteorological conditions, and differences in the colocation of the multitemporal images. We will study some of these factors, adapt and improve existing algorithms, and develop new methods for analyzing time series of SAR data. In particular, we will:

- Demonstrate the capability of more advanced non-Gaussian matrix-variate distributions for representing the variable texture observed in multilooked, multitemporal PolSAR data sets. This includes studies of the added value of combining non-Gaussian modeling and Markov random field (MRF) contextual smoothing with respect to improving accuracy and reliability of PolSAR image segmentation.

- Investigate the impact of DEM-based radiometric terrain correction and polarization orientation angle compensation of multitemporal multipolarization SAR data acquired with different imaging conditions on land cover change detection in multilooked, multitemporal PolSAR data.

- Study and evaluate the appropriateness of new and existing test statistics for unsupervised direct change detection in multilooked, multitemporal PolSAR images.

The application will be related to the surveillance of polar areas, with emphasis on detecting and monitoring changes in glacier surfaces. Glacier variability is known to be important indicator of global climate change, and hence of great importance for climate research.

1.2 Organization of the Thesis

Chapter 2 provides an introduction to the SAR imagery. We describe the geometry in side-looking SAR, SAR spatial resolution in range and azimuth directions, speckle phenomenon, and geometric distortions. The material presented here is a review of earlier work.

Chapter 3 gives an introduction to the PolSAR concepts and data models for PolSAR. We look at the data formats delivered by polarimetric radars, starting with single-look complex (SLC) data and moving on to multilook complex (MLC) data, while explaining the concept of multilooking. The chapter ends with definition of the doubly stochastic product model for SLC and MLC data and finally some matrix-variate statistical distributions for multilook data.

Chapter 4 an overview on the contextual MRF-based classification of PolSAR data. This chapter address the problem of contextual PolSAR image clustering by combining pixelwise statistical distributions for multilook data and contextual information. In the context of this chapter, we first detail the MRF theory and then the algorithm steps of the Bayesian classification are described.

Chapter 5 describes the geometric and radiometric terrain correction of PolSAR data. Surface topography has influence on the geometric and radiometric quality of SAR images. In this chapter, precise SAR geocoding using digital elevation model (DEM) and orbital information is described. Then, radiometric correction of multilook PolSAR data is addressed that utilizes the pixel size area normalization on each element of covariance matrix data. Finally, polarimetric orientation angle compensation which is an important correction for PolSAR data in rugged terrain areas is introduced.

Chapter 6 describes the concept of direct change detection in multipolarizarion SAR data, and also reviews some test statistics useful for polarimetric changes detection in addition to our detector.

Chapter 7-9 contains three manuscripts, forming the basis of the thesis. These are separately described in the next section, where we summarize the main findings and highlight the original contribution of the authors.

Chapter 10 gives the summary and the concluding remarks and points out some suggestions for future work based on the work documented in the thesis.

Appendix A is a conference paper (referenced as Paper 11 in the list of Section 1.4) which contains the clustering method used in the processing chain of the post-classification change detection described in the Chapter 8.

Appendix B is a supporting document for Paper 3 (referenced as Paper 10 in the list of Section 1.4). It introduces the complex Hotelling-Lawley (HL) trace statistic and expressions for the moments of the HL statistic under the complex Wishart distribution.

1.2.1 Publication Review

The main body of this thesis is presented as three journal publications which are included as Chapters 7, 8 and 9. A summary is given for each article, describing the key findings and highlighting the original contributions of the authors. The papers appear in chronological order to reflect the progress of the research itself. Paper 1 introduces the non-Gaussian contextual clustering algorithm for multilook covariance matrix data. Paper 2 introduces a processing chain for the post-classification change detection of Arctic glaciers from multitemporal multipolarization SAR data, where Paper 1 makes a foundation of the clustering method used in Paper 2. Paper 3 introduces a novel direct change detection in multilook PolSAR data under the complex Wishart distribution .

Paper 1

V. Akbari, A. P. Doulgeris, G. Moser, T. Eltoft, S. N. Anfinsen, and S. B. Serpico, "**A Textural-Contextual Model for Unsupervised Segmentation of Multipolarization Synthetic Aperture Radar Images**," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 4, pp. 2442-2453, Apr. 2013.

In this paper, a novel unsupervised, non-Gaussian, and contextual segmentation algorithm for PolSAR imagery has been developed by combining an advanced statistical distribution with spatial contextual information. This extends on previous studies that have shown the added value of both non-Gaussian modeling and contextual smoothing individually or for intensity channels only. We use the non-Gaussian $\mathcal{K}$-Wishart distribution, which accounts for potential textural differences in the classes, to represent the individual pixelwise statistical properties. The classifier based on pixel statistics only is severely affected by overlapping class statistics due to speckle noise. This problem is reduced by incorporating spatial contextual information of the associated image data in the analysis process by MRF modeling. The proposed contextual clustering method uses a specific Markovian energy function for integrating the $\mathcal{K}$-Wishart distribution for the PolSAR data statistics conditioned to each image cluster and a Potts model for the spatial context. Specifically, the proposed algorithm is constructed based upon the iterative stochastic expectation maximization (SEM) algorithm. A new formulation of SEM is developed to jointly address both the data clustering and parameter estimation of the $\mathcal{K}$-Wishart distribution and the MRF model.

The clustering algorithm requires that some parameters are given in advance. One key parameter in unsupervised image segmentation is the appropriate number of clusters which was determined in a preclustering process. The clustering also should be initial-

ized using either K-Means clustering of the logarithm intensities of the individual polarimetric channels or random initialization. The parameter of Potts MRF model needs also to be initialized. Moreover, the effective or equivalent number of looks (ENL) is a key parameter in all pixelwise distributions for multilook PolSAR data and had to be estimated using the method of matrix log-cumulants (MoMLC) in a preanalysis of the image.

The added value of combining the flexible non-Gaussian $\mathcal{K}$-Wishart distribution and the Potts MRF model was tested on three simulated and real data. The segmentation results before and after MRF modeling for both the standard Wishart and the $\mathcal{K}$-Wishart classifier have been obtained. The segmentations have been compared in terms of discriminability of non-Gaussian regions with $\mathcal{K}$-Wishart with respect to standard Wishart model and contextual smoothing with MRF. The effectiveness of MRF models in improving the accuracy (quantified for simulated data) and reliability of PolSAR image clustering has been remarked for all examples. The results show improvement with respect to segmentation of pixelwise clustering. With regard to the computation time, the whole process is slightly slower than the original pixelwise SEM algorithm due to the additional MRF stage in the clustering scheme. Even on the basis of data with a low number of looks (and therefore a high degree of speckle), the proposed approach is able to generate homogeneous and reliable clustering results.

Paper 2

V. Akbari, A. P. Doulgeris, and T. Eltoft, "**Monitoring Glacier Changes by Multitemporal Multipolarization SAR Images**," submitted June 2012, and under review in *IEEE Transactions on Geoscience and Remote Sensing*.

This paper presents a processing chain for post-classification change detection of Arctic glaciers from multitemporal multipolarization SAR images acquired with different polarization configurations, different satellite fight paths and different look directions. The algorithm has been tested on dual polarization ENVISAT ASAR images for the period 2004-2006 over the Arctic glacier, Konsvegen, Svalbard. We first produced terrain corrected multilook complex covariance data by reducing the effects of topography on both geolocation and SAR radiometry as well as azimuth slope variations on polarization signature. Terrain correction is a prerequisite for intercomparisions of multitemporal SAR images.

We showed in Paper 1 that the $\mathcal{K}$-Wishart distribution can be used to model SAR image texture. However, the analysis shows that this model does not always represent the data well. The $\mathcal{G}_d^0$ has already been included as a choice of model for extremely heterogeneous area. The Kummer-U distribution has been introduced in [Bombrun et al., 2011] to represent PolSAR vector data . We use its multilook extension, named the multivariate $\mathcal{U}_d$-distribution, for multilook covariance matrix data. The flexibility of this model with respect to $\mathcal{K}$-Wishart and $\mathcal{G}_d^0$ with an extra texture parameter is evident

that covers more of the space of matrix log-cumulant observed in multitemporal multi-polarization SAR data. The $\mathcal{U}_d$ distribution covers the manifold between the $\mathcal{K}$-Wishart distribution and the $\mathcal{G}_d^0$ distribution, but not below either of them. It is expected to yield improved results because of its flexibility to model more varied textures. The matrix log-cumulant diagram was demonstrated for each scene to visualize the capability of the $\mathcal{U}_d$-distribution to model texture in the multitemporal dual polarization SAR data over this glacier.

Unsupervised contextual non-Gaussian clustering method was then performed using the $\mathcal{U}$-MRF classifier over the terrain corrected SAR scenes. The basis of the textural-contextual classifier was made in Paper 1/Chapter 7. More details about clustering algorithm can be found in Appendix A. The contextual smoothing yields homogeneous segmentation which leads to more robust change detection results. Ground truth data are used to label segmented images into the three major classes of glacier facies, i.e., firn, glacier ice (GI), and superimposed ice (SI) and to investigate the classification accuracies.

We then characterized the consistency of the classification as the total variation of firn/SI boundary between two no change images to obtain the expected variation just due to processing errors in the processing chain. Finally, we did post-classification change detection analysis based on the classified images on a pixel-by-pixel based analysis. The variations of the boundaries between glacier facies were clearly detected within the period of study. The variation for the two-year period, 2004-2006, exceeds the measured classification variation and thus shows significant change for this period although one year differences were not significant. These procedures may form the basis for more operational monitoring of Arctic areas.

Paper 3

V. Akbari, S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, G. Moser, and S. B. Serpico, "**Change Detection for Polarimetric SAR Data with the Hotelling-Lawley Trace Statistic under the Complex Wishart Distribution**," submitted May 2013, and under review in *IEEE Transactions on Geoscience and Remote Sensing*.

In this paper, we propose a new test statistic for unsupervised change detection in multilook PolSAR data under the complex Wishart distribution. We apply the complex HL trace statistic as a test statistic on multitemporal PolSAR images for measuring the similarity of two covariance matrices. Moments of the HL statistic under the complex Wishart model were derived in the paper of Apendix B. The sampling distribution of the HL statistic is then approximated by a Fisher-Snedecor (FS) distribution. The model parameters of FS distribution is calculated by estimated ENL for each image and polarimetric dimension. ENL estimation is an important input to the proposed change detection method and needs to be estimated accurately in preanalysis. The proposed method is to match the population moments of the FS distribution with those of the

HL statistic. Eventually, a binary decision can be made to get a final change map at a predefined false alarm rate.

In brief, the proposed change detection algorithm is made up of four main steps: 1) separate ENL estimation of the input PolSAR images, 2) generation of the HL test statistic image, 3) FS modeling of the test statistic image, and 4) thresholding. We compare the change detection results obtained from our proposed method with the Wishart likelihood ratio test (LRT) statistic proposed in [Conradsen et al., 2003] in terms of detection accuracy, false alarm rate, overall error rate and receiver operating characteristic (ROC) plots.

We have found that FS can model the null hypothesis which corresponds to no change hypothesis. This was confirmed by testing on synthesized Gaussian polarimetric pairs with different number of looks and dimensions. This experiment was then extended on this pair with simulated change in terms of polarimetry and intensity variations. In this case, the HL statistic performed slightly better performance than the LRT statistic to detect changes. The HL and LRT change detectors are then compared in detail on two real PolSAR pairs, with modified and real change in the data sets. Compared to the LRT statistic, the HL statistic represents higher sensitivity to the differences in polarimetric information, and hence a better performance in detecting changes. In cases of a bad fit of the histograms to the estimated PDF, it may be due to presence of texture that make deviation from the Wishart distribution. To reduce this problem, we suggest to multilook the original SLC data with high degree of smoothing as a simple solution.

1.3 My Contributions to the Journal Publications

- **Paper 1 (Chapter 7)**: The suitability of non-Gaussian modeling in PolSAR classification has been reported in [Doulgeris et al., 2008]. The idea of incorporating spatial contextual information in image classification has for example been investigated in the research of the IPRS research group at the University of Genoa (e.g., [Serpico and Moser, 2006]), where I worked on the problem of contextual MRF-based classification. My contribution to this paper is on combining non-Gaussian modeling and contextual smoothing for clustering of multilook PolSAR data. The experimental setup was designed by me, and I have conducted all the practical experiments. The clustering algorithm in this paper is a basis for the Paper 2.

- **Paper 2 (Chapter 8)**: My contribution to this paper is the development of the workflow for analyzing multitemporal multipolarizarion SAR data for glacier change detection. The workflow consists of five major steps: 1) SAR data selection and multilooking, 2) terrain correction, 3) PDF selection, 4) unsupervised segmentation, and 5) post-classification change detection. The experiments have been implemented by myself. Another major contribution of mine in this paper is the extension of the theory of RTC to the polarimetric case. Based on the projection cosine approach in [Ulander, 1996] for the radiometric slope correction of SAR im-

agery, I extended the method to multilook PolSAR data to allow running the RTC on all covariance matrix elements and project them to the ground range area. I also implemented the algorithm for the OAC of PolSAR data proposed in [Lee et al., 2000].

To choose an appropriate PDF to better model the variable texture in multitemporal PolSAR data, in the third step I used the implemented matrix log-cumulant diagram proposed in [Anfinsen and Eltoft, 2011]. Paper 2 represents an extension with respect to Paper 1 by applying a more flexible distribution to model the variable texture in multitemporal PolSAR data in the segmentation algorithm. This was worked out in collaboration with the second author.

- **Paper 3 (Chapter 9)**: The complex HL trace statistic was first introduced by the first author of the paper in Appendix B. The first three moments of the HL trace statistic were also derived by the first author. My contribution to this paper is to study the usefulness of the HL trace statistic as a new test statistic for change detection in polarimetric radar images. Thus, the experimental setup was completely done by myself. I started to simulate multilook polarimetric pairs with different number of looks and dimensions under the null (no change) hypothesis. I tested if the FS distribution can model the null hypothesis. Then I moved on to simulating a change in the data by introducing polarimetry and intensity variations to mimic both land cover and seasonal change. I should mention that the HL statistic can detect any type of changes in the data. Finding suitable real PolSAR data was a big challenge for me in the experiments. In order to show how the proposed test statistic works on real data, I created a change in PolSAR data from a rainforest area in the Amazon, and I could show that the proposed change detection test statistic has higher sensitivity in detecting polarimetric changes than the Wishart LRT.

1.4 Other Publications and Presentations

As first author:

1. V. Akbari, S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, "The Hotelling-Lawley trace Statistic for change detection in polarimetric SAR data under the complex Wishart distribution", submitted to *IEEE Geoscience and Remote Sensing Symposium (IGARSS2013)*, Melbourne, Australia, 21-26 Jul. 2013.

2. V. Akbari, Y. Larsen, A. P. Doulgeris, and T. Eltoft, "The impact of terrain correction of polarimetric SAR data on glacier change detection", *Proc. IEEE Geoscience and Remote Sensing Symposium (IGARSS2012)*, Munich, Germany, pp. 5129-5132, 22-27 Jul. 2012.

3. V. Akbari and M. Motagh, "Improved ground subsidence monitoring using small baseline SAR interferograms and a weighted least squares inversion algorithm", *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 3, pp. 437-441, May, 2012.

4. V. Akbari, A. P. Doulgeris, and T. Eltoft, "Time series analysis of multi-polarisation synthetic aperture radar images with a textural-contextual model", *Proc. 9th Eur. Conf. Synthetic Aperture Radar (EUSAR 2012)*, Nuremberg, Germany, pp. 726-729, 23-26 April, 2012, in press.

5. V. Akbari, G. Moser, A. P. Doulgeris, S. N. Anfinsen, T. Eltoft, and S. B. Serpico "A K-Wishart Markov random field model for clustering of polarimetric SAR imagery", *Proc. IEEE Geoscience and Remote Sensing Symposium (IGARSS2011)*, Vancouver, Canada, 1357-1360, 24-29 July, 2011, in press.

6. V. Akbari, A. P. Doulgeris, and T. Eltoft, "Glacier change detection from SAR data by contextual non-Gaussian clustering", *Nordic Remote Sensing Days (NRSD) Conf.*, Tromø, 30-31 Aug. 2011, presentation only.

7. V. Akbari, "Investigating ground deformation with spaceborne synthetic aperture radar (SAR) interferometry", *Norsk forening for bildebehandling og mønstergjenkjenning (NOBIM2010)*, Tromsø, Norway, Jun. 2010, oral presentation.

8. V. Akbari, A. P. Doulgeris, and T. Eltoft, "Non-Gaussian clustering of SAR images for glacier change detection", *Proc. European Space Agency (ESA) Symposium 2010*, Bergen, Norway, 4 pp., 28 June-2 July, 2010, in press.

9. V. Akbari, M. Motagh, M. A. Rajabi and Y. Djamour, "Time series investigation of land subsidence using a weighted least squares adjustment based on image mode interferometric data", *Proc. 8th Eur. Conf. Synthetic Aperture Radar (EUSAR 2010)*, Aachen, Germany, pp. 770-773, 7-10 June, 2010, in press.

As coauthor:

10. S. N. Anfinsen and V. Akbari, "Statistical characterisation of the complex Hotelling-Lawley trace statistic", in preparation for submission to the *IEEE Transactions on Signal Processing*.

11. M. Mahdian Pari, V. Akbari, F. Mohammadimanesh, and M. A. Fazel, "Polarimetric SAR data for urban land cover classification using finite mixture model", *European Geosciences Union (EGU 2013)*, Vienna, Austria, 7-12 April, 2013, poster presentation.

12. A. P. Doulgeris, V. Akbari, and T. Eltoft, "Automatic PolSAR segmentation with the $\mathcal{U}$-distribution and Markov random fields", *Proc. 9th Eur. Conf. Synthetic Aperture Radar (EUSAR 2012)*, pp. 183-186, 23-26 April, 2012, in press.

13. M. A. Fazel, S. Homayouni, V. Akbari, and M. Mahdian Pari, “Speckle reduction of SAR images using curvelet and wavelet transforms based on spatial features characteristics", *Proc. IEEE Int. Geosci. Remote Sens. Symp (IGARSS 2012) Conf.*, Munich, Germany, 22-27 Jul. 2012.

14. M. Mahdian Pari, M. Motagh, and V. Akbari, “Speckle reduction and restoration of synthetic aperture radar data with an adoptive Markov random field model", *Proc. IEEE Int. Geosci. Remote Sens. Symp (IGARSS 2012) Conf.*, Munich, Germany, 22-27 Jul. 2012.

Chapter 2

Synthetic Aperture Radar Imaging

In this chapter, we give an introduction to the synthetic aperture radar (SAR), followed by its geometric configuration and SAR spatial resolution. A detailed introduction to SAR can be found in text books such as [Oliver and Quegan, 2004], [Cumming and Wang, 2005], [Elachi and Van Zyl, 2006], [Curlander and McDonough, 1991], and [Massonnet and Souyris, 2008].

2.1 Synthetic Aperture Radar

SAR is a coherent and microwave imaging radar to obtain high spatial resolution two-dimensional (2-D) reflectivity images of the Earth's surface in nearly all weather conditions and independently of the day-night cycle. The imaging SAR system is an active radar system operating in the microwave region of the electromagnetic spectrum, usually between P-band and Ka-band, as presented in Table 2.1. The radar is usually mounted on a flying platform such as an airplane or a satellite and operates in a side-looking geometry with an illumination perpendicular to the flight line direction. Similar to other active systems, it emits microwave radiation to the ground and measures the electromagnetic signal backscattered from the illuminated area. The aim of SAR signal processing is to synthesize a 2-D high spatial resolution image of the Earth's surface reflectivity from all the received signals.

2.1.1 Imaging Geometry

Figure 2.1 illustrates the geometry of the imaging radar often employed for Earth observation. The radar antenna illuminates a surface trip to one side of the nadir track. The direction in which the platform moves is called the *azimuth* direction. The direction in which the radar transmits and receives radiation is called *range*. The radar transmits a cone-shaped microwave beam to the ground continuously with a side-looking angle θ in the direction perpendicular to the flying track (azimuth direction). This side-

Table 2.1: The microwave bands for radar remote sensing.

Band	Wavelength (cm)	Frequency (GHz)
K_α	0.75-1.1	40–26.5
K	1.1-1.67	26.5-18
K_u	1.67-2.4	18-12.5
X	2.4-3.75	12.5-8
C	3.75-7.5	8-4
S	7.5-15	4-2
L	15-30	2-1
P	30-100	1-0.3

looking geometry is important to eliminate right-left ambiguities from two symmetric equidistant points [Elachi and Van Zyl, 2006, Swart, 2000]: if they would see points on either side of the ground track, the radar would not be able to distinguish them, because they have the same range. The radial axis or radar-line-of-sight (RLOS) is referred to as *slant-range* (R). The area covered by the antenna beam in the *ground range* (Y) and azimuth (X) directions is the *antenna footprint*. This area consists of many small cells. The echo backscattered from each ground cell within the footprint is received and recorded as a pixel in the image plane according to the slant range between the antenna and the ground cell.

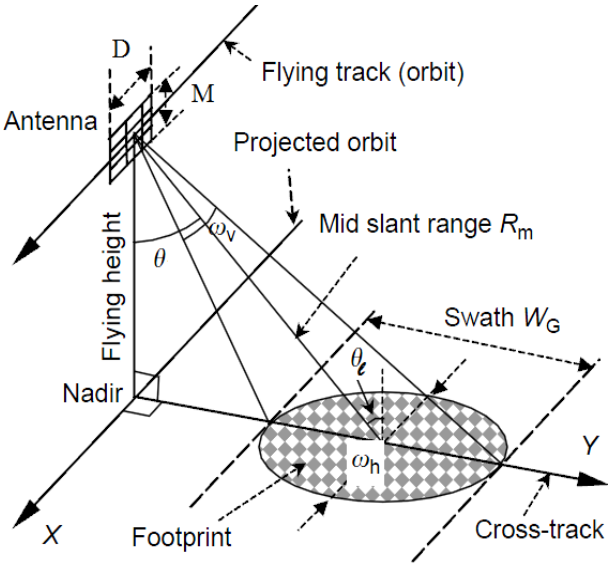


Figure 2.1: Radar imaging geometry-the sensor flies in the flying (azimuth) direction and looks in the cross-track (range) direction (figure taken from [Li et al., 2004]).

The platform moving along the azimuth direction provides the scanning. The area scanned by the antenna beam is called the *radar swath*. The antenna apertures (or named

the antenna beam width) in the cross-track (range) direction and the flying (azimuth) direction are given by

$$\omega_v \approx \frac{\lambda}{M} \qquad \omega_h \approx \frac{\lambda}{D} \tag{2.1}$$

where M and D correspond to the physical dimensions of the antenna in the range and azimuth directions, respectively, and λ is the wavelength corresponding to the carrier frequency of the transmitted signal. The swath W_G can be then approximated by:

$$W_G \approx \frac{\lambda R_m}{M \cos\theta} = \frac{\lambda H_{\text{sat}}}{M \cos^2\theta} \tag{2.2}$$

where R_m is the slant range from the center of the antenna to the center of the footprint; H_{sat} is the height of the satellite orbit above the Earth; and θ is the radar look angle.

2.1.2 SAR Spatial Resolution

The resolution is expressed as the minimum distance that two scatter points must have in order to be solved. The resolution of a radar image for Earth observation is defined by the azimuth resolution in the flying direction and the ground range resolution in the range direction, as illustrated in Figure 2.2. If an infinitely short pulse is transmitted toward a point target a distance R away, an infinitely short echo will be received at time $t = 2R/c$, where c is the speed of light. The factor 2 represents the fact the radar signal travels two times the distance R. If the pulse has a length τ, the echo will have a length τ. If there are two targets separated by a distance δ_r, the shortest separation δ_r which is measured as the range achievable resolution is given by [Elachi and Van Zyl, 2006]

$$\delta_r = \frac{c\tau}{2} = \frac{c}{2B} \tag{2.3}$$

where B is the bandwidth of the signal. If we project this onto the ground with incidence angle θ_ℓ, we get the ground range resolution which is coarser than the slant range resolution.

$$\delta_g = \frac{\delta_r}{\sin\theta_\ell} = \frac{c\tau}{2\sin\theta_\ell} \tag{2.4}$$

Thus, in order to achieve a resolution as high as possible, a short pulse or a wide-bandwidth pulse is required. The energy in a pulse is equal to

$$E = P\tau \tag{2.5}$$

where P is the instantaneous peak power. The energy in a pulse characterizes the capability of the pulse to detect a target, and a high pulse energy is desired. This can be obtained by increasing the peak power P. However, in particular for radar imaging satellites the maximum power is limited by the sensor hardware. The other possibility is to increase pulse duration to transmit sufficient energy to receive a certain backscattered

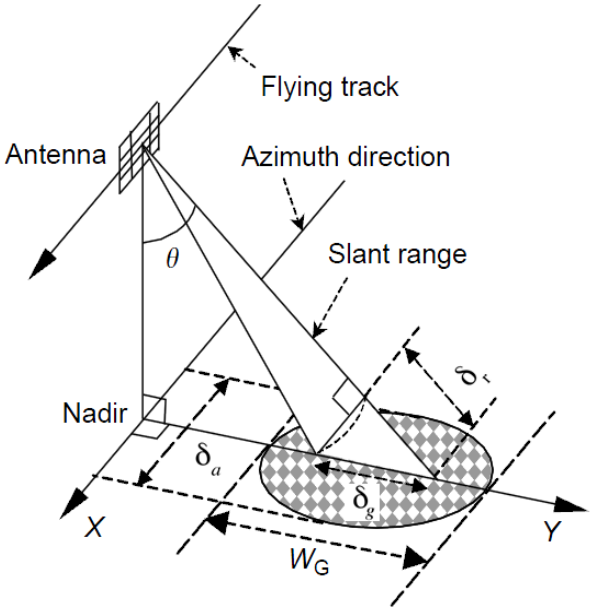


Figure 2.2: Resolution of radar images (figure taken from [Li et al., 2004]).

energy. But, according to (2.3) a long pulse, corresponds to a narrow bandwidth B, results in a poor range resolution. Thus, in order to have a high detection ability and a high resolution, a pulse with characteristics of large τ and large B is needed [Elachi and Van Zyl, 2006, Cumming and Wang, 2005]. This is made possible to change the nature of the pulse by modulating it during its transmission. In other word, a linear frequency modulated signal called the *chirp* is transmitted. Using a chirp the energy of the signal is spread over a large bandwidth B. The energy is compressed again in a short time interval in the receiver using a matching filter. By this way, the pulse duration and signal-to-noise ratio (SNR) are increased without decreasing the range resolution.

