



Universidade Federal de Pernambuco  
Centro de Ciências Exatas e da Natureza

Programa de  
Pós-graduação em Estatística

**Teoria Estatística da Informação para  
Dados de *Radar de Abertura Sintética*  
Univariados e Polarimétricos**

Abraão David Costa do Nascimento

Tese de Doutorado

Recife  
29 de fevereiro de 2012

Universidade Federal de Pernambuco  
Centro de Ciências Exatas e da Natureza

Abraão D. C. do Nascimento

**Teoria Estatística da Informação para Dados de *Radar de*  
*Abertura Sintética* Univariados e Polarimétricos**

*Trabalho apresentado ao Programa de Pós-graduação em  
Estatística do Centro de Ciências Exatas e da Natureza da  
Universidade Federal de Pernambuco como requisito para  
obtenção do grau de Doutor em Estatística.*

Orientador: *Dr. Alejandro C. Frery*  
Co-orientador: *Dr. Renato J. Cintra*

Recife  
29 de fevereiro de 2012



**Catálogo na fonte**  
**Bibliotecária Jane Souto Maior, CRB4-571**

**Nascimento, Abraão David Costa do**  
**Teoria estatística da informação para dados de radar**  
**de abertura sintética univariados e polarimétricos /**  
**Abraão David Costa do Nascimento - Recife: O Autor,**  
**2012.**

**xix, 263p.: il., fig., tab.**

**Orientador: Alejandro César Frery.**  
**Tese (doutorado) - Universidade Federal de Pernambuco.**  
**CCEN, Estatística, 2012.**

**Inclui bibliografia anexo e apêndice.**

**1. Estatística matemática. 2. Teoria assintótica. 3. Teoria da**  
**Informação. 4. Processamento de sinais. I. Frery, Alejandro**  
**César (orientador). II. Título.**

**519.9**

**CDD (23. ed.)**

**MEI2012 – 031**

Universidade Federal de Pernambuco  
Pós-Graduação em Estatística

**29 de fevereiro de 2012**

Nós recomendamos que a tese de doutorado de autoria de  
**Abraão David Costa do Nascimento**

intitulada

“Teoria Estatística da Informação para Dados de *Radar de Abertura Sintética* Univariados e Polarimétricos”

seja aceita como cumprimento parcial dos requerimentos para o grau de Doutor em Estatística.

---

Coordenador da Pós-Graduação em Estatística

Banca Examinadora:

---

**Renato José Sobral Cintra** co-orientador

---

**Klaus Leite Pinto Vasconcellos**

---

**Gauss Moutinho Cordeiro**

---

**Hélio Magalhães de Oliveira**

---

**Corina da Costa Freitas (INPE)**

Este documento será anexado à versão final da dissertação.

*Como adepto do criacionismo, dedico esta tese à Deus.  
Adicionalmente, ainda que não distintamente, dedico-a à  
minha mãe, Marluce Maria do Nascimento, à minha  
esposa, Rosângela Avelino do Nascimento e ao  
complemento de minha família, que represento nas figuras  
de minha irmã, Rebeca Abi do Nascimento, de minha  
sobrinha, Ludmila Abi do Nascimento, de meu irmão, João  
Francisco do Nascimento, e de minha mãe de criação,  
Maria Luiza Gonzaga de Oliveira.*

# Agradecimentos

Agradeço aos meus orientadores, Alejandro C. Frery e Renato J. Cintra, por me ensinarem lições que realmente são não triviais de se aprender sozinho, a saber: (i) valoração científica de minhas idéias, (ii) vivência no desenvolvimento de manuscritos científicos e (iii) uso de acessórios computacionais para elaboração de *revisões da literatura*.

Especialmente, lembro com gratidão dos simpósios organizados pelo Prof. Renato, em comemorações aos artigos aceitos e/ou para discussões de tópicos de pesquisas. A cada comemoração, o Prof. Renato realmente adicionava à nossa *maturidade acadêmica* o amor que ele dispensa à atividade científica. Adicionalmente, recordo da viagem a Hong Kong, China, com o Prof. Alejandro para participação na conferência internacional de processamento de imagens (*International Conference of Image Processing - ICIP*). Muito embora a viagem tenha sido cansativa, as experiências que tive vou levar como “*lições de bolso*” (aquelas que insistimos em lembrar para suavizar os desafios inerente à nossa profissão) por toda a minha vida.

À minha esposa, Rosângela Nascimento, pelo amor, companheirismo e palavras de encorajamento nesta fase final do doutorado. Seus cuidados forneceram, sem sombra de variação, um ambiente necessário para o desenvolvimento desta tese. À minha mãe, que por acreditar em mim, investiu seu “suor” no fornecimento da melhor educação possível e cujo referencial me ensinou que o que dignifica o homem é o carinho e dedicação que dispensa ao seu trabalho.

Aos amigos e companheiros de trabalho: Fábio Bayer, Hemílio Coelho, Luz Mileza, Manoel Santos, Rodrigo Bernardo, Tatiene, Tarciana, Marcelo Bourguignon e Laércio Dias (no doutorado); Raphael, Maria Lídia, Juliana, Fábio Verissimo, Marcelo Rodrigo, Larissa e Maria do Carmo (no mestrado); Adriano Victor, Juliana Kátia, Vanessa Kelly, Raul Elias, Lucas, Luiz Henrique e Valmir (na graduação).

Aos companheiros de conversa, Diego Hernane, Cleice Santos, Herbert Diógenes e Ramsés, pelos momentos de descontração.

À Valéria, secretária do programa, que sempre teve muito cuidado e carinho no momento de me orientar em burocracias acadêmicas e me lembrar da semana de matrícula. Além de suas palavras de incentivo, que certamente me nortearam a prosseguir consistentemente.

À todos os professores do Departamento de Estatística da UFPE. Em caráter especial, aos professores Manoel Raimundo, Cláudia Regina e Maria Cristina por partici-

parem de todo os meus dez anos de formação no DE-UFPE.

À CAPES, pelo apoio financeiro. E, finalmente, aos membros da banca examinadora pela abertura de visão a novos trabalhos a partir deste documento.

# Resumo

A tecnologia de radar de abertura sintética polarimétrico (*Polarimetric Synthetic Aperture Radar* – PolSAR) vem sendo apontada como um indispensável ferramental para capturar e analisar informações aeroespaciais, em particular, da superfície da terra [Freitas, C. C., Frery, A. C. & Correia, A. H. (2005), ‘The polarimetric G distribution for SAR data analysis’. *Environmetrics* 16(1), 13–31]. O objetivo desta tese é a proposta de métodos inferenciais baseados em Teoria da Informação para dados polarimétricos. Assim, o trabalho de investigação desta tese está dividido em duas áreas de estudo. Primeiramente, propor distâncias estocásticas bem como investigar seus comportamentos assintóticos para as diferentes distribuições voltadas a imagens PolSAR. Adicionalmente, visa propor outras medidas dentro da Teoria da Informação direcionadas aos contextos de imagens PolSAR, tal como entropia [Zografos, K. & Nadarajah, S. (2005), ‘Expressions for Rényi and Shannon entropies for multivariate distributions’, *Statistics & Probability Letters* 71(1), 71–84]. Em segundo lugar, realizar aplicações dos ferramentais derivados aos contextos de classificação, segmentação e detecção de bordas em tais imagens.

Inicialmente, na Seção “*Testes paramétricos e não paramétrico para dados speckle*”, uma imagem SAR é o resultado de observar uma variável aleatória com distribuição  $\mathcal{G}_1^0$ . Desta forma, recursos de Inferência Estatística (tais como estimação por máxima verossimilhança - EMV, teste de hipótese e medidas descritivas das estimativas) tentam mapear comportamentos e comparações entre as ferramentas desenvolvidas dentro de uma imagem SAR. Ainda, os resultados são submetidos a um modelo de contaminação paramétrico espelho do fenômeno de reflexão de canto.

Como um segundo estudo, a Seção “*Estimações melhoradas no modelo Wishart complexo escalonado*” apresenta dois métodos de EMV melhorados para o número de equivalente de looks. Tais metodologias são baseadas na expressão para o viés de segunda ordem em EMV e em uma fórmula para a verossimilhança perfilada modificada. Finalmente, seus desempenhos são comparados com outros métodos clássicos.

Um terceiro estudo é apresentado na Seção “*Expressões para distâncias em dados PolSAR*”. Neste caso, os pixels são matrizes aleatórias pertencentes a família de distribuições Wishart complexa escalonada. Nesta perspectiva, esta seção propõe ferramentas estocásticas baseadas na Teoria da Informação (com respeito ao funcional  $(h, \phi)$ ), proposto por Salicrú, Menéndez, Pardo & Morales [(1994), ‘On the applications of divergence type measures in testing statistical hypothesis’, *Journal of Multivariate Analysis*

51, 372–391]) para o contexto da distribuição Wishart complexa escalonada.

Posteriormente, na Seção “*Novas propostas de detecção de bordas em imagens PolSAR*”, nós aplicamos o ferramental derivado na Seção “*Expressões para distâncias em dados PolSAR*” ao contexto de detecção de borda em imagens PolSAR. Como resultado principal, as ferramentas derivadas são propostas como quatro novos detectores de bordas e comparadas a uma abordagem usual.

Na Seção “*Análise estatística para dados PolSAR baseada em medidas de entropia*”, um quinto estudo considera a proposta de testes de homogeneidade em imagens PolSAR por meio de entropias estocásticas. Esta metodologia é uma especificação da classe  $(h, \phi)$  de entropias proposta por Salicrú, Mendéndez & Pardo [(1993), ‘Asymptotic distribution of  $(h, \phi)$ -entropy’, *Communications in Statistics - Theory Methods* 22(7), 2015–2031] para a distribuição Wishart complexa escalonada e as entropias de Shannon e Rényi.

Finalmente, na Seção “*Comparação entre testes de homogeneidade em dados PolSAR*”, discute-se a comparação de três testes de homogeneidade envolvendo ferramentas da Teoria da Informação com a abordagem baseada na razão de verossimilhanças proposta por Conradsen, Nielsen, Schou & Skriver [(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]. O desempenho dos testes é avaliado tendo como critérios de comparação seus poderes e tamanhos estimados assim como outras medidas descritivas.

**Palavras-chave:** PolSAR; Teoria Estatística da Informação; Correção de Viés; Medidas de Entropias; Medidas de Divergências; Detecção de Borda.

# Abstract

The technology of Polarimetric Synthetic Aperture Radar (PolSAR) has been indicated as an essential tool for capturing and analyzing aerospace informations, in particular, obtained from the earth's surface [Freitas, C. C., Frery, A. C. & Correia, A. H. (2005), 'The polarimetric G distribution for SAR data analysis'. *Environmetrics* 16(1), 13–31]. The aim of this thesis is the proposal of statistical inference methods based on Information Theory for polarimetric data. Thus, its research work is organized into two segments. Firstly, propose stochastic distances and investigate its asymptotic behaviors for different distributions of PolSAR images. Additionally, others Information Theory measures were considered and applied on PolSAR scenarios, such as entropy [Zografos, K. & Nadarajah, S. (2005), 'Expressions for Rényi and Shannon entropies for multivariate distributions', *Statistics & Probability Letters* 71(1), 71–84]. On the other hand, propose methodologies of PolSAR image processing by means of the derived tools, such as classification, segmentation, and edge detection methods.

Initially, in Section *"Parametric and nonparametric tests for data speckle"*, a SAR image is a realization of a scalar random variable with  $\mathcal{G}_1^0$  distribution. In this context, statistical inference techniques (such as maximum likelihood estimation - MLE, hypothesis test, and performance descriptive measures) are performed in order to identify behavior as well as to compare different methods on SAR images. Moreover, the results are submitted to a parametric contamination model which represents the corner reflection phenomenon.

As a second study, two improved MLE methods for the equivalent number of looks are presented in the section *"Modified profile likelihood and bias correction on the scaled complex Wishart distribution"*. Such methodologies are based on the expression for second order bias correction in MLE and on a formula for modified profile likelihood. Finally, their performance are compared with others classical methods.

A third study is presented in Section *"Expressions for distances in PolSAR data"*. In this case, the pixels are random matrices which belong the family of scaled complex Wishart distributions. In this perspective, such section proposes stochastic tools based on Information Theory (with respect to the  $(h, \phi)$  functional, proposed by Salicrú, Menéndez, Pardo & Morales [(1994), 'On the applications of divergence type measures in testing statistical hypothesis', *Journal of Multivariate Analysis* 51, 372–391]) for the context of the scaled complex Wishart distribution.

Subsequently, in the section *"New proposals of edge detection in PolSAR images"*, we



apply the derived tools in Section “*Expressions for distances in PolSAR data*” to edge detection in PolSAR images. As main result, we propose four new edge detectors and compare with an usual approach.

In Section “*Statistical analysis for data PolSAR based on entropy measures*”, a fifth study considers the proposal of homogeneity tests in PolSAR images by means of stochastic entropies. This methodology is a specification of the  $(h, \phi)$  class of entropies proposed by Salicrú, Mendéndez & Pardo [(1993), ‘Asymptotic distribution of  $(h, \phi)$ -entropy’, *Communications in Statistics - Theory Methods* 22(7), 2015–2031] to the scaled complex Wishart distribution and Shannon and Rényi entropies.

Finally, in Section “*Homogeneity hypothesis tests on PolSAR images: Information Theory vs. Likelihood Ratio*”, it discusses a comparison of three homogeneity tests based on Theory Information measures with the approach based on likelihood ratio proposed by Conradsen, Nielsen, Schou & Skriver [(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]. The performance of tests is quantified and compared, considering as comparison criteria their estimated sizes and powers as well as others descriptive measures.

**Keywords:** PolSAR; Statistical Information Theory; Bias Correction; Entropy Measures; Divergence Measures; Boundary Detection.