According to the electromagnetic wave theory, the azimuth resolution for conventional radar systems is given by [Curlander and McDonough, 1991]

$$\delta_a \approx \frac{R\lambda}{D} = \frac{H_{\text{sat}}\lambda}{D\cos\theta} \tag{2.6}$$

where D, as already denoted, is the length of the aperture of the radar antenna. To illustrate, if $H_{\text{sat}} = 785$ km, $\lambda = 5.66$ cm, $\theta = 23°$, and $D = 10$ m, then $\delta_a = 4.8$ km, which is considered a low resolution for imaging applications. To get an azimuth resolution of 10 m from 785 km away, the required length of its aperture is longer than 3 km. This is impossible for any flying platform to carry such a long antenna and when high resolution is desired, the real-aperture technique is not appropriate for such applications. In order to improve the azimuth resolution, a *synthetic-aperture* technique developed in the 1960s is used which is based on the construction of a longer effective antenna by moving the real sensor antenna along the flight direction [Oliver and Quegan, 2004]. It is based on the principle of the *Doppler frequency shift* [Wiley, 1954] caused by the relative

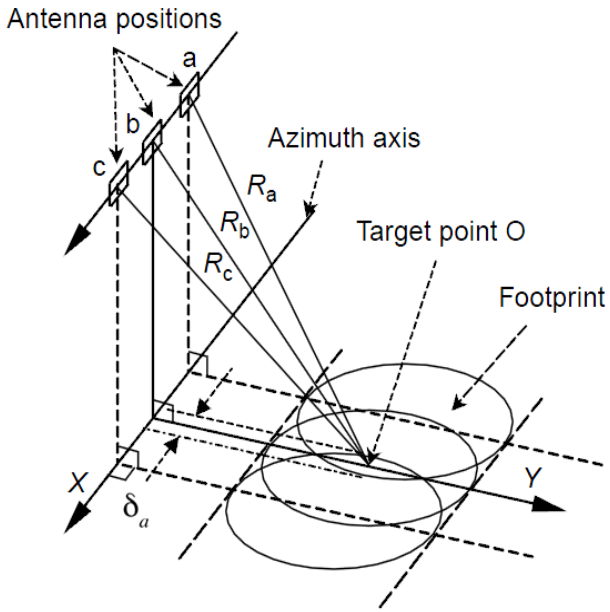


Figure 2.3: Imaging geometry of SAR-The scatterer point is seen by the antenna from different positions (figure taken from [Li et al., 2004]).

movement between the antenna and the target. This technique is based on the fact that the response of a scatterer is contained in more than one single radar echo. A scatterer point, in fact, remains in the antenna beam for a significant amount of time. So a scatter point is observed by the radar from different positions during the movement of the antenna on its orbit (Figure 2.3). Therefore, instead of usage of a large antenna, coherent combination of different echoes relative to a scatterer point realize a *synthetic enlarged antenna*, a sort of antenna array. Compared to the azimuth resolution of a real aperture radar, the azimuth resolution of the SAR is much improved and is given as:

$$\delta_a = \frac{D}{2} \tag{2.7}$$

This means that the azimuth resolution of an SAR is only determined by the length of the real aperture, independent of the distance between the sensor and the area being imaged. It is clear that a finer resolution can be obtained by making the real antenna length very small. This is true, but there some are ambiguities which place some certain limits on this, in particular, related to the pulse repetition frequency (PRF) and the ground coverage capability of the SAR [Elachi and Van Zyl, 2006]. The corresponding azimuthal resolution expression for an orbital SAR imaging system, which considers neither the Earth's curvature nor the curved flight path of the satellite.

$$\delta_a = \frac{R_E}{R_E + H_{\text{sat}}} \frac{D}{2} \tag{2.8}$$

where R_E is the Earth's radius.

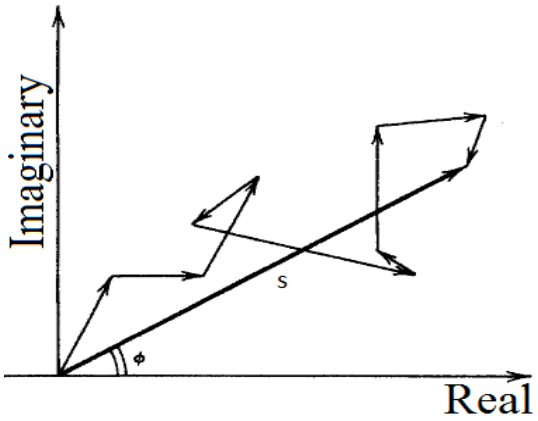


Figure 2.4: Composite return from an area with multiple scatters.

2.2 SAR Complex Images

Some specific signal processing operations are required to convert the collected raw data into into a well focused image. A SAR raw data is not an image yet since point targets are spread out in range and in azimuth. The echo of a target point is received from the moving antenna for a time defined as integration time. The SAR processor is to combine all these echoes coherently referred to the same target point received during the integration time. The focusing is needed both in azimuth and in range dimension, in order to create the image. After processing of SAR raw data with some advanced techniques such as *Omega-k* and *Range-Doppler* algorithms [Cumming and Wang, 2005], each pixel of the SAR image contains not only the gray value (i.e., amplitude image) but also the phase value related to the radar slant range. These two components can be expressed by a complex number. Therefore, the SAR image can also be called a radar complex image.

2.2.1 Speckle

The coherent interference of waves reflected from many small elementary scatterers[1] generates the so-called *speckle*. This effect causes a pixel-to-pixel variation in intensity even over homogeneous areas, and gives the SAR image its noisy appearance. This effect is a sort of *salt and pepper* screen superimposed on a uniform amplitude image and a phase randomly distributed [Ferretti et al., 2007]. In distributed targets, each resolution cell contains a number of discrete scatterers (Figure 2.4). As the wave interacts with the target, each scatterer contributes a backscattered wave with a phase and amplitude change, so the total returned modulation of the incident wave is

$$S = |S|e^{j\phi} = \sum_{i=1}^{N} |S_i|e^{j\phi_i} \tag{2.9}$$

[1] Those with a dimension on the scale of the radar wavelength

i being the index of individual scatterer with amplitude $|S_i|$ and phase ϕ_i, and N the number of scatterers in the resolution cell [Sarbandi, 1992, Jakeman and Pusey, 1976, Elachi and Van Zyl, 2006, Oliver and Quegan, 2004].

Figure 2.5: ERS-2 SAR detected image of the Linate airport. The speckle effect is clearly visible on the homogeneous fields nearby the airport (figure taken from [Ferretti et al., 2007]).

Figure 2.6: Average of multiple ERS SAR images of the Linate airport. The speckle is suppressed on the homogeneous fields nearby the airport (figure taken from [Ferretti et al., 2007]).

An example of speckle is shown in Figure 2.5 in which the salt and pepper effect is visible by visual inspection on homogeneous fields surrounding the Linate Airport. The speckle in SAR images complicates the interpretation of images and influences the usefulness of SAR images. Generally, image classification suffers severely from speckle. One optimum solution to reduce speckle is taking more images of the same area at different times or from slightly different look angles, provided that land cover change

does not occur significantly. Figure 2.6 shows an example of speckle reduction. The average of 60 SAR images acquired bye ERS-1 and ERS-2 of the area surrounding the Linate airport in Milan is illustrated. Another way of speckle reduction is an incoherent averaging of neighboring pixels together to spread the aberration out at the cost of resolution. Several algorithms have been developed for speckle reduction in the last decade [Lee et al., 1991], [Lopes et al., 1993], [Touzi and Lopes, 1994].

2.2.2 SAR Geometrical Effects

When terrain is imaged using a radar sensor, each SAR resolution cell depends on the look angle and the local topography. The ground objects are imaged as a function of their distance from the antenna - not as a function of their distance from each other on the ground. Figure 2.7 shows how slant-range is projected onto the ground. Side-looking geometry of the radar causes the three inherent distortions: *foreshortening*, *layover* and *shadowing*.

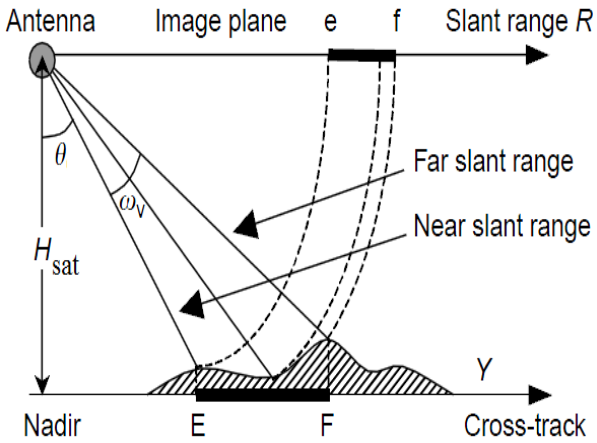


Figure 2.7: Projection of radar image. The part of the terrain imaged in each resolution cell clearly depends on the surface topography (figure taken from [Li et al., 2004]).

As the terrain slope increases with respect to a flat horizontal surface (forslope areas, i.e., the normal to the ground moves toward the RLOS direction), the ground resolution cell dimension in range increases. This effect is called *foreshortening*. Suppose that points A, B, and C, which are equally spaced on the ground, are imaged by the radar sensor. Because the SAR is viewing from an angle, a cross-track compression of the radiometric information backscattered from foreslope areas is inevitable [Lee and Pottier, 2009]. As seen in Figure 2.8, the length of $A^{'}B^{'}$ is considerably shortened compared to the length of $B^{'}C^{'}$ on the ground providing a tilting of the top of the mountain towards the sensor.

When the terrain slope exceeds the radar look angle, the scatterers are imaged in reverse order and superimposed on the contribution coming from other areas. This effect is called *layover* and is illustrated in Figure 2.8. The point B on the top of the mountain is closer to the radar sensor, therefore it is imaged on the image plane earlier

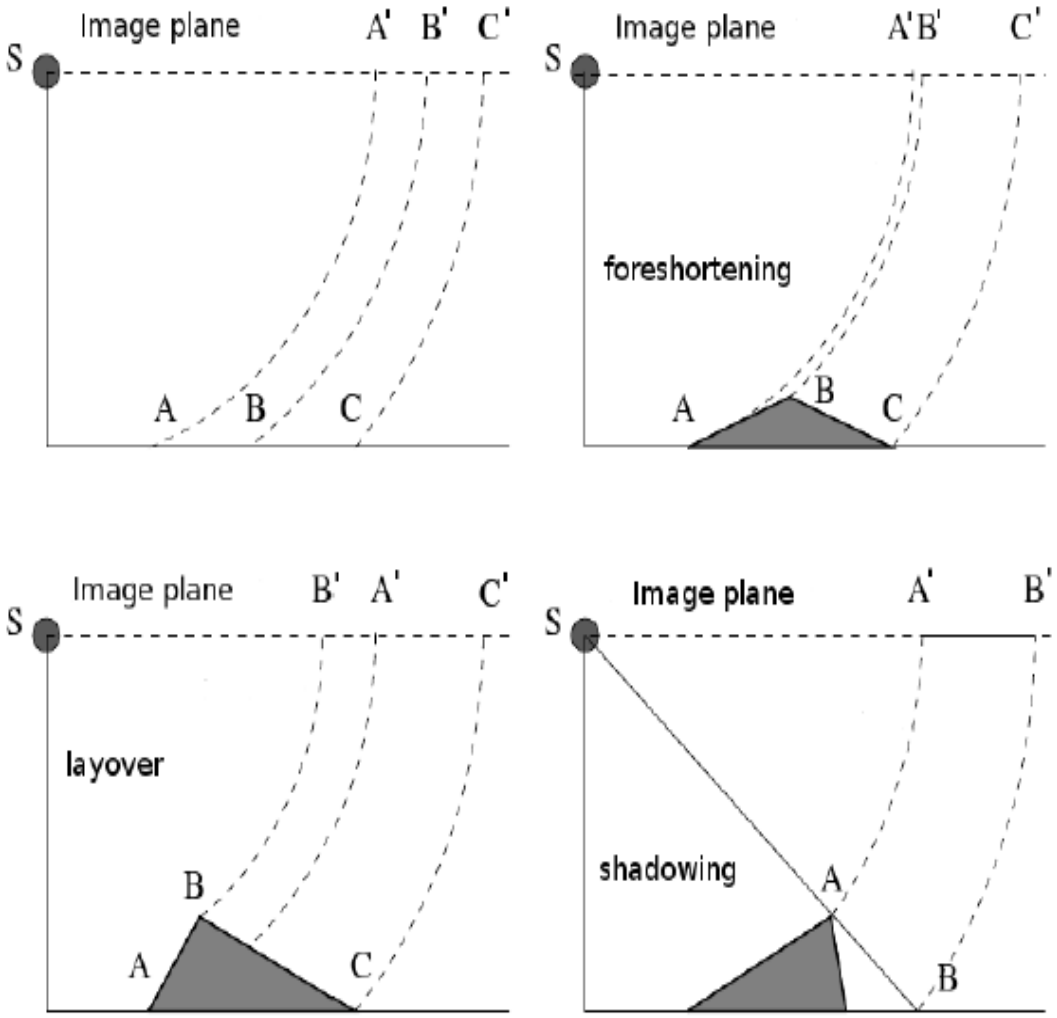


Figure 2.8: SAR geometrical distortions - Foreshortening, layover and shadowing effects in areas with strong topography.

than the point A at the bottom of the mountain. Consequently, the echo from AB is distributed in a reverse order on the image plane ($A^{'}B^{'}$), as shown in Figure 2.8..

The other effect of vertical structures is to produce shadows in the SAR image. Because of the side-looking geometry there will be a region of ground behind the vertical structure that the beam cannot reach. Thus for the time period corresponding to that ground area, no echoes will be returned. This results in a black area on the image which is called *shadow*. As seen in Figure 2.8, Backscattered information is lost and no signal is collected from point A to point B. This causes an area $A^{'}B^{'}$ on the image plane containing only systems noise. The minimum slant range resolution is achieved when the terrain is parallel to the RLOS which is the lower slope limit for a SAR system, since below this angle the terrain is in shadow [Ferretti et al., 2007].

Chapter 3

SAR Polarimetry

In this chapter, the principle of polarimetric SAR imaging is briefly introduced to give the necessary keys for understanding polarimetric measurements. Many studies have concerned only with the information carried by a single SAR image. SAR systems with multiple frequencies or polarizations provide a much enhanced capacity for extracting information from the images.

3.1 Scattering Matrix

SAR polarimetry is concerned with exploring the target properties from the behaviors of backscattered polarized electromagnetic waves. The effects of the interactions between the electromagnetic waves and the observed targets are associated with imaging systems such as the frequency of the radar signal, polarization, incident angle or orientation of the target with respect to the radar antenna, scattering directions and target characteristics such as geometrical structure and dielectric properties [Elachi and Van Zyl, 2006].

Polarimetric SAR systems use antennae designed to transmit and receive electromagnetic waves of a specific polarization, being the two most common ones the horizontal linear or H, and vertical linear or V. Due to the possible change in polarization of the scattered wave, radar antennae are designed to receive the different polarization components simultaneously and, therefore, HH, VV, HV and VH data will be available in a full polarimetric system [Lee and Pottier, 2009], see Figure 3.1. Any polarization state of the transmitted wave can be described by an electric field vector of the form

$$\underline{\mathbf{E}}^s = E_h^s \widehat{u}_h + E_v^s \widehat{u}_v, \tag{3.1}$$

or equivalently

$$\underline{\mathbf{E}}^s = \begin{bmatrix} E_h^s \\ E_v^s \end{bmatrix}, \tag{3.2}$$

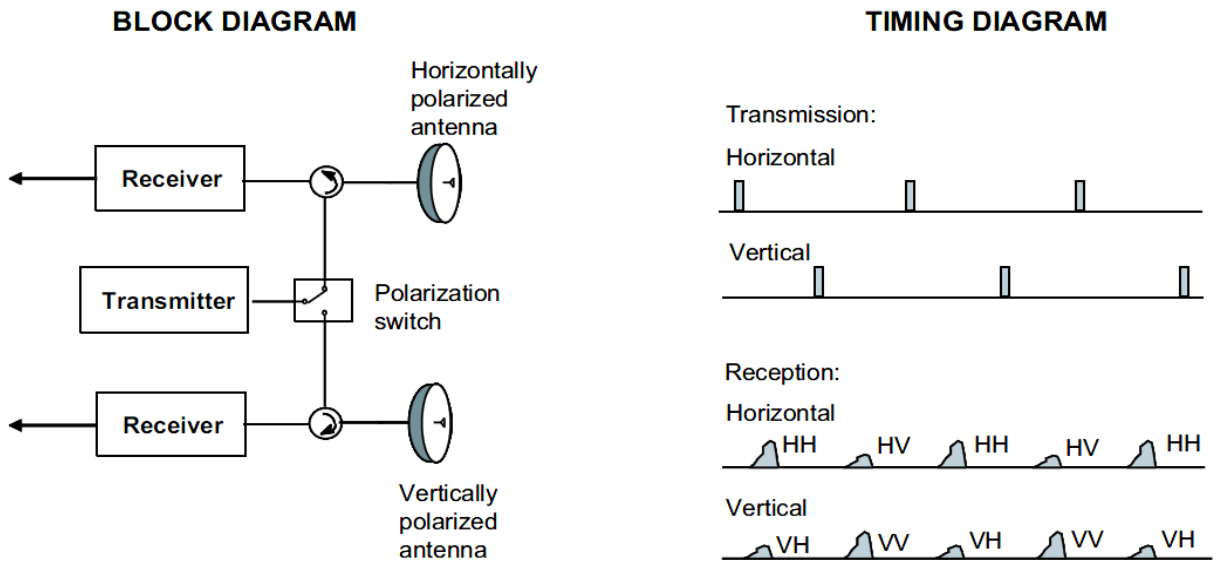


Figure 3.1: A polarimetric radar is implemented by alternatively transmitting signals out of horizontally and vertically polarized antennas, and receiving at both polarizations simultaneously. Two pulses are needed to measure all the elements in the scattering matrix (figure taken from [Elachi and Van Zyl, 2006]).

where the subscript i denotes incident and $(\widehat{u}_h, \widehat{u}_v)$ are two unit vectors, i.e., the linear horizontal and vertical polarization basis: H, V [Oliver and Quegan, 2004]. The electric fields $\underline{\mathbf{E}}^s$ and $\underline{\mathbf{E}}^i$ of the scattered wave and the incident wave are related by a complex 2×2 scattering matrix $[\mathbf{S}]$ associated with each resolution cell in the image according to [Lee and Pottier, 2009, Touzi et al., 2004]

$$\begin{aligned} \underline{\mathbf{E}}^s &= \frac{e^{-jkR}}{R} \cdot [\mathbf{S}] \cdot \underline{\mathbf{E}}^i \\ &= \begin{bmatrix} E_h^s \\ E_v^s \end{bmatrix} = \frac{e^{-jkR}}{R} \cdot [\mathbf{S}] \cdot \begin{bmatrix} E_h^i \\ E_v^i \end{bmatrix}, \end{aligned} \tag{3.3}$$

Here k denotes the wavenumber, R is the distance between radar and target, and $j = \sqrt{-1}$ is the imaginary unit. The superscript of the electromagnetic field components indicates incident (i) or scattered (s) wave. The term $\frac{e^{-jkR}}{R}$ takes into account the propagation effects both in amplitude and phase. Expression (3.3) is only valid for the far field zone, where the planar wave assumption is considered for the incident and the scattered fields [Lee and Pottier, 2009]. The $[\mathbf{S}]$ matrix is called *Sinclair matrix* [Sinclair, 1950] or target scattering matrix. This matrix for a monostatic SAR system, in which the same antenna transmits and receives, is given in the horizontal-vertical polarization basis as

follows:

$$\begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix}, \tag{3.4}$$

where thescattering coefficients S_{rt} are subscripted with the associated receive (r) and transmit (t) polarization, which is horizontal (H) or vertical (V). It is worth to note that the backscattered alignment (BSA) convention is the preferred system in the area of monostatic SAR polarimetry. Matrix elements denote the corresponding components of the backscattered electric fields, with the first subscript indicating the polarization of the transmitted electric field and the second subscript indicating the polarization of the detected component of the backscattered electric field. Terms S_{hh} and S_{vv} are named “co-polar", whereas terms S_{hv} and S_{vh} are named “cross-polar".

If the observed media contain reciprocal material and if the wave propagation between the radar and the ground does not involve non-reciprocal phenomena (such as when low frequencies cross the ionosphere), then applying the *reciprocity* theorem [Kong, 1990] to the target-radar system implies that the off-diagonal terms of $[\mathbf{S}]$ are equal, $S_{hv} = S_{vh}$ [Massonnet and Souyris, 2008, Lee et al., 1994b, Cloude and Pottier, 1996]. The scattering matrix is consequently a symmetric matrix, defined by six parameters, three amplitude terms and three phase terms. If we factor out a term of absolute phase, the number of independent parameters is reduced to five parameters: the three amplitudes and the two relative phases [Lee and Pottier, 2009, Massonnet and Souyris, 2008]:

$$\begin{aligned} [\mathbf{S}] &= \begin{bmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{bmatrix} \\ &= e^{j\cdot\phi_{hh}} \cdot \begin{bmatrix} |S_{hh}| & |S_{hv}| \cdot e^{j(\phi_{hv}-\phi_{hh})} \\ |S_{hv}| \cdot e^{j(\phi_{hv}-\phi_{hh})} & |S_{vv}| \cdot e^{j(\phi_{vv}-\phi_{hh})} \end{bmatrix}, \end{aligned} \tag{3.5}$$

3.2 Scattering Vector

The scattering matrix $[\mathbf{S}]$ can be vectorized and represented as the target scattering vector $\underline{\mathbf{k}} = [k_0, k_1, k_2, k_3]^T$ by:

$$k_i = \mathrm{tr}([\mathbf{S}] \cdot [\mathbf{\Psi}_i]), \tag{3.6}$$

where the superscript T is the matrix transpose and $\mathrm{tr}(\cdot)$ denotes the trace operator and $[\mathbf{\Psi}_i]$ is a complete set of 2×2 basis matrices which are constructed as an orthogonal set under the Hermitian inner product [Lee and Pottier, 2009]. There are two groups for choosing the structure of $[\mathbf{\Psi}_i]$, from either the $\{\mathbf{\Psi}_P\}$ or the $\{\mathbf{\Psi}_L\}$ families. The first group is made of from the Pauli basis matrices set $\{\mathbf{\Psi}_P\}$ given by

$$\{\mathbf{\Psi}_P\} = \left\{ \frac{1}{\sqrt{2}} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \frac{1}{\sqrt{2}} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \frac{1}{\sqrt{2}} \cdot \begin{bmatrix} 0 & -j \\ j & 0 \end{bmatrix} \right\}, \tag{3.7}$$

and the corresponding *4-D Pauli scattering vector* becomes

$$\underline{\mathbf{k}} = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{hh} + S_{vv} \\ S_{hh} - S_{vv} \\ S_{hv} + S_{vh} \\ j(S_{hv} - S_{vh}) \end{bmatrix}. \tag{3.8}$$

When Assuming that the reciprocity, $S_{hv} = S_{vh}$, the Pauli basis will only contain the first three matrices from $\{\mathbf{\Psi}_P\}$ family given in (3.7), and we obtain *3-D Pauli scattering vector*

$$\underline{\mathbf{k}} = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{hh} + S_{vv} \\ S_{hh} - S_{vv} \\ 2S_{hv} \end{bmatrix}. \tag{3.9}$$

The second group is the Lexicographic matrix basis set $\{\mathbf{\Psi}_L\}$ given by

$$\{\mathbf{\Psi}_L\} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}, \tag{3.10}$$

and the corresponding *4-D Lexicographic scattering vector* becomes

$$\underline{\mathbf{\Omega}} = \begin{bmatrix} S_{hh} \\ S_{hv} \\ S_{vh} \\ S_{vv} \end{bmatrix}, \tag{3.11}$$

For a reciprocal medium, in the monostatic backscattering case, the Lexicographic matrix basis set, $\{\mathbf{\Psi}_L\}$ is given as

$$\{\mathbf{\Psi}_L\} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \sqrt{2} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}, \tag{3.12}$$

and the corresponding *3-D Lexicographic basis scattering vector* becomes

$$\underline{\mathbf{\Omega}} = \begin{bmatrix} S_{hh} \\ \sqrt{2}S_{hv} \\ S_{vv} \end{bmatrix}, \tag{3.13}$$

where $\sqrt{2}$ in (3.12) and (3.13) arises from the requirement to conserve the total scattered power, so that

$$\begin{aligned} \text{Span} &= |S_{hh}|^2 + 2|S_{hv}|^2 + |S_{vv}|^2 \\ &= \underline{\mathbf{k}}^H \cdot \underline{\mathbf{k}} = |\underline{\mathbf{k}}|^2 \\ &= \underline{\mathbf{\Omega}}^H \cdot \underline{\mathbf{\Omega}} = |\underline{\mathbf{\Omega}}|^2 \end{aligned} \tag{3.14}$$

where the superscript $(\cdot)^H$ stands for the Hermitian or complex conjugate transpose. The vectors $\underline{\mathbf{k}}$ and $\underline{\mathbf{\Omega}}$ are single-look complex (SLC) format representations of polarimetric SAR data. The transformation between these two vectors is given as [Cloude, 1986, Lee and Pottier, 2009]

$$\underline{\mathbf{k}} = U_{4(L\rightarrow P)}\underline{\mathbf{\Omega}}, \tag{3.15}$$

where $U_{4(L\rightarrow P)}$ is a special unitary SU(4) transformation from the Lexicographic scattering vector to the Pauli scattering vector as follows

$$\mathbf{U}_{4(L\rightarrow P)} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & j & -j & 0 \end{bmatrix} \tag{3.16}$$

and satisfies $|\mathbf{U}_{4(L\rightarrow P)}| = +1$ and $\mathbf{U}^{-1}_{4(L\rightarrow P)} = \mathbf{U}^{H}_{4(L\rightarrow P)}$. The unitary transformation matrix between the two 3-D polarimetric scattering vectors is given as

$$\mathbf{U}_{3(L\rightarrow P)} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \end{bmatrix} \tag{3.17}$$

3.3 Radar Brightness

This section provides the default product calibrations that we will need later for radiometric terrain correction of the multilook radar data in Section 5.2.

The radar brightness (RB), β^0, is defined as the radar cross section (RCS) per unit image area in range-azimuth coordinates [Frey et al., 2013]. The RB of a given target measured in the slant range plane for linear polarizations $r, t \in \{H, V\}$ can be obtained from the elements of target scattering matrix as follows [Freeman, 1992, Oliver and Quegan, 2004]:

$$\beta^0_{rt} = \frac{4\pi|S_{rt}|^2}{A_\beta}, \tag{3.18}$$

where A_β represents the reference area of β^0 backscatter coefficient which contributes to the recorded signal. Both S_{rt} and β^0_{rt} are functions of wave frequency, viewing geometry, wave polarization, geometrical structure and dielectric properties of the object [Lee and Pottier, 2009]. The area normalized scattering vector in the slant range plane (subscript β^0) is defined as [Akbari et al., 2012a]:

$$\underline{\mathbf{\Omega}}_{\beta^0} = \frac{\underline{\mathbf{\Omega}}_\beta}{\sqrt{A_\beta}}. \tag{3.19}$$

3.4 Multilook Complex Data

Usually, polarimetric data are transformed into the form of multilooked sample covariance matrices in order to reduce the statistical variation due to speckle effects and data compression of SAR data by spatial averaging of several neighboring SLC image pixels. From the target scattering vector, the polarimetric coherency matrix and the covariance matrix are generated from the complex product of the associated target vector with its conjugate transpose. The L-look sample covariance matrix is then defined as

$$\begin{aligned}\mathbf{C} &= \frac{1}{L}\sum_{\ell=1}^{L}\underline{\boldsymbol{\Omega}}_{\ell}\cdot\underline{\boldsymbol{\Omega}}_{\ell}^{H} \\ &= \begin{bmatrix} \langle|S_{hh}|^2\rangle & \langle S_{hh}S_{hv}^*\rangle & \langle S_{hh}S_{vh}^*\rangle & \langle S_{hh}S_{vv}^*\rangle \\ \langle S_{hv}S_{hh}^*\rangle & \langle|S_{hv}|^2\rangle & \langle S_{hv}S_{vh}^*\rangle & \langle S_{hv}S_{vv}^*\rangle \\ \langle S_{vh}S_{hh}^*\rangle & \langle S_{vh}S_{hv}^*\rangle & \langle|S_{vh}|^2\rangle & \langle S_{vh}S_{vv}^*\rangle \\ \langle S_{vv}S_{hh}^*\rangle & \langle S_{vv}S_{hv}^*\rangle & \langle S_{vv}S_{vh}^*\rangle & \langle|S_{vv}|^2\rangle \end{bmatrix},\end{aligned} \tag{3.20}$$

where $(\cdot)^*$ means the complex conjugation, L is the nominal number of looks used for averaging and $\langle\cdots\rangle$ denotes temporal and spatial ensemble averaging, and $|\cdot|$ refers to absolute magnitude. The diagonal elements of the covariance matrix represent the multilook intensities of the linear polarizations and the off-diagonal elements are the complex cross-correlation between channels and may also hold valuable polarimetric information. Hence, after multilooking, each pixel in the image is a realization of the $d\times d$ stochastic matrix variable denoted $\mathbf{C}$, and the image is referred to as the multilook complex (MLC) covariance image. For a reciprocal medium, the 4-D covariance matrix reduces to the 3-D covariance matrix

$$\begin{aligned}\mathbf{C} &= \frac{1}{L}\sum_{\ell=1}^{L}\underline{\boldsymbol{\Omega}}_{\ell}\cdot\underline{\boldsymbol{\Omega}}_{\ell}^{H} \\ &= \begin{bmatrix} \langle|S_{hh}|^2\rangle & \langle\sqrt{2}S_{hh}S_{hv}^*\rangle & \langle S_{hh}S_{vv}^*\rangle \\ \langle\sqrt{2}S_{hv}S_{hh}^*\rangle & \langle 2|S_{hv}|^2\rangle & \langle\sqrt{2}S_{hv}S_{vv}^*\rangle \\ \langle S_{vv}S_{hh}^*\rangle & \langle\sqrt{2}S_{vv}S_{hv}^*\rangle & \langle|S_{vv}|^2\rangle \end{bmatrix}.\end{aligned} \tag{3.21}$$

And the polarimetric coherency matrix is constructed from the spatial average of $\underline{\mathbf{k}}$:

$$\mathbf{T} = \frac{1}{L}\sum_{\ell=1}^{L}\underline{\mathbf{k}}_{\ell}\cdot\underline{\mathbf{k}}_{\ell}^{H} \tag{3.22}$$

The two matrices $[\mathrm{C}]$ and $[\mathbf{T}]$ carry the same information about polarimetric scattering amplitudes, phase angles, and correlations; both are Hermitian positive semidefinite and both have the identical eigenvalues but different eigenvectors [Cloude and Pottier, 1996]. It is worth to note that it is always feasible to generate the MLC data from the SLC data by multilook averaging, but the averaging is not reversible, so it is not

always possible to re-generate the unique SLC scattering matrix back from the MLC covariance matrix, because the multilook averaging is a many-to-one operation and thus irreversible [Elachi and Van Zyl, 2006].

There is also another representation of the polarimetric covariance matrix given in (3.21) [Ziegler et al., 1992]

$$\mathbf{C} = \sigma_{hh} \cdot \begin{bmatrix} 1 & \sqrt{2} \cdot \varepsilon \cdot \sqrt{e} & \varrho \cdot \sqrt{g} \\ \sqrt{2} \cdot \varepsilon^* \cdot \sqrt{e} & 2 \cdot e & \sqrt{2} \cdot \varkappa \cdot \sqrt{e \cdot g} \\ \varrho^* \cdot \sqrt{g} & \sqrt{2} \cdot \varkappa^* \cdot \sqrt{e \cdot g} & g \end{bmatrix}, \tag{3.23}$$

where this parametric form of covariance matrix $[\mathrm{C}]$ includes nine parameters:

- an energy parameter
$$\sigma_{hh} = \langle S_{hh} \cdot S_{hh}^* \rangle, \tag{3.24}$$

- two real parameters (e and g)
$$g = \frac{\langle S_{vv} \cdot S_{vv}^* \rangle}{\langle S_{hh} \cdot S_{hh}^* \rangle} \quad e = \frac{\langle S_{hv} \cdot S_{hv}^* \rangle}{\langle S_{hh} \cdot S_{hh}^* \rangle}, \tag{3.25}$$
where g and e characterize the ratios of co-polarized energies [Massonnet and Souyris, 2008, Lee and Pottier, 2009]

- three complex numbers (thus six real parameters) ϱ, ε and $\varkappa$,
$$\varrho = \frac{\langle S_{hh} \cdot S_{vv}^* \rangle}{\sqrt{\langle S_{hh} \cdot S_{hh}^* \rangle \cdot \langle S_{vv} \cdot S_{vv}^* \rangle}} \quad \varepsilon = \frac{\langle S_{hh} \cdot S_{hv}^* \rangle}{\sqrt{\langle S_{hh} \cdot S_{hh}^* \rangle \cdot \langle S_{hv} \cdot S_{hv}^* \rangle}} \tag{3.26}$$
$$\varkappa = \frac{\langle S_{vv} \cdot S_{hv}^* \rangle}{\sqrt{\langle S_{vv} \cdot S_{vv}^* \rangle \cdot \langle S_{hv} \cdot S_{hv}^* \rangle}}$$
which are the complex cross-correlation between channels (polarimetric intercorrelation) and quantify the amount of similarity of responses backscattered by a target under different transmission and reception conditions [Massonnet and Souyris, 2008].

In the case of a natural media, such as soil and forest, the correlation between co- and cross-polarized channels is assumed to be zero, referred to as reflection symmetry [van Zyl, 1993], and the corresponding covariance matrix $[\mathrm{C}]$ can then be expressed as

$$\mathbf{C} = \begin{bmatrix} \langle |S_{hh}|^2 \rangle & 0 & \langle S_{hh} S_{vv}^* \rangle \\ 0 & \langle 2|S_{hv}|^2 \rangle & 0 \\ \langle S_{vv} S_{hh}^* \rangle & 0 & \langle |S_{vv}|^2 \rangle \end{bmatrix} = \sigma_{hh} \cdot \begin{bmatrix} 1 & 0 & \varrho \cdot \sqrt{g} \\ 0 & 2 \cdot e & 0 \\ \varrho^* \cdot \sqrt{g} & 0 & g \end{bmatrix}, \tag{3.27}$$

Therefore, the only remaining cross-correlation is ϱ, which expresses the correlation between the co-polarized channels.

3.5 Models for Polarimetric Data

It has been verified that the product (or named multiplicative) model [Lee et al., 1994b, Oliver and Quegan, 2004, Freitas et al., 2005, Eltoft et al., 2006] is proper to represent the statistical properties of SAR data which decomposes the target scattering vector as a product of two independent stochastic variables

$$\underline{\mathbf{\Omega}} = \sqrt{Z}\underline{\mathbf{Y}}, \tag{3.28}$$

where the strictly positive, unit mean scalar random variable Z models texture, and represents the backscatter variability due to heterogeneity of the sensed area. The texture term is scalar because of the assumption of equal textural variations for all polarimetric channels. The random vector $\underline{\mathbf{Y}}$ represents the polarimetric speckle noise and follows a circular complex multivariate Gaussian distribution [van den Bos, 1995], denoted as $\mathcal{N}_d^{\mathbb{C}}(\mathbf{0}, \mathbf{\Sigma})$, with zero mean and covariance matrix $\mathbf{\Sigma}$, and polarimetric dimension d. The probability density function (PDF) of $\underline{\mathbf{Y}}$ is given as

$$p_{\underline{\mathbf{Y}}}(\underline{\mathbf{Y}}) = \frac{1}{\pi^d |\mathbf{\Sigma}|} \exp\left(-\underline{\mathbf{Y}}^H \mathbf{\Sigma}^{-1} \underline{\mathbf{Y}}\right), \tag{3.29}$$

where $|\cdot|$ is the determinant of a matrix. As the random variable Z models the variance of the signal rather than its amplitude, the texture component is introduced with a square root [Oliver and Quegan, 2004]. Using the definition given in (3.21) with the decomposition presented in (3.28), the multilook covariance matrix becomes

$$\mathbf{C} = \frac{1}{L}\sum_{\ell=1}^{L} \underline{\mathbf{\Omega}}_\ell \cdot \underline{\mathbf{\Omega}}_\ell^H = \frac{1}{L}\sum_{\ell=1}^{L} Z_\ell \cdot (\underline{\mathbf{Y}}_\ell \cdot \underline{\mathbf{Y}}_\ell^H). \tag{3.30}$$

The texture variable is assumed constant over each area for which the pixels are being averages. Therefore, for a small number of looks, we can assume that Z_ℓ is independent of L for pixels included in the multilook averaging i.e., $Z_\ell = Z$ for every ℓ [Lee et al., 1994b, Lopes and Sery, 1997, Frery et al., 2010, Doulgeris et al., 2008] and then (3.30) becomes

$$\mathbf{C} = \frac{Z}{L}\sum_{l=1}^{L} \underline{\mathbf{Y}}_\ell \cdot \underline{\mathbf{Y}}_\ell^H = Z\mathbf{W}. \tag{3.31}$$

It follows from the Gaussian assumption that if $L \geq d$ and the $\{Y_\ell\}_{\ell=1}^L$ are independent, then $\mathbf{Y} = \sum_{\ell=1}^L \underline{\mathbf{Y}}_\ell \cdot \underline{\mathbf{Y}}_\ell^H$, follows the complex Wishart distribution [Goodman, 1963, Lee et al., 1994a], denoted $\mathcal{W}(d, L, \mathbf{\Sigma})$, with parameters L, the nominal number of looks, and $\mathbf{\Sigma} = E\{\underline{\mathbf{Y}}_\ell \cdot \underline{\mathbf{Y}}_\ell^H\} = E\{\mathbf{Y}\}$. The PDF of $\mathbf{Y}$ is given as

$$p_{\mathbf{Y}}(\mathbf{Y}) = \frac{|\mathbf{Y}|^{L-d}}{\Gamma_d(L)|\mathbf{\Sigma}|^L} \exp\left(-\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{Y})\right), \tag{3.32}$$

Table 3.1: Texture and covariance matrix distributions under the multilook product model given in (3.31) [Anfinsen et al., 2011]

$p_Z(z)$ of texture variable Z		$p_{\mathbf{C}}(\mathbf{C})$ of covariance matrix $\mathbf{C}$	
Cons.	$\delta(z-1)$	$s\mathcal{W}_d^{\mathbb{C}}(L,\boldsymbol{\Sigma})$	$\frac{L^{Ld}\lvert\mathbf{C}\rvert^{L-d}}{\Gamma_d(L)\lvert\boldsymbol{\Sigma}\rvert^L}\exp\left(-L\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C})\right)$
$\bar{\gamma}(\alpha)$	$\frac{\alpha^\alpha}{\Gamma(\alpha)}z^{\alpha-1}\exp\left(-\alpha z\right)$	$\mathcal{K}_d(L,\boldsymbol{\Sigma},\alpha)$	$\frac{2\lvert\mathbf{C}\rvert^{L-d}}{\lvert\boldsymbol{\Sigma}\rvert^L\Gamma_d(L)\Gamma(\alpha)}\left(L\alpha\right)^{\frac{\alpha+Ld}{2}}\left(\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C})\right)^{\frac{\alpha-Ld}{2}}K_{\alpha-Ld}\left(2\sqrt{\alpha L\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C})}\right)$
$\bar{\gamma}^{-1}(\eta)$	$\frac{(\eta-1)^\eta}{\Gamma(\eta)}\frac{1}{z^{\eta+1}}\exp\left(-\frac{\eta-1}{z}\right)$	$\mathcal{G}_d^0(L,\boldsymbol{\Sigma},\eta)$	$\frac{L^{Ld}\lvert\mathbf{C}\rvert^{L-d}}{\Gamma_d(L)\lvert\boldsymbol{\Sigma}\rvert^L}\frac{\Gamma(Ld+\eta)(\eta-1)^\eta}{\Gamma(\eta)}\left(L\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C})+\eta-1\right)^{-\eta-Ld}$
$\overline{\mathcal{FS}}(\xi,\zeta)$	$\frac{\Gamma(\xi+\zeta)}{\Gamma(\xi)\Gamma(\zeta)}\frac{\xi}{\zeta-1}\frac{\left(\frac{\xi}{(\zeta-1)}z\right)^{\xi-1}}{\left(\frac{\xi}{\zeta-1}z+1\right)^{\xi+\zeta}}$	$\mathcal{U}_d(L,\boldsymbol{\Sigma},\xi,\zeta)$	$\frac{L^{Ld}}{\Gamma_d(L)}\frac{\lvert\mathbf{C}\rvert^{L-d}}{\boldsymbol{\Sigma}^L}\frac{\Gamma(\xi+\zeta)}{\Gamma(\xi)\Gamma(\zeta)}\left(\frac{\xi}{\zeta-1}\right)^{Ld}\Gamma(Ld+\zeta)\times U\left(Ld+\zeta,Ld-\xi+1,L\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C})\xi/(\zeta-1)\right)$

where $\Gamma_d(L)$ is a normalization constant

$$\Gamma_d(L)=\pi^{\frac{d(d-1)}{2}}\prod_{i=1}^{d}\Gamma(L-i+1), \tag{3.33}$$

named the complex-kind multivariate gamma function in [Anfinsen et al., 2009], while $\Gamma(\cdot)$ is the standard Euler gamma function. $L\geq d$ ensures that $\mathbf{W}$ is nonsingular. Due to normalization by L, the speckle noise term $\mathbf{W}$ follows a *scaled complex Wishart* distribution [Anfinsen and Eltoft, 2011], denoted $s\mathcal{W}(d,L,\boldsymbol{\Sigma})$, whose PDF is

$$p_{\mathbf{W}}(\mathbf{W})=\frac{L^{Ld}\lvert\mathbf{W}\rvert^{L-d}}{\Gamma_d(L)\lvert\boldsymbol{\Sigma}\rvert^L}\exp\left(-L\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{W})\right). \tag{3.34}$$

The marginal distribution for $\mathbf{C}$ may be obtained by integrating the conditional PDF over the prior distribution of Z, that is

$$p_{\mathbf{C}}(\mathbf{C})=\int_0^\infty p_{\mathbf{C}|Z}(\mathbf{C}|z)p_Z(z)dz,\qquad \mathbf{C}|z\sim s\mathcal{W}(L,z\boldsymbol{\Sigma}). \tag{3.35}$$

The PDF of $\mathbf{C}$ depends on the specific model for the scalar texture variable Z. Several distributions have been proposed in the literature, for the backscatter Z, in order to model *homogeneous*, *heterogeneous* and *extremely heterogeneous* clutter. Tabel 3.1 lists the Dirac delta function, $\delta(z-1)$, gamma ($\bar{\gamma}$), inverse gamma ($\bar{\gamma}^{-1}$) and Fisher-Snedecor ($\overline{\mathcal{FS}}$) distribution as possible choices of $p_Z(z)$ and the resulting distributions for $\mathbf{C}$ as follows, depending on the homogeneity of the target.

- Homogeneous surfaces: the texture random variable is constant, i.e., the Dirac delta function $\delta(z-1)$ and the PDF for $\mathbf{C}$ is the scaled Wishart distribution.
- Heterogeneous surfaces: in this case the texture is assumed to be a gamma distribution, denoted $\bar{\gamma}(\alpha)$, leading to the $\mathcal{K}_d$ model [Jakeman and Pusey, 1976,Jakeman, 1980,Jakeman and Tough, 1987] for $\mathbf{C}$. Usefulness of of the multivariate multilook

polarimetric $\mathcal{K}_d$ model and its use in analysis and clustering are presented in [Yueh et al., 1989, Lee et al., 1994b, Doulgeris et al., 2008]. Note that the PDF is parameterized by the shape parameter $\alpha > 0$, the number of looks L and the scale matrix $\mathbf{\Sigma}$.

- Extremely heterogeneous surfaces: in this case the texture is assumed to be an inverse gamma, denoted $\bar{\gamma}^{-1}(\eta)$, leading to the $\mathcal{G}_d^0$ model [Freitas et al., 2005, Frery et al., 2007, Frery et al., 2010] for $\mathbf{C}$, where $\eta < 0$.

- All the mentioned distributions above are special cases of the multivariate $\mathcal{U}_d$-distribution from FS distributed texture, such that

$$\lim_{\zeta \to \infty} \mathcal{U}_d(L, \mathbf{\Sigma}, \xi, \zeta) = \mathcal{K}_d(L, \mathbf{\Sigma}, \alpha) \tag{3.36}$$

$$\lim_{\xi \to \infty} \mathcal{U}_d(L, \mathbf{\Sigma}, \xi, \zeta) = \mathcal{G}_d^0(L, \mathbf{\Sigma}, \eta) \tag{3.37}$$

$$\lim_{\xi \to \infty, \zeta \to \infty} \mathcal{U}_d(L, \mathbf{\Sigma}, \xi, \zeta) = s\mathcal{W}_d^{\mathbb{C}}(L, \mathbf{\Sigma}) \tag{3.38}$$

  Bombrun *et al.* in [Bombrun et al., 2011, Bombrun and Beaulieu, 2008]have shown the potential of the $\mathcal{U}_d$ PDF to model both extremely heterogeneous, heterogeneous, and homogeneous surfaces. Therefore, $\mathcal{U}_d$ distribution with extra texture parameter provides flexible model with less restrictions with respect to $\mathcal{K}_d$ and $\mathcal{G}_d^0$.

Parameter estimation of the above distributions is achieved with the method of matrix log-cumulants (MoMLC) because it has been proved to be a feasible and effective estimation technique with multilook polarimetric product model [Anfinsen and Eltoft, 2011].

Chapter 4

Contextual MRF-based Classification for PolSAR Data

Most of the classification techniques for PolSAR data only consider information contained within a pixel, even though intensity levels of neighboring pixels of images are known to have significant correlation. This is our motivation for the current chapter to address the problem of contextual polarimetric SAR image clustering by combining pixelwise matrix-variate statistical distributions for the multilook data, described in the previous chapter, and contextual information.

4.1 Bayesian Classification Scheme

Bayesian estimation theory has been widely used for the classification of remotely sensed data. In order to achieve maximum a posteriori (MAP) estimation (i.e., maximize the posterior probability), it is obligatory to model both class-conditional and prior probabilities. We model each class in PolSAR data by the matix-variate statistical distributions. We propose to use the pixelwise statistical information in a Bayesian classification scheme and to improve the classification by using the spatial context in a Markovian framework. The Markov random field (MRF) theory is first detailed; then, the algorithm steps of the classification method are explained.

4.1.1 Markov Random Fields Theory

MRF theory have been widely used in remote sensing to address many image-analysis problems, including (supervised and unsupervised) classification, segmentation, texture extraction, denoising, and change detection (see, e.g., [Schou and Skriver, 2001, Tison et al., 2004,Serpico and Moser, 2006,Li, 2006,Moser and Serpico, 2009,Akbari et al., 2013,Doulgeris et al., 2012]). Huge number of applications have concentrated especially on statistical applications regarding contextual modeling and spatial data analysis. To

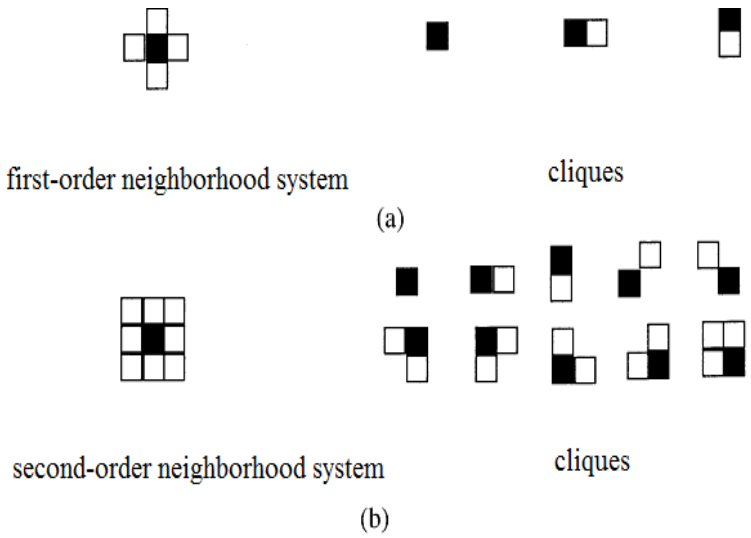


Figure 4.1: Neighborhood systems and corresponding cliques, first and second-order neighborhood systems and cliques of different orders.

introduce the MRF is necessary to first explain the concepts of neighborhood systems and cliques (figure 4.1).

Let $\mathcal{S} = \{s_{i,j}; 1 \leq i \leq R, 1 \leq j \leq M\}$ be regarded as a 2-D pixel lattice, where $s_{i,j}$ is site (i, j), R and M are the number of rows and columns of the image, respectively, and let $\mathcal{L} = \{1, 2, ..., J\}$ denote the set of all possible labels in the clustering map. A neighborhood system for the set $\mathcal{S}$ is defined as

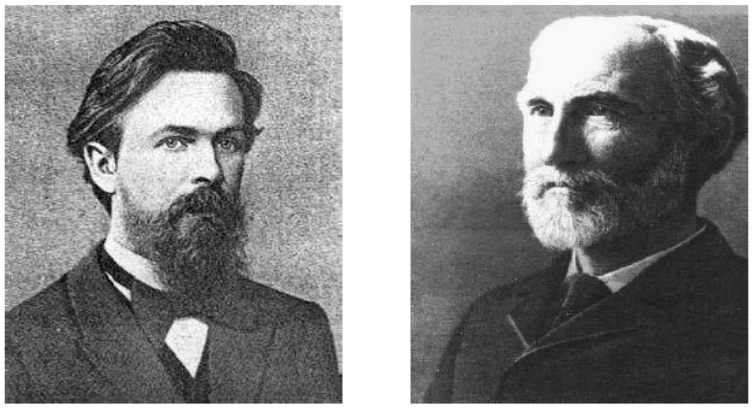
$$\mathcal{N} = \{\mathcal{N}_s | \forall s \in \mathcal{S}\} \tag{4.1}$$

where $\mathcal{N}_s$ is the set of neighbors of the site s. The neighboring relationship has two properties:

- A site s is not neighbor to itself: $s \notin \mathcal{N}_s; \forall s \in \mathcal{S}$;
- The relation is mutual: $s \in \mathcal{N}_j \Leftrightarrow j \in \mathcal{N}_s$;

Commonly two structures are used for $\mathcal{N}_s$ in case of a 2-dimensional lattice $\mathcal{S}$: the first-order neighborhood system and the second-order neighborhood system. As seen in figure 4.1, the first-order neighborhood system is usually defined as the four pixels surrounding a given pixel, and higher orders are defined by adding corner pixels to a lower order neighborhood system.

A *clique* is a set of sites where each member of the set is a neighbor of all the other members. A clique can be of different order. Normally, first-order and second-order cliques are used, because higher order cliques cause too much computational burden. First-order and second-order cliques over $\mathcal{S}$ are denoted by $c_1 = \{s | s \in \mathcal{S}\}$ and $c_2 = \{s, j | s \in \mathcal{S} \quad j \in \mathcal{N}_s\}$. The interaction between sites in a clique is referred as *pixel interaction* and is modeled by a *clique potential*. The collection of all cliques with different orders with respect to the adopted neighborhood system is denoted by $\mathcal{C} = c_1 \cup c_2 \cup \cdots$.



(a) (b)

Figure 4.2: (a) Andrei Andreyevich Markov (1856 - 1922); (b) Josiah Willard Gibbs (1839 - 1903).

A label random field $\mathbf{X} = \{X_s; X_s \in \mathcal{L}, s \in \mathcal{S}\}$ defined on $\mathcal{S}$ is an MRF with respect to a given neighborhood system if, and only if, the following two conditions hold [Li, 2006, Celeux and Diebolt, 1985]:

$$p(\mathbf{X}) > 0 \tag{4.2}$$

$$p(X_s|\mathbf{X}_{\mathcal{S}\setminus s}) = p(X_s|\mathbf{X}_{\mathcal{N}_s}) \tag{4.3}$$

where $\mathcal{S}\setminus s$ denotes the set containing all sites in $\mathcal{S}$ except s. The first condition is related to the positivity of the joint distribution. The second condition, named Markovianity, implies that the the information of one site s depends only on the information carried by its neighbors belonging to the considered neighborhood system. This second property allows to model the spatial interactions between sites.

According to the Hammersley-Clifford theorem, an MRF can equivalently be characterized by a Gibbs distribution [Hammersley and Clifford, 1971]. This enables the field to be characterized by its local, instead of its global, properties. The local characteristics of the MRF can be transformed to a brief energy function $U(\mathbf{X})$ of the Gibbs random field (GRF); hence, it is possible to implement the MRF. A random field $\mathbf{X}$ on $\mathcal{S}$ is called a GRF if and only if the joint distribution is

$$p(\mathbf{X}) = \frac{1}{W} \cdot \exp\left(-U(\mathbf{X}/T)\right), \tag{4.4}$$

where T is the temperature controlling the shape. In practice, T can be taken as a constant and thus be omitted. U is an *energy function*

$$U(\mathbf{X}) = \sum_{c \in \mathcal{C}} V_c(\mathbf{X}_c). \tag{4.5}$$

c is a clique and $\mathcal{C}$ is the collection of all cliques with respect to the adopted neighborhood system, $\mathbf{X}_c$ is the set of samples $\mathbf{X}_s$ such that s belongs to c, and V_c is the potential

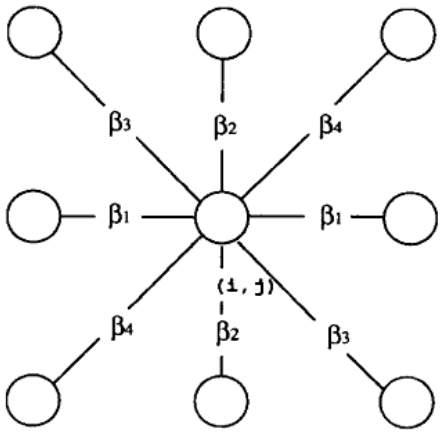


Figure 4.3: An example of non-isotropic second-order system which allows to change the β-parameter in different directions.

associated with clique c. $W = \sum_{\mathbf{X}} \exp\left(-U(\mathbf{X}/T)\right)$ is a normalization constant called the *partition function* [Li, 2006].

Consequently, the local spatial correlations of X_s with respect to its neighbors are modeled by defining suitable potential functions V_c. This actually is an essential problem in all MRF models. For simplicity of computations, we may confine the neighborhood system to an isotropic second-order system, with the related set of pairwise cliques depicted in Figure 4.1.