# Sumário

|          |                                                           |           |
|----------|-----------------------------------------------------------|-----------|
| <b>1</b> | <b>Introdução</b>                                         | <b>1</b>  |
| 1.1      | Discussão geral                                           | 1         |
| 1.2      | Contribuições                                             | 4         |
| 1.3      | Plataforma computacional                                  | 8         |
| 1.4      | Objetivo e organização da tese                            | 9         |
| <b>2</b> | <b>Radar Polarimétrico</b>                                | <b>11</b> |
|          | O que é polarização?                                      | 11        |
| 2.1      | Sistemas de radar                                         | 14        |
| 2.1.1    | Sistemas de satélites e areotransportados                 | 14        |
| 2.1.2    | Geometria de imageamento                                  | 15        |
| 2.2      | Comportamento dos alvos                                   | 17        |
| 2.3      | Formação de imagens PolSAR                                | 20        |
| <b>3</b> | <b>Modelos para dados PolSAR</b>                          | <b>22</b> |
| 3.1      | Física PolSAR e modelos básicos                           | 24        |
| 3.1.1    | Distribuição gaussiana complexa                           | 26        |
| 3.1.2    | Distribuição Wishart complexa escalonada                  | 28        |
| 3.1.3    | Estimação dos parâmetros usuais em PolSAR                 | 29        |
| 3.2      | Modelos multilook SAR bivariados                          | 35        |
| 3.2.1    | Densidade para a diferença de fase                        | 37        |
| 3.2.2    | Densidade para a magnitude do produto                     | 38        |
| 3.2.3    | Densidade conjunta para o par de intensidades             | 38        |
| 3.2.4    | Densidades para as razões entre intensidades e amplitudes | 39        |

|          |                                                             |           |
|----------|-------------------------------------------------------------|-----------|
| 3.3      | Modelos multiplicativos multilook PolSAR                    | 40        |
| 3.3.1    | Ruído PolSAR                                                | 41        |
| 3.3.2    | Retroespalhamento (backscatter) PolSAR                      | 41        |
| 3.3.3    | Distribuições dos Retornos PolSAR                           | 44        |
| 3.3.3.1  | Classes simples:                                            | 46        |
| 3.3.3.2  | Classes moderadas:                                          | 48        |
| 3.3.3.3  | Classes amplas:                                             | 50        |
| 3.3.4    | Retornos SAR                                                | 52        |
| <b>4</b> | <b>Teoria Estatística da Informação</b>                     | <b>55</b> |
| 4.1      | A classe de entropias $(h, \phi)$                           | 61        |
| 4.1.1    | Teste de hipóteses                                          | 65        |
| 4.1.2    | Intervalo de confiança                                      | 66        |
| 4.2      | A classe de divergências $(h, \phi)$                        | 68        |
| <b>5</b> | <b>Estudos Desenvolvidos</b>                                | <b>75</b> |
| 5.1      | Testes paramétricos e não paramétrico para dados speckle    | 77        |
| 5.1.1    | Contribuições                                               | 79        |
| 5.1.2    | Considerações da Seção 5.1                                  | 92        |
| 5.1.3    | Trabalhos futuros                                           | 94        |
| 5.2      | Estimações melhoradas no modelo Wishart complexo escalonado | 117       |
| 5.2.1    | Contribuições                                               | 118       |
| 5.2.2    | Considerações da Seção 5.2                                  | 121       |
| 5.2.3    | Trabalhos futuros                                           | 123       |
| 5.3      | Expressões para distâncias em dados PolSAR                  | 134       |
| 5.3.1    | Contribuições                                               | 136       |
| 5.3.2    | Considerações da Seção 5.3                                  | 143       |
| 5.3.3    | Trabalhos futuros                                           | 143       |
| 5.4      | Novas propostas de detecção de bordas em imagens PolSAR     | 165       |
| 5.4.1    | Detectores de bordas                                        | 165       |

|          |                                                                   |            |
|----------|-------------------------------------------------------------------|------------|
| 5.4.2    | Performance dos detectores                                        | 168        |
| 5.4.3    | Método B-spline                                                   | 168        |
| 5.4.4    | Aplicação a dados reais                                           | 171        |
| 5.4.5    | Considerações da Seção 5.4                                        | 174        |
| 5.4.6    | Trabalhos futuros                                                 | 176        |
| 5.5      | Análise estatística via entropias                                 | 177        |
| 5.5.1    | Contribuições                                                     | 178        |
| 5.5.2    | Considerações da Seção 5.5                                        | 188        |
| 5.5.3    | Trabalhos futuros                                                 | 191        |
| 5.6      | Comparação entre testes de homogeneidade em dados PolSAR          | 207        |
| 5.6.1    | Razão de verossimilhanças para Wishart complexa                   | 208        |
| 5.6.1.1  | Discussão para Wishart escalonada                                 | 208        |
| 5.6.1.2  | Comparação com a estatística proposta por Conradsen               | 210        |
| 5.6.2    | Estudo de simulação                                               | 210        |
| 5.6.2.1  | Estudo com dados sintéticos                                       | 211        |
| 5.6.2.2  | Estudo com dados reais                                            | 215        |
| 5.6.3    | Considerações da Seção 5.6                                        | 215        |
| 5.6.4    | Trabalhos futuros                                                 | 219        |
| <b>6</b> | <b>Conclusões</b>                                                 | <b>220</b> |
| <b>A</b> | <b>Software PolSARpro</b>                                         | <b>223</b> |
| A.1      | Sistemas SAR aerotransportados                                    | 226        |
| A.1.1    | Sensor AIRSAR                                                     | 226        |
| A.1.2    | Sensor E-SAR                                                      | 227        |
| A.1.3    | Sensor EMISAR                                                     | 228        |
| <b>B</b> | <b>Expressões para as integrais da Seção 5.1</b>                  | <b>231</b> |
| <b>C</b> | <b>Propriedades estatísticas da distribuição Wishart complexa</b> | <b>232</b> |
| C.1      | Correção de Viés                                                  | 232        |

|          |                                                                       |            |
|----------|-----------------------------------------------------------------------|------------|
| C.2      | Derivadas de primeira e segunda ordem                                 | 234        |
| C.3      | Derivadas de terceira ordem                                           | 237        |
| C.4      | Expressão para o viés de segunda ordem                                | 239        |
| <b>D</b> | <b>Resultados para distribuição <math>\mathcal{G}_m^{0(L)}</math></b> | <b>241</b> |
| <b>E</b> | <b>Derivação das distâncias</b>                                       | <b>244</b> |

# Lista de Figuras

|      |                                                                        |    |
|------|------------------------------------------------------------------------|----|
| 2.1  | Comprimentos de onda e frequências para um espectro.                   | 12 |
| 2.2  | Diagrama de sensores.                                                  | 13 |
| 2.3  | Satélite SEASAT.                                                       | 15 |
| 2.4  | Sistemas aerotransportados.                                            | 16 |
| 2.5  | Imageamento via SAR.                                                   | 18 |
| 2.6  | Reflexividade em diferentes terrenos.                                  | 20 |
| 2.7  | Formatos em uma imagem SAR.                                            | 21 |
| 3.1  | Natureza do speckle.                                                   | 23 |
| 3.2  | Padrão do ruído speckle em uma imagem SAR.                             | 23 |
| 3.3  | Efeito multilook.                                                      | 25 |
| 3.4  | Imagem sintética Wishart distribuída.                                  | 30 |
| 3.5  | Imagem E-SAR (canal HH) com regiões selecionadas.                      | 32 |
| 3.6  | Densidades empíricas e ajustadas para o canal de polarização HH.       | 34 |
| 3.7  | Imagem AIRSAR de São Francisco (canal HH) com três áreas selecionadas. | 36 |
| 3.8  | Diagrama de modelos para o backscatter.                                | 43 |
| 3.9  | Curvas para os modelos backscatter mais gerais.                        | 44 |
| 3.10 | Ajustes para as distribuições das classes simples.                     | 49 |
| 3.11 | Diagrama de modelos para o retorno.                                    | 52 |
| 3.12 | Diagrama de modelos para o ruído speckle.                              | 53 |
| 4.1  | Utilização de distâncias estocástica em processamento de imagem.       | 56 |
| 5.1  | Efeito de uma reflexão de canto sobre um cenário homogêneo.            | 78 |
| 5.2  | Imagem sintética $\mathcal{G}_I^0$ distribuída.                        | 81 |

|      |                                                                               |     |
|------|-------------------------------------------------------------------------------|-----|
| 5.3  | Valores para o tamanho empírico médio.                                        | 85  |
| 5.4  | Poderes empíricos dos testes para apenas brilhos diferentes.                  | 88  |
| 5.5  | Poderes empíricos dos testes para brilhos e/ou rugosidades diferentes.        | 89  |
| 5.6  | Estimativas para o tamanho do teste sob uma contaminação de grau $\epsilon$ . | 90  |
| 5.7  | Imagem E-SAR e regiões selecionadas.                                          | 91  |
| 5.8  | Funções de distribuição acumulada empírica para diferentes regiões.           | 93  |
| 5.9  | Imagem AIRSAR de São Francisco com regiões selecionadas.                      | 121 |
| 5.10 | Imagens PolSAR real e sintética.                                              | 136 |
| 5.11 | Imagem PolSAR como quatro regiões selecionadas.                               | 141 |
| 5.12 | Análise de agrupamento envolvendo imagem PolSAR real.                         | 145 |
| 5.13 | Imagem PolSAR sintética com duas sub-regiões.                                 | 166 |
| 5.14 | Matrizes utilizadas em experimento.                                           | 168 |
| 5.15 | Avaliação da capacidade dos detectores.                                       | 169 |
| 5.16 | Passo a passo da obtenção dos pontos de controle.                             | 172 |
| 5.17 | Passos do método de segmentação para a região de florestas.                   | 173 |
| 5.18 | Definição dos segmentos de controle.                                          | 173 |
| 5.19 | Região $150 \times 150$ de São Francisco.                                     | 174 |
| 5.20 | Resultados com as regiões da imagem São Francisco.                            | 175 |
| 5.21 | Curvas para as entropias de Shannon, Tsallis e Rényi.                         | 179 |
| 5.22 | Ajustes para dados reais.                                                     | 182 |
| 5.23 | Tamanho empírico do teste sob testes de entropias                             | 187 |
| 5.24 | Imagem PolSAR com regiões selecionadas e densidades empíricas.                | 189 |
| 5.25 | Poder empírico do teste para diferentes regiões.                              | 190 |
| 5.26 | Imagem AIRSAR de Flevoland (canal HH).                                        | 212 |
| 5.27 | Tamanho empírico dos testes                                                   | 214 |
| 5.28 | Poder empírico dos testes                                                     | 216 |
| 5.29 | Ajustes sob dados reais.                                                      | 217 |
| 5.30 | Tamanhos empíricos do testes para dados reais.                                | 218 |
| A.1  | Software PolSARpro.                                                           | 224 |

|     |                                    |     |
|-----|------------------------------------|-----|
| A.2 | Imagem do sistema AIRSAR.          | 226 |
| A.3 | Imagem AIRSAR de São Francisco.    | 227 |
| A.4 | Imagem AIRSAR de Flevoland.        | 227 |
| A.5 | Imagem do sistema E-SAR.           | 228 |
| A.6 | Imagem E-SAR de Munique.           | 228 |
| A.7 | Imagem do sistema EMISAR.          | 229 |
| A.8 | Imagem EMISAR de Follum.           | 230 |
| B.1 | Identities para algumas integrais. | 231 |



# Lista de Tabelas

|      |                                                                             |     |
|------|-----------------------------------------------------------------------------|-----|
| 2.1  | Evolução de radares                                                         | 17  |
| 3.1  | Estimativas dos parâmetros                                                  | 33  |
| 3.2  | AICs para distribuições Wishart                                             | 35  |
| 3.3  | Módulos das correlações médias entre os canais de polarizações              | 36  |
| 3.4  | Módulos das correlações médias entre os canais de polarizações              | 48  |
| 3.5  | Fatores para densidades PolSAR                                              | 51  |
| 3.6  | Fatores para densidades SAR                                                 | 54  |
| 4.1  | $(h, \phi)$ -entropia e funções relacionadas                                | 63  |
| 4.2  | Distâncias tipo $(h, \phi)$ e funções $h$ e $\phi$ relacionadas             | 72  |
| 5.1  | Estimativas e medidas descritivas associadas                                | 84  |
| 5.2  | Taxas de rejeição dos testes sob brilhos diferentes e rugosidades similares | 86  |
| 5.3  | Taxas de rejeição dos testes sob brilhos e/ou rugosidades diferentes        | 87  |
| 5.4  | Estimativas dos parâmetros                                                  | 91  |
| 5.5  | Estimativas para o tamanho do teste em dados reais                          | 92  |
| 5.6  | Estimativas para o poder do teste em dados reais                            | 93  |
| 5.7  | Estimativas MV melhoradas dados sintéticos                                  | 122 |
| 5.8  | Estimativas MV melhoradas dados reais                                       | 122 |
| 5.9  | Tamanho empírico do teste para a região “forest-1”                          | 140 |
| 5.10 | Resultados baseados sobre coeficiente de correlação                         | 142 |
| 5.11 | Estimativas MV                                                              | 181 |
| 5.12 | Intervalos assintóticos para entropias                                      | 181 |
| 5.13 | Tamanhos empíricos do teste                                                 | 184 |

|                                                                 |     |
|-----------------------------------------------------------------|-----|
| 5.14 Média e coeficiente de variação sob a hipótese alternativa | 185 |
| 5.15 estimativas MV                                             | 188 |
| 5.16 Tamanhos empíricos dos testes                              | 212 |
| 5.17 Estimativas dos parâmetros                                 | 213 |

## Capítulo 1

# Introdução

### 1.1 Discussão geral

A tecnologia de radar de abertura sintética polarimétrico (*Polarimetric Synthetic Aperture Radar* – PolSAR) vem sendo apontada como um indispensável ferramental para capturar e analisar informações aeroespaciais, em particular, da superfície da terra (Freitas, Frery & Correia 2005). Dispositivos PolSAR trabalham com a emissão de pulsos polarizados ortogonalmente em direção a um alvo, cenário real sob estudo. Neste processo, as informações fornecidas pelo sinal retornado são registradas para cada polarização. Aparelhos PolSAR fornecem meios para uma melhor captação de atributos relativos a uma cena, quando comparados com sua versão univariada bem como com outras modalidades de sensoriamento remoto (Lee & Pottier 2009, López-Martínez, Fábregas & Pottier 2005).