Potts MRF Model

We will use the classical homogeneous Potts model to model the spatial correlation between pixels. According to this model, a single global parameter $\beta > 0$, which is known as the spatial interaction parameter, regulates the pairwise pixel interactions.[1] With cliques consisting of up to two sites, the second-order energy function of the homogeneous MRF model is given as [Li, 2006, Celeux and Diebolt, 1985]

$$U(\mathbf{X}) = \sum_{s \in S} V_1(X_s) + \sum_{s \in S} \sum_{r \in \mathcal{N}_s} V_2(X_s, X_r), \tag{4.6}$$

where X_s is the label of the central pixel s, X_r is the label of a neighboring pixel in the neighborhood system of s. The first term of the energy function, $\sum_{s \in S} V_1(X_s)$, is determined by the pixelwise matrix-variate statistical distribution, as introduced in the previous chapter, and the second term conveys the contextual information. For a single pixel, s, the contextual energy term is the sum over the second-order cliques of the

[1] Spatially inhomogeneous MRF models allow the β-parameter to change over the field, in which case the models are defined by a set of local conditional density functions [Levada et al., 2008].

neighborhood, i.e.,

$$U(X_s|\mathbf{X}_{\mathcal{N}_s}) = \sum_{r\in\mathcal{N}_s} V_2(X_s, X_r), \tag{4.7}$$

which for the global Potts model is given as

$$V_2(X_s, X_r) = -\delta(X_s, X_r) = \begin{cases} -\beta & \text{if } X_s = X_r \\ 0 & \text{otherwise} \end{cases} \tag{4.8}$$

This potential function results in the following conditional probability mass function (PMF) of X_s, given its neighbors [Yamazaki and Gingras, 1995]:

$$\begin{aligned} p(X_s|\mathbf{X}_{\mathcal{N}_s};\beta) &= \frac{\exp\left(\sum_{r\in\mathcal{N}_s} V_2(X_s, X_r)\right)}{\sum_{X_s\in\mathcal{L}} \exp\left(\sum_{r\in\mathcal{N}_s} V_2(X_s, X_r)\right)} \\ &= \frac{\exp\left(\beta m_{X_s}(s)\right)}{\sum_{l\in\mathcal{L}} \exp\left(\beta m_l(s)\right)}, \end{aligned} \tag{4.9}$$

where $m_{X_s}(s)$ is the number of neighbors of pixel s with label equal to X_s.

4.1.2 MRF Parameter Estimation

The main difficulty of MRF parameter estimation is that the ML method is computationally intractable for most MRF models. An alternative is to adopt the maximum pseudo-likelihood (MPL) technique, which is computationally feasible and simple to implement. The pseudo-likelihood (PL) approach consists in approximating the likelihood in (4.4) as follows [Besag, 1977, Wang et al., 2000]:

$$PL(\mathbf{X};\beta) = \prod_{s\in S} p(X_s|\mathbf{X}_{\mathcal{N}(s)};\beta), \tag{4.10}$$

By substituting the local conditional probabilities from (4.9) into (4.10), the mode-field approximation of $p(\mathbf{X})$ in (4.4) is obtained by

$$\begin{aligned} PL(\mathbf{X};\beta) &= \prod_{s\in S} p(X_s|\mathbf{X}_{\mathcal{N}(s)};\beta) \\ &= \prod_{s\in S} \frac{\exp\left(\beta m_{X_s}(s)\right)}{\sum_{l\in\mathcal{L}} \exp\left(\beta m_l(s)\right)}, \end{aligned} \tag{4.11}$$

Taking the logarithm, the above equation leads to the maximization of

$$\Phi(\beta) = \sum_{s\in S} \left[\beta m_{X_s}(s) - \log\left(\sum_{l\in\mathcal{L}} \exp\left(\beta m_l(s)\right)\right)\right]. \tag{4.12}$$

This function is optimized by a simulated annealing algorithm [Geman and Geman, 1984] to estimate the MRF parameter β, a computationally intensive global minimization approach [Wang et al., 2000, Yu and Cheng, 2003]. Simulated annealing procedure to estimate β is implemented by following steps:

- **Step 0**. Initialization of temperature T_0, β_0, t=1, γ and maximum number of iterations. n.
- **Step 1**. Sampling $\beta^{'} \sim \mathcal{N}(\cdot|\beta_t, 1)$
- **Step 2**. $\beta^{'}$ is accepted with probability

$$\alpha(\beta^{'}, \beta_t) = \min\left\{1, \exp\left(\frac{\Phi(\beta^{'}) - \Phi(\beta^t)}{T_t}\right)\right\} \tag{4.13}$$

  If β' accepted then $\beta_{t+1} = \beta^{'}$, otherwise $\beta_{t+1} = \beta_t$.
- **Step 3**. Incrementing t; if $t \geq 0$ set $T_t = \gamma T_{t-1}$ and goto **Step 1**, otherwise stop.
  By choosing γ and n we thus define the final temperature $T_n = T_0\gamma^n$
  The final estimate $\beta^{\star}$ is chosen by averaging the estimates on the last t itations.

$$\beta^{\star} = \frac{1}{t}\sum_{i=n-1}^{n} \beta_i. \tag{4.14}$$

4.1.3 MRF-MAP Framework

The image classification problem involves assigning to each pixel a class label taking a value from the set $\mathcal{L}$. Let $\mathbf{C}$=$\{\mathbf{C}_s; s \in S\}$ be an MLC image, and let $\mathbf{X} = \{X_s; X_s \in \mathcal{L}, s \in S\}$ be the class labels of $\mathbf{C}$. We seek a labeling $\hat{\mathbf{X}}$, which is an estimate of the true labeling $\mathbf{X}$, according to MAP criterion

$$\hat{\mathbf{X}} = \arg\max_{\mathbf{X}}\{p(\mathbf{X}|\mathbf{C})\} = \arg\max_{\mathbf{X}}\{p(\mathbf{C}|\mathbf{X})p(\mathbf{X})\} \tag{4.15}$$

where $\hat{\mathbf{X}}$ is referred as a MAP estimate of the field of class labels that maximizes the posterior probability $p(\mathbf{X}|\mathbf{C})$. We need to compute the prior probability of the class $p(\mathbf{X})$ and the likelihood probability of the observation $p(\mathbf{C}|\mathbf{X})$. Note that the estimate of (4.15) becomes the *noncontextual pixelwise classifier* if the prior probability does not consider contextual information in the expression (4.15). In order to develop the *contextual classifier* using the MAP criterion, $\mathbf{X}$ is considered as a realization of an MRF, then its prior probability can be derived from (4.4). The unobserved class labels, $\mathbf{X}$, are now considered as a discrete Potts MRF with the energy function $U(\mathbf{X})$ depending on the parameter β [Celeux and Diebolt, 1985]. The MLC image is statistically modeled as a mixture of matrix-variate distributions, where the pixels are assumed to be conditionally independent and identically distributed given the label field $\mathbf{X}$, i.e.,

$$p(\mathbf{C}|\mathbf{X}) = \prod_{s\in S} p(\mathbf{C}_s|X_s), \tag{4.16}$$

where the pixel covariance matrix $\mathbf{C}_s$ follows a matrix-variate distribution with vector of the parameters $\boldsymbol{\theta}$, given the class label $X_s = l$

$$p(\mathbf{C}_s|X_s) = p(\mathbf{C}_s; \boldsymbol{\theta}_l) \tag{4.17}$$

Together with the joint class-conditional distribution and prior distribution of (4.4), the MAP estimates of true class labels as given by (4.15) become

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \{U(\mathbf{X}|\mathbf{C})\} = \arg\min_{\mathbf{X}} \{-\log p(\mathbf{C}|\mathbf{X}) + U(\mathbf{X})\} \tag{4.18}$$

where

$$U(\mathbf{X}|\mathbf{C}) = \sum_{s\in S} -\log p(\mathbf{C}_s|X_s) + \sum_{s\in S}\sum_{r\in\mathcal{N}_s} V_c(X_s, X_r), \tag{4.19}$$

The minimization of (4.18) is necessary to derive a MAP estimate of $\mathbf{X}$. In order to approximate $\hat{\mathbf{X}}$, a feasible and computationally affordable approach is the iterated conditional mode (ICM) technique for the energy minimization [Li, 2006,Jackson and Landgrebe, 2002]. Accordingly, we need to obtain the conditional probability mass function of the label random field $\mathbf{X}$ given an observation of the MLC image $\mathbf{C}$. In the homogeneous Potts MRF model, for each individual site s, X_s is then estimated by

$$\begin{aligned} \hat{X_s} &= \arg\max_{X_s\in\mathcal{L}} \{p(X_s|\mathbf{C}_s, \mathbf{X}_{\mathcal{N}_s}; \mathbf{\Theta})\} \\ &= \arg\max_{X_s\in\mathcal{L}} \{p(\mathbf{C}_s|X_s; \boldsymbol{\theta}) p(X_s|\mathbf{X}_{\mathcal{N}_s}; \beta)\}, \end{aligned} \tag{4.20}$$

and $\mathbf{\Theta} = (\boldsymbol{\theta}, \beta)$ has to be estimated in the iterative segmentation process. The posterior probability of X_s given $\mathbf{C}_s$ becomes

$$p(X_s|\mathbf{C}_s, \mathbf{X}_{\mathcal{N}_s}; \mathbf{\Theta}) = \frac{\exp\left(-U(X_s|\mathbf{C}_s, \mathbf{X}_{\mathcal{N}_s}; \mathbf{\Theta})\right)}{\sum_{X_s\in\mathcal{L}} \exp\left(-U(X_s|\mathbf{C}_s, \mathbf{X}_{\mathcal{N}_s}; \mathbf{\Theta})\right)}, \tag{4.21}$$

and the associated local posterior energy function [Jackson and Landgrebe, 2002]

$$\begin{aligned} U(X_s|\mathbf{C}_s, \mathbf{X}_{\mathcal{N}_s}; \mathbf{\Theta}) &= U(X_s|\mathbf{X}_{\mathcal{N}_s}; \beta) + U(\mathbf{C}_s|X_s; \boldsymbol{\theta}) \\ &= -\beta m_{X_s}(s) + \log\left(\sum_{l\in\mathcal{L}} \exp(\beta m_l(s))\right) - \log p(\mathbf{C}_s|X_s; \boldsymbol{\theta}). \end{aligned} \tag{4.22}$$

And the total posterior energy is given by

$$U(\mathbf{X}|\mathbf{C}) = \prod_{s\in S} U(X_s|\mathbf{C}_s, \mathbf{X}_{\mathcal{N}_s}; \mathbf{\Theta}). \tag{4.23}$$

4.1.4 Incomplete Data Problem

Since both the class labels and the parameters are unknown and must be estimated from the observations, we have a so-called *incomplete data* problem. Many techniques have been proposed to solve this problem, among which the expectation maximization (EM) algorithm [Dempster et al., 1977], [McLachlan and Krishnan, 1996] is the one most widely used. The EM is an iterative parameter estimation technique developed

for parametric modeling problems characterized by data incompleteness. The EM algorithm starts with an initial segmentation with J classes and parameter vector $\mathbf{\Theta}^0$ and works in an iterative manner where, in each iteration, the current cluster parameter estimates are used to segment the image and the new segments are used to recompute the cluster parameters. We attempt to maximize the expectation of the complete-data log likelihood $E\{\log p(\mathbf{X}, \mathbf{C}; \mathbf{\Theta})\}$. Each iteration of the EM algorithm consists of two processes:

- Expectation step (E-step): The missing data are estimated given the observed data and current estimate of the model parameters. This is achieved using the conditional expectation of the unobservable labels $\mathbf{X}$ given the observed data $\mathbf{C}$ and the current estimation $\mathbf{\Theta}^n$ based on the information collected on the previous iteration.

$$Q(\mathbf{\Theta}|\mathbf{\Theta}^n) = E\{\log p(\mathbf{C}; \mathbf{\Theta}|\mathbf{X}, \mathbf{\Theta}^n)\} \tag{4.24}$$

- Maximization step (M- step): Find the parameter that maximizes this quantity:

$$\mathbf{\Theta}^{n+1} = \arg\max_{\mathbf{\Theta}} Q(\mathbf{\Theta}|\mathbf{\Theta}^n) \tag{4.25}$$

Chapter 5

Geometric and Radiometric Terrain Correction of PolSAR Data

Spaceborne SAR instruments have proven to be a very useful tool for multitemporal image analysis. However, terrain topography has significant impact on the geometric and radiometric quality of SAR images [Wivell et al., 1992, Goering et al., 1995, Loew and Mauser, 2007]. To utilize SAR data fully in the study of terrestrial processes, these topographical effects must be characterized and corrected. The objective of this chapter is extension of the theory of SAR terrain correction to polarimetric case, where correction is performed explicitly based on multilooked PolSAR data, and therefore utilize the all polarimetric signature available.

For the geometric correction of topographic effect, precise terrain geocoding of SAR data is required. This can be accomplished by using a DEM and precise satellite state vector. In addition to geometric distortions, surface slopes modulates SAR backscatter which can be split into two main effects. The first effect is changes in radar cross sections (RCS) per unit image area [van Zyl et al., 1993, Luckman, 1998, Ulander, 1996, Shimada, 2010,Small, 2011,Frey et al., 2013] and the second effect is that polarization states are also affected since the terrain slopes specially azimuthal slopes induce polarization orientation changes [Lee et al., 2000,Sabry et al., 2011]. Hence, before analyzing the data, terrain correction is a prerequisite for intercomparisions of multitemporal SAR images. This chapter first addresses the precise geocoding and geometric terrain correction (GTC) of SAR data and then radiometric terrain correction (RTC) of PolSAR data that utilizes the pixel size normalization on each element of the covariance matrix [Atwood et al., 2012]. Moreover, a discussion of orientation angle compensation (OAC) [Lee et al., 2002, Lee and Ainsworth, 2011] is followed after radiometric normalization.

Accurate backscatter estimates enable more robust use of the retrieved values for some applications such as land cover classification, change detection, urban area monitoring, and retrieval of geophysical parameters [Schuler et al., 1999, Hajnsek, 2001]. We have studied the impact of terrain correction of multitemporal PolSAR data on post-classification change detection of Arctic glaciers in [Akbari et al., 2012b]. The procedure

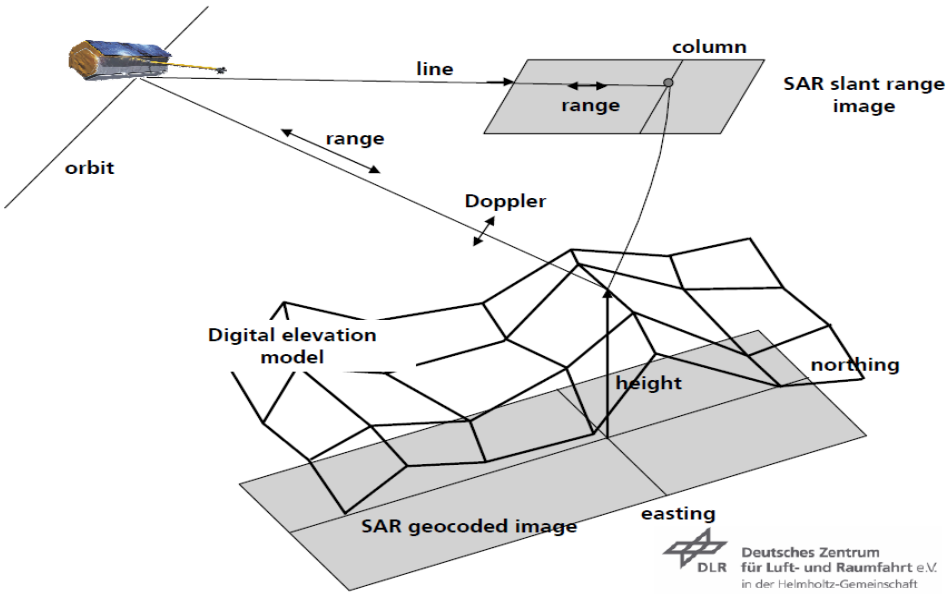


Figure 5.1: Precise terrain Geocoding is performed using a high resolution DEM and additional information about the orbit of the sensor platform.

for terrain correction of PolSAR data is therefore implemented in three steps:

- Precise geocoding and geometric terrain correction (GTC)
- Radiometric terrain correction (RTC)
- Orientation angle compensation (OAC) .

5.1 Precise Geocoding and Geometric Terrain Correction

The objective of the SAR geocoding is to find for each image pixel, the corresponding position on the Earth. With the known satellite state vector, the position of each SAR pixel is estimated on a given Earth model (possibly including a DEM) which has a given slant range from the satellite, and yields zero Doppler shift. If we assume an ellipsoidal earth model such as WGS-84, the problem may be formulated mathematically by solving the so called range-Doppler-Earth model (RDE) equations [Curlander and McDonough, 1991, Kampes, 1999] as follows:

$$\frac{X_{\text{tar}}^2 + Y_{\text{tar}}^2}{(R_e + H_{\text{tar}})^2} + \frac{Z_{\text{tar}}^2}{R_p^2} = 1 \qquad \text{(Earth model equation)} \tag{5.1}$$

$$||\underline{\mathbf{r}}_{\text{tar}} - \underline{\mathbf{r}}_{\text{sat}}|| = R(i,j) \qquad \text{(Range equation)} \tag{5.2}$$

$$[\underline{\mathbf{v}}^{\text{sat}}]^T \cdot (\underline{\mathbf{r}}_{\text{tar}} - \underline{\mathbf{r}}_{\text{sat}}) = 0 \qquad \text{(Zero Doppler centroid)}, \tag{5.3}$$

where (i, j) is the row-column number of the pixels in the original SAR image, R_e and R_p are the semimajor axis (or equatorial radius) and semimnor axis (or polar radius) of the given Earth model, $\underline{\mathbf{r}}_{\text{tar}} = (X_{\text{tar}}, Y_{\text{tar}}, Z_{\text{tar}})$ is target position vector in the Earth Centered Earth Fixed (ECEF) reference frame [Zhu, 1994], $\underline{\mathbf{r}}_{\text{sat}}$ is the sensor position vector, $\underline{\mathbf{v}}_{\text{sat}}$ is the sensor velocity vector, H_{tar} is the local target elevation relative to the assumed model, and $R(i, j)$ is the slant range of the pixel.

The three unknown target position parameters as given by $(X_{\text{tar}}, Y_{\text{tar}}, Z_{\text{tar}})$ is determined from the simultaneous solution of the RDE equations in a iterative manner. Geometrically, the solution of RDE equations corresponds to finding the intersection between the Earth (geoid) surface, a sphere with radius R around the radar antenna, and the zero-Doppler plane described by (5.3) [Elachi and Van Zyl, 2006]. When the slant range distance R is greater than the flying hight above the Earth, there are exactly two points in this intersection. Using the knowledge of the sensor's pointing direction we may exclude one of these, such that a unique solution exists [Larsen, 2011]. The accuracy of this geo-location procedure depends on the accuracy of the available orbit data and velocity vectors, the SAR system parameters, the measurement accuracy of the pulse delay time, and knowledge of the target height relative to the assumed Earth model [Wegmuller, 1999]. From the first-step geolocation procedure, the initial geocoding lookup tables (LUTs) are obtained and any errors are corrected in a refinement step. To refine initial LUTs, the input SAR image and a corresponding DEM are used to create a simulated SAR image. Then, an automatic cross correlation analysis between the simulated and the real SAR image is done to determine the geocoding refinement [Wivell et al., 1992, Wegmuller, 1999, Wang et al., 2011]. At highest cross correlation, the radiometric values of the real SAR image are placed by inverting the mapping from the DEM to the SAR image [Wang et al., 2011].

5.2 Radiometric Terrain Correction

The radiometric values of the geocoded-terrain corrected imagery are usually on the Earth model. In other word, although the position of the backscatter estimate has been corrected by the GTC, the radiometry of the goecoded image remains in ellipsoid-model based. Therefore, RTC is applied to correct distortions due to the side-looking geometry of SAR systems and hill-slope modulations. This section details the RTC of SAR imagery for topographically induced changes in the pixel scattering area. In areas affected by terrain relief, the scattering area is dependent on the local slope as well as the radar look angle [van Zyl et al., 1993]. Hence, the effective pixel area varies due to the variation in ground surface area contributing to each pixel backscatter. The SAR image is commonly calibrated to σ^0 backscatter coefficient, which is defined as the average RCS per unit ground area, or to γ^0, which is defined as the average RCS per unit area obtained by the projecting the ground area into the plane perpendicular to the RLOS. The reference areas of backscatter conventions β^0, σ^0, and γ^0 are illustrated in Figure 5.2. The σ_0 and γ_0

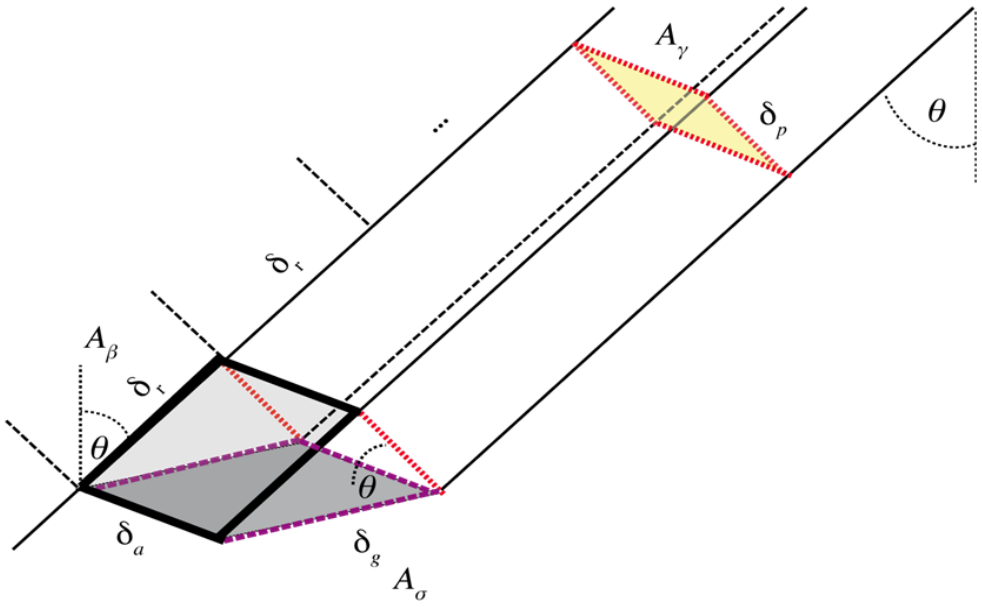


Figure 5.2: The reference areas for the three radar backscatter conventions: standard ellipsoidal β^0, σ^0, and γ^0 (figure taken from [Small, 2011].)

backscatter coefficients are obtained by relating the RB given in (3.18) to the respective reference areas A_β, A_σ, and A_γ as follows [Small, 2011]:

$$\sigma^0 = \beta^0 \frac{A_\beta}{A_\sigma} \tag{5.4}$$

$$\gamma^0 = \beta^0 \frac{A_\beta}{A_\gamma} \tag{5.5}$$

For the standard ellipsoid-based products, the ellipsoidal incidence angle θ_E makes the relationships between these reference areas [Luckman, 1998, Frey et al., 2013]:

$$A_{\sigma_E} = \frac{A_\beta}{\sin(\theta_E)} = \frac{\delta_a \delta_r}{\sin(\theta_E)} \tag{5.6}$$

$$A_{\gamma_E} = A_{\sigma_E} \cos(\theta_E) = \delta_a \delta_g \cos(\theta_E) = \frac{A_\beta}{\tan(\theta_E)} \tag{5.7}$$

where δ_r and δ_a are the dimensions of the scattering area in the slant range and azimuth directions, respectively. The ellipsoidal incidence angles θ_E in the above equations are calculated based on the knowledge of the Earth model in the geocoding step, which are used to calculate the σ^0 and γ^0. In the presence of rugged topography, obviously the σ^0_E and γ^0_E may be terrain-geocoded and a high resolution DEM is required to give an accurate estimate of the projected area on the ground. Ulander [Ulander, 1996] proposed a projection cosine (PC) approach, which was applied for terrain corrected SAR

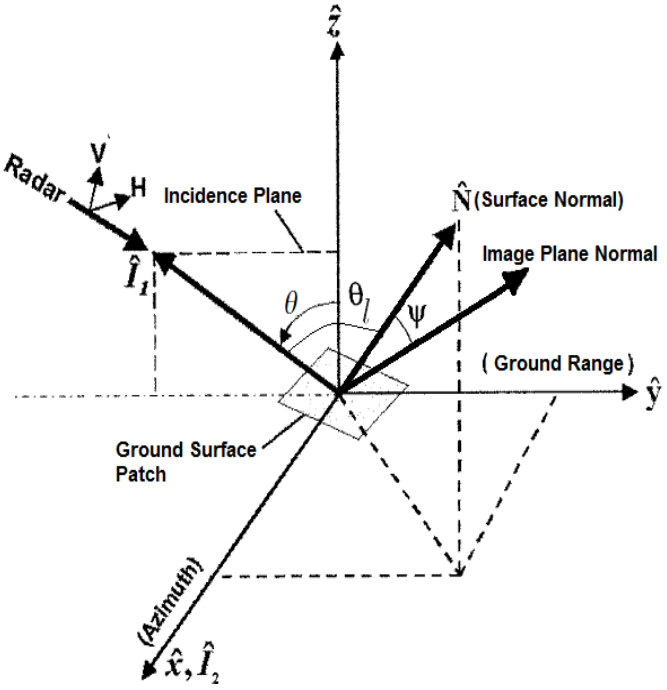


Figure 5.3: The radar imaging geometry which relates the orientation angle to ground slopes and the projection angle Ψ relates the unit image area to the unit ground area. θ is the radar look angle and θ_ℓ is the local incidence angle (figure taken and modified from [Lee et al., 2002]).

geocoding in [Loew and Mauser, 2007]. Area reference areas are related to each other by a projection angle (PA) Ψ:

$$A_{\sigma_T} = \frac{A_\beta}{\cos(\Psi)} \tag{5.8}$$

$$A_{\gamma_T} = A_{\sigma_T} \cos(\theta_\ell) \tag{5.9}$$

where Ψ, as depicted in Figure 5.3, is the angle between the surface normal and the image plane normal, which is assumed to vary between $0\,^\circ$ and $90\,^\circ$ and exclude layover areas [Ulander, 1996]. The terrain corrected radar backscattering coefficient σ_T^0 is related to the radar brightness as

$$\sigma_T^0 = \beta^0 \cos(\Psi) \tag{5.10}$$

Ulander parametrized this angle by using spherical angles u and v as follows:

$$\cos(\Psi) = \sin(\theta)\cos(u) + \cos(\theta)\sin(u)\sin(v), \tag{5.11}$$

where u and v are the terrain slope and aspect angles with respect to the vertical and sensor azimuth directions, respectively, and θ is the radar look angle. We now extend the theory of radiometric slope correction using the PC method to reduce the influence of the terrain topography on the multilook PolSAR data. We first normalize the scattering coefficients with respect to the β^0 reference area (A_β) according to (3.19). To perform

a precise radiometric correction, we need to estimate the changing scattering area in the radar geometry, caused by the terrain topography, parameterized by the PA. Provided that the data are already calibrated with respect to the elevation antenna pattern, range spreading loss, and channel to channel calibration, all elements of covariance matrix in (3.21) can be simultaneously radiometrically corrected to the ground range area by [Akbari et al., 2012b]

$$\mathbf{C}_{\sigma_T^0} = \frac{\cos(\Psi)}{L} \sum_{\ell=1}^{L} (\underline{\mathbf{\Omega}}_{\beta^0})_\ell (\underline{\mathbf{\Omega}}_{\beta^0})_\ell^H = \cos(\Psi) \cdot \mathbf{C}_{\beta^0}, \tag{5.12}$$

All covariance matrix values now correspond to the σ^0 backscatter coefficient, which is equivalent to (5.10). The conventional radiometric normalization method, which relies on the local incidence angle only, is adequate for flatlands or for pixels with zero slope ($u = 0$) in (5.11). The expected results are radiometrically "flattened" SAR images. The required parameters for performing the RTC are estimated in the geocoding step using an available DEM and the satellite state vector. In cases of an unavailable DEM, ellipsoid-based radiometric correction is achieved over an Earth model.

5.3 Orientation Angle Compensation

When applying the RTC to PolSAR data, the polarimetric components associated with a pixel are compensated equally by the same factor. This section addresses the need for PolSAR data compensation for the second effect on the RCS as function of polarization states. For *reflection symmetrical* media, such as horizontal surface, the polarization orientation angle (POA) [1] is about zero and the surface normal is in the incidence plane. For rugged terrain areas, the POA is shifted from zero and the surface normal is no longer in the incidence plane. These shifts are induced by surfaces with nonzero azimuth slopes and also by man-made targets such as buildings that are not aligned in the azimuth direction [Lee and Ainsworth, 2011]. For the distributed media, in general, the higher is the radar frequency, the lower is the POA sensitivity to the surface slopes. The induced POA shift ϑ is the angle that rotates the incidence plane about the RLOS to the surface normal [Lee et al., 2002]. Lee *et al.* showed in [Lee et al., 2000] that the induced POA ϑ is related to the azimuth and range slopes by the following equation

$$\tan(\vartheta) = \frac{\tan(\omega)}{-\tan(\gamma)\cos(\theta) + \sin(\theta)} \tag{5.13}$$

where ω is the azimuth slope angle and γ is the slope angle in the ground range direction. These shifts produce higher cross-polarization (HV) intensity and make coherency or covariance matrix *reflection asymmetrical* [Lee et al., 2000]. Several methods have been proposed in the literature to extract the POA. The most accurate approach is use of the

[1] is the angle of rotation about the RLOS

DEM to compute the POA in (5.13). In cases of unavailable DEM, the POA is derived directly from either coherency matrix or covariance matrix [Lee et al., 2000]. The most promising results have been obtained by the phase difference between right-right and left-left circular polarizations. The POA is computed using either single look complex or multilook data within the range of $\pi/4$ to $\pi/4$ as [Lee and Ainsworth, 2011]

$$\begin{aligned}\vartheta &= \frac{1}{4}\left[\arg(S_{RR}S_{LL}^{*}) + \pi\right] \\ &= \frac{1}{4}\left[\arctan\left(\frac{-4\Re(\langle(S_{hh}-S_{vv})S_{hv}^{*}\rangle)}{-\langle|S_{hh}-S_{vv}|^{2}\rangle + 4\langle|S_{hv}|^{2}\rangle}\right) + \pi\right],\end{aligned} \tag{5.14}$$

where $\Re(A)$ denotes the real part of A. (5.14) is equivalent to:

$$\vartheta = \arctan\left(\frac{2\Re\{T_{23}\}}{T_{22}-T_{33}}\right). \tag{5.15}$$

To account for negative orientation angles the above equation has to be modified

$$\text{For} \quad \vartheta > \frac{\pi}{4}, \quad \vartheta = \vartheta - \frac{\pi}{2}. \tag{5.16}$$

Once the POA induced by the azimuth slope is obtained, the compensation for this slope effect is straightforward, which involves rotating the data along the RLOS by the negative of the induced POA. For the radiometrically terrain corrected multilook coherency matrix, the compensation can be done by [Lee et al., 2002, Yamaguchi et al., 2011]

$$\mathbf{T}_{\text{OAC}} = \mathbf{R}(\vartheta)\mathbf{T}_{\sigma_T^0}\mathbf{R}^{T}(\vartheta), \quad \text{with} \quad \mathbf{R}(\vartheta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(2\vartheta) & \sin(2\vartheta) \\ 0 & -\sin(2\vartheta) & \cos(2\vartheta) \end{bmatrix}. \tag{5.17}$$

where $\mathbf{T}_{\sigma_T^0}$ and $\mathbf{T}_{\text{OAC}}$ are the coherency matrix before (after the RTC) and after rotation, respectively, and $\mathbf{R}(\vartheta)$ is a unitary rotation matrix [2]. The compensation reduces the cross-polarization power, T_{33}, but increases T_{22} power by the same amount.

For the radiometrically terrain corrected multilook covariance matrix given in (5.12), the data compensation can be achieved by [Lee et al., 2000]

$$\mathbf{C}_{\text{OAC}} = \mathbf{R}(\vartheta)\mathbf{C}_{\sigma_T^0}\mathbf{R}^{T}(\vartheta), \quad \text{with} \quad \mathbf{R}(\vartheta) = \frac{1}{2}\begin{bmatrix} 1+\cos(2\vartheta) & \sqrt{2}\sin(2\vartheta) & 1-\cos(2\vartheta) \\ -\sqrt{2}\sin(2\vartheta) & 2\cos(2\vartheta) & \sqrt{2}\sin(2\vartheta) \\ 1-\cos(2\vartheta) & -\sqrt{2}\sin(2\vartheta) & 1+\cos(2\vartheta) \end{bmatrix}. \tag{5.18}$$

where $\mathbf{C}_{\sigma_T^0}$ and $\mathbf{C}_{\text{OAC}}$ are the covariance matrix before (after the RTC) and after rotation, respectively.

[2] $\mathbf{R}^{T}(\vartheta) = \mathbf{R}^{-1}(\vartheta)$

Chapter 6

Direct Change Detection in Multilook PolSAR Data

Change detection is a particular application in multitemporal SAR image analysis. Two distinct approaches have been proposed: post-classification comparison and direct change detection.

The post-classification change detection classifies the images separately, after they have been co-registered and co-calibrated. Then the classified images are compared and analyzed to form a change matrix which described the mapping of classes between the images. From this matrix we can extract a simple map of change versus no change, but also more detailed information on the nature of the change. Paper 2 of Chapter 8 is an example of post-classification change detection over glaciers.

In direct change detection, two co-registered and co-calibrated MLC images are compared directly on a pixel-by-pixel basis by a desired test statistic, without going through any classification stage. This only yields a binary map of change versus no change. The objective of this chapter is discussion on the direct change detection in multilook PolSAR data, which is also the topic of Paper 3 in Chapter 9.

6.1 Problem Formulation

Let $\mathcal{A} = \{\mathbf{A}(i,j); 1 \leq i \leq I, 1 \leq j \leq J\}$ and $\mathcal{B} = \{\mathbf{B}(i,j); 1 \leq i \leq I, 1 \leq j \leq J\}$ be two equal-sized co-registered MLC images acquired over the same geographical area at times t_0 and t_1 ($t_1 > t_0$), where I and J are the number of rows and columns of the images, respectively. It is assumed that $\mathcal{A}$ and $\mathcal{B}$ are geometrically and radiometrically terrain corrected (as discussed in Chapter 5) and registered. Our objective is to produce a map representing the changes occurring in the scene between t_0 and t_1. The final goal of a change detection analysis is to produce a binary map corresponding to the two classes: change and no change. The problem can be decomposed into two steps: the generation of a test statistic image which compacts the matrix-variate polarimetric

information into a scalar feature and the thresholding of the test statistic image in order to produce the binary change map. Figure 6.1 shows a block diagram describing a general change detection processing chain in multilook PolSAR data. The overall performance of the polarimetric change detector depends on both the quality of the chosen test statistic image and the quality of the thresholding.

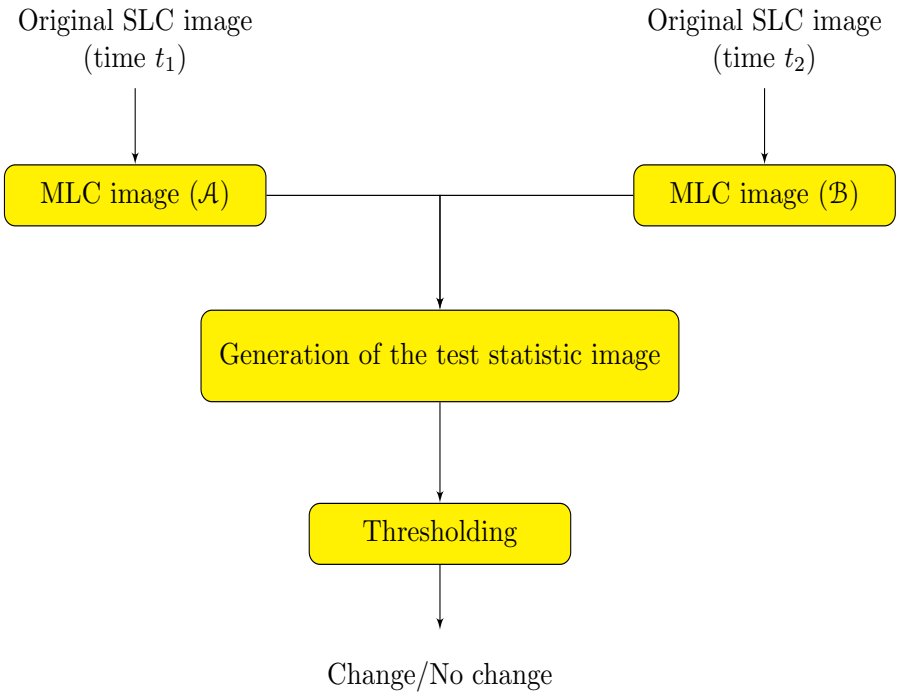


Figure 6.1: General block diagram of the direct change detection algorithms for mutlilook PolSAR data.

Let the $d \times d$ Hermitian and positive definite matrices $\mathbf{A}$ and $\mathbf{B}$ be independent, nonsingular, and scaled complex Wishart distributed, with the PDF given in (3.34), i.e.,

$$\mathbf{A} \sim s\mathcal{W}(d, L_a, \mathbf{\Sigma}_a) \text{ and } \mathbf{B} \sim s\mathcal{W}(d, L_b, \mathbf{\Sigma}_b). \tag{6.1}$$

We use a desired test statistic between the two covariance matrices $\mathbf{A}$ and $\mathbf{B}$, denoted as $\tau(\mathbf{A}, \mathbf{B})$, to decide between the hypotheses:

$$\begin{aligned} H_0 &: \mathbf{\Sigma}_a = \mathbf{\Sigma}_b, \\ H_1 &: \mathbf{\Sigma}_a \neq \mathbf{\Sigma}_b, \end{aligned} \tag{6.2}$$

H_0 and H_1 correspond to the no change hypothesis and the change hypothesis, respectively. A threshold selection process can be applied to the test statistic to distinguish quantitatively between H_0 and H_1. Several unsupervised change image thresholding algorithms have been proposed in the literature, such as Otsu's method [Otsu, 1979],

the Kittler and Illingworth (K&I) algorithm [Kittler and Illingworth, 1986], the entropy-based Kapur algorithm [Kapur et al., 1985], and Liu's method [Liu and Feng, 2006]. It is worth to note that some of these algorithms require statistical modeling of the test statistic.

6.2 Test Statistics for Polarimetric Change Detection

This section reviews well-known test statistics, whereof some have previously been used for change detection in PolSAR images. It is commonly preferred to use matrix distance measures as polarimetric test statistics for change detection. Typically different matrix distance measures used as polarimetric change detectors fall into one of two classes: *metric* and *semimetric*. To be classified as metric, a distance between two matrices $\mathbf{A}$ and $\mathbf{B}$ must obey the following rules:

1. The distance must be *non-negative*, i.e., $\tau(\mathbf{A}, \mathbf{B}) \geq 0$.

2. The distance must be *symmetric*, so that the distance from $\mathbf{A}$ to $\mathbf{B}$ is the same as the distance from $\mathbf{B}$ to $\mathbf{A}$, i.e., $\tau(\mathbf{A}, \mathbf{B}) = \tau(\mathbf{B}, \mathbf{A})$. This is sometimes called the *symmetry rule*.

3. Two identical covariance matrices have zero distance value, i.e., $\tau(\mathbf{A}, \mathbf{B}) = 0 \Leftrightarrow \mathbf{A} = \mathbf{B}$. This is called the *identity of indiscernibles*.

4. When considering three matrices, $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, the distance from $\mathbf{A}$ to $\mathbf{C}$ is always less than or equal to the sum of the distance from $\mathbf{A}$ to $\mathbf{B}$ and the distance from $\mathbf{B}$ to $\mathbf{C}$, i.e., $\tau(\mathbf{A}, \mathbf{C}) \leq \tau(\mathbf{A}, \mathbf{B}) + \tau(\mathbf{B}, \mathbf{C})$. This is called either the *triangle inequality* or the *subadditivity*. The triangle inequality also implies that if matrices $\mathbf{A}$ and $\mathbf{B}$ are similar and $\mathbf{B}$ and $\mathbf{C}$ are similar, then $\mathbf{A}$ and $\mathbf{C}$ must also be similar.

Distance measures that obey the first three rules, but fail to obey rule 4 are referred to as semimetric. We now group matrix distance measures for multilook data into two types: those that cannot (Type I) and those that can (Type II) be directly related to the statistical distribution of the covariance matrix through a probabilistic or an information theoretical interpretation.

6.2.1 Matrix Distance Measures of Type I

In this section, we review some matrix distance measures that meet the requirements for being a metric. However, they cannot be directly related to the statistical distribution of the covariance matrix through probability or information theory. For matrix distance measures of this type introduced below, under the null (no change) hypothesis both covariance matrices $\mathbf{A}$ and $\mathbf{B}$ are identical, so $\tau(\mathbf{A}, \mathbf{B}) = 0$. The null hypothesis is rejected if $\tau(\mathbf{A}, \mathbf{B}) > Threshold$.

Contrast Ratio

The contrast ratio (CR) (or Rayleigh Quotient) is defined in terms of the eigenvalues of the following matrix [Kersten et al., 2005a]:

$$\mathbf{C}_{\text{ab}} = (\mathbf{A}^{-\frac{1}{2}})^H \mathbf{B} \mathbf{A}^{-\frac{1}{2}}, \tag{6.3}$$

$\mathbf{C}_{\text{ab}}$ is also a positive definite Hermitian matrix since both $\mathbf{A}$ and $\mathbf{B}$ are. Then the CR is defined as

$$\tau_{\text{CR}}(\mathbf{A}, \mathbf{B}) = \ln\big(\max(\lambda_{\max}, \lambda_{\min}^{-1})\big), \tag{6.4}$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are the largest and smallest eigenvalue of $\mathbf{C}_{\text{ab}}$ in (6.3). Under the null hypothesis, the expectation of $\mathbf{C}_{\text{ab}}$ is an identity matrix, whose eigenvalues are all unity. The performance of the CR for polarimetric change detection has been shown in [Kersten et al., 2005a].

Ellipticity

The ellipticity [Muirhead, 2009] is also defined in terms of the eigenvalues of the matrix $\mathbf{C}_{\text{ab}}$, given in (6.3). It is defined as

$$\tau_{\text{E}}(\mathbf{A}, \mathbf{B}) = -\ln \prod_{i=1}^{d} \lambda_i + d \ln \left(\frac{1}{d} \sum_{i=1}^{d} \lambda_i \right). \tag{6.5}$$

Also the ellipticity has been assessed in [Kersten et al., 2005a] as a test statistic for change detection.

Minkowski Distance

The Minkowski distance of order ν between two complex covariance matrices is defined as:

$$\tau_{\nu}(\mathbf{A}, \mathbf{B}) = \left(\sum_{i \leqslant j} \text{abs}\,(a_{i,j} - b_{i,j})^{\nu} \right)^{1/\nu}, \tag{6.6}$$

where the variables $a_{i,j}$ and $b_{i,j}$ refer to the elements of the matrices $\mathbf{A}$ and $\mathbf{B}$. In order to avoid confusion with the determinant, the absolute value of the scalar x is denoted $\text{abs}(x)$. The Minkowski distance is typically used with ν being 1 or 2. The latter is the Euclidean distance, while the former is known as the Manhattan distance. The Euclidean matrix distance is a natural candidate to contrast matrices that follow the matrix-variate Gaussian distribution, but neither the Euclidean nor the Manhattan distance match the geometry of the covariance matrix domain. This is seen as they cannot be related to the complex Wishart distribution by probabilistic consideration or information theory. Accordingly, they are also reported to produce poor results when applied to clustering of PolSAR data ([Anfinsen et al., 2007]; [Dabboor et al., 2013b]).That is why that there is a less interest to use this distance measure for change detection.

Euclidean Distance

Th ℓ_2 metric refers to the Euclidean distance between two matrices $\mathbf{A}$ and $\mathbf{B}$, which is defined as

$$\tau_2(\mathbf{A},\mathbf{B}) = \sqrt{\sum_{i\leqslant j} \text{abs}\left(a_{i,j}-b_{i,j}\right)^2}. \tag{6.7}$$

The value of the Euclidean distance belongs to the interval of $[0,\infty]$.

Manhattan Distance

The Manhattan distance is given as

$$\tau_1(\mathbf{A},\mathbf{B}) = \sum_{i\leqslant j} \text{abs}\left(\Re(a_{i,j}-b_{i,j}\right) + \sum_{i\leqslant j} \text{abs}\left(\Im(a_{i,j}-b_{i,j})\right), \tag{6.8}$$

where $\Re$ and $\Im$ are the real and imaginary parts, respectively. This distance is also known as the rectilinear distance, Minkowski's ℓ_1 distance, the taxi cab metric, or the city block distance.

Chebyshev Distance

In the limiting case of ν reaching infinity in the Minkowski distance, we obtain the Chebyshev distance. It is also called the Tchebychev distance, the Maximum metric, or the ℓ_∞ metric [Cantrell, 2000].

$$\tau_{\text{CH}}(\mathbf{A},\mathbf{B}) = \lim_{\nu\to\infty}\left(\sum_{i\leqslant j} \text{abs}\left(a_{i,j}-b_{i,j}\right)^\nu\right)^{1/\nu} = \max_{i\leqslant j}\left(\text{abs}\left(a_{i,j}-b_{i,j}\right)\right). \tag{6.9}$$

Canberra Distance

The Canberra distance (CAD) was introduced in [Lance and Williams, 1966] and refined in [Lance and Williams, 1967]. The CAD between two covariance matrices is given as follows:

$$\tau_{\text{CAD}}(\mathbf{A},\mathbf{B}) = \sum_{i\leqslant j} \frac{\text{abs}\left(a_{i,j}-b_{i,j}\right)}{\text{abs}\left(a_{i,j}\right)+\text{abs}\left(b_{i,j}\right)}, \tag{6.10}$$

6.2.2 Matrix Distance Measures of Type II

In this section, we present some matrix distance measures that are related to the statistical distribution of the covariance matrix. The matrix distance measures of this section are based on the assumption that PolSAR data follow the scaled complex Wishart distribution, characterized by the density given in (3.34).

Wishart Distance

Lee *et al.* derived a distance measure called the Wishart distance under the complex Wishart distribution and defined by [Lee et al., 1994a]

$$\tau_{\mathrm{W}}(\mathbf{A},\mathbf{B}) = \ln|\mathbf{B}| + \mathrm{tr}(\mathbf{B}^{-1}\mathbf{A}). \tag{6.11}$$

The revised Wishart distance is derived by [Kersten et al., 2005b]

$$\tau_{\mathrm{RW}}(\mathbf{A},\mathbf{B}) = \ln\frac{|\mathbf{A}|}{|\mathbf{B}|} + \mathrm{tr}(\mathbf{A}^{-1}\mathbf{B}) - d. \tag{6.12}$$

The revised Wishart distance is asymmetric, which makes the use of the distance measure more difficult for change detection. Anfinsen *et al.* derived a symmetric measure [Anfinsen et al., 2007] as follows:

$$\begin{aligned}\tau_{\mathrm{SRW}}(\mathbf{A},\mathbf{B}) &= \frac{1}{2}(\tau_{\mathrm{RW}}(\mathbf{A},\mathbf{B}) + \tau_{\mathrm{RW}}(\mathbf{B},\mathbf{A})) \\ &= \frac{\mathrm{tr}(\mathbf{A}^{-1}\mathbf{B}) + \mathrm{tr}(\mathbf{B}^{-1}\mathbf{A})}{2} - d.\end{aligned} \tag{6.13}$$

It satisfies all conditions, except the triangle inequality, an is thus a semimetric and the form of the rejection region is similar to the matrix distance measures of Type I. The application of the symmetrized Wishart distance for spectral clustering of PolSAR data has been reported in [Anfinsen et al., 2007]. It can be used as a test statistic for change detection and we may threshold the test statistic by the algorithms mentioned in Section 6.1 to discriminate change and no change areas.

Bartlett Distance

The Bartlett distance or dissimilarity measure is given as in [Kersten et al., 2005b]

$$\begin{aligned}\tau_{\mathrm{B}}(\mathbf{A},\mathbf{B}) &= \log\frac{|\mathbf{A}+\mathbf{B}|}{2|\mathbf{A}|} + \log\frac{|\mathbf{A}+\mathbf{B}|}{2|\mathbf{B}|} - 2d\log 2 \\ &= 2\log\frac{|\mathbf{A}+\mathbf{B}|}{\sqrt{|\mathbf{A}||\mathbf{B}|}} - 2d\log 2\end{aligned} \tag{6.14}$$

This dissimilarity represents the ratio of two averaging methods (the geometric mean of the determinants to the determinant of the arithmetic mean). The distance is called Bartlett, since Bartlett first proposed the ratio of the arithmetic and geometric means of the sample variances as a robust test for the equality of two distributions [Hampel et al., 2011]. It is easy to prove that the distance satisfies symmetry. However, triangle inequality does not hold, so the Bartlett distance is a semimetric only. Under the null hypothesis both $\mathbf{A}$ and $\mathbf{B}$ are identical, so $\tau_{\mathrm{B}}(\mathbf{A},\mathbf{B}) = 0$. The efficacy of the Bartlett distance in polarimetric change detection was first demonstrated in [Kersten et al., 2005a].

Hellinger Distance

The Hellinger distance is also used to quantify the similarity between two probability distributions [Beran, 1977] and is defined as

$$\tau_{\mathrm{H}}(\mathbf{A},\mathbf{B}) = 1 - \int \sqrt{p_{\mathbf{A}} p_{\mathbf{B}}}, \tag{6.15}$$

where $p_{\mathbf{A}}$ and $p_{\mathbf{B}}$ are the PDF of two populations. The values of $\tau_{\mathrm{H}}(\mathbf{A},\mathbf{B})$ lie between zero and unity. The Hellinger distance of two scaled complex Wishart distributions is derived in [Frery et al., 2011] as:

$$\tau_{\mathrm{H}}(\mathbf{A},\mathbf{B}) = 1 - \frac{\left|\left(\frac{L_a\mathbf{A}^{-1}+L_b\mathbf{B}^{-1}}{2}\right)^{-1}\right|^{\frac{L_a+L_b}{2}}}{|\mathbf{A}|^{\frac{L_a}{2}}|\mathbf{B}|^{\frac{L_b}{2}}} \sqrt{L_a^{dL_a} L_b^{dL_b}} \prod_{i=1}^{d-1} \frac{\Gamma(\frac{L_a+L_b}{2}-i)}{\sqrt{\Gamma(L_a-i)\Gamma(L_b-i)}}. \tag{6.16}$$

If $L_a = L_b = L$, we get

$$\tau_{\mathrm{H}}(\mathbf{A},\mathbf{B}) = \frac{\left|\left(\frac{\mathbf{A}^{-1}+\mathbf{B}^{-1}}{2}\right)^{-1}\right|^{L}}{|\mathbf{A}|^{\frac{L}{2}}|\mathbf{B}|^{\frac{L}{2}}}. \tag{6.17}$$

It is possible to prove that the distance satisfies all conditions in Section 6.2, so the Hellinger distance is a metric. The Hellinger distance is equivalent to the Bartlett distance for practical purposes, such as implementation as a test statistic in change detection, as seen from (6.14). Therefore, it should produce exactly the same results as the Bartlett distance.

Bhattacharyya Distance

The Bhattacharyya distance [Bhattacharyya, 1943] measures the similarity of two discrete or continuous probability distributions and is defined as

$$\begin{aligned} \tau_{\mathrm{BH}}(\mathbf{A},\mathbf{B}) &= -\log \int \sqrt{p_{\mathbf{A}} p_{\mathbf{B}}} \\ &= -\log(1-\tau_{\mathrm{H}}(\mathbf{A},\mathbf{B})). \end{aligned} \tag{6.18}$$

Clearly $0 \leq \tau_{\mathrm{BH}}(\mathbf{A},\mathbf{B}) \leq \infty$, and therefore the distance is non-negative. The Bhattacharyya distance of two scaled complex Wishart populations with unequal covariance matrices was derived in [Frery et al., 2011] as follows:

$$\begin{aligned} \tau_{\mathrm{BH}}(\mathbf{A},\mathbf{B}) = &\sum_{i=0}^{d-1} \log \frac{\sqrt{\Gamma(L_a-i)\Gamma(L_b-i)}}{\Gamma(\frac{L_a+L_b}{2}-i)} + \frac{L_a \log|\mathbf{A}|}{2} + \frac{L_b \log|\mathbf{B}|}{2} \\ &- \frac{d}{2}(L_a \log L_a + L_b \log L_b) - \frac{L_a+L_b}{2} \log \left|\left(\frac{L_a\mathbf{A}^{-1}+L_b\mathbf{B}^{-1}}{2}\right)^{-1}\right|. \end{aligned} \tag{6.19}$$

If $L_a = L_b = L$, then the distance is simplified to the following derivation,

$$\tau_{\text{BH}}(\mathbf{A},\mathbf{B}) = L\left[\frac{\log|\mathbf{A}| + \log|\mathbf{B}|}{2} - \log\left|\left(\frac{\mathbf{A}^{-1}+\mathbf{B}^{-1}}{2}\right)^{-1}\right|\right]. \tag{6.20}$$

It can be shown that the distance satisfies symmetry. However, it does not obey the triangle inequality, so the Bhattacharyya distance is a semimetric.