Devido à utilização de um processo de iluminação coerente<sup>1</sup> em sua configuração e à grande abertura sintética de sua antena, PolSAR pode alcançar alta resolução espacial em suas imagens resultantes (Oliver & Quegan 1998). Um benefício no uso de tal sistema se deve à sua capacidade de operar independentemente das condições climáticas bem como de ultrapassar coberturas vegetais (dependendo de diversos fatores, tais como das bandas de operação dos sensores), com vistas à compreensão da natureza do relevo. A caracterização de uma cena pelo PolSAR é realizada sob a perspectiva de múltiplos canais de polarização, isto é, uma mesma informação é analisada segundo várias dimensões. Além disso, esta metodologia considera a modelagem das estruturas de covariâncias entre tais canais. Pela significância do uso de imagens PolSAR, Boerner & Lee (2007) publicaram um artigo elencando os principais trabalhos (artigos, livros e teses) realizados nesta área e incentivando novos desenvolvimentos. Nessa linha, além de contribuições teórico computacionais, esta tese

---

<sup>1</sup>Em função da característica coerente do sensor, a amplitude e a fase do sinal retornado (de natureza complexa) são resultantes de uma soma vetorial (*coerente*) dos diversos elementos de um alvo natural (Freitas, Sant’Anna, Rennó & Correia 2003).

- ✎ avança neste sentido uma vez que elenca um grande número de livros e artigos voltados à modelagem e ao processamento de dados PolSAR em suas referências bibliográficas.

Sistemas de imageamento que empregam iluminação coerente estão sujeitos a um padrão de interferência nas imagens resultantes, o “ruído speckle” (Oliver & Quegan 1998). Este ruído tem uma natureza não aditiva e não gaussiana, o que impossibilita o uso de um grande número de técnicas de processamento que supõem comportamento aditivos gaussianos nos dados. Além disso, este fenômeno dificulta significativamente a modelagem, a interpretação e o processamento de imagens PolSAR (López-Martínez et al. 2005). Assim, métodos especializados de análise do sinal retornado são normalmente requeridos no tratamento de imagens PolSAR.

Técnicas de processamento, tais como segmentação (Beaulieu & Touzi 2004), classificação (Kersten, Lee & Ainsworth 2005) e detecção de bordas (Schou, Skriver, Nielsen & Conradsen 2003) necessitam frequentemente de medidas confiáveis de discriminação de dados como uma etapa de pré-processamento. Mais uma aplicação da discriminação de dados é a detecção de mudanças, tal como apresentada por Inglada & Mercier (2007). Medidas de discriminação também são fundamentais em interferometria (*Interferometric Synthetic Aperture Radar* – InSAR), conforme Gierull & Sikaneta (2002), para a identificação de regiões em que não houve alterações, pois elas servem de referência para a obtenção de diversos produtos. Tais medidas têm sido usadas para quantificar diferença entre as regiões de uma imagem PolSAR (Goudail, Réfrégier & Delyon 2004, Frery, Cintra & Nascimento 2011), sendo chamadas de “contraste” (Nascimento, Cintra & Frery 2010). A natureza de tais medidas tem um importante papel na proposta de algoritmos eficientes nos diversos ramos de processamento em tais imagens (Goudail & Réfrégier 2004). Assim, a busca por ferramentas analíticas e robustas de contraste é uma linha de pesquisa por si só.

Em anos recentes, o interesse em adaptar técnicas da Teoria Estatística da Informação aos contextos de processamento de imagens (Inglada & Mercier 2007, Nascimento, Cintra & Frery 2009b, Nascimento, Frery & Cintra 2009, Nascimento, Cintra & Frery 2009a) tem crescido bastante. Paralelamente ao tratamento de imagens, alguns trabalhos no campo da Estatística Matemática têm endereçado a proposta de testes paramétricos baseados em classes de divergências (Ali & Silvey 1996, Menéndez, Pardo & Pardo 2007, Broniatowski & Keziou 2009). Adicionalmente, outros trabalhos têm utilizado medidas de entropias como uma eficiente alternativa de quantificar variação (ou a quantidade de informação média em um evento), sendo a comparação destas medidas em eventos diferentes uma proposta de contraste. Neste contexto, Zografos e Nadarajah propuseram expressões fechadas para as entropias de Shannon e Rényi para várias distribuições univariadas (Nadarajah & Zografos 2003), bivariadas (Nadarajah & Zografos 2005) e multivariadas (Zografos & Nadarajah 2005). Além disso, medidas de entropia têm sido aplicadas a problemas de processamento de imagens, tais como em mineração de dados, detecção, segmentação e classificação (Basseville 1989, Maybank 2007, Liuni, Robel, Romito & Rodet 2011).

Esta tese tem os seguintes objetivos:

- ✎ apresentar um resumo sobre modelos para imagens polarimétricas (incluindo também detalhes sobre situações uni e bivariadas),
- ✎ estudar a estimação por máxima verossimilhança (MV) dos parâmetros envolvidos nas metodologias de testes de hipóteses consideradas e interpretar suas estimativas quanto à física dos sistemas PolSAR,
- ✎ aplicar os conceitos de testes de hipóteses baseados na classe de divergências  $(h, \phi)$  proposta por Salicrú, Menéndez, Pardo & Morales (1994) e outros métodos baseados em medidas de entropias (Salicrú, Mendéndez & Pardo 1993, Pardo, Morales, Salicrú & Menéndez 1997) às famílias de distribuições cujos parâmetros estão relacionados com a natureza de uma imagem PolSAR e
- ✎ aplicar as ferramentas derivadas a métodos de processamento de imagens: classificação e detecção de bordas.

Goudail & Réfrégier (2004) utilizaram essas medidas estocásticas para caracterizar o desempenho de algoritmos de detecção de bordas bem como de segmentação em processamento de imagens PolSAR. Neste estudo, tanto a distância de Kullback-Leibler quanto a de Bhattacharyya foram consideradas como instrumentos para quantificar as diferenças entre distribuições gaussianas complexas. Como resultado, a medida de Bhattacharyya revelou as melhores capacidades de contraste. Entretanto, as propriedades estatísticas de tais medidas não foram consideradas explicitamente.

Teorias em processamento de imagens PolSAR prescrevem que o sinal retornado é representado por uma matriz de covariância hermitiana positiva definida. Goodman (1963b) apresentou uma análise abrangente sobre o modelo gaussiano complexo, juntamente com a relação entre o estimador de MV para sua matriz de covariância complexa e a distribuição Wishart complexa. Esta distribuição tem alcançado bastante êxito como uma forma de modelagem para dados polarimétricos (Conradsen, Nielsen, Schou & Skriver 2003).

Conradsen et al. (2003) propuseram uma metodologia baseada no teste da razão de verossimilhança, objetivando a discriminação de duas imagens PolSAR distribuídas segundo um modelo Wishart complexo. Testes de hipóteses para dados speckle *monopolarizados* (informações relacionadas a apenas um canal de polarização em dados PolSAR) foram propostos por Nascimento et al. (2010). Como complemento aos objetivos anteriormente mencionados, considera-se

- ✎ propor testes estatísticos via Teoria da Informação para avaliar homogeneidade em imagens PolSAR distribuídas segundo uma lei Wishart complexa escalonada e
- ✎ comparar estes testes com a metodologia proposta por Conradsen et al. (2003).

A seguir, são apresentadas as principais contribuições desta tese.

## 1.2 Contribuições

Esta tese avança no tratamento de dados SAR polarimétricos, bivariados e univariados através dos seguintes resultados:

1. Comparação da capacidade de discriminação entre testes paramétricos baseados sobre a classe de divergências  $(h, \phi)$  e sobre o teste não-paramétrico de Kolmogorov-Smirnov (Cintra, Frery & Nascimento 2011). Neste caso, considera-se dados tipo speckle modelados conforme a família de distribuições  $\mathcal{G}_I^0$ . Os desempenhos dos testes são avaliados tendo como critério as estimativas de seus tamanhos e poderes. Além disso, um breve estudo de robustez é realizado para avaliar tanto os tamanhos dos testes quanto os estimadores de MV para os parâmetros da distribuição  $\mathcal{G}_I^0$ , considerando diferentes graus de contaminação. Uma expressão analítica para a distância de Kolmogorov-Smirnov entre funções de distribuição acumulada do tipo  $\mathcal{G}_I^0$  também é apresentada. No tocante à análise monopolarizada (univariada) dados PolSAR, esta tese avançou na forma dos seguintes trabalhos:

.....  
Artigos para periódicos [testes  $(h, \phi)$  e estudos comparativos]

- 2011 Cintra, R. J., Frery, A. C. & Nascimento, A. D. C. (2011), 'Parametric and nonparametric tests for speckled imagery', *Pattern Analysis and Applications*, 1-21.
- 2010 Nascimento, A. D. C.; Cintra, R. J. & Frery, A. C. (2010), 'Hypothesis testing in speckled data with stochastic distances', *IEEE Transactions on Geoscience and Remote Sensing*, **48**(1), 373–385.

Artigos para congressos

- 2010 Frery, A. C., Nascimento, A. D. C. & Cintra, R. J. (2010), Contrast in speckled imagery with stochastic distances. in 'International Conference on Image Processing (ICIP)', Hong Kong, China, September, pp. 26-29.
- 2009 Nascimento, A. D. C., Cintra, R. J. & Frery, A. C. (2009), Stochastic distances and hypothesis testing in speckled data, in 'XIV Brazilian Remote Sensing Symposium (SBSR)', **14**, RN, Natal, pp. 7353-7360.
- 2009 Nascimento, A. D. C., Cintra, R. J. & Frery, A. C. (2009), Measuring synthetic aperture radar target differences with stochastic distances, in 'TIC-STH Symposium on Advances in Systems and Sensors', Toronto, Canada, September, pp. 587-592.
- 2009 Nascimento, A. D. C.; Frery, A. C. & Cintra, R. J. (2009), Analysis of speckled imagery with parametric and nonparametric tests, in 'Sibgrapi 2009 - Theses and Dissertations', RJ, Brazil, October.

- .....
2. Correção de viés e o uso de verossimilhança perfilada modificada em distribuições Wishart complexa escalonada (Nascimento, Cintra & Frery 2012). Neste trabalho generalizamos a abordagem de estimação MV para o número de looks proposta por Anfinson, Doulgeris & Eltoft (2009). Apresentamos a expressão para o viés de segunda ordem conforme a proposta de Cox & Snell (1968) e, consequentemente, propomos um novo estimador com viés da ordem  $O(N^{-2})$  em que  $N$  representa o tamanho de amostra. Além disso, a expressão para log-verossimilhança perfilada modificada segundo Barndorff-Nielsen foi derivada, implicando em um outro novo estimador com viés de ordem  $O(N^{-3/2})$ . Os estimadores obtidos são comparados com a abordagem de Anfinson, Doulgeris & Eltoft (2009) por meio de simulação de Monte Carlo, considerando como critério de comparação três medidas decritivas: erro quadrático médio, viés e coeficiente de variação. Adicionalmente, os métodos são aplicados a dados reais.
- .....

Sugestão de artigo em produção [Inferência melhorada em dados PolSAR]

- 2012 Nascimento, A. D. C., Cintra, R. J. & Frery, A. C. (2012), 'Modified profile likelihood and bias correction in Wishart complex distribution'.
- .....

3. Apresentação de expressões analíticas para as distâncias Kullback-Leibler, Rényi (de ordem  $\beta$ ), Bhattacharyya e Hellinger entre variáveis modeladas segundo a lei Wishart complexa escalonada relaxada e em importantes casos particulares (Frery, Nascimento & Cintra 2010a). Verificamos também que estas distâncias apresentam invariância de escala em relação às suas matrizes de covariância. Usando estas medidas, obtivemos desigualdades que dependem de matrizes de covariância; duas dentre elas (as derivadas das distâncias de Kullback-Leibler e Hellinger) proporcionam formas alternativas para a obtenção das distâncias clássicas conhecidas como "revised Wishart" (Anfinson, Jenssen & Eltoft 2007) e "Bartlett" (Kersten et al. 2005), respectivamente. Baseando-se em uma reparametrização da distribuição Wishart proposta por Lee, Hoppel, Mango & Miller (1994), propomos também medidas de contraste dependentes do coeficiente de correlação complexo. Através de duas destas medidas (baseadas nas distâncias de Kullback-Leibler e Hellinger), propuseram-se regiões de confiança para o coeficiente de correlação. Finalmente, as medidas de dissimilaridade propostas são incorporadas ao método  $k$ -médias, redundando em quatro novos métodos de agrupamento em dados PolSAR.
- .....

Artigo submetido a IEEE TGRS [Medidas de similaridades em dados PolSAR]

- 2012 A. C. Frery, A. D. C. Nascimento, and R. J. Cintra (2012), 'Analytic Expressions for Stochastic Distances Between Relaxed Complex Wishart Distributions'. Under review on *IEEE Transactions on Geoscience and Remote Sensing*, 2012.