Kullback-Leibler Distance

The Kullback-Leibler(KL) distance in a symmetric form is defined as [Jeffreys and Jeffreys, 1946]

$$\tau_{\text{KL}}(\mathbf{A},\mathbf{B}) = \frac{1}{2}\int (p_{\mathbf{A}} - p_{\mathbf{B}}) \log\left(\frac{p_{\mathbf{A}}}{p_{\mathbf{B}}}\right). \tag{6.21}$$

The value of $\tau_{\text{KL}}(\mathbf{A},\mathbf{B})$ is zero if $\mathbf{A} = \mathbf{B}$, which implies that the two covariance matrices are identical. The KL distance of two scaled complex Wishart populations is given as [Frery et al., 2011]:

$$\begin{aligned}\tau_{\text{KL}}(\mathbf{A},\mathbf{B}) = \frac{L_a - L_b}{2}\Bigg[&\log\frac{|\mathbf{A}|}{|\mathbf{B}|} + d\big[\Psi^{(0)}(L_a - d + 1) - \Psi^{(0)}(L_b - d + 1)\big] \\ &- d\log\frac{L_a}{L_b} + (L_b - L_a)\sum_{i=1}^{d-1}\frac{i}{(L_a - i)(L_b - i)}\Bigg] \\ &- \frac{d(L_a + L_b)}{2} + \frac{\text{tr}(L_b\mathbf{B}^{-1}\mathbf{A} + L_a\mathbf{A}^{-1}\mathbf{B})}{2}.\end{aligned} \tag{6.22}$$

If $L_a = L_b = L$, we get

$$\tau_{\text{KL}}(\mathbf{A},\mathbf{B}) = L\left(\frac{\text{tr}(\mathbf{B}^{-1}\mathbf{A} + \mathbf{A}^{-1}\mathbf{B})}{2} - d\right). \tag{6.23}$$

Then τ_{KL}/L becomes the revised Wishart distance given in (6.13). It satisfies all requirements of the distance measure, except the triangle inequality, an is also a semimetric. It was shown in [Inglada and Mercier, 2007] that the KL distance can be applied for change detection in mulitemporal single-channel SAR data, but no literature is available but for PolSAR data.

Wishart Chernoff Distance

The Wishart-Chernoff (WC) distance was proposed in [Dabboor et al., 2013a] for the agglomerative clustering of PolSAR data and is given as:

$$\tau_{\text{WC}}(\mathbf{A},\mathbf{B}) = -\ln\left(\frac{|\Upsilon\mathbf{A}^{-1} + (1-\Upsilon)\mathbf{B}^{-1}|^{-1}}{|\mathbf{A}|^{\Upsilon}|\mathbf{B}|^{(1-\Upsilon)}}\right), \tag{6.24}$$

where Υ is a parameter, $0 \leq \Upsilon \leq 1$. The optimum WC distance that best describes the similarity between the two scaled complex Wishart distributions is obtained by finding the optimum value of Υ that minimizes $g(\Upsilon) = \exp\{-\tau_{\text{WC}}(\mathbf{A}, \mathbf{B})\}$ [Dabboor et al., 2013b]. The WC distance is symmetric. The proposed distance measures the similarity between two complex Wishart distributed covariance matrices. This distance measure is also a potenial test statistic for change detection applications, but it has not been reported anywhere.

6.3 Proposed Polarimetric Change Detector

We now introduce our test statistic called the complex Hotelling-Lawley (HL) trace statistic that can be used for polarimetric change detection.

$$\tau_{\text{HL}} = \text{tr}(\mathbf{A}^{-1}\mathbf{B}). \tag{6.25}$$

The proposed test statistic does not obey all requirements for being a metric in Section 6.2. However, it maps the no change case to values centered around the polarimetric dimension d, and the change case to values much smaller or much larger than that, such that indications of change can be clearly identified.

Moments of the HL statitics under the scaled complex Wishart distribution was derived in [Anfinsen and Akbari, 2013]. The expressions of the first three moments are given below, and the derivation is given in Appendix B.

$$m_1^{(\text{HL})} = E\{\tau_{\text{HL}}\} = \frac{dL_a}{Q_a}. \tag{6.26}$$

where $Q_a = L_a - d$.

$$\begin{aligned} m_2^{(\text{HL})} &= E\{\tau_{\text{HL}}^2\} \\ &= \frac{L_a^2}{Q_a^3 - Q_a}\left(d^2\left(Q_a + \frac{1}{L_b}\right) + d\left(\frac{Q_a}{L_a} + 1\right)\right). \end{aligned} \tag{6.27}$$

$$\begin{aligned} m_3^{(\text{HL})} &= E\{\tau_{\text{HL}}^3\} \\ &= \frac{L_a^3}{Q_a^5 - 5Q_a^3 + 4Q_a} \times \left[d^3\left((Q_a^2 - 1) + \frac{3Q_a}{L_b} + \frac{4}{L_b^2}\right)\right. \\ &\qquad + d^2\left(3Q_a + \frac{3(Q_a^2 + 2)}{L_b} + \frac{6Q_a}{L_b^2}\right) \\ &\qquad \left. + d\left(4 + \frac{6Q_a}{L_b} + \frac{2Q_a^2}{L_b^2}\right)\right]. \end{aligned} \tag{6.28}$$

It is seen that these moments are functions of the number of looks and the number of polarimetric channels under the null hypothesis. We then approximate the sampling

distribution of the complex HL statistic by a FS distribution [Galland et al., 2009], denoted.

$$\tau_{\text{HL}} \sim \mathcal{FS}(\xi, \zeta, \mu) \tag{6.29}$$

where $\mu = E\{\tau_{\text{HL}}\} > 0$ is the mean and $\xi > 0$ and $\zeta > 0$ are two shape parameters. The FS distribution is given by [Anfinsen, 2010]

$$p_\tau(t) = \frac{\Gamma(\xi+\zeta)}{\Gamma(\xi)\Gamma(\zeta)} \frac{\xi}{\mu(\zeta-1)} \frac{\left(\frac{\xi}{\mu(\zeta-1)}t\right)^{\xi-1}}{\left(\frac{\xi}{\mu(\zeta-1)}t+1\right)^{\xi+\zeta}}. \tag{6.30}$$

The flexibility of the FS distribution is controlled with the shape parameters to steer between heavy-headed and heavy-tailed distributions [Bombrun et al., 2011]. The ν^{th} order moment relations, in terms of the PDF parameters, are shown in [Galland et al., 2009] to be

$$m_\nu^{(\text{FS})}\{\tau\} = \left(\frac{(\zeta-1)\mu}{\xi}\right)^\nu \frac{\Gamma(\xi+\nu)}{\Gamma(\xi)} \frac{\Gamma(\zeta-\nu)}{\Gamma(\zeta)}. \tag{6.31}$$

The proposed method is to match the population moments of the FS distribution in (6.31) with the population moments of the τ_{HL}, i.e.,

$$m_\nu^{(\text{FS})}(\xi, \zeta, \mu) = m_\nu^{(\text{HL})}(L_a, L_b, d). \tag{6.32}$$

The location parameter μ of the FS distribution is computed analytically, but, to match the second and third-order moments, we use minimum distance optimization [Parr and Schucany, 1982] for shape parameters ξ and ζ as.

$$\epsilon^2 = \sum_{\nu=2}^{3} (m_\nu^{(\text{HL})} - m_\nu^{(\text{FS})})^2, \tag{6.33}$$

$$(\hat{\xi}, \hat{\zeta}) = \arg\{\min_{(\xi,\zeta)} \{\epsilon^2\}\}. \tag{6.34}$$

In this numerical problem, ξ and ζ are the unknown variables, while the values of L_a and L_b are assumed to be given as inputs and need to be replaced with the estimated ENLs of the input PolSAR images.

By assuming nontextured data, the ENL can be estimated from the first-order matrix log-cumulant equation of the scaled complex Wishart distribution as follows [Anfinsen et al., 2009]:

$$\langle \kappa_1\{\mathbf{C}\} \rangle = \ln|\mathbf{\Sigma}| + \psi_d^0(\hat{L}) - d\ln\hat{L} \tag{6.35}$$

where $\psi_d^0(\cdot)$ is the multivariate digamma function [Anfinsen and Eltoft, 2011] and $\hat{L}$ is the ML estimator of the number of looks. The Newton-Raphson numerical optimization

technique is used for solving (6.35). We must insert the first-order sample matrix log-cumulant given by

$$\langle \kappa_1\{\mathbf{C}\}\rangle = \frac{1}{N}\sum_{i=1}^{N}\log|\mathbf{C}_i| \tag{6.36}$$

and an estimate of $\boldsymbol{\Sigma}$ before solving for $\hat{L}$. The ML estimator of $\boldsymbol{\Sigma}$ is given by the sample mean as follows:

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{C}_i. \tag{6.37}$$

In order to avoid manual selection of regions of interest, we apply the method proposed in [Anfinsen et al., 2009] to estimate the ENLs in an unsupervised manner. This method is based on (6.35), and the ENL estimate is extracted from the empirical density of small sample estimates calculated in a sliding window which traverses the whole image. The overall distribution of estimates is expected to be dominated by estimates computed from truly Wishart distributed samples, because of the no-texture assumption. Therefore, the greatest mode value can be used as an estimate of the ENL [Anfinsen et al., 2009].

Using the approximated distribution of the test statistic, the change detection is performed at a given probability of false alarm, and as the approximated distribution is dependent of the parameters contained in the HL moment expressions, the proposed change detector belongs to the class of constant false alarm rate (CFAR) detectors, making it easy to use for multilook PolSAR images. Paper 3 demonstrates the usefulness of the HL statistic for polarimetric change detection. It is compared to the Wishart LRT statistic developed in [Conradsen et al., 2003]. The superior sensitivity to changes of the HL test statistic compared to the LRT statistic is pointed out on real PolSAR data sets in Paper 3.

Chapter 7

Paper 1: published

A Textural-Contextual Model for Unsupervised Segmentation of Multipolarization Synthetic Aperture Radar Images

Chapter 8

Paper 2: under review

Monitoring Glacier Changes by Multitemporal Multipolarization SAR images

Chapter 9

Paper 3: under review

Change Detection for Polarimetric SAR Data with the Hotelling-Lawley Trace Statistic under the Complex Wishart Distribution

Chapter 10

Conclusions and Future Research

In this chapter, we give summary, concluding remarks, and suggestions for future research.

10.1 Summary

The research topic of this study has been multitemporal analysis of multipolarization SAR data. In Chapter 2 and 3 of this thesis, we presented a brief background on the principles behind SAR imaging and an introduction to the PolSAR concept.

Chapter 4 described a contextual MRF-based classification for PolSAR data. The classification scheme is combing the pixelwise matrix-variate statistical distributions for mutilook data with contextual information.

Chapter 5 explained the terrain correction of PolSAR data in three steps. The first step is geometric terrain correction to correct the position of the backscatter value before analyzing the data. The next step is the radiometric terrain correction due to rugged terrain areas and the SAR geometry. The last step of terrain correction is to compensate the polarization signature which is due to azimuthal and range slopes.

Chapter 6 discussed the concept of direct change detection in multilook PolSAR data and reviewed some matrix distance measures that can be used as test statistics for change detection in addition to our proposed test statistic.

Chapters 7-9 include the 3 papers that make up the core of the thesis. The first paper, which studied the use of the textual-contextual classification algorithm, makes a foundation for the second paper. The second paper described a processing chain for post-classification change detection in multitemporal PolSAR data over glaciers. The third paper and paper in Appendix B are linked to each other. In the Appendix paper, the complex Hotelling-Lawley (HL) trace statistic was introduced and the moments of complex-kind HL trace statistics were derived. Paper 3 demonstrated the capability of the HL test statistics for change detection in PolSAR data.

This thesis described two separate approaches for change detection in multipolariz-

ation SAR data; (1) comparative analysis of independently produced classifications for different dates called post-classification change detection (PCCD) and (2) simultaneous analysis of multitemporal SAR data called direct change detection (DCD).

The first approach involves a methodology for a PCCD of Arctic glaciers from multitemporal polarimetric images with different configurations. Using the matrix log-cumulant diagrams, we demonstrate the capability of the matrix-variate non-Gaussian-based $\mathcal{U}_d$-distribution to model the variable texture in the multitemporal PolSAR data set. The post-classification comparison change detection requires the comparison of independently produced classified images. Therefore, the DEM-based terrain corrected multilook images are separately clustered into thematic maps, using fully unsupervised, contextual non-Gaussian image segmentation described in Appendix A. The contextual smoothing by the MRF model yields homogeneous segmentations and minimizes the impact of speckle on the change detection results. The ground truth data are then used to label the segmented images into glacier facies and to investigate the classification accuracies. Finally, the classified images are compared pixel by pixel to produce change maps which show a complete matrix of changes. The variations of the boundaries between glacier facies can be detected within the period of study from the change matrix. The proposed method shows promising potential for operational firn-line monitoring from multitemporal multipolarization images.

The second approach describes a new approach for DCD in polarimetric SAR data under the complex Wishart distribution. We employ the complex HL trace statistic for equality of two complex covariance matrices. Then the distribution of the HL statistic is approximated by a Fisher-Snedecor distribution. Thus, we tested whether two covariance matrices belong to the same population and changes are detected when the hypothesis is rejected. Experiments with simulated and real data sets demonstrate the effectiveness of the algorithm.

10.2 Concluding Remarks

It may be appreciated that when a difference in radar backscattering between multitemporal data is taken as a change indicator, the difference may be due to several factors such as actual land cover change, viewing geometry (like satellite orbit and look angle), surface topography and other external factors (like meteorological conditions, that are minimized by using radar frequencies and choosing the scenes wisely). These considerations were the main reason for developing the workflow in Paper 2, and where we also choose our SAR images to avoid the wet weather conditions. Models accounting for temporal variation in multitemporal SAR images should separate useful temporal variation, i.e., land cover or seasonal changes from variation arising due to other factors mentioned above.

As discussed in Chapter 5, challenges for multitemporal SAR image analysis increase when topography is introduced. So of the various requirements of pre-processing for change detection in SAR images, precise terrain geocoding accounting for geomet-

rical distortions and radiometric terrain correction between multitemporal images are the most important and critical. The importance of accurate geometric registration of multitemporal imagery is obvious because largely spurious results of change detection will be produced if there is misregistration. In cases of misregistration, a number of false alarms, especially in the region of rapid intensity change such as edges, occur. This necessitates the use of precise terrain geocoding of SAR data that not only registers the images to a standard map projection by using a high resolution DEM and precise orbital information, but also performs the geometric terrain correction. Regarding the effects of both terrain topography and SAR geometry on radar backscattering, we refer readers to the results of Figure 4 in Paper 2, where it plots the effect of the terrain correction on four candidate images of dual polarization ASAR scenes. It shows the suppression of the topographical effect and the viewing geometry on the covariance matrix data sets such that the images are comparable for the monitoring purpose.

In general, change detection techniques can be grouped into two types: (1) those detecting binary change/no change information; and (2) those detecting detailed "from-to" change.

The PCCD approach falls into type (1) which can provide a complete matrix of change information. The major advantage of the developed workflow for the PCCD in Paper 2 is that the unsupervised segmentation algorithm together with the DEM-based terrain correction are reliable and robust enough to give consistent change detection results. It should be mentioned that the example glacier was only a very slow moving glacier and the annual change is small. A faster glacier, or other more drastic change examples like deforestation, would likely work much better. Another important point about this approach is that the method is based on the classified images, in which the quality and quantity of the ground truth data are crucial to label the segmentation results and produce high-quality classification results. However, selection of high-quality and sufficient ground truth data for labeling image segmentation results is often difficult, in particular for historical image data classification. When high-quality ground truth data are not available, production of highly accurate classifications from the segmented scenes is a difficult task and post-classification comparison from classified historical image data often seriously affects the change detection results. The key is to create accurate thematic classification images. The errors of individual-date thematic images will affect the final change detection accuracy. High accuracy GPR-based ground truth data and consistent segmentation results in the glacier example of Paper 2 guarantee to obtain high classification accuracies, which was evaluated on classification of the year 2005, and thus robust change detection.

The DCD approach falls into type (2) which can only provide change/no change information. On the other hand, one advantage of the DCD approach compared to the PCCD approach is that it does not need any ground truth data to label every class. So we get "less" information, but we need "less" inputs to get it. The approach belongs to the class of CFAR detectors which makes it easy to use for multilook PolSAR images. One critical step in using the DCD for change/no change detection is to select an appropriate false alarm rate to threshold the test statistic in order to separate areas of change from

those of no change.

The comparison of the results of the proposed test statistic with those of the Wishart LRT statistic for Experiments II and III of Paper 3 reports higher sensitivity of the HL statistic in detecting changes, and thus better detection performance with respect to the LRT detector. For Experiment III, in particular, the HL statistic represents higher sensitivity to the differences in polarimetric information with respect to the LRT detector. In order to better evaluate the change detection test statistics quantitatively, receiver operating characteristic (ROC) curves were plotted for both detectors versus the reference change maps. For both experiments, the ROC plot of the HL is above the LRT indicating better detection performance obtained from the HL statistic.

The statistical modeling of the test statistic in the direct change detection algorithm requires the ENL estimation for each image in the pair of multitemporal PolSAR images in advance. Underestimation of the ENL results in a mismatch of the HL histogram and the estimated PDF of no change class (after thresholding at a predefined false alarm rate). Another source of the mismatch of the HL histogram to the estimated PDF could be due to presence of texture that makes deviation from the Wishart distribution. To reduce this problem, we multilook the original SLC data with high degree of smoothing.

10.3 Suggestions for Future Work

There are still some possibilities to be investigated. Here we list some suggestions for future research which extends the work in this thesis.

- We proposed a method for the post-classification change detection by mutlitemporal multipolarization SAR data acquired over glaciers in the second paper. Future studies will consider how to relate the variations of the boundaries between glacier facies to the glacier mass-balance which is an important indicator of global climate change, and hence wider scientific importance for climate studies.

- The change detection algorithm proposed in the third paper assumes Gaussian-based Wishart distribution for covariance matrices and the distribution of HL trace statistic was approximated under the complex Wishart model. We have found some sets of data with an extreme texture provide partially bad fit between the test statistic histogram and the estimated PDF. Further developments of this study will include extending the HL statistic moments to include texture under the multilook polarimetric product model.

- The change detection method in the third paper is aimed at extracting a scalar feature that accurately discriminates changed and unchanged areas. We have currently only performed this method for CFAR detection. Further possible extensions of this work could be the combination of this feature with automatic Bayesian thresholding techniques, based for example on generalizations of the

K&I approach [Moser and Serpico, 2006], or with contextual unsupervised segmentation methods proposed in Paper 1 for minimizing the impact of speckle on the resulting change map.

Appendix A

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

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 α and λ, 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^{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
$\mathcal{W}_d^{\mathbb{C}}(L, \mathbf{\Sigma})$	$\frac{L^{Ld}\lvert\mathbf{C}\rvert^{L-d}}{\Gamma_d(L)\lvert\mathbf{\Sigma}\rvert^{L}} \exp\left(-L\text{tr}(\mathbf{\Sigma}^{-1}\mathbf{C})\right)$	$\kappa_1\{\mathbf{C}\} = \ln\lvert\mathbf{\Sigma}\rvert + \psi_d^0(L) - d\ln L$ $\kappa_{\nu>1}\{\mathbf{C}\} = \psi_d^{\nu-1}(L)$
$\mathcal{K}_d(L, \mathbf{\Sigma}, \alpha)$	$\frac{2\lvert\mathbf{C}\rvert^{L-d}}{\Gamma_d(L)\Gamma(\alpha)} (L\alpha)^{\frac{\alpha+Ld}{2}} \left(\text{tr}(\mathbf{\Sigma}^{-1}\mathbf{C})\right)^{\frac{\alpha-Ld}{2}}$ $\times K_{\alpha-Ld}\left(2\sqrt{L\alpha\text{tr}(\mathbf{\Sigma}^{-1}\mathbf{C})}\right)$	$\kappa_1\{\mathbf{C}\} = \ln\lvert\mathbf{\Sigma}\rvert + \psi_d^0(L) + d\left(\psi^0(\alpha) - \ln(\alpha L)\right)$ $\kappa_{\nu>1}\{\mathbf{C}\} = \psi_d^{\nu-1}(L) + d^{\nu}\psi^{\nu-1}(\alpha)$
$\mathcal{G}_d^0(L, \mathbf{\Sigma}, \lambda)$	$\frac{L^{Ld}\lvert\mathbf{C}\rvert^{L-d}}{\Gamma_d(L)\lvert\mathbf{\Sigma}\rvert^{L}} \frac{\Gamma(Ld+\lambda)(\lambda-1)^{\lambda}}{\Gamma(\lambda)}$ $\left(L\text{tr}(\mathbf{\Sigma}^{-1}\mathbf{C}) + \lambda - 1\right)^{-\lambda-Ld}$	$\kappa_1\{\mathbf{C}\} = \ln\lvert\mathbf{\Sigma}\rvert + \psi_d^0(L) + d\left(\ln(\frac{\lambda-1}{L}) - \psi^0(\lambda)\right)$ $\kappa_{\nu>1}\{\mathbf{C}\} = \psi_d^{\nu-1}(L) + (-d)^{\nu}\psi^{\nu-1}(\lambda)$
$\mathcal{U}_d(L, \mathbf{\Sigma}, \alpha, \lambda)$	$\frac{L^{Ld}}{\Gamma_d(L)} \frac{\lvert\mathbf{C}\rvert^{L-d}}{\mathbf{\Sigma}^{L}} \frac{\Gamma(\alpha+\lambda)\Gamma(Ld+\lambda)}{\Gamma(\alpha)\Gamma(\lambda)} \left(\frac{\alpha}{\lambda-1}\right)^{Ld}$ $\times U\left(Ld+\lambda, Ld-\alpha+1, L\text{tr}(\mathbf{\Sigma}^{-1}\mathbf{C})(\frac{\alpha}{\lambda-1})\right)$	$\kappa_1\{\mathbf{C}\} = \ln\lvert\mathbf{\Sigma}\rvert + \psi_d^0(L) + d\left(\psi^0(\alpha) - \psi^0(\lambda) + \ln(\frac{\lambda-1}{\alpha L})\right)$ $\kappa_{\nu>1}\{\mathbf{C}\} = \psi_d^{\nu-1}(L) + d^{\nu}\left(\psi^{\nu-1}(\alpha) + (-1)^{\nu-1}\psi^{\nu-1}(\lambda)\right)$

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.

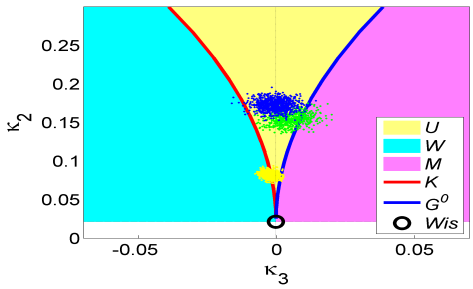


Figure 1: Log-cumulant diagram for real dual-polarisation ASAR scene of an Arctic glacier, showing clusters for firn, superimposed and glacier ice regions. Note that the clusters are clearly within the yellow $\mathcal{U}_d$-distribution region in the diagram.

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.

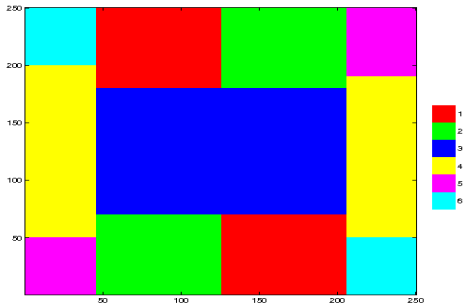


Figure 2: Perfect clustering result of a six-class simulated test image.

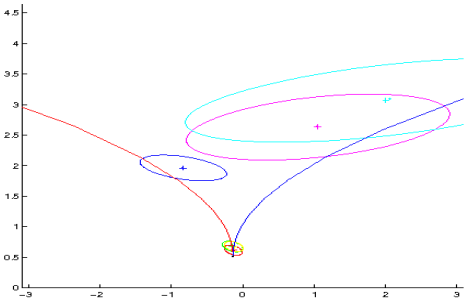


Figure 3: Log-cumulant diagram of simulated clustering result. The coloured ellipses indicate the six different classes and their expected sample scatter, and the red and blue curved lines indicate the $\mathcal{K}_d$ and $\mathcal{G}_d^0$ models, respectively. Note that several classes are nearly $\mathcal{W}_d$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

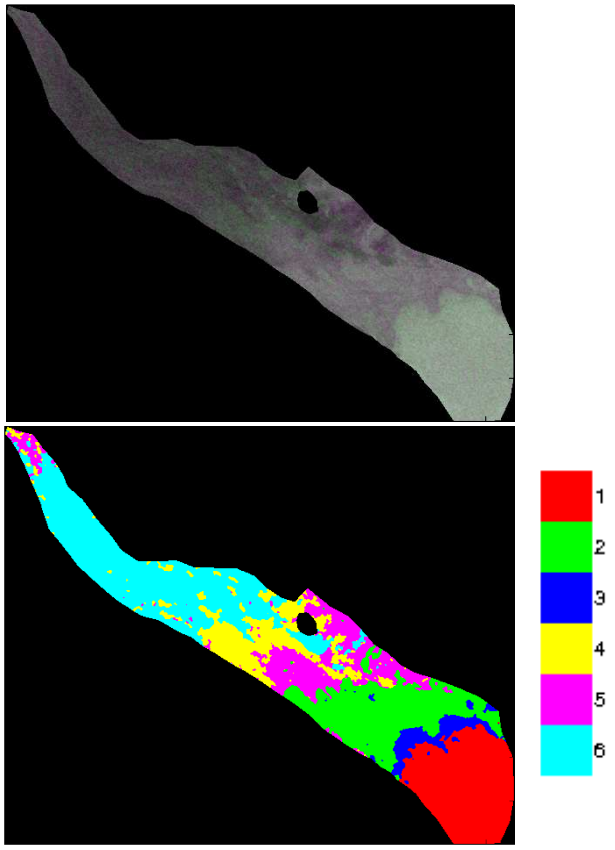


Figure 4: Real ENVISAT, dual-pol, 24-look example of an Arctic glacier. Pseudo-Pauli RGB image (top) and the automatically clustered image (below) which found six clusters in good (visual) correspondence to real glacier classes.

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.

Appendix B

Statistical Characterisation of the Complex Hotelling-Lawley Trace Statistic

Statistical Characterisation of the Complex Hotelling-Lawley Trace Statistic

Stian Normann Anfinsen and Vahid Akbari

***Abstract*—The paper introduces a complex version of the Hotelling-Lawley (HL) trace statistic, whose real case equivalent is an important test statistic in the field of multivariate analysis and in multivariate analysis of variance in particular. The first, second and third-order moments of the HL trace statistic are derived. These are used to specify the parameters of a Fisher-Snedecor distribution, which approximates the sampling distribution of the test statistic. The application of the model to change detection in polarimetric radar images is discussed.**

***Index Terms*—Complex Wishart distribution, Hotelling-Lawley trace statistic, matrix moments, change detection**

I. Introduction

The Hotelling-Lawley (HL) trace statistic is an important test statistic in real multivariate analysis, where it is used to test two or more multidimensional populations for equality of means. This is done in a procedure known as multivariate analysis of variance (MANOVA) [1], [2], where the sample covariance matrices are used to test the statistical significance of differences in the mean vectors, under the assumption that the populations are multivariate Gaussian. The HL trace is one of the common choices for the test statistic, together with Wilks' lambda, the Pillai-Bartlett trace statistic, and Roy's greatest root [3].

By its definition, the HL trace statistic directly compares two sample covariance matrices, that are usually assumed to follow a central or noncentral Wishart distribution. It can thus be used to test if the sample covariance matrices come from the same population, under the preassumption that their means are equal. This is our motivation for studying the complex HL trace statistic: It can be used for change detection in polarimetric synthetic aperture radar (PolSAR) images, where each pixel is represented by a sample covariance matrix commonly modelled with the complex Wishart distribution [4], [5]. To be useful in this context, the HL trace statistic must be extended to the complex case. The purpose of this paper is thus to characterise the complex kind HL trace statistic in terms of its moments and its sampling distribution.

The signal processing community has witnessed an increasing interest in complex matrix-variate statistics, after these have become important tools in the analysis of wireless communications systems [6] and polarimetric radar images [7], [8]. As specific examples, complex Wishart matrices occur frequently in the performance analysis of wireless networks, such as derivations of channel and receiver capacity [9]–[11]. In polarimetric radar imaging, the complex Wishart distribution is the default model for the sample covariance of the scattering vector, which is the measurement most commonly used in monitoring of natural surfaces. In both areas, the phase information of the information carrying signal requires a complex-valued description, and leads to problems that require analysis of complex random matrices, their determinants, traces and eigenvalues.

Returning to the comparison of covariance matrices, we believe that the complex HL trace makes a good test statistic for change detection in PolSAR images, both in terms of accuracy, computational cost and ease of implementation. In parallel with this work, we have therefore proposed an algorithm which uses this statistic to compare sample covariance matrices representing multitemporal polarimetric measurements [12]. The current paper provides the underlying theory for the proposed change detector. Most importantly, it provides a sampling distribution for the test statistic, which is needed to determine an appropriate threshold for testing the hypotheses of change versus no change.

For all of the test statistics commonly used in MANOVA, it is difficult to derive an exact sampling distributions in the the most general case specified by a finite data dimension and degrees of freedom of the sample covariance matrix [13]. The true sampling distribution of the real case HL trace statistic was first derived for two-dimensional data by Hotelling [14]. Expressions for a general data dimension were then given in [15] and [16], although the complexity of the solutions limit their applicability. An approximation for the general sampling distribution based on matching moments of the F-distribution and the HL trace statistic was first proposed in [17], and then improved in [18], before McKeon [19] gave an alternative which is reported to be the superior one [3].

Whereas much of the literature of multivariate statistics has been extended from the real to the complex case, little work has been done on the HL trace statistic. An exception is Kabe's discussion of what is referred to as Hotelling's complex T^2 statistic [20], which is proportional to the complex HL trace statistic. In the absence of other results, we therefore copy the approach used in the real scenario [17]–[19] to produce an approximation for the sampling distribution of the complex HL trace. The Fisher-Snedecor (FS) distribution [21], which is an F-distribution extended with a location parameter, is taken as the approximate density model for the complex HL trace. We then match three of its population moments with the corresponding theoretical moments of the HL trace statistic, and solve the equation system for the parameters of the FS distribution.

The paper proceeds as follows. Section II introduces the central complex Wishart distribution and reviews previous results on moments of the trace of a Wishart matrix. A new derivation of cumulants of the Wishart matrix trace is also presented. Section III defines the complex HL trace

statistic. Moments of the HL trace are then derived from the moments of the trace of complex Wishart matrices and inverse complex Wishart matrices. The section further introduces the Fisher-Snedecor distribution and its moment expressions, and explains procedure used to solve for its parameters. This concludes the theoretical results.

The notational convention of the paper is that matrices are written as boldface upper case letters and vectors as boldface lower case letters. Scalars are written as letters with standard weight. Ordinary and Hermitian transposition of vectors and matrices are denoted as $(\cdot)^T$ and $(\cdot)^H$, respectively. The trace function is denoted as $\mathrm{tr}(\cdot)$, the determinant function as $|\cdot|$, and the expectation operator with respect to the random variate X as $\mathrm{E}_X\{\cdot\}$. The distinction between random variates and their realisations must be ascertained through context.

II. The complex Wishart matrix

This section presents the complex Wishart and its density function, the complex Wishart distribution. We make the distinction between the true complex Wishart matrix, as defined in the seminal study of Goodman [22], and the scaled complex Wishart matrix [23], a term used to distinguish the sample covariance matrix normalised (or scaled) by the number of vector samples used in its computation. We then summarise moment expression for the trace of products of complex Wishart matrices and inverse complex Wishart matrices, that are needed in subsequent derivations.

A. The complex Wishart distribution

Assume that $\mathcal{Z}=\{\mathbf{z}_i\}_{i=1}^n$ is a sample of size $d\times 1$ vectors drawn from a circular, complex and zero-mean multivariate Gaussian distribution, given as [24]

$$p_{\mathbf{z}}(\mathbf{z};\mathbf{0},\boldsymbol{\Sigma}) = \frac{1}{\pi|\boldsymbol{\Sigma}|}\exp\left\{-\mathbf{z}^H\boldsymbol{\Sigma}^{-1}\mathbf{z}\right\}. \tag{1}$$

This is denoted as $\mathbf{z}_i \sim \mathcal{N}_d^{\mathbb{C}}(\mathbf{0},\boldsymbol{\Sigma})$, where $\mathbf{0}$ is a zero vector and $\boldsymbol{\Sigma} = \mathrm{E}\{\mathbf{z}_i\mathbf{z}_i^H\}$ is the population covariance matrix. Further let the random matrix $\mathbf{W}$ be defined as

$$\mathbf{W} = \sum_{i=1}^{L}\mathbf{z}_i\mathbf{z}_i^H. \tag{2}$$

We then refer to $\mathbf{W}$ as a complex Wishart matrix whose size is $d\times d$. It is defined on the cone of positive definite and Hermitian matrices, a domain denoted as Ω_+, and follows the nonsingular central complex Wishart distribution, whose probability density function (pdf) is [22]

$$p_{\mathbf{W}}(\mathbf{W};L,\boldsymbol{\Sigma}) = \frac{|\mathbf{W}|^{L-d}}{|\boldsymbol{\Sigma}|^L\Gamma_d(L)}\exp\left\{-\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{W})\right\}, \tag{3}$$

provided that $L \geq d$. This is written as $\mathbf{W} \sim \mathcal{W}_d^{\mathbb{C}}(L,\boldsymbol{\Sigma})$, where L is the shape parameter (or the degrees of freedom) and $\boldsymbol{\Sigma} = \mathrm{E}\{\mathbf{W}\}/L$ is the scale matrix. The complex Wishart pdf contains the multivariate gamma function of the complex kind,

$$\begin{aligned}\Gamma_d(L) &= \int_{\Omega_+}|\mathbf{W}|\exp\{-\mathrm{tr}(\mathbf{W})\}\,d\mathbf{W}\\ &= \pi^{d(d-1)/2}\prod_{i=0}^{d-1}\Gamma(L-i),\end{aligned} \tag{4}$$

as a normalisation constant.

Let $\mathbf{C}$ be a sample covariance matrix, which is computed from $\mathcal{Z}$ by the sample mean estimator as

$$\mathbf{C} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{z}_i\mathbf{z}_i^H,$$

which is also the maximum likelihood estimate of $\boldsymbol{\Sigma}$ based on $\mathcal{Z}$. It is then known that $\mathbf{C} \sim \mathcal{W}_d^{\mathbb{C}}(L,\boldsymbol{\Sigma}/L)$, and the pdf of $\mathbf{C}$ becomes [23]

$$p_{\mathbf{C}}(\mathbf{C};L,\boldsymbol{\Sigma}) = \frac{L^{Ld}}{\Gamma_d(L)}\frac{|\mathbf{C}|^{L-d}}{|\boldsymbol{\Sigma}|^L}\exp\left\{-L\,\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C})\right\}. \tag{5}$$

This is denoted $\mathbf{C} \sim s\mathcal{W}_d^{\mathbb{C}}(L,\boldsymbol{\Sigma})$, and the pdf is referred to as the scaled complex Wishart distribution, to distinguish it from (3). Accordingly, we refer to $\mathbf{C}$ as the scaled complex Wishart matrix. From here on, we shall use $\mathbf{C}$ instead of $\mathbf{W}$ in our derivations, since it represents the sample covariance matrix which is the relevant observation in most practical cases, including polarimetric radar imaging.

We finally note that mean of the scaled complex Wishart matrix and its inverse are given by [25], [26]

$$\mathrm{E}\{\mathbf{C}\} = \boldsymbol{\Sigma}, \tag{6}$$

$$\mathrm{E}\{\mathbf{C}^{-1}\} = \frac{L}{Q}\boldsymbol{\Sigma}^{-1}. \tag{7}$$

B. Trace moments of the complex Wishart matrix

Let $\mathbf{C}^n = \mathbf{C}\cdots\mathbf{C}$ denote the matrix power, defined as a repeated matrix product of n scaled complex Wishart matrices. The matrix power may also be extended by pre-multiplication with a constant matrix $\mathbf{A}\in\Omega_+$, such that $(\mathbf{AC})^n = (\mathbf{AC})\cdots(\mathbf{AC})$. In the following, we are interested in population moments of the trace of such matrix products, $\mathrm{E}_{\mathbf{C}}\{\mathrm{tr}((\mathbf{AC})^n)\}$, and also in the moments of compositions of such trace expressions, for instance $\mathrm{E}_{\mathbf{C}}\{\mathrm{tr}((\mathbf{AC})^m)\,\mathrm{tr}((\mathbf{AC})^n)\}$. These expressions are referred to as trace moments of order $m+n$, where the order is the multiplicity of the trace terms. They are needed to derive moments of the complex HL trace statistic.

The derivation of the required trace moments was explained and exemplified for low orders by Maiwald and Kraus [25], [27], but only some of the moments were given explicitly, and sometimes in a truncated version. We therefore turn to Graczyk et al., who used theory of the symmetric group of permutations to derive elegant formulas for trace moments of general orders [26]. We here provide explicit results of their method for the convenience of the reader.

The first-order trace moment is

$$\mathrm{E}_{\mathbf{C}}\{\mathrm{tr}(\mathbf{AC})\} = \mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}). \tag{8}$$

The second-order moments are

$$\mathrm{E}_{\mathbf{C}}\{\mathrm{tr}(\mathbf{AC})^2\} = \mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma})^2 + \frac{1}{L}\mathrm{tr}((\mathbf{A}\boldsymbol{\Sigma})^2), \tag{9}$$

$$\mathrm{E}_{\mathbf{C}}\{\mathrm{tr}((\mathbf{AC})^2)\} = \mathrm{tr}((\mathbf{A}\boldsymbol{\Sigma})^2) + \frac{1}{L}\mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma})^2. \tag{10}$$

The third-order moments are

$$\begin{aligned}\mathrm{E}_{\mathbf{C}}\{\mathrm{tr}(\mathbf{AC})^3\} &= \mathrm{tr}(\mathbf{A\Sigma})^3 \\ &+ \frac{3}{L}\,\mathrm{tr}(\mathbf{A\Sigma})\,\mathrm{tr}((\mathbf{A\Sigma})^2) + \frac{2}{L^2}\,\mathrm{tr}((\mathbf{A\Sigma})^3)\,.\end{aligned} \tag{11}$$

$$\begin{aligned}\mathrm{E}_{\mathbf{C}}\{\mathrm{tr}(\mathbf{AC})\,\mathrm{tr}((\mathbf{AC})^2)\} &= \mathrm{tr}(\mathbf{A\Sigma})\,\mathrm{tr}((\mathbf{A\Sigma})^2) \\ &+ \frac{1}{L}(2\,\mathrm{tr}((\mathbf{A\Sigma})^3) + \mathrm{tr}(\mathbf{A\Sigma})^3) \\ &+ \frac{2}{L}\,\mathrm{tr}(\mathbf{A\Sigma})\,\mathrm{tr}((\mathbf{A\Sigma})^2)\,.\end{aligned} \tag{12}$$

$$\begin{aligned}\mathrm{E}_{\mathbf{C}}\{\mathrm{tr}(\mathbf{AC})\,\mathrm{tr}((\mathbf{AC})^2)\} &= \mathrm{tr}((\mathbf{A\Sigma})^3) \\ &+ \frac{3}{L}\,\mathrm{tr}(\mathbf{A\Sigma})\,\mathrm{tr}((\mathbf{A\Sigma})^2) \\ &+ \frac{1}{L^2}(\mathrm{tr}(\mathbf{A\Sigma})^3 + \mathrm{tr}((\mathbf{A\Sigma})^3))\,.\end{aligned} \tag{13}$$

The linearity of the trace operator together with $\mathbf{W} = L\mathbf{C}$ asserts that the equivalent trace moments of the complex Wishart matrix $\mathbf{W}$ can be found as

$$\mathrm{E}_{\mathbf{W}}\{f(\mathbf{W})\} = L^n\,\mathrm{E}_{\mathbf{C}}\{f(\mathbf{C})\} \tag{14}$$

where n is the moment order and $f(\cdot)$ is any of the trace expressions subjected to the expectation operator in Eqs. (8)-(13).

C. Trace moments of the inverse complex Wishart matrix

Trace moments of the inverse scaled complex Wishart matrix, $\mathbf{C}^{-1}$, are also derived in [25]. However, to obtain full versions of all the required expressions, we must also in this case use the methods introduced in [26].

The first-order moment is

$$E_{\mathbf{C}}\{\mathrm{tr}(\mathbf{C}^{-1}\mathbf{A})\} = \frac{L}{Q}\,\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{A})\,. \tag{15}$$

For the second-order moments, Theorem 2 in [26] provides the equation system

$$\begin{bmatrix} \mathrm{tr}(\mathbf{C}^{-1}\mathbf{A})^2 \\ \mathrm{tr}((\mathbf{C}^{-1}\mathbf{A})^2) \end{bmatrix} = \frac{1}{L^2}\begin{bmatrix} Q^2 & -Q \\ -Q & Q^2 \end{bmatrix}\mathrm{E}\left\{\begin{bmatrix} \mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{A})^2 \\ \mathrm{tr}((\mathbf{\Sigma}^{-1}\mathbf{A})^2) \end{bmatrix}\right\}, \tag{16}$$

where we define $Q = L - d$. The system is inverted to yield

$$E\{\mathrm{tr}(\mathbf{C}^{-1}\mathbf{A})^2\} = \frac{L^2}{Q^3 - Q}(Q\,\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{A})^2 + \mathrm{tr}((\mathbf{\Sigma}^{-1}\mathbf{A})^2)\,, \tag{17}$$

$$E\{\mathrm{tr}((\mathbf{C}^{-1}\mathbf{A})^2)\} = \frac{L^2}{Q^3 - Q}(Q\,\mathrm{tr}((\mathbf{\Sigma}^{-1}\mathbf{A})^2) + \mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{A})^2\,. \tag{18}$$

In the same manner, we find that the third order moments are defined by the equation system

$$\begin{bmatrix} \mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{A})^3 \\ \mathrm{tr}((\mathbf{\Sigma}^{-1}\mathbf{A})^2)\,\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{A}) \\ \mathrm{tr}((\mathbf{\Sigma}^{-1}\mathbf{A})^3) \end{bmatrix} = \begin{bmatrix} Q^3 & -3Q^2 & 2Q \\ -Q^2 & Q^3{+}2Q & -2Q^2 \\ Q & -3Q^2 & Q^3{+}Q \end{bmatrix} \times \frac{1}{L^3}\,\mathrm{E}\left\{\begin{bmatrix} \mathrm{tr}(\mathbf{C}^{-1}\mathbf{A})^3 \\ \mathrm{tr}((\mathbf{C}^{-1}\mathbf{A})^2)\,\mathrm{tr}(\mathbf{C}^{-1}\mathbf{A}) \\ \mathrm{tr}((\mathbf{C}^{-1}\mathbf{A})^3) \end{bmatrix}\right\}, \tag{19}$$

which can be inverted into

$$\begin{aligned}\mathrm{E}\{\mathrm{tr}(\mathbf{C}^{-1}\mathbf{A})^3\} &= \frac{L^3}{Q^5 - 5Q^3 + 4Q} \\ &\times \Big((Q^2 - 2)\,\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{A})^3 \\ &+ 3Q\,\mathrm{tr}((\mathbf{\Sigma}^{-1}\mathbf{A})^2)\,\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{A}) \\ &+ 4\,\mathrm{tr}((\mathbf{\Sigma}^{-1}\mathbf{A})^3)\Big)\,,\end{aligned} \tag{20}$$

$$\begin{aligned}\mathrm{E}\{\mathrm{tr}((\mathbf{C}^{-1}\mathbf{A})^2)\,\mathrm{tr}(\mathbf{C}^{-1}\mathbf{A})\} &= \frac{L^3}{Q^5 - 5Q^3 + 4Q} \\ &\times \Big(Q\,\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{A})^3 \\ &+ (Q^2 + 2)\,\mathrm{tr}((\mathbf{\Sigma}^{-1}\mathbf{A})^2)\,\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{A}) \\ &+ 2Q\,\mathrm{tr}((\mathbf{\Sigma}^{-1}\mathbf{A})^3)\Big)\,,\end{aligned} \tag{21}$$

$$\begin{aligned}\mathrm{E}\{\mathrm{tr}((\mathbf{C}^{-1}\mathbf{A})^3)\} &= \frac{L^3}{Q^5 - 5Q^3 + 4Q} \\ &\times \Big(2\,\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{A})^3 \\ &+ 3Q\,\mathrm{tr}((\mathbf{\Sigma}^{-1}\mathbf{A})^2)\,\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{A}) \\ &+ Q^2\,\mathrm{tr}((\mathbf{\Sigma}^{-1}\mathbf{A})^3)\Big)\,.\end{aligned} \tag{22}$$

Again, the equivalent nth-order trace moments of $\mathbf{W}^{-1}$ are found as

$$\mathrm{E}_{\mathbf{W}}\{f(\mathbf{W}^{-1})\} = \frac{1}{L^n}\,\mathrm{E}_{\mathbf{C}}\{f(\mathbf{C}^{-1})\}\,. \tag{23}$$

III. The complex Hotelling-Lawley trace statistic

Let $\mathbf{X} \sim s\mathcal{W}_d^{\mathbb{C}}(L_x, \mathbf{\Sigma})$ and $\mathbf{Y} \sim s\mathcal{W}_d^{\mathbb{C}}(L_y, \mathbf{\Sigma})$ be two random matrices that follow the scaled complex Wishart distribution in (5). Further assume that $\mathbf{X}$ and $\mathbf{Y}$ are statistically independent, which is a reasonable assumption if the matrices represent time separated polarimetric measurements of a target which exhibits distributed scattering. The complex HL trace statistic is then defined as

$$\tau_{HL} = \mathrm{tr}(\mathbf{X}^{-1}\mathbf{Y})\,. \tag{24}$$

The order of the matrices in (24) can of course be reversed. In the context of change detection, it can be relevant to consider both $\mathrm{tr}(\mathbf{X}^{-1}\mathbf{Y})$ and $\mathrm{tr}(\mathbf{Y}^{-1}\mathbf{X})$ if we want to apply these test statistics sequentially and thus use the same procedure to test for changes both in the direction of increase and decrease of the radar backscatter. In the following, we analyse only $\mathrm{tr}(\mathbf{X}^{-1}\mathbf{Y})$, since the results are equivalent.

A. Moment derivation

We now want to derive some lower-order moments of τ_{HL}. The derivation is aided by the assumed independency of $\mathbf{X}$ and $\mathbf{Y}$, since we can apply the expectation operator sequentially with respect to $\mathbf{X}$ and $\mathbf{Y}$. With the moment expressions listed in sections II-B and II-C, we are now equipped for the derivation.

The mean of τ_{HL} is found by use of Eqs. (6) and (7) as

$$\mathrm{E}\{\mathrm{tr}(\mathbf{X}^{-1}\mathbf{Y})\} = \mathrm{tr}(\mathrm{E}_{\mathbf{X}}\{\mathbf{X}^{-1}\}\,\mathrm{E}_{\mathbf{Y}}\{\mathbf{Y}\}) = \frac{L_x}{Q_x}d\,, \quad (25)$$

where $Q_x = L_x - d$. The second-order moment of τ_{HL} is found from Eqs. (9), (17) and (10) as

$$\begin{aligned}\mathrm{E}\{\mathrm{tr}(\mathbf{X}^{-1}\mathbf{Y})^2\} &= \mathrm{E}_{\mathbf{X}}\{\mathrm{E}_{\mathbf{Y}}\{\mathrm{tr}(\mathbf{X}^{-1}\mathbf{Y})^2\}\} \\ &= \mathrm{E}_{\mathbf{X}}\{\mathrm{tr}(\mathbf{X}^{-1}\boldsymbol{\Sigma})^2\} + \frac{1}{L_y}\mathrm{E}_{\mathbf{X}}\{\mathrm{tr}(\mathbf{X}^{-1}\boldsymbol{\Sigma}\mathbf{X}^{-1}\boldsymbol{\Sigma})\} \\ &= \frac{L_x^2}{Q_x^3 - Q_x}\left(d^2\left(Q_x + \frac{1}{L_y}\right) + d\left(\frac{Q_x}{L_y} + 1\right)\right).\end{aligned} \quad (26)$$

From Eqs. (11), (20), (21) and (22), we obtain the third-order moment as

$$\begin{aligned}&\mathrm{E}\{\mathrm{tr}(\mathbf{X}^{-1}\mathbf{Y})^3\} = \mathrm{E}_{\mathbf{X}}\{\mathrm{E}_{\mathbf{Y}}\{\mathrm{tr}(\mathbf{X}^{-1}\mathbf{Y})^3\}\} \\ &= \mathrm{E}_{\mathbf{X}}\left\{\mathrm{tr}(\mathbf{X}^{-1}\boldsymbol{\Sigma})^3\} + \frac{3}{L_y}\,\mathrm{tr}((\mathbf{X}^{-1}\boldsymbol{\Sigma})^2)\,\mathrm{tr}(\mathbf{X}^{-1}\boldsymbol{\Sigma})\right. \\ &\qquad \left. + \frac{2}{L_y^2}\,\mathrm{tr}((\mathbf{X}^{-1}\boldsymbol{\Sigma})^3)\right\} \\ &= \frac{L_x^3}{Q_x^5 - 5Q_x^3 + 4Q_x} \times \left[d^3\left((Q_x^2 - 1) + \frac{3Q_x}{L_y} + \frac{4}{L_y^2}\right)\right. \\ &\qquad + d^2\left(3Q_x + \frac{3(Q_x^2+2)}{L_y} + \frac{6Q_x}{L_y^2}\right) \\ &\qquad \left. + d\left(4 + \frac{6Q_x}{L_y} + \frac{2Q_x^2}{L_y^2}\right)\right].\end{aligned} \quad (27)$$

The moments above represent the hypothesis that $\mathbf{X}$ and $\mathbf{Y}$ come from populations with the same scale matrix $\boldsymbol{\Sigma}$. This would be the null hypothesis in a change detection algorithm for PolSAR data on covariance matrix format, where $\boldsymbol{\Sigma}$ characterises the polarimetric radar reflectivity of the imaged surface. We may further assume that $L_x = L_y$, such that we are testing for totally identical populations. This is an even more realistic case for change detection in PolSAR images, as the multitemporal images are likely to be processed with the same parameters, including the level of averaging and noise suppression, which is effectively expressed as the shape parameter L.

Note that the moments we obtain in this final case are identical to those we get by assuming that $\mathbf{X}$ and $\mathbf{Y}$ follow the true complex Wishart distribution in (3).

B. The Fisher-Snedecor distribution

We shall approximate the distribution of the complex kind Hotelling-Lawley statistic (τ_{HL}) by a FS distribution. The motivation is that this is the true distribution in the univariate case [18], [19], and therefore it is assumed to be an acceptable model also in the general matrix-variate case.

The FS distribution is given as [21], [28]

$$f_\tau(\tau;\mu,\alpha,\lambda) = \frac{1}{B(\alpha,\lambda)}\,\frac{\alpha}{(\lambda-1)\mu}\,\frac{\left(\frac{\alpha\tau}{(\lambda-1)\mu}\right)^{\alpha-1}}{\left(\frac{\alpha\tau}{(\lambda-1)\mu}+1\right)^{\alpha+\lambda}}\,; \quad \tau \geq 0\,, \quad (28)$$

where $\alpha > 0$ and $\lambda > 0$ are shape parameters, $\mu = \mathrm{E}\{\tau\} > 0$ is a location parameter, and $B(\alpha,\lambda)$ is the beta function defined as

$$\begin{aligned}B(x,y) &= \int_0^1 u^{x-1}(1-x)^{y-1}\,du \\ &= \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}\,.\end{aligned} \quad (29)$$

If τ follows the FS distribution, denoted as $\tau \sim \mathcal{FS}(\mu,\alpha,\lambda)$, then the νth-order moment of τ is given by

$$m_\nu\{\tau;\mu,\alpha,\lambda\} = \left(\frac{(\lambda-1)\mu}{\alpha}\right)^\nu \frac{\Gamma(\alpha+\nu)}{\Gamma(\alpha)}\,\frac{\Gamma(\lambda-\nu)}{\Gamma(\lambda)}\,. \quad (30)$$

From (30), we can extract the first three moments and simplify them as

$$m_1\{\tau\} = \mu\,, \quad (31)$$

$$m_2\{\tau\} = \left(\frac{\alpha+1}{\alpha}\right)\left(\frac{\lambda-1}{\lambda-2}\right)\mu^2\,, \quad (32)$$

$$m_3\{\tau\} = \frac{(\alpha+1)(\alpha+2)}{\alpha^2}\,\frac{(\lambda-1)^2}{(\lambda-2)(\lambda-3)}\mu^3\,. \quad (33)$$

C. Distribution parameters

The method used to specify the FS distribution which approximates the sampling distribution of τ_{HL} is to match the population moments of the FS distribution from Eqs. (31)-(33) with the population moments of τ_{HL} found in Eqs. (25)-(27). This results in a system of equations which we have solved with the Mathematica software package [29] to obtain Eqs. (34)-(36) on page 5.

REFERENCES

[1] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 3rd ed. New York, USA: John Wiley & Sons, 2003.

[2] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis*. London, UK: Academic Press, 1979.

[3] G. A. F. Seber, *Multivariate Observations*. Hoboken, USA: John Wiley & Sons, 1984.

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

[5] K. Conradsen, A. A. Nielsen, J. Schou, and H. Skriver, "A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 1, pp. 4–19, Jan. 2003.

[6] A. L. Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications*. Delft, The Netherlands: now Publishers, 2004.

[7] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*, ser. Optical Science and Engineering. Boca Raton, USA: CRC Press, 2009, no. 143.

[8] S. R. Cloude, *Polarisation: Applications in Remote Sensing*. Oxford, UK: Oxford University Press, 2010.

[9] M. Kang and M.-S. Alouini, "Largest eigenvalue of complex Wishart matrices and performance analysis of MIMO MRC systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 418–428, Apr. 2003.

[10] L. G. Ordóñez, D. P. Palomar, and J. R. Follonosa, "Ordered eigenvalues of a general class of Hermitian random matrices with application to the performance analysis of MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 672–689, Feb. 2009.