Artigo para congresso

- 2011 Nascimento, A. D. C., Cintra, R. J. & Frery, A. C. (2011), Polarimetric image understanding by statistical information theory, in 'Sibgrapi 2011 - Theses and Dissertations', Maceió, Brazil, October.

4. Propostas alternativas para detecção de bordas em imagens PolSAR baseadas em distâncias estocásticas (Nascimento, Horta & Frery 2011). Os desempenhos de cinco métodos de detecção de borda em uma imagem PolSAR, modelada pela distribuição Wishart complexa escalonada, foram comparados: em que um é o método clássico de detecção proposto por Gambini, Mejail, Jacobo-Berlles & Frery (2008) baseado na comparação de log-verossimilhanças; enquanto, os outros quatro são novas metodologias baseadas em distâncias estocásticas. Um estudo das performances dos detectores é realizado com dados PolSAR reais, utilizando o algoritmo de  $k$ -médias modificado por Cao, Hong, Wu & Pottier (2007) e a abordagem de detecção proposta por Gambini et al. (2008).

Artigos para congressos [Detecção de borda em dados PolSAR]

- 2011 Nascimento, A. D. C., Horta, M. M. & Frery, A. C. (2011), Boundary detection with stochastic distance in polarimetric imagery, in 'the 5th International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLinSAR'2011)', Frascati, Italy.

5. Proposta de uma estatística baseada na distância de Kullback-Leibler para avaliação de homogeneidade em imagem PolSAR multilook (Frery, Cintra & Nascimento 2011). As influências de fatores importantes em imagem PolSAR sobre os estimadores para os tamanhos de cinco testes de hipóteses são avaliadas. Esta simulação foi baseada em replicações de Monte Carlo. Um breve estudo de robustez é ainda realizado para mensurar o efeito de um modelo de contaminação sobre as estimativas para os tamanhos dos testes e sobre os estimadores de MV para os parâmetros da distribuição Wishart complexa escalonada. Em caráter complementar ao estudo de simulação, uma aplicação com dados PolSAR reais é fornecida.

Artigo para periódico [Teste de hipótese e robustez em dados PolSAR]

- 2011 Frery, A. C., Nascimento, A. D. C. & Cintra, R. J. (2011), Information Theory and Image Understanding: An Application to Polarimetric SAR Imagery, *Chilean Journal of Statistics*, 2, pp. 81-99.



Artigo para congresso

- 2011 Frery, A. C., Cintra, R. J. & Nascimento, A. D. C. (2011). Hypothesis test in complex Wishart distributions, *in* 'the 5th International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLinSAR'2011)', Frascati, Italy, January.
- .....

6. Análise Estatística baseada em entropia (Frery, Cintra & Nascimento 2012). Nós derivamos expressões analíticas para as entropias de Shannon, Rényi e Tsallis para um escalonamento da distribuição Wishart complexa. Adicionalmente, as variâncias assintóticas das entropias de Shannon e Rényi são obtidas, quando substituímos seus parâmetros pelos estimadores de MV. As relações das medidas derivadas com os parâmetros da distribuição Wishart são quantificadas. Subsequentemente, intervalos de confiança baseados nas medidas de entropias também são propostos e indicados como uma nova metodologia para quantificar contraste. Adicionalmente, novas metodologias de teste de hipóteses em homogeneidade são proposta e avaliadas via simulação de Monte Carlo.
- .....

Artigo submetido a IEEE TGRS

- 2012 A. C. Frery, A. D. C. Nascimento, and R. J. Cintra (2012), 'Entropy-based statistical analysis of PolSAR data'. Under review on *IEEE Transactions on Geoscience and Remote Sensing*, 2012.
- .....

7. Comparação entre testes sobre teoria da informação e testes da razão de verossimilhança (Nascimento, Frery & Cintra 2012b). Considerou-se também um estudo da performance de testes baseados na razão de verossimilhança e de testes baseados em medidas de distância e entropias para a distribuição Wishart complexa escalonada. Utilizado experimentos de Monte Carlo, os testes são comparado adotando como critério de comparação seus tamanhos e poderes estimados.
- .....

Sugestão de artigo em produção [Comparação entre testes em dados PolSAR]

- 2012 Nascimento, A. D. C., Frery, A. C. & Cintra, R. J. (2012), 'Parametric hypothesis tests in PolSAR imagery: Information theory vs. likelihood ratio statistics'.
- .....

8. Proposta de novas distribuições para imagens SAR monopolarizadas (Nascimento, Frery & Cintra 2012a). Neste artigo, nós estendemos a classe de distribuições  $\mathcal{G}_I$  que foi proposta por Frery, Muller, Yanasse & Sant'Anna (1997). Dois dos casos

particulares da distribuição proposta estendem as distribuições clássicas  $\mathcal{G}_I^0$  e  $\mathcal{K}_I$ . Finalmente, aplicações a dados SAR reais dão evidências de que as distribuições propostas superam as clássicas.

.....  
Sugestão de artigo em produção [Novas distribuições para imagens SAR]

2012 Nascimento, A. D. C., Frery, A. C. & Cintra, R. J. (2012a), ‘An extension for the family of  $\mathcal{G}_I$  distributions’.

.....

Estes estudos foram realizados utilizando as plataformas computacionais discutidas a seguir.

### 1.3 Plataforma computacional

Esta tese utilizou o programa `PolSARpro` e as linguagens de programação `R` e `Oct`, como acessórios computacionais em seu desenvolvimento.

O software educacional `PolSARpro` (*Polarimetric SAR Data Processing and Education Tool*) foi utilizado para obter imagens polarimétricas a fim de as utilizar nos estudos práticos e para estimar os parâmetros nos estudos com dados gerados por simulação de Monte Carlo. Este programa fornece ferramentas para visualização, análise e processamento de imagens PolSAR. Este projeto foi desenvolvido pela Agência Espacial Européia (*European Space Agency* – ESA). Uma versão deste programa bem como tutoriais e dados podem ser obtidos através do domínio <http://earth.eo.esa.int/polsarpro/>.

A produção de gráficos foi realizada utilizando a linguagem de programação `R` (Chambers 2008). O `R` também foi utilizado para fazer as aplicações e simulações no contexto polarimétrico bem como produzir programas iterativos envolvendo medidas de entropias e distâncias em dados PolSAR. Esta linguagem tem sido bastante difundida, principalmente no campo da Estatística, por comportar funções para uma grande diversidade de linhas de estudos (Cribari-Neto & Zarkos 1999). Ihaka & Gentleman (1996) propuseram esta linguagem objetivando produzir um ambiente de uso similar ao `S`, porém se desviando de seus problemas de demanda de memória e desempenho. Esta tese utilizou a versão `R` 2.13.2 desta linguagem e uma versão mais atual pode ser obtida pelo sítio <http://www.r-project.org/>.

Algumas das simulações e cálculos foram realizados empregando a versão 5.10 da linguagem de programação `Oct` em um computador Intel Core 2 Duo sob o sistema operacional Windows XP. Elas também utilizaram o gerador de números pseudo-aleatórios “multiply-with-carry” com 52 bits e período aproximado de  $2^{8222}$ , desenvolvido por George Marsaglia. Baseada na linguagem de programação `C`, esta flexível

linguagem foi proposta por Doornik (1999). Ela vem encontrando espaço entre estatísticos, economistas e engenheiros, devido a sua eficiência quanto a velocidade e manipulação. A versão mais atual desta linguagem de programação está disponível gratuitamente para uso acadêmico no domínio <http://www.doornik.com/index.html>.

Um panorama completo desta tese é apresentado na seguinte seção.

## 1.4 Objetivo e organização da tese

A seguinte discussão visa agrupar os objetivos específicos destacados na primeira seção deste capítulo na forma de objetivos gerais.

O escopo desta tese é a modelagem e proposta de métodos inferenciais baseados em Teoria da Informação para dados polarimétricos, com ênfase especial nas aplicações de radar de abertura sintética. Assim, o trabalho de investigação desta tese está dividido em duas áreas de estudo. Primeiramente, propor distâncias estocásticas bem como investigar seus comportamentos assintóticos para as diferentes distribuições voltadas a imagens PolSAR. Adicionalmente, visa propor outras medidas dentro da Teoria da Informação direcionadas aos contextos de imagens PolSAR, tal como entropia (Zografos & Nadarajah 2005). Em segundo lugar, realizar aplicações dos ferramentas derivados aos contextos de classificação, segmentação e detecção de bordas em tais imagens.

O restante desta tese está organizado na forma dos seguintes capítulos. O **Capítulo 2** apresenta uma discussão sobre radar polarimétrico. A geometria de obtenção de uma imagem SAR é explicada através de definições e ilustrações. O **Capítulo 3** fornece uma discussão dos principais tipos de modelagem estatística para dados SAR nas perspectivas de polarização simples e múltipla. As relações de “convergência em distribuição” entre essas leis são diagramadas e a interpretabilidade de seus parâmetros é comentada e relacionada com processamento de imagens PolSAR. O **Capítulo 4** descreve a metodologia de teste de hipóteses baseado na classe de divergências  $(h, \phi)$  proposta por Salicrú et al. (1994). Algumas distâncias clássicas são contextualizadas em aplicações no campo da Inferência Estatística. Adicionalmente, medidas de entropias são consideradas bem como o estudo de seus comportamentos assintóticos. O **Capítulo 5** aborda uma discussão a respeito dos principais estudos realizados nesta tese. As distribuições  $\mathcal{G}_I^0$  e Wishart complexa escalonada são submetidas a estudos com dados obtidos via simulação de Monte Carlo e reais, obedecendo à estrutura estatística de uma imagem PolSAR. Em termos gerais, este capítulo fornece (i) estruturas analíticas pertinentes ao contexto de discriminação em imagens PolSAR e (ii) conclusões importantes na linha de modelagem e métodos inferenciais em dados PolSAR. Finalmente, no **Capítulo 6**, nós concatenamos toda a tese na forma de um texto com

as principais discussões e conclusões.

No tocante à notação utilizada nesta tese, empregaram-se as seguintes regras:

- Matrizes e vetores estocásticos são representados por letras e símbolos negritos maiúsculos e minúsculos, respectivamente.

*“Nas páginas 27 e 28, um vetor  $\mathbf{y}$  distribuído com uma lei gaussiana complexa multivariada e uma matriz com distribuição Wishart complexa escalonada são representados por  $\mathbf{y} \sim \mathcal{N}_m^c(\mathbf{\Sigma}, \boldsymbol{\mu})$  e  $\mathbf{Y} \sim \mathcal{W}(\mathbf{\Sigma}, L)$ , respectivamente.”*

- As realizações das matrizes e vetores aleatórios se distinguem dos mesmos pelo uso do apóstrofo ( $\cdot$ )'. Exemplificando, seja um vetor aleatório  $\mathbf{x}$  e uma matriz aleatória  $\mathbf{X}$ , então as realizações simples destas variáveis são representadas por  $\mathbf{x}'$  e  $\mathbf{X}'$ , respectivamente.

*“Nas páginas 27 e 28, as densidades relativas às variáveis  $\mathbf{y}$ ,  $f_{\mathbf{y}}(\mathbf{y}': \cdot)$ , e  $\mathbf{Y}$ ,  $f_{\mathbf{Y}}(\mathbf{Y}': \cdot)$ , têm como realizações  $\mathbf{y}'$  e  $\mathbf{Y}'$ , respectivamente.”*

- Mantendo a mesma lógica, uma variável aleatória escalar  $X$  tem por realização a variável  $X'$ .

*“Na página 42, a variável  $X$  com distribuição gaussiana inversa generalizada e densidade  $f_X(X'; \cdot)$  tem como realização a variável  $X'$ .”*

- Ainda com relação ao apóstrofo ( $\cdot$ )', quando relacionado a funções, ele representará a derivada de primeira ordem. Mantendo o mesmo raciocínio, ( $\cdot$ )'' indicará a derivada de segunda ordem de uma função.

*“Na página 58,  $\phi'$  e  $\phi''$  representam as derivadas de primeira e segunda ordem da função  $\phi$  núcleo das medidas de divergências, respectivamente.”*

## Capítulo 2

# Radar Polarimétrico

Começaremos este capítulo tentando responder à seguinte pergunta:

### O que é polarização?

Segundo Lee & Pottier (2009), três autores desenvolveram as primeiras contribuições para compreensão de tal termo. Em 1669, o primeiro trabalho quantitativo sobre a observação da luz foi publicado por Erasmus Bartholinus. Seguidamente, Christian Huygens, em 1677, propôs importantes contribuições quanto à natureza da luz e, mais pontualmente, sobre luz polarizada. Finalmente, Étienne-Louis Malus, em 1808, provou a conjectura de Newton, que a polarização é uma propriedade intrínseca da luz (Lee & Pottier 2009).

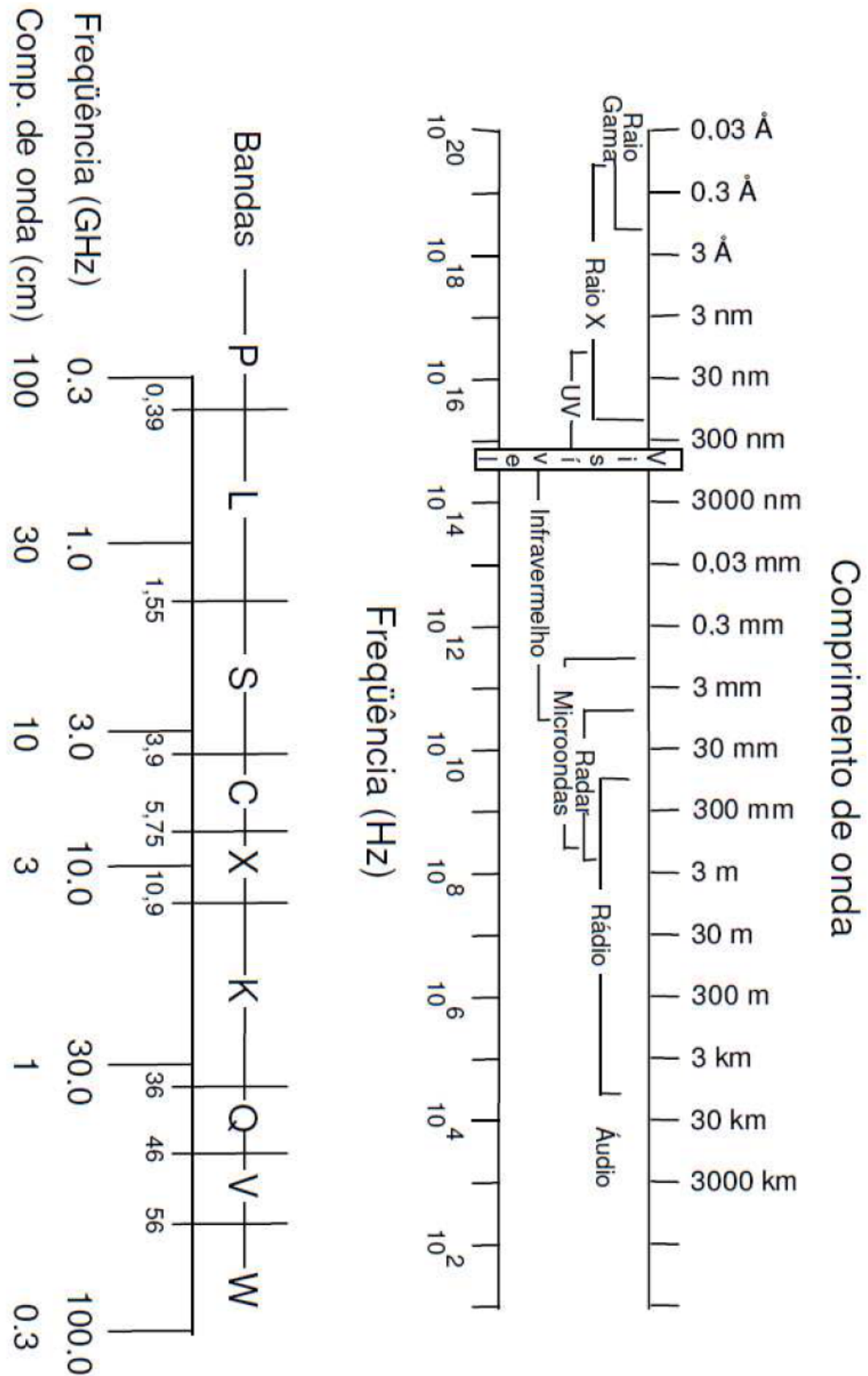
Esta tese entende este termo tal como proposto por Freitas et al. (2003),

*‘a orientação segundo a qual oscila, no tempo, o vetor de campo elétrico de onda eletromagnética’.*

A compreensão deste termo é importante, pois as propriedades das imagens obtidas pelos aparelhos de sensoriamento remoto estão relacionadas a este efeito.

O sensoriamento remoto tem se mostrado uma boa ferramenta para mapear a superfície da terra. As primeiras formas de sensoriamento, tais como fotografias aéreas apresentaram altos custos de obtenção (como discutido por Freitas et al. 2003). Além disso, o levantamento aerofotogramétrico era realizado apenas em pequenas áreas e demandava muito tempo. A consolidação de sensores orbitais permitiu que grandes áreas pudessem ser imageadas com uma certa periodicidade. A partir deste instante, os sensores foram se refinando e se organizando em diferentes classes. A seguinte discussão objetiva delimitar o tipo de sensor trabalhado no decurso desta tese.

Diferentemente dos sensores ópticos (que trabalham na faixa óptica de comprimento de onda, isto é,  $0.3\ \mu\text{m}$  a  $15\ \mu\text{m}$  conforme a Figura 2.1), os que trabalham na faixa do espectro de microondas, de  $1\ \text{mm}$  até cerca de  $1\ \text{m}$ , são mais procurados pelas seguintes vantagens:



**Figura 2.1** Faixas do comprimento de onda e da frequência para um espectro eletromagnético (Fonte: Freitas et al. 2003, p. 2).

- a. Pela sua capacidade de adquirir dados independentemente da iluminação solar bem como das condições atmosféricas (por exemplo, opera em camadas atmosféricas nebulosas);
- b. Realce do relevo do terreno, em virtude da possibilidade da visada inclinada;
- c. Produção de “Modelo Digital do Terreno” ou do “Modelo Digital de Superfície” por meio do uso das técnicas interferométricas.

Quanto a tais sensores, os radares (conhecidos como sistemas *ativos*) têm uma fonte própria de energia e funcionam a emitindo em direção a um alvo sob estudo e, seguidamente, capturando seu retorno. Por outro lado, os radiômetros (*passivos*) dependem de recursos de iluminação naturais para detectarem a energia enviada pelos cenários.

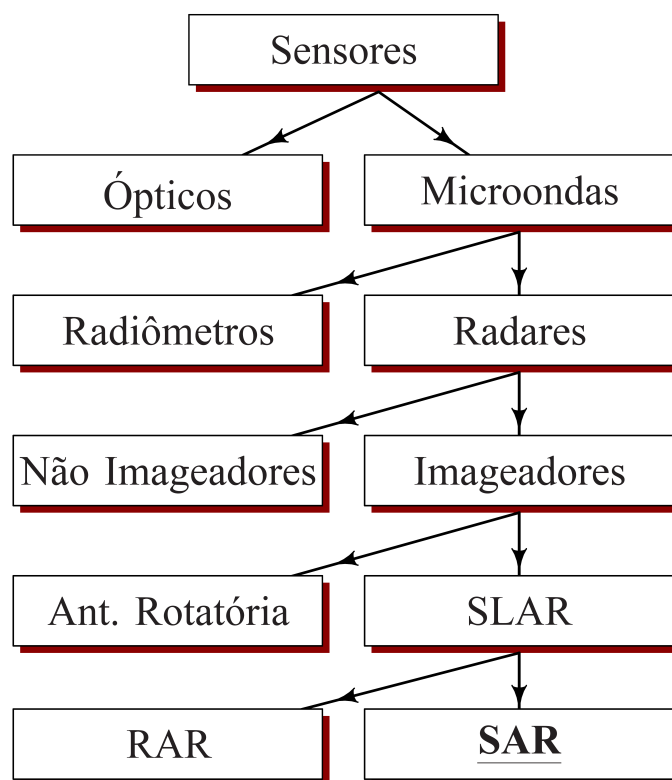


Figura 2.2 Diagrama de sensores.

Os radares imageadores são um sensores de rádio-freqüência que permite analisar informações sobre propriedades físicas e elétricas dos cenários em estudos. Tais sistemas, por sua vez, são divididos naqueles com antena rotatória e os de visada lateral (*Side Looking Airbone Radar* – SLAR). Finalmente, o SLAR se ramifica no radar de abertura real (*Real Aperture Radar* – RAR) e naquele de abertura sintética (*Synthetic Aperture Radar* – SAR). Esta tese faz contribuições no contexto de imagens obtidas a partir de

um SAR. Buscando situar esta tese na tecnologia de radares, o diagrama da Figura 2.2 traça um panorama das classes de sensores discutidas acima.

A seguir, a estrutura deste capítulo é apresentada. A Seção 2.1 mostra os conceitos relativos aos sistemas de radares. O comportamento dos cenários sob estudos é observado na Seção 2.2. Finalmente, o processo de formação de uma imagem SAR é analisado na Seção 2.3.

## 2.1 Sistemas de radar

O sistema SAR opera com o envio e captura de pulsos, polarizados circular (Rizki Akbar, Tetuko S. S. & Kuze 2010) e linearmente (Woelder & Granholm 2010), em direção a um cenário real. Esta tese se baseia na polarização linear:

*‘Quando a orientação do vetor de onda eletromagnética varia segundo uma linha reta (horizontal ou vertical), a polarização é dita ser linear’.*

Os sistemas de radares comumente utilizam polarizações paralelas, denominadas por HH e VV, ou cruzadas, HV e VH (nestas notações, a primeira e a segunda letra se referem às polarizações da radiação transmitida e recebida pela antena, respectivamente). A seguir, discute-se sobre (i) os tipos de plataformas que transportam os radares e (ii) a geometria de obtenção de suas imagens.

### 2.1.1 Sistemas de satélites e areotransportados

Datando o ano de 1978, o primeiro satélite orbitando a terra com SAR foi o SEASAT, que estava designado para o sensoriamento remoto dos oceanos e mares gelados (mostrado na Figura 2.3). Como resultado da experiência, verificou-se sua forte capacidade de discriminar terrenos e de detectar formas. O SEASAT SAR operou na banda L (com comprimento de onda 23.5 cm) com um canal de polarização simples, HH (transmissão e recepção horizontais). Com o sucesso do seu lançamento, o imageamento de radar se estabeleceu como uma ferramenta indispensável para o sensoriamento remoto espacial.

A Tabela 2.1 traça uma retrospectiva da evolução destes sistemas (Maiores detalhes são dados por Filho & Crósta 2003). Portanto, o SEASAT SAR originou o desenvolvimento do radar polarimétrico, o qual é uma extensão natural do sistema SAR. No tocante ao envolvimento do Brasil em missões de sensoriamento remoto, destaca-se a missão SSR (Satélite de Sensoriamento Remoto) e SABIA (Satélite Argentino-Brasileiro de Informações sobre Águas, Alimento e Ambiente)





**Figura 2.3** Satélite SEASAT (Fonte: European Space Agency 2009).

Além dos sistemas em que sensores são acoplados a satélites, existem aqueles que trabalham fixados em aviões, chamados de *aerotransportados*. Esta tese trabalha com imagens fornecidas por estes sistemas. A Figura 2.4 exibe os sistemas aerotransportados com suas principais informações (no sentido de cima para baixo, (i) o nome do sensor, (ii) o órgão gerenciador, (iii) o nome da aeronave transportadora e (iv) as bandas utilizadas pelo sensor). O Apêndice A detalha três destes sistemas que serão utilizados nesta tese: AIRSAR, EMISAR e ESAR.

### 2.1.2 Geometria de imageamento

Um sistema de obtenção de imagens SAR consiste de um transmissor de microondas em forma de pulsos, uma unidade receptora, uma antena (que é usada comumente tanto para transmissão como para recepção, *sistemas monoestáticos*)<sup>1</sup>, um modulador e um processador. Tais radares são montados em uma plataforma móvel operando conforme uma configuração geométrica ilustrada na Figura 2.5(a). O sistema de imageamento SAR está situado a uma altura 'H' e se move com uma velocidade ' $V_{\text{SAR}}$ '. A emissão de pulsos é feita por uma antena ajustada perpendicularmente à direção de voo. O feixe da antena cujo eixo é a reta  $r$  se dirige ao chão numa incidência de inclinação  $\theta_0$ . A abertura física do radar, de largura  $\ell_X$  e comprimento  $\ell_Y$ , gera um feixe elíptico centrado em  $P_0$  e cujos comprimentos dos eixos são  $\Delta X$  e  $\Delta Y$ . A área imageda

<sup>1</sup>Alguns casos utilizam *sistemas biestáticos*, que trabalham com duas antenas de configurações distintas (Ulaby, Moore & Fung 1986).

|                                                                                                                |                                                                                      |                                                                                                                  |                                                                                                                                   |
|----------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------|
|                             |   |                                |                                                  |
| <b>AES1</b><br>InterMap Technologies (D)<br><a href="#">GulfStream Commander</a><br>X-Band (HH), P-Band (Quad) | <b>AIRSAR</b><br>NASA / JPL (USA)<br><a href="#">DC8</a><br>P, L, C-Band (Quad)      | <b>AuSAR - INGARA</b><br>D.S.T.O (Aus)<br><a href="#">DC3 (97) KingAir 350 (00) Beach 1900C</a><br>X-Band (Quad) | <b>DOSAR</b><br>EADS / Dornier GmbH (D)<br><a href="#">DO 228 (89), C160 (98), G222 (00)</a><br>S, C, X-Band (Quad), Ka-Band (VV) |
|                             |   |                                |                                                  |
| <b>ESAR</b><br>DLR (D)<br><a href="#">DO 228</a><br>P, L, S-Band (Quad)<br>C, X-Band (Sngl)                    | <b>EMISAR</b><br>DCRS (DK)<br><a href="#">G3 Aircraft</a><br>L, C-Band (Quad)        | <b>MEMPHIS / AER II-PAMIR</b><br>FGAN (D)<br><a href="#">Transal C160</a><br>Ka, W-Band (Quad) / X-Band (Quad)   | <b>STORM</b><br>UVSQ / CETP (F)<br><a href="#">Merlin IV</a><br>C-Band (Quad)                                                     |
|                            |  |                               |                                                 |
| <b>PHARUS</b><br>TNO - FEL (NL)<br><a href="#">CESSNA - Citation II</a><br>C-Band (Quad)                       | <b>PISAR</b><br>NASDA / CRL (J)<br><a href="#">GulfStream</a><br>L, X-Band (Quad)    | <b>RAMSES</b><br>ONERA (F)<br><a href="#">Transal C160</a><br>P, L, S, C, X, Ku, Ka, W-Band (Quad)               | <b>SAR580</b><br>Environnement Canada (CA)<br><a href="#">Convair CV-580</a><br>C, X-Band (Quad)                                  |

Figura 2.4 Sistemas aerotransportados (Fonte: [http://earth.eo.esa.int/polsarpro/course\\_material.html](http://earth.eo.esa.int/polsarpro/course_material.html)).