[11] B. Nosrat-Makouei, J. G. Andrews, and J. Robert W. Heath, "MIMO interference alignment over correlated channels with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2783–2794, Jun. 2011.

$$\mu = \frac{L}{Q} d \,, \tag{34}$$

$$\alpha = -\frac{2d(-2d + 2d^3 + 4L - 5d^2L - 4dL^2 + 2d^3L^2 + 8L^3 - 4d^2L^3 + 2dL^4)}{2d^4 + 12dL - 12d^3L + d^5L - 24L^2 + 16d^2L^2 - 2d^4L^2 - 4dL^3 + d^3L^3} \tag{35}$$

$$\lambda = \frac{-6d^2 + 6d^4 + 26dL - 27d^3L + d^5L - 28L^2 + 23d^2L^2 + d^4L^2 + 18dL^3 - 9d^3L^3 - 20L^4 + 11d^2L^4 - 4dL^5}{-2d^2 + 2d^4 + 14dL - 13d^3L + d^5L - 20L^2 + 21d^2L^2 - 3d^4L^2 - 6dL^3 + 3d^3L^3 - 4L^4 - d^2L^4} \tag{36}$$

[12] V. Akbari, S. N. Anfinsen, A. P. Doulgeris, T. Eltoft, G. Moser, and S. B. Serpico, "Change detection for polarimetric SAR data with the Hotelling-Lawley trace statistic under the complex Wishart distribution," *IEEE Trans. Geosci. Remote Sens.*, p. 15, Apr. 2013, submitted. [Online]. Available: http://eo.uit.no/publications/VA-TGRS-13sub.pdf

[13] A. M. Mathai, "A review of the different techniques used for deriving the exact distributions of multivariate test criteria," *Sankhyā, Indian J. Statist., ser. A*, vol. 35, no. 1, pp. 39–60, Mar. 1973.

[14] H. Hotelling, "A generalized T test and measure of multivariate dispersion," in *Proc. Second Berkeley Symp. Math. Statist. Prob.*, J. Neyman, Ed. Berkeley, USA: University of California Press, 31 Jul-12 Aug 1951, pp. 23–41.

[15] K. C. S. Pillai and D. L. Young, "On the exact distribution of Hotelling's generalized T_0^2," *J. Multivar. Anal.*, vol. 1, no. 1, pp. 90–107, Apr. 1971.

[16] P. R. Krishnaiah and T. C. Chang, "On the exact distributions of the traces of $S_1(S_1 + S_2)^{-1}$ and $S_1 S_2^{-1}$," *Sankhyā, Indian J. Statist., ser. A*, vol. 34, no. 2, pp. 153–160, Jun. 1972.

[17] K. C. S. Pillai and P. Samson, Jr., "On Hotelling's generalization of T^2," *Biometrika*, vol. 46, no. 1/2, pp. 160–168, Jun. 1959.

[18] D. T. Hughes and J. G. Saw, "Approximating the percentage points of Hotelling's generalized T_0^2 statistic," *Biometrika*, vol. 59, no. 1, pp. 224–226, Apr. 1972.

[19] J. J. McKeon, "F approximations to the distribution of Hotelling's T_0^2," *Biometrika*, vol. 61, no. 2, pp. 381–383, Aug. 1974.

[20] D. G. Kabe, "Some aspects of analysis of variance and covariance theory for a certain multivariate complex Gaussian distribution," *Metrika*, vol. 13, no. 1, pp. 86–97, Dec. 1968.

[21] S. N. Anfinsen, "Statistical analysis of multilook polarimetric radar images with the Mellin transform," Ph.D. dissertation, University of Tromsø, Tromsø, Norway, May 2010.

[22] 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, Mar. 1963.

[23] 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.

[24] A. van den Bos, "A Cramér-Rao lower bound for complex parameters," *IEEE Trans. Signal Process.*, vol. 42, no. 10, pp. 2859–2859, Oct. 1994.

[25] D. Maiwald and D. Kraus, "Calculation of moments of complex Wishart and complex inverse Wishart distributed matrices," *IEE Proc. Radar, Sonar, Navigation*, vol. 147, no. 4, pp. 162–168, 2000.

[26] P. Graczyk, G. Letac, and H. Massam, "The complex Wishart distribution and the symmetric group," *Ann. Statist.*, vol. 31, no. 1, pp. 287–309, Feb. 2003.

[27] D. Maiwald and D. Kraus, "On moments of complex Wishart and complex inverse Wishart distributed matrices," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP'97)*, vol. 5, Munich, Germany, Apr. 1997, pp. 3817–3820.

[28] C. Tison, J.-M. Nicolas, F. Tupin, and H. Maître, "A new statistical model for Markovian classification of urban areas in high-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2046–2057, Oct. 2004.

[29] Wolfram Research, Inc., "Mathematica, Version 9.0.1.0," Champaign, USA, 2013.

Bibliography

[Akbari et al., 2013] Akbari, V., Doulgeris, A., Moser, G., Eltoft, T., Anfinsen, S., and Serpico, S. (2013). A textural–contextual model for unsupervised segmentation of multipolarization synthetic aperture radar images. *IEEE Transactions on Geoscience and Remote Sensing*, 51(4):2442–2453.

[Akbari et al., 2012a] Akbari, V., Doulgeris, A. P., and Eltoft, T. (submitted2012a). Monitoring glacier changes by multitemporal multipolarization SAR images. *submitted to IEEE Transactions on Geoscience and Remote Sensing*.

[Akbari et al., 2012b] Akbari, V., Larsen, Y., Doulgeris, A., and Eltoft, T. (2012b). The impact of terrain correction of polarimetric SAR data on glacier change detection. In *IEEE International Geoscience and Remote Sensing Symposium*.

[Anfinsen et al., 2007] Anfinsen, S., Jenssen, R., and Eltoft, T. (2007). Spectral clustering of polarimetric SAR data with Wishart-derived distance measures. In *International POLinSAR Workshop*, 8 pp., Frascati, Italy.

[Anfinsen, 2010] Anfinsen, S. N. (2010). *Statistical Analysis of Multilook Polarimetric Radar Images with the Mellin Transform*. University of Tromsø, Tromsø.

[Anfinsen and Akbari, 2013] Anfinsen, S. N. and Akbari, V. (2013). Statistical characterisation of the complex Hotelling-Lawley trace statistic. In preparation for submission to the IEEE Trans. on Signal Process.

[Anfinsen et al., 2009] Anfinsen, S. N., Doulgeris, A. P., and Eltoft, T. (2009). Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 47(11):3795–3809.

[Anfinsen et al., 2011] Anfinsen, S. N., Doulgeris, A. P., and Eltoft, T. (2011). Goodness-of-fit tests for multilook polarimetric radar data based on the Mellin transform. *IEEE Transactions on Geoscience and Remote Sensing*, 49(7):2764–2781.

[Anfinsen and Eltoft, 2011] Anfinsen, S. N. and Eltoft, T. (2011). Application of the matrix-variate Mellin transform to analysis of polarimetric radar images. *IEEE Transactions on Geoscience and Remote Sensing*, 49(6):2281–2295.

[Atwood et al., 2012] Atwood, D. K., Small, D., and Gens, R. (2012). Improving PolSAR land cover classification with radiometric correction of the coherency matrix. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(3):848–856.

[Beran, 1977] Beran, R. (1977). Minimum hellinger distance estimates for parametric models. *The Annals of Statistics*, 5(3):445–463.

[Besag, 1977] Besag, J. (1977). Efficiency of pseudo-likelihood estimation for simple gaussian fields. *Biometrika 64*, 64(3):616–618.

[Bhattacharyya, 1943] Bhattacharyya, A. (1943). On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc*, 35(99-109):4.

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

[Bombrun et al., 2011] Bombrun, L., Vasile, G., Gay, M., and F., T. (2011). Hierarichal segmentation of polarimetric SAR images using heterogeneous clutter models. *IEEE Transactions on Geoscience and Remote Sensing*, 49(2):726–737.

[Cantrell, 2000] Cantrell, C. D. (2000). *Modern mathematical methods for physicists and engineers*. Cambridge University Press.

[Celeux and Diebolt, 1985] Celeux, G. and Diebolt, J. (1985). The SEM algorithm: A probabilistic teacher algorithm derived from the EM algorithm for the mixture problem. *Comput. Stat. Q.*, 2(1):73–82.

[Cloude, 1986] Cloude, S. (1986). Group theory and polarization algebra. *OPTIK*, 75(1):26–36.

[Cloude and Pottier, 1996] Cloude, S. R. and Pottier, E. (1996). A Review of Target Decomposition Theorems in Radar Polarimetry. *IEEE Transactions on Geoscience and Remote Sensing*, 34(2):498–518.

[Conradsen et al., 2003] Conradsen, K., Nielsen, A. A., Schou, J., and Skriver, H. (2003). A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 41(1):4–19.

[Cumming and Wang, 2005] Cumming, G. I. and Wang, H. F. (2005). *Digital Processing Of Synthetic Aperture Radar Data: Algorithms And Implementation*. Artech House, first edition.

[Curlander and McDonough, 1991] Curlander, J. and McDonough, R. (1991). *Synthetic Aperture Radar- Systems and Signal Processing*. New York: John Wiley & Sons, Inc, 1991.

[Dabboor et al., 2013a] Dabboor, M., Collins, M. J., Karathanassi, V., and Braun, A. (2013a). An unsupervised classification approach for polarimetric SAR data based on the Chernoff distance for complex Wishart distribution. *IEEE Transactions on Geoscience and Remote Sensing*, Early Access.

[Dabboor et al., 2013b] Dabboor, M., Yackel, J., Hossain, M., and Braun, A. (2013b). Comparing matrix distance measures for unsupervised POLSAR data classification of sea ice based on agglomerative clustering. *International Journal of Remote Sensing*, 34(4):1492–1505.

[Dempster et al., 1977] Dempster, A. P., Laird, N. M., and Bubin, D. B. (1977). Maximum likelihood from incomplete data via EM algorithm. *Journal of the Royal Statistical Society. Series B*, 39(1):1–38.

[Doulgeris et al., 2012] Doulgeris, A. P., Akbari, V., and Eltoft, T. (23-26 April, 2012). Automatic PolSAR segmentation with the $\mathcal{U}$-distribution and Markov random fields. In *9th European Conference on Synthetic Aperture Radar (EUSAR)*, pages 183–186, Nuremberg, Germany.

[Doulgeris et al., 2008] Doulgeris, A. P., Anfinsen, S., and Eltoft, T. (2008). Classification with a non-Gaussian model for PolSAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 46(10):2999–3009.

[Elachi and Van Zyl, 2006] Elachi, C. and Van Zyl, J. (2006). *Introduction to the Physics and Techniques of Remote Sensing*. A John Wiley & Sons, Inc., Hoboken, New Jersey, 2nd edition.

[Eltoft et al., 2006] Eltoft, T., Kim, T., and Lee, T. (2006). A multivariate Laplace distribution. *IEEE Signal Processing Letters*, 13, 5:300–303.

[Ferretti et al., 2007] Ferretti, A., Monti-Guarnieri, A., Prati, C., Rocca, F., and Massonet, D. (2007). *InSAR Principles-Guidelines for SAR Interferometry Processing and Interpretation*, volume 19. European Space Agency.

[Freeman, 1992] Freeman, A. (1992). SAR calibration: an overview. 30(6):1107–1121.

[Freitas et al., 2005] Freitas, C. C., Frery, A. C., and Correia, A. H. (2005). The polarimetric $\mathcal{G}$ distribution for SAR data analysis. *Environmetrics*, 16:13–31.

[Frery et al., 2007] Frery, A. C., Correia, A. H., and Freitas, C. d. C. (2007). Classifying multifrequency fully polarimetric imagery with multiple sources of statistical evidence and contextual information. *IEEE Trans. Geoscience and Remote Sensing*, 45(10):3098–3109.

[Frery et al., 2010] Frery, A. C., Jacobo-Berlles, J., Gambini, J., and Mejail, M. E. (2010). Polarimetric SAR image segmentation with B-splines and a new statistical model. *Multidimensional Systems and Signal Processing*, 21(4):319–342.

[Frery et al., 2011] Frery, D. C., Nascimento, A., and Cintra, R. J. (2011). Information theory and image understanding : An application to polarimetric SAR imagery. *Chilean Journal of Statistics*, 2(2):81–100.

[Frey et al., 2013] Frey, O., Santoro, M., Werner, C. L., and Wegmuller, U. (2013). DEM-based SAR pixel-area estimation for enhanced geocoding refinement and radiometric normalization. *IEEE Geoscience and Remote Sensing Letters*, 10(1):48–52.

[Galland et al., 2009] Galland, F., Nicolas, J.-M., Sportouche, H., Roche, M., Tupin, F., and Refregier, P. (2009). Unsupervised synthetic aperture radar image segmentation using Fisher distributions. *IEEE Trans. Geoscience and Remote Sensing*, 47(8):2966–2972.

[Geman and Geman, 1984] Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the Bayesian restoration of images. PAMI-6(6):721–741.

[Goering et al., 1995] Goering, D., Hinzman, L., and Kane, D. (1995). Removal of terrain effects from SAR satellite imagery of Arctic tundra. *IEEE Transactions on Geoscience and Remote Sensing*, 33(1):185–194.

[Goodman, 1963] Goodman, N. (1963). Statistical analysis based on certain multivariate complex Gaussian distribution. In *Ann. Math. Statist.*, volume 34, pages 152–177.

[Hajnsek, 2001] Hajnsek, I. (2001). *Inversion of surface parameters using polarimetric SAR*. Dt. Zentrum für Luft-und Raumfahrt, Bibliotheks-und Informationswesen.

[Hammersley and Clifford, 1971] Hammersley, J. M. and Clifford, P. (1971). Markov field on finite graphs and lattices. *Technical Report, unpublished*.

[Hampel et al., 2011] Hampel, F., Ronchetti, E., Rousseeuw, P., and Stahel, W. (2011). *Robust statistics: the approach based on influence functions*, volume 114. Wiley Series in Probability and Statistics.

[Inglada and Mercier, 2007] Inglada, J. and Mercier, G. (2007). A New Statistical Similarity Measure for Change Detection in Multitemporal SAR Images and Its Extension to Multiscale Change Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 45(5):1432–1445.

[Jackson and Landgrebe, 2002] Jackson, Q. and Landgrebe, D. A. (2002). Adaptive Bayesian contextual classification based on Markov random fields. *IEEE Transactions on Geoscience and Remote Sensing*, 40(11):2454–2463.

[Jakeman, 1980] Jakeman, E. (1980). On the statistics of K-distributed noise. *J. Phys. C*, 13:31–48.

[Jakeman and Pusey, 1976] Jakeman, E. and Pusey, P. N. (1976). A model for non-Rayleigh sea echo. *IEEE Trans. Antennas Propagat.*, 24, 6:806–814.

[Jakeman and Tough, 1987] Jakeman, E. and Tough, R. J. A. (1987). Generalized K distribution: a statistical model for weak scattering. *J. Opt. Soc. Am. A*, 4, 9:1764–1772.

[Jeffreys and Jeffreys, 1946] Jeffreys, H. and Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 186(1007):453–461.

[Kampes, 1999] Kampes, B. (1999). Delft object-oriented radar interferometric software: Users manual and technical documentation. *Delft University of Technology, Delft*, 1.

[Kapur et al., 1985] Kapur, J. N., Sahoo, P. K., and C., W. A. K. (1985). A new method for gray-level picture thresholding using the entropy of the histogram. *Computer Vision, Graphics, and Image Processing*, 29:273–285.

[Kersten et al., 2005a] Kersten, P. R., Lee, J. S., and Ainsworth, T. L. (2005a). A comparison of change detection statistics in POLSAR images. In *Proc. IEEE Int. Geoscience and Remote Sensing Symp. IGARSS '05*, volume 7, pages 4836–4839.

[Kersten et al., 2005b] Kersten, P. R., Lee, J.-S., and Ainsworth, T. L. (2005b). Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and em clustering. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3):519–527.

[Kittler and Illingworth, 1986] Kittler, J. and Illingworth, J. (1986). Minimum error thresholding. *Pattern Recognition*, 19(1):41–47.

[Kong, 1990] Kong, J. A. (1990). *Electromagnetic Wave Theory,*. JohnWiley & Sons, Inc., 2nd edition.

[Lance and Williams, 1966] Lance, G. and Williams, W. (1966). Computer programs for hierarchical polythetic classification (similarity analyses). *The Computer Journal*, 9(1):60–64.

[Lance and Williams, 1967] Lance, G. and Williams, W. (1967). Mixed-data classificatory programs i - agglomerative systems. *Australian Computer Journal*, 1(1):15–20.

[Larsen, 2011] Larsen, Y. (2011). SAR processing geometry. *Unpublished Technical Note*.

[Lee et al., 1994a] Lee, J., Grunes, M., and Kwok, R. (1994a). Classification of multi-look polarimetric SAR imagery based on the complex Wishart distribution. In *Int. J. Remote Sensing*, volume 15.

[Lee and Ainsworth, 2011] Lee, J.-S. and Ainsworth, T. L. (2011). The effect of orientation angle compensation on coherency matrix and polarimetric target decompositions. 49(1):53–64.

[Lee et al., 1991] Lee, J.-S., Grunes, M. R., and Mango, S. A. (1991). Speckle reduction in multipolarization, multifrequency SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 29(4):535–544.

[Lee and Pottier, 2009] Lee, J.-S. and Pottier, E. (2009). *Polarimetric radar imaging: from basics to applications*. CRC Press, Taylor & Francis Group, Boca Raton, USA, 2nd edition.

[Lee et al., 2000] Lee, J.-S., Schuler, D., and Ainsworth, T. (2000). Polarimetric SAR data compensation for terrain azimuth slope variation. *IEEE Transactions on Geoscience and Remote Sensing*, 38(5):2153–2163.

[Lee et al., 2002] Lee, J.-S., Schuler, D. L., Ainsworth, T. L., Krogager, E., Kasilingam, D., and Boerner, W.-M. (2002). On the estimation of radar polarization orientation shifts induced by terrain slopes. *IEEE Transactions on Geoscience and Remote Sensing*, 40(1):30–41.

[Lee et al., 1994b] Lee, J. S., Schuler, D. L., Lang, R. H., and Ranson, K. J. (1994b). K-distribution for multi-look processed polarimetric SAR imagery. In *IEEE Int. Geosci. Remote Sensing Symp.*, pages 2179–2181.

[Levada et al., 2008] Levada, A., Mascarenhas, N., and Tannus, A. (2008). Pseudolikelihood equations for potts mrf model parameter estimation on higher order neighborhood systems. *Pattern Recognition Letters* 5(3):522–526.

[Li, 2006] Li, L. S. (2006). *Markov Random Field Modeling in Image Analysis*. Springer-Verlag, London, U.K., 3rd edition.

[Li et al., 2004] Li, Z., Zhu, Q., and Gold, C. (2004). *Digital terrain modeling: principles and methodology*. CRC.

[Liu and Feng, 2006] Liu, J. D. Jiang, Z. and Feng, H. (2006). A novel fuzzy classification entropy approach to image thresholding. *Pattern Recognition Letters*, 27:1968–1975.

[Loew and Mauser, 2007] Loew, A. and Mauser, W. (2007). Generation of geometrically and radiometrically terrain corrected SAR image products. *Remote Sensing of Environment*, 106(3):337–349.

[Lopes et al., 1993] Lopes, A., Nezry, E., Touzi, R., and Laur, H. (1993). Structure detection and statistical adaptive speckle filtering in sar images. *International Journal of Remote Sensing*, 14(9):1735–1758.

[Lopes and Sery, 1997] Lopes, A. and Sery, F. (1997). Optimal speckle reduction for the product model in multilook polarimetric SAR imagery and the Wishart distribution. *IEEE Transactions on Geoscience and Remote Sensing*, 35(3):632–647.

[Luckman, 1998] Luckman, A. J. (1998). Correction of SAR imagery for variation in pixel scattering area caused by topography. *IEEE Transactions on Geoscience and Remote Sensing*, 36(1):344–350.

[Massonnet and Souyris, 2008] Massonnet, D. and Souyris, J.-C. (2008). *Imaging with Synthetic Aperture Radar*. EPFL Press, Raleigh, USA, 1st edition.

[McLachlan and Krishnan, 1996] McLachlan, g. and Krishnan, T. (1996). *The EM Algorithm and Extensions*.

[Moser and Serpico, 2006] Moser, G. and Serpico, S. B. (2006). Generalized minimum-error thresholding for unsupervised change detection from SAR amplitude imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 44(10):2972–2982.

[Moser and Serpico, 2009] Moser, G. and Serpico, S. B. (2009). Unsupervised change detection from multichannel SAR data by Markovian data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 47(7):2114–2128.

[Muirhead, 2009] Muirhead, R. J. (2009). *Aspects of multivariate statistical theory*, volume 197. Wiley-Interscience.

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

[Otsu, 1979] Otsu, N. (1979). A threshold selection method from gray level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9:62–66.

[Parr and Schucany, 1982] Parr, W. C. and Schucany, W. R. (1982). Minimum distance estimation and components of goodness-of-fit statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44(2):178–189.

[Sabry et al., 2011] Sabry, R., Vachon, P. W., and Cole, M. (2011). Prediction of polarimetric-SAR field-orientation rotation due to topographical slope variation for squint operations. *IEEE on Geoscience and Remote Sensing Letters*, 8(3):570–574.

[Sarbandi, 1992] Sarbandi, K. (1992). Derivations of phase statistics from the müller matrix. *Radio Science*, 27(5):553–560.

[Schou and Skriver, 2001] Schou, J. and Skriver, H. (2001). Restoration of polarimetric sar images using simulated annealing. *IEEE Transactions on Geoscience and Remote Sensing*, 39(9):2005–2016.

[Schuler et al., 1999] Schuler, D., Lee, J., and Ainsworth, T. (1999). Compensation of terrain azimuthal slope effects in geophysical parameter studies using polarimetric sar data. *Remote sensing of environment*, 69(2):139–155.

[Serpico and Moser, 2006] Serpico, S. B. and Moser, G. (2006). Weight parameter optimization by the Ho–Kashyap algorithm in MRF models for supervised image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 44(12):3695–3705.

[Shimada, 2010] Shimada, M. (2010). Ortho-Rectification and Slope Correction of SAR Data Using DEM and Its Accuracy Evaluation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 3(4):657–671.

[Sinclair, 1950] Sinclair, G. (1950). Transmission and reception of elliptically polarized waves. In *Proceedings of the IRE*, pp. 148–151.

[Small, 2011] Small, D. (2011). Flattening Gamma: radiometric terrain correction for SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 49(8):3081–3093.

[Swart, 2000] Swart, L. (2000). Spectral filtering and oversampling for radar interferometry. *M.Sc. thesis Delft University of Technology*.

[Tison et al., 2004] Tison, C., Nicolas, J.-M., Tupin, F., and Maitre, H. (2004). A new statistical model for Markovian classification of urban areas in high-resolution SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 42(10):2046–2057.

[Touzi et al., 2004] Touzi, R., Boerner, W. M., Lee, J. S., and Lueneburg, E. (2004). A review of polarimetry in the context of synthetic aperture radar: concepts and information extraction. *Canadian Journal of Remote Sensing*, 30(3):380–407.

[Touzi and Lopes, 1994] Touzi, R. and Lopes, A. (1994). The principle of speckle filtering in polarimetric SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 32(5):1110–1114.

[Ulander, 1996] Ulander, L. (1996). Radiometric slope correction of synthetic aperture radar images. *IEEE Transactions on Geoscience and Remote Sensing*, 34(5):1115–1122.

[van den Bos, 1995] van den Bos, A. (1995). The multivariate complex normal distribution - a generalization. *IEEE Transactions on Information Theory*, 41(2):537–539. .

[van Zyl, 1993] van Zyl, J. (1993). Application of cloude's target decomposition theorem to polarimetric imaging radar data. In *San Diego'92*, pages 184–191. International Society for Optics and Photonics.

[van Zyl et al., 1993] van Zyl, J., Chapman, B., Dubois, P., and Shi, J. (1993). The effect of topography on SAR calibration. *IEEE Transactions on Geoscience and Remote Sensing*, 31(5):1036–1043.

[Wang et al., 2011] Wang, G., Zi, W., Xie, C., and Zhang, F. (2011). Dual-aspect geometric and radiometric terrain correction method for high-resolution SAR data. In *Proc. IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS)*, pages 1894–1897.

[Wang et al., 2000] Wang, L., Liu, J., and Li, S. Z. (2000). MRF parameter estimation by MCMC method. *Pattern Recognition*, 33(11):1919–1925.

[Wegmuller, 1999] Wegmuller, U. (1999). Automated terrain corrected SAR geocoding. *IEEE 1999 International Geoscience and Remote Sensing Symposium. IGARSS'99 (Cat. No.99CH36293)*, 3:1712–1714.

[Wiley, 1954] Wiley, C. (1954). Pulsed doppler radar methods and apparatus. US Patent 3,196,436.

[Wivell et al., 1992] Wivell, C., Steinwand, D., Kelly, G., and Meyer, D. (1992). Evaluation of terrain models for the geocoding and terrain correction, of synthetic aperture radar SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 30(6):1137–1144.

[Yamaguchi et al., 2011] Yamaguchi, Y., Sato, A., Boerner, W.-M., Sato, R., and Yamada, H. (2011). Four-component scattering power decomposition with rotation of coherency matrix. *IEEE Transactions on Geoscience and Remote Sensing*, 49(6):2251–2258.

[Yamazaki and Gingras, 1995] Yamazaki, T. and Gingras, D. (1995). Image classification using spectral and spatial information based on MRF models. *IEEE Transactions on Image Processing*, 4(9):1333–1339.

[Yu and Cheng, 2003] Yu, Y. and Cheng, Q. (2003). MRF parameter estimation by an accelerated method. *Pattern Recognition Letters*, 24(9-10):1251–1259.

[Yueh et al., 1989] Yueh, S. H., Kong, J. A., Jao, J. K., Shin, R. T., and Novak, L. M. (1989). K-Distribution and polarimetric terrain radar clutter. *J. Electro. Waves Applic.*, 3:747–768.

[Zhu, 1994] Zhu, J. (1994). Conversion of earth-centered earth-fixed coordinates to geodetic coordinates. *IEEE Transactions on Aerospace and Electronic Systems*, 30(3):957–961.

[Ziegler et al., 1992] Ziegler, V., Luneburg, E., and Schroth, A. (1992). Mean backscattering properties of random radar targets: A polarimetric covariance matrix concept. In *Proc. Int. Geoscience and Remote Sensing Symp. IGARSS '92*, pages 266–268.

About the author — *Vahid Akbari* was born in 1983 in Shabestar, Iran where he grew up. He received the M.Sc. degree in Geomatics Engineering in February 2009 from the University of Tehran, Tehran, Iran. He then took up the position as Research Fellow at the University of Tromsø, Department of Physics and Technology, working towards the Ph.D. degree. From March 2011 to July 2011, he was a visiting researcher with the Signal Processing and Telecommunications Laboratory of the Department of Electrical, Electronic, Telecommunications Engineering, and Naval Architecture, University of Genoa, Genoa, Italy. In June 2013, he defended the current Ph.D. thesis, entitled "Multitemporal Analysis of Multipolarization Synthetic Aperture Radar Images for Robust Surface Change Detection". His current research interests lies in the field of statistical analysis of polarimetric synthetic aperture radar images and radar interferometry. In particular, his work concerns statistical modeling for segmentation and multitemporal change detection of multipolarization radar imagery.

Trykk: Tromsprodukt: 40 00 72 00

ISBN 978-82-8236-092-0

ISBN 978-82-8236-092-0