**Tabela 2.1** Evolução de radares

| Orgão                                                | Ano  | Radar      |
|------------------------------------------------------|------|------------|
| National Aeronautics and Space Administration (NASA) | 1980 | SIR-A      |
| NASA                                                 | 1990 | SIR-B      |
| European Space Agency (ESA) ESA                      | 1992 | ERS-1      |
| Japanese Earth Resources Satellite (JERS)            | 1992 | JERS-1     |
| ESA                                                  | 1995 | ERS-2      |
| Canadian Space Agency (CSA)                          | 1995 | RADARSAT-1 |
| ISRO (Indian Space Research Organization)            | 1996 | IRS-P3     |
| KARI (Korea Aerospace Research Institute)            | 1999 | Kompsat-1  |
| NASA                                                 | 1999 | QuikSCAT   |
| CSA                                                  | 2001 | RADARSAT-2 |
| SPOT (Système Pour l'Observation de la Terre)        | 2002 | SPOT-5     |
| SSTL (Surrey Satellite Technology Ltd)               | 2005 | TopSat     |

é definida pelas aberturas da antena,  $\theta_x$  e  $\theta_y$ , que são dadas pelas aproximações

$$\theta_x \approx \frac{\lambda}{\ell_x} \quad \text{e} \quad \theta_y \approx \frac{\lambda}{\ell_y},$$

em que  $\lambda$  é o tamanho de onda correspondente à frequência do sinal transmitido (Para maiores detalhes, consultar Lee & Pottier 2009).

As Figuras 2.5(b) e 2.5(c) apresentam secções transversais da Figura 2.5(a) em termos dos eixos  $x$  e  $y$ , respectivamente. A partir delas, é possível definir expressões aproximadas para  $\Delta X$  e  $\Delta Y$ :

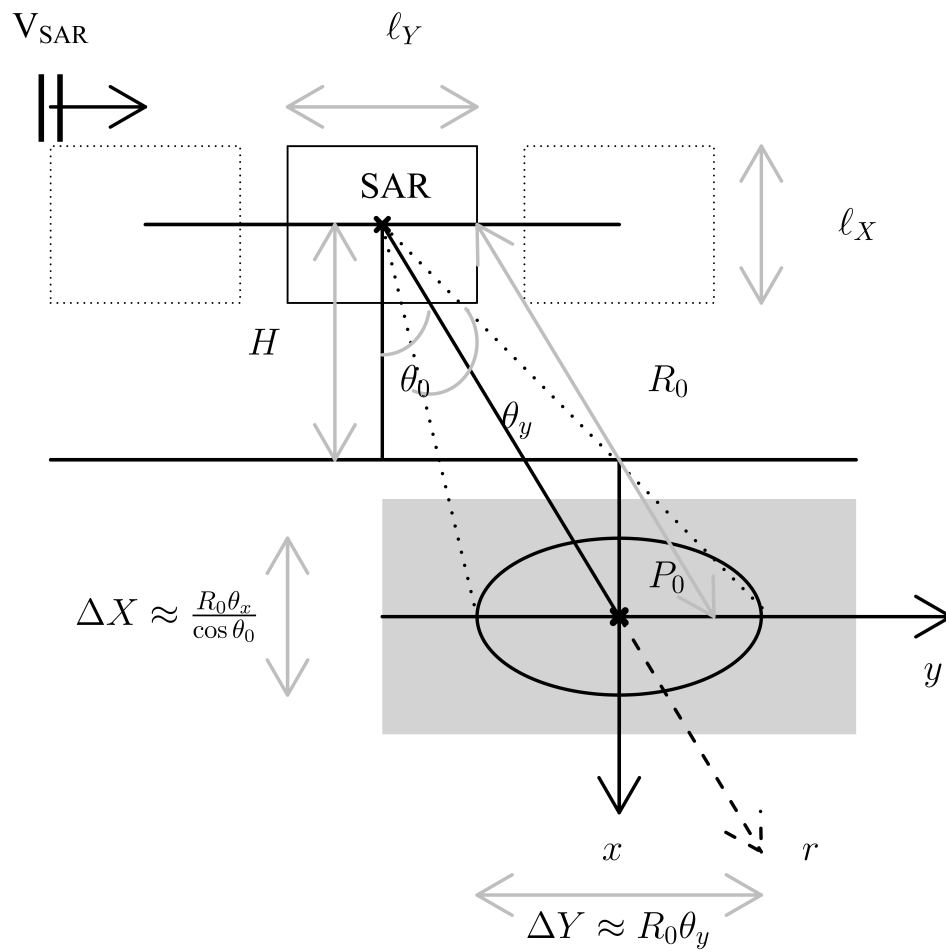
$$\Delta X \approx \frac{R_0 \theta_x}{\cos \theta_0} \quad \text{e} \quad \Delta Y \approx R_0 \theta_y,$$

em que  $R_0$  é a distância entre o radar e o centro do cenário imageado,  $P_0$ .

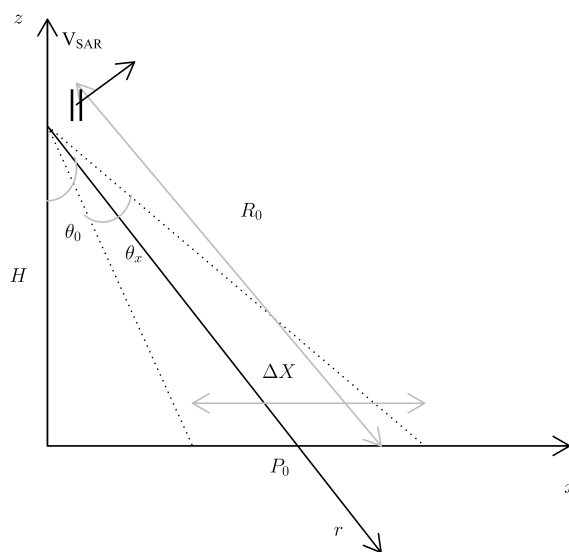
Por fim, um modulador e um processador instalados na plataforma móvel proporcionam a obtenção dos dados e digitalização da imagem, respectivamente. Assim, a área coberta pelo feixe da antena nas direções  $x$  e  $y$  é, então, representada pela antena.

## 2.2 Comportamento dos alvos

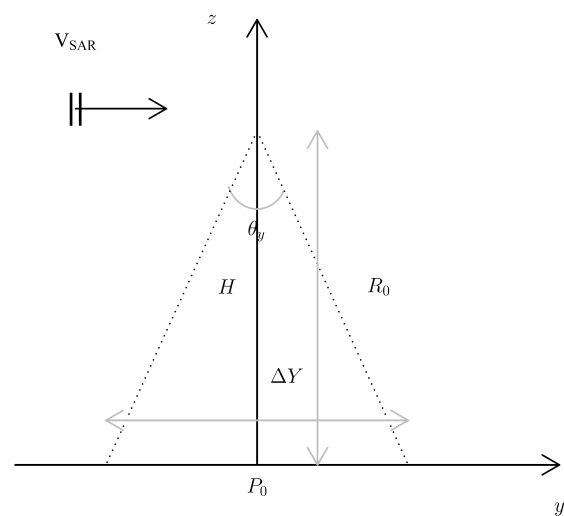
A física de formação de uma imagem SAR é modelada pela equação de radar (Ulaby & Elachi 1990). Essa expressão resume a relação entre (i) os elementos que compõem



(a) Geometria



(b) Plano para ordenada constante



(c) Plano para abscissa constante

**Figura 2.5** Imageamento por SAR.

o sistema de radar, (ii) o alvo sob análise e (iii) o sinal transmitido pela antena:

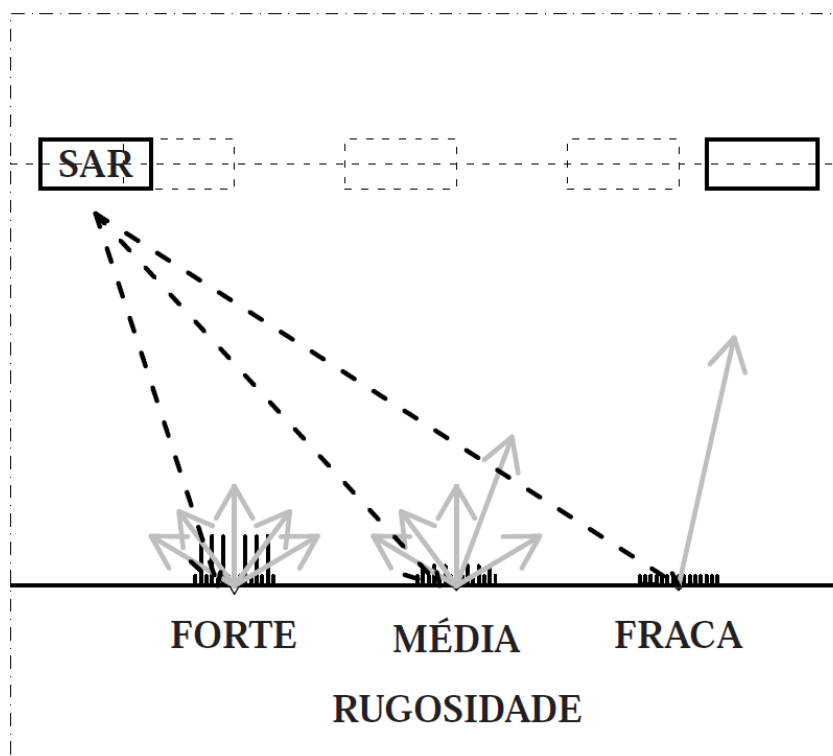
$$\bar{P}_r = \frac{P_t \lambda^2 G_0^2}{(4\pi)^3 R^4} \sigma, \quad (2.1)$$

em que  $\bar{P}_r$  é a potência média recebida pela antena (em watts),  $P_t$  é a potência transmitida pela antena,  $\lambda$  é o comprimento de onda (em nanômetros,  $10^{-9}$  metros),  $G_0$  é o ganho máximo da antena (é uma quantidade adimensional representando a razão entre as tensões de entrada e saída, medida, as vezes, em volt por volt),  $R$  é a distância entre a antena e o alvo (em metros) e  $\sigma$  é o coeficiente de retroespalhamento, “backscatter” (metros quadrados). Primeiramente, note que o único elemento, do lado direito desta igualdade, que se relaciona com o alvo é  $\sigma$ , sendo o restante dos elementos destinados a descrever o sistema de radar. Assim, a física de formação de uma imagem SAR tem uma natureza multiplicativa, cujos fatores procuram modelar a estrutura do radar e a resposta do relevo do cenário sob estudo aos pulsos emitidos.

Quanto ao estudo de alvos em processamento de imagens SAR, existe uma relação biunívoca entre os termos *célula de resolução* e *pixel* de uma imagem. O primeiro deles trata de um conjunto de difusores presentes na unidade especificada do terreno. Por outro lado, um pixel pode ser definido como uma aproximação da realidade de uma dada célula de resolução. Desta forma, o dado associado a um pixel é resultante da soma das contribuições dos difusores de uma célula de resolução e das reflexões múltiplas entre eles. Maiores detalhes são descritos por Ulaby et al. (1986).

Ainda quanto à resposta dos alvos aos pulsos emitidos pelo SAR, os tipos de terrenos apresentam diferentes retornos. Estes efeitos são ilustrados na Figura 2.6.

As superfícies mais rugosas dispersam o sinal recebido para várias direções. A medida que as superfícies vão ficando mais lisas, os sinais retornados ficam menos dispersos. Pode-se perceber ainda que as superfícies lisas refletem quase toda a energia incidente em direção oposta ao sistema imageador. Esta reflexão é denominada por *especular* e dá as regiões homogêneas (tais como oceanos e pastos) de uma imagem um tom mais escuro. Consequentemente, quanto mais rugosa é a superfície mais energia é retornada ao radar, o que caracteriza as regiões heterogêneas (como, por exemplo, áreas urbanas) com um maior brilho. Nestes casos, os pulsos encontram grandes obstáculos e são retornados ao radar, gerando um efeito conhecido como *reflexão de canto*. No caso de florestas, há uma dispersão dos raios, mas uma quantidade maior de energia é perdida com relação às áreas de cidades. Tal reflexão é conhecida como *difusa*. As regiões de florestas em uma imagem PolSAR são tipicamente caracterizadas por um brilho menor do que aquele de áreas urbanas e maior relativamente às regiões de oceanos e pastos. Além disso, algumas regiões são caracterizadas pelo fenômeno de “double bounce”: situações em que os pixels exibem altos valores associados aos pulsos retornados. Este evento é causado por comportamentos naturais ou resultado de erros de mensuração feito pelo homem e estão presentes em cenários de áreas urbanas e florestas densas. O estudo de pixels afetados por *double bounce* é importante pois eles tendem a modificar as propriedades estatísticas da área em torno



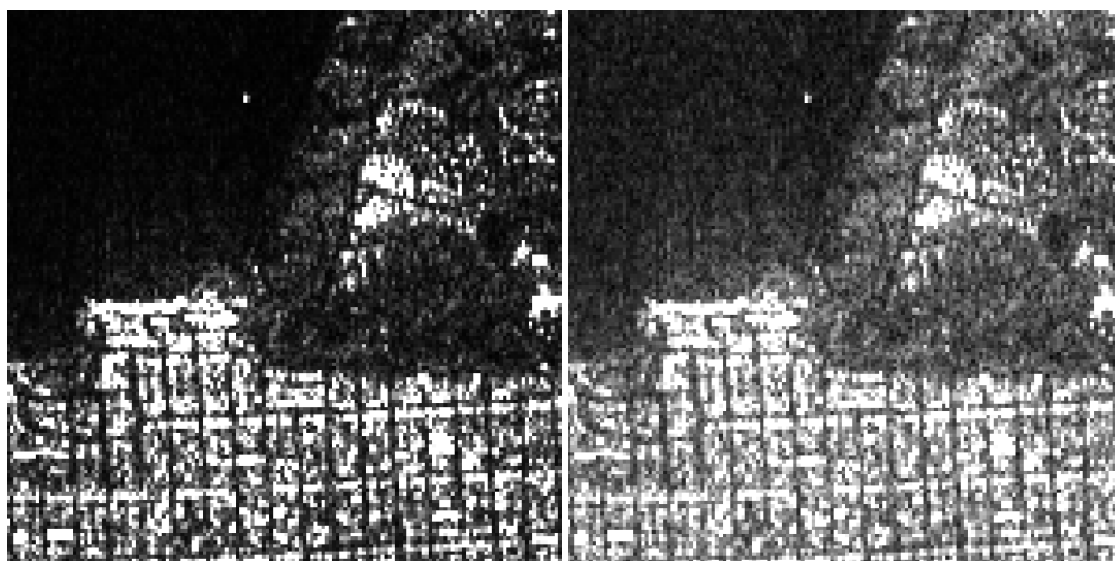
**Figura 2.6** Reflexividade em diferentes terrenos.

deles, produzindo estimativas e inferências mais viesadas do que o usual. Finalmente, em complemento a discussão anterior, é importante mencionar que a quantidade de rugosidade depende do comprimento de onda; isto é, uma área rugosa na banda C, pode não o ser na banda L.

A Figura 2.7 apresenta duas imagens de uma mesma área em formatos diferentes com tamanhos  $150 \times 150$  pixels de uma imagem AIRSAR da cidade de São Francisco, discutida no Apêndice A. A área considerada exibe três regiões com diferentes graus de rugosidade: oceano (parte superior a esquerda do leitor), floresta (superior a direita) e urbana (parte inferior). Note que quanto mais heterogênea é a região mais clara é a sua representação em uma imagem SAR. Por fim, a região de floresta apresenta um ponto central destoante do conjunto, caracterizando o fenômeno de *double bounce*.

## 2.3 Formação de imagens PolSAR

O processamento do sinal coerente recebido é requerido para a formação de uma imagem SAR. O sinal retornado (de natureza complexa) é submetido às seguintes eta-



(a) Intensidade

(b) Amplitude

**Figura 2.7** Formatos em uma imagem SAR.

pas: (i) demodulado em fase ( $V$ ) e quadratura ( $Q$ ), (ii) discretizado (dada uma função amostra) e (iii) convertido para o formato digital visando obter uma sequência de dados. Neste passo, a taxa de amostragem empregada no sinal (iii.a) determina o número de pixels que comporá as linhas da imagem resultante e (iii.b) define o espaçamento entre os centros dos pixels. À medida que o sensor se desloca na direção de  $V_{\text{SAR}}$  na Figura 2.5(a), o processamento dos pulsos emitidos compõe as colunas da imagem resultante. Quanto aos formatos em que uma imagem SAR pode ser tratada, têm-se a intensidade ( $V^2 + Q^2$ ) e a amplitude ( $\sqrt{V^2 + Q^2}$ ) além dos componentes  $V$  e  $Q$ . A Figura 2.7 exhibe os aspectos de uma imagem nos formatos de intensidade e amplitude.

## Modelos Multilook para Dados SAR Polarimétricos, Bivariados e Univariados

Os sistemas PolSAR têm alcançado uma posição de destaque dentre as tecnologias de sensoriamento remoto (Lee & Pottier 2009). Este fato é justificável pois, em resumo às várias vantagens sobre os sensores ativos elencadas no Capítulo 2, tais sistemas (i) têm a capacidade de operar independentemente das condições atmosféricas e (ii) são capazes de prover imagens com altas resoluções espaciais. A seguir, apresenta-se uma discussão do sistema SAR quanto ao número de polarizações.

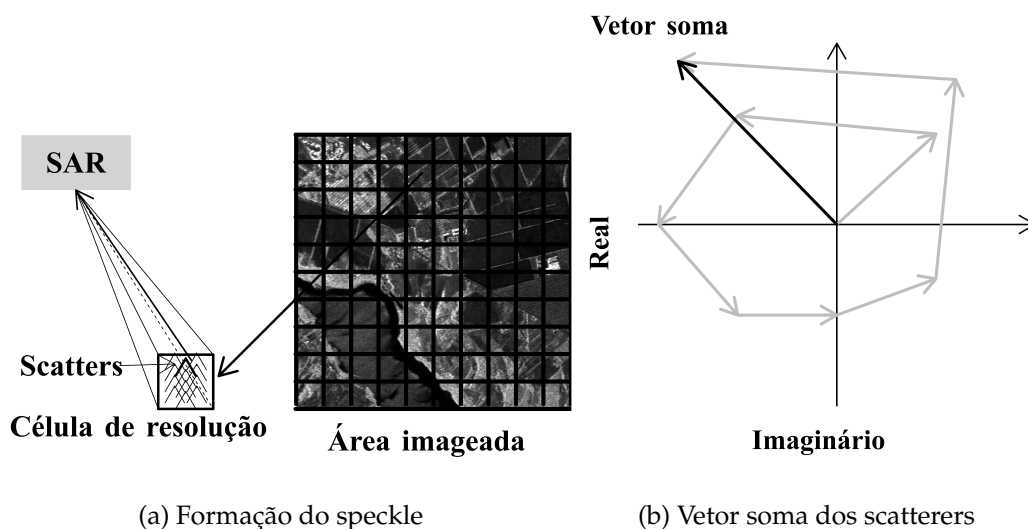
A versão univariada do sistema SAR considera o mapeamento de uma área sob estudo (frequentemente chamada de *cena* ou *cenário*) através da emissão e do registro de pulsos sob uma direção (ou polarização) de onda, tendo como resultado imagens *monopolarizadas*. Em anos recentes, o interesse em se estudar tais sistemas numa perspectiva multidimensional tem crescido. Os dados obtidos nesta abordagem são referenciados por *dados PolSAR*. Neste caso, as informações polarimétricas são produzidas pela emissão e obtenção de sinais multivariados e complexos. Os dados resultantes registram a amplitude e a fase dos sinais retornados por uma cena, considerando possíveis combinações das polarizações lineares de transmissão e recepção: HH, HV, VH e VV (H e V são representações para as polarizações horizontais e verticais, respectivamente).

Devido à utilização de iluminação coerente no seu processamento, as imagens PolSAR são contaminadas com flutuações sobre as intensidades registradas. Este fenômeno é chamado de ruído speckle.

O aparecimento do speckle em tais imagens se deve à interferência coerente do sinal refletido a partir das muitas unidades elementares da área sob estudo. Isto é, quando um radar ilumina uma superfície áspera, o sinal retornado consiste de ondas refletidas pelos muitos elementos dentro de uma célula de resolução (tal como ilustrado na Figura 3.1(a)), denominados de “scatters”.

As distâncias entre as posições dos scatters e a antena do radar variam aleatoriamente de célula para célula, produzindo um vetor soma (vide Figura 3.1(b)) para cada célula de resolução. Assim, uma imagem pode ser encarada como um conjunto de variáveis aleatórias (Nascimento et al. 2010), sendo cada célula de resolução repre-





**Figura 3.1** Natureza do speckle.

sentada por um pixel, que é modelado por uma variável estocástica. Essa variação pixel-a-pixel manifesta um padrão de ruído granular sobre as imagens do tipo SAR. A Figura 3.2 exibe uma área da imagem AIRSAR (sensor discutido no Apêndice A.1.2) de Flevoland, Holanda, em que se observa a presença deste padrão de ruído.



**Figura 3.2** Padrão do ruído speckle em uma imagem SAR.

Além de dificultar a interpretação e análise de imagens SAR, o comportamento não aditivo e estocasticamente não gaussiano do ruído speckle (Frery et al. 1997) reduz a eficiência das técnicas usuais de processamento de imagens, tais como, classificação (Frery, Correia & Freitas 2007), segmentação (Beaulieu & Touzi 2004) e detecção de borda (Schou et al. 2003). Desta forma, a proposta de modelos adequados para descrever dados contaminados pelo ruído speckle é um passo fundamental de pré-processamento (Nascimento et al. 2010). As próximas seções objetivam satisfazer este

requisito através da apresentação e discussão dos principais modelos estatísticos para dados PolSAR.

Este capítulo está organizado na seguinte forma. Uma explanação sobre efeito “multilook” e modelos polarimétricos básicos abre este capítulo na forma da Seção 3.1. Subsequentemente, modelos particulares bivariados conforme a proposta de Lee, Hopel, Mango & Miller (1994) são considerados na Seção 3.2. Baseando-se na abordagem de modelagem multiplicativa (Freitas et al. 2005, Gao 2010), a Seção 3.3 finaliza este capítulo apresentando um modelo compacto para várias distribuições PolSAR, incluindo suas versões univariadas.

### 3.1 Física PolSAR e modelos básicos

Conforme a apresentação inicial, imagens polarimétricas são representadas por dados multivariados e complexos associados com as configurações de polarização da antena (HH – horizontal/horizontal, HV – horizontal/vertical, VH – vertical/horizontal e VV – vertical/vertical). Neste caso, costuma-se atribuir, a cada célula da imagem, uma matriz de elementos complexos,

$$\begin{bmatrix} S_{VV} & S_{VH} \\ S_{HV} & S_{HH} \end{bmatrix} \quad (3.1)$$

em que  $S_{ij}$  é o elemento de transmissão  $i$  e recepção  $j$  tal que  $i, j \in \{V, H\}$ . Tais elementos são definidos por  $S_{ij} = |S_{ij}|e^{i\psi_{ij}}$ , em que  $|S_{ij}|$  representa sua amplitude e  $\psi_{ij}$  indica sua fase (para maiores detalhes, consultar Sarabandi 1992). Adicionalmente, outras formas de parametrizar células de dados polarimétricos são a matriz de Stokes, a matriz de covariância e a matriz de Mueller (Ulaby & Elachi 1990). Na prática, sob as condições do teorema de reciprocidade (Ulaby & Elachi 1990), (3.1) pode ser simplificada no vetor complexo

$$\mathbf{s} = [S_{VV} \quad S_{VH} \quad S_{HH}]^t, \quad (3.2)$$

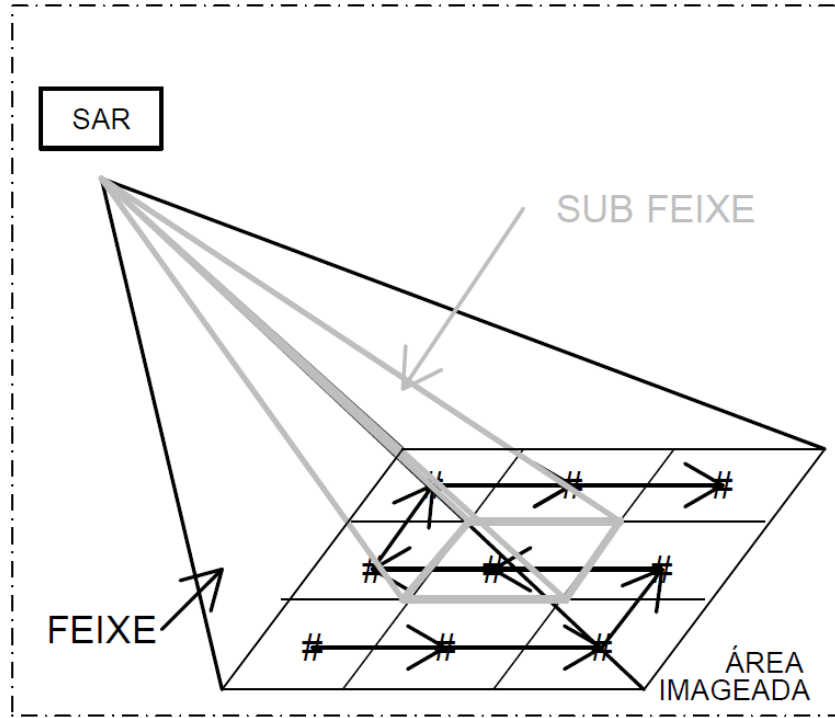
em que  $(\cdot)^t$  representa o operador de transposição. Uma boa descrição destes três elementos foi apresentada por Freeman & Durden (1998). O vetor de coeficientes complexos  $\mathbf{s}$  é conhecido na literatura de radar como “dado PolSAR para um look”.

É comum se trabalhar com uma extensão desta metodologia denominada de processamento “multilook”. Considerando o domínio espacial, o processo multilook é definido como a computação do valor médio de uma quantidade  $L$  de dados referen-

tes a pixels adjacentes (Lee & Pottier 2009), resultando na forma

$$\begin{aligned} \mathbf{S} &= L^{-1} \sum_{\ell=1}^L \mathbf{s}(\ell) \mathbf{s}(\ell)^H \\ &= L^{-1} \sum_{\ell=1}^L \begin{bmatrix} |S_{VV}(\ell)|^2 & S_{VV}(\ell)S_{VH}^*(\ell) & S_{VV}(\ell)S_{HH}^*(\ell) \\ S_{VH}(\ell)S_{VV}^*(\ell) & |S_{VH}(\ell)|^2 & S_{VH}(\ell)S_{HH}^*(\ell) \\ S_{HH}(\ell)S_{VV}^*(\ell) & S_{HH}(\ell)S_{VH}^*(\ell) & |S_{HH}(\ell)|^2 \end{bmatrix}, \end{aligned} \quad (3.3)$$

em que  $(\cdot)^H$  representa o transposto conjugado de uma matriz e  $(\cdot)^*$  indica o conjugado de uma variável complexa. Note que para o caso em que temos uma única polarização, a equação (3.3) resulta em  $\sum_{\ell=1}^L |S_i(\ell)|^2 / L$  para  $i = HH, VV, VH$ ; isto é,  $\mathbf{S}$  se reduz a uma variável real positiva. A Figura 3.3 exhibe uma ilustração deste princípio. Um feixe de radar foi dividido em  $L = 9$  sub feixes caracterizando um processamento multilook sob uma determinada área. Assim, o dado capturado pelo pulso retornado da área imageada é representado pela média dos retornos das células adjacentes, iluminadas pelos sub feixes. Embora esta ilustração sugira a presença de algum processo estocástico, para os resultados desta tese, os pixels são independentes e o efeito multilook é representado pela computação dada em (3.3).



**Figura 3.3** Efeito multilook.

Este efeito é incorporado aos modelos descritos neste capítulo. Nas próximas subseções, nós apresentamos as distribuições básicas para ajustar dados polarimétricos sem e com o efeito multilook. Estas situações serão denominadas por “1-look” e “L-looks”, respectivamente.

### 3.1.1 Distribuição gaussiana complexa

Com o objetivo de construir um tratamento estocástico da subseção anterior, admita um sistema PolSAR com  $m$  elementos de polarização e representado por

$$\mathbf{y} = [S_1 \ S_2 \ \cdots \ S_m]^t. \quad (3.4)$$

No tratamento de suas imagens,  $\mathbf{y}$  é frequentemente descrito por uma distribuição gaussiana complexa  $m$  variada com vetor de médias nulo cuja função densidade de probabilidade (fdp) é dada por

$$f_{\mathbf{y}}(\mathbf{y}'; \Sigma) = \frac{1}{\pi^m |\Sigma|} \exp(-\mathbf{y}'^H \Sigma^{-1} \mathbf{y}'), \quad (3.5)$$

em que  $|\cdot|$  é o determinante de uma matriz para argumentos matriciais e o módulo para variáveis escalares complexas,  $\Sigma$  é a matriz de covariância de  $\mathbf{y}$  dada por

$$\Sigma = E(\mathbf{y}\mathbf{y}^H) = \begin{bmatrix} E(S_1 S_1^*) & E(S_1 S_2^*) & \cdots & E(S_m S_m^*) \\ E(S_2 S_1^*) & E(S_2 S_2^*) & \cdots & E(S_m S_m^*) \\ \vdots & \vdots & \ddots & \vdots \\ E(S_m S_1^*) & E(S_m S_2^*) & \cdots & E(S_m S_m^*) \end{bmatrix}$$

e  $E(\cdot)$  representa o operador de esperança estatística. Apresentando uma natureza hermitiana positiva definida, a matriz de covariância  $\Sigma$  contém toda a informação necessária para caracterizar os dados retornados sob análise (López-Martínez et al. 2005). Dada a importância da classe de matrizes hermitianas na tecnologia PolSAR, consideremos sua definição e algumas propriedades.

**Definição 1 (Matrizes hermitianas: Seber (2007)).** *Considere uma matriz  $\mathbf{A}$ . Se ela é complexa, então ela pode ser escrita como  $\mathbf{A} = \mathbf{A}_1 + i\mathbf{A}_2$ , em que  $\mathbf{A}_1$  e  $\mathbf{A}_2$  são matrizes reais. A matriz é dita ser hermitiana se  $\mathbf{A}^H = \mathbf{A}$ .*

*Para esta classe de matrizes, as seguintes propriedades se verificam: Seja  $\mathbf{A}$  uma matriz hermitiana  $m \times m$ , então*

1. *para  $\mathbf{A}$  não singular,  $\mathbf{A}^{-1}$  é uma matriz hermitiana.*
2.  *$\mathbf{A}^k$  é hermitiana, para  $k = 1, 2, \dots$*
3. *os autovalores de  $\mathbf{A}$  são reais.*
4.  *$\mathbf{x}^H \mathbf{A} \mathbf{x}$  é real para todo  $\mathbf{x} \in \mathbb{C}^n$ .*
5.  *$\mathbf{S}^H \mathbf{A} \mathbf{S}$  é hermitiana para toda matriz  $m \times m$   $\mathbf{S}$ .*

Em particular, se (i)  $\mathbf{A}$  é uma matriz hermitiana e (ii)  $\mathbf{x}^H \mathbf{A} \mathbf{x} > 0$  para todo  $\mathbf{x} \in \mathbb{C}^n$ , então  $\mathbf{A}$  é dita ser hermitiana positiva definida. Neste caso, uma característica importante de  $\mathbf{A}$  é que todos os seus autovalores e menores principais são positivos. (O uso de autovalores destas matrizes para caracterizar propriedades estatísticas de dados POLSAR foi discutido por Cloude & Pottier 1997).

A distribuição gaussiana complexa foi estudada por Goodman (1963b) para trabalhar com certas funções e com a transformada de Fourier (função característica), sendo, posteriormente, destinada a modelar o ruído speckle em cenários não multi-look (Goodman 1985). Sua definição pode ser dada a partir da lei gaussiana multivariada real, conforme discussão a seguir.

Considere, inicialmente, que um vetor aleatório  $\mathbf{h} = [V_1 \ Q_1 \ V_2 \ Q_2 \ \cdots \ V_m \ Q_m]^t$  tem uma distribuição gaussiana real  $2m$ -variada com parâmetros  $\boldsymbol{\mu} = [\mu_{V_1} \ \mu_{Q_1} \ \mu_{V_2} \ \mu_{Q_2} \ \cdots \ \mu_{V_m} \ \mu_{Q_m}]$  e  $\boldsymbol{\Sigma}$ ,  $\mathbf{h} \sim \mathcal{N}_{2m}(\boldsymbol{\Sigma}, \boldsymbol{\mu})$ . Neste caso, os termos  $V_k$  e  $Q_\ell$ , para  $\ell, k = 1, 2, \dots, m$ , serão usados para representar as partes real (em fase, vide Seção 2.3) e imaginária (em quadratura), respectivamente, de uma variável complexa “ $V_k + \mathbf{i}Q_\ell$ ” (sinal retornado), em que  $\mathbf{i} = \sqrt{-1}$  é a unidade imaginária. Admita ainda a seguinte estrutura para a matriz de covariância  $\boldsymbol{\Sigma}$ :

$$(\boldsymbol{\Sigma})_{k,\ell} = \mathbb{E} \begin{pmatrix} (V_k - \mu_{V_k})(V_\ell - \mu_{V_\ell}) & (V_k - \mu_{V_k})(Q_\ell - \mu_{Q_\ell}) \\ (Q_k - \mu_{Q_k})(V_\ell - \mu_{V_\ell}) & (Q_k - \mu_{Q_k})(Q_\ell - \mu_{Q_\ell}) \end{pmatrix} \\ = \begin{cases} \frac{\sigma_k^2}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, & \text{se } k = \ell \\ \frac{\sigma_k \sigma_\ell}{2} \begin{bmatrix} a_{k,\ell} & -b_{k,\ell} \\ b_{k,\ell} & a_{k,\ell} \end{bmatrix}, & \text{se } k \neq \ell, \end{cases}$$

tal que (i)  $\sigma_k/\sqrt{2}$  é o desvio padrão associado aos elementos  $V_k$  e  $Q_k$  do vetor  $\mathbf{h}$ , (ii)  $a_{k,\ell} = \text{Corr}(V_k, V_\ell) = \text{Corr}(Q_k, Q_\ell)$  representa o coeficiente de correlação definido nos pares  $(V_k, V_\ell)$  e  $(Q_k, Q_\ell)$  e (iii)  $b_{k,\ell} = \text{Corr}(V_k, Q_\ell) = \text{Corr}(Q_k, V_\ell)$  é o coeficiente de correlação definido nos pares  $(V_k, Q_\ell)$  e  $(V_k, Q_\ell)$ .

Satisfeitas as condições acima, a variável gaussiana complexa multivariada é definida pelo vetor aleatório

$$\mathbf{y} = [T_1 \ T_2 \ \cdots \ T_m]^t,$$

em que  $T_k = V_k + \mathbf{i}Q_k$  (para  $k = 1, 2, \dots, m$ ). Esta distribuição é denotada por  $\mathbf{y} \sim \mathcal{N}_m^c(\boldsymbol{\Sigma}, \boldsymbol{\mu})$  e sua densidade é

$$f_{\mathbf{y}}(\mathbf{y}'; \boldsymbol{\Sigma}, \boldsymbol{\mu}) = \frac{1}{\pi^m |\boldsymbol{\Sigma}|} \exp(-(\mathbf{y}' - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{y}' - \boldsymbol{\mu})),$$

em que  $\boldsymbol{\mu}$  é o vetor de médias e  $\boldsymbol{\Sigma}$  é a matriz de covariância hermitiana dada por

$$(\boldsymbol{\Sigma})_{k,\ell} = \begin{cases} \sigma_k^2, & \text{se } k = \ell, \\ \sigma_k \sigma_\ell (a_{k,\ell} + \mathbf{i}b_{k,\ell}), & \text{se } k \neq \ell. \end{cases}$$

Esta discussão da obtenção da gaussiana complexa multivariada a partir de sua versão real (frequentemente disponível em software estatísticos, tais como R e OX) é bastante útil para a geração de vetores complexos com distribuição  $\mathcal{N}_m^c$ .

### 3.1.2 Distribuição Wishart complexa escalonada

Com o objetivo de melhorar a razão sinal-ruído,  $L$  vetores aleatórios independentes e identicamente distribuídos (iid) são considerados e sua média é obtida visando formar a matriz de retroespalhamento  $L$ -looks (Anfinsen, Doulgeris & Eltoft 2009):

$$\mathbf{Y} = \frac{1}{L} \sum_{i=1}^L \mathbf{y}_i \mathbf{y}_i^H, \quad (3.6)$$

em que  $\mathbf{y}_i$ , para  $i = 1, 2, \dots, L$ , representa um vetor estocástico tal como na equação (3.4) com densidade dada por (3.5). A suposição de independência é assumida com base no fato que o tempo de captura de um pixel para o outro é praticamente negligível.

Satisfeitas as hipóteses mencionadas acima,  $\mathbf{Z}$  tem uma distribuição *Wishart complexa escalonada* e suporte dado pelo cone ( Um cone é definido como um subconjunto de um espaço vetorial que é fechado em termos da operação de multiplicação por escalar positivo. ) das matrizes hermitianas positivas definidas discutidas na Definição 1:  $\mathcal{A} = \{\mathbf{Y} : \mathbf{Y} \succ 0, \mathbf{Y} = \mathbf{Y}^H \in \mathbb{C}^{m \times m}\}$ , em que " $\mathbf{Y} \succ 0$ " significa que a matriz  $\mathbf{Y}$  é positiva definida. Definindo  $\mathbf{\Sigma}$  e  $L$  como seus parâmetros, essa lei é caracterizada pela seguinte fdp:

$$f_{\mathbf{Y}}(\mathbf{Y}'; \mathbf{\Sigma}, L) = \frac{L^m |\mathbf{Y}'|^{L-m}}{|\mathbf{\Sigma}|^L \Gamma_m(L)} \text{etr}(-L \mathbf{\Sigma}^{-1} \mathbf{Y}'), \quad (3.7)$$

em que  $L \geq m$ ,  $\text{etr}(\cdot) = \exp(\text{tr}(\cdot))$ ,  $\Gamma_m(L)$  é a função gamma multivariada para o caso complexo com expressão dada por

$$\Gamma_m(L) = \int_{\mathcal{A}} |\mathbf{Y}'|^{L-m} \text{etr}(-\mathbf{Y}') d\mathbf{Y}' = \pi^{m(m-1)/2} \prod_{i=0}^{m-1} \Gamma(L-i),$$

como definido por Mathai (1997, definição 6.1). A distribuição de  $\mathbf{Y}$  é denotada por  $\mathcal{W}(\mathbf{\Sigma}, L)$  e tem  $E(\mathbf{Y}) = \mathbf{\Sigma}$ , que é uma matriz positiva definida hermitiana (Anfinsen, Doulgeris & Eltoft 2009).

A distribuição para  $\mathbf{Y}$  é um escalonamento da lei Wishart complexa original estudada por Goodman (1963b) cuja densidade é dada por

$$f_{\mathbf{C}}(\mathbf{C}'; \mathbf{\Sigma}_{\mathbf{Y}}, L) = \frac{|\mathbf{C}'|^{L-m} \text{etr}(-\mathbf{\Sigma}_{\mathbf{Y}}^{-1} \mathbf{C}')}{\Gamma_m(L) |\mathbf{\Sigma}_{\mathbf{Y}}|^L}. \quad (3.8)$$

É possível provar que  $f_Y = f_C(LY')|J_{C \rightarrow Y}|$ , em que  $|J_{C \rightarrow Y}| = L^{m^2}$  é o determinante do Jacobiano da transformação  $C = LY$  (Mathai 1997, Capítulo 3).

Os parâmetros  $\Sigma$  e  $L$  possuem uma interpretabilidade voltada à física de formação de imagens PolSAR:  $|\Sigma|$  está relacionado ao brilho de uma imagem, enquanto  $L$  se relaciona com a razão sinal-ruído. A título de elucidação dos efeitos destes parâmetros na imagem formada, foi gerada uma imagem sintética, variando  $\Sigma$  e  $L$ . A Figura 3.4 exibe a influência destes parâmetros sobre imagens sintéticas de tamanhos  $100 \times 100$  pixels com  $L \in \{3, 12\}$  (representando o menor valor considerado em casos com três canais de polarização e um valor alto em situações práticas) e  $k \in \{0, 5, 10, 15\}$  tal que  $\Sigma = (1 + k)B$ , em que

$$B = \begin{bmatrix} 360932 & 11050 + 3759i & 63896 + 1581i \\ & 98960 & 6593 + 6868i \\ & & 208843 \end{bmatrix}, \quad (3.9)$$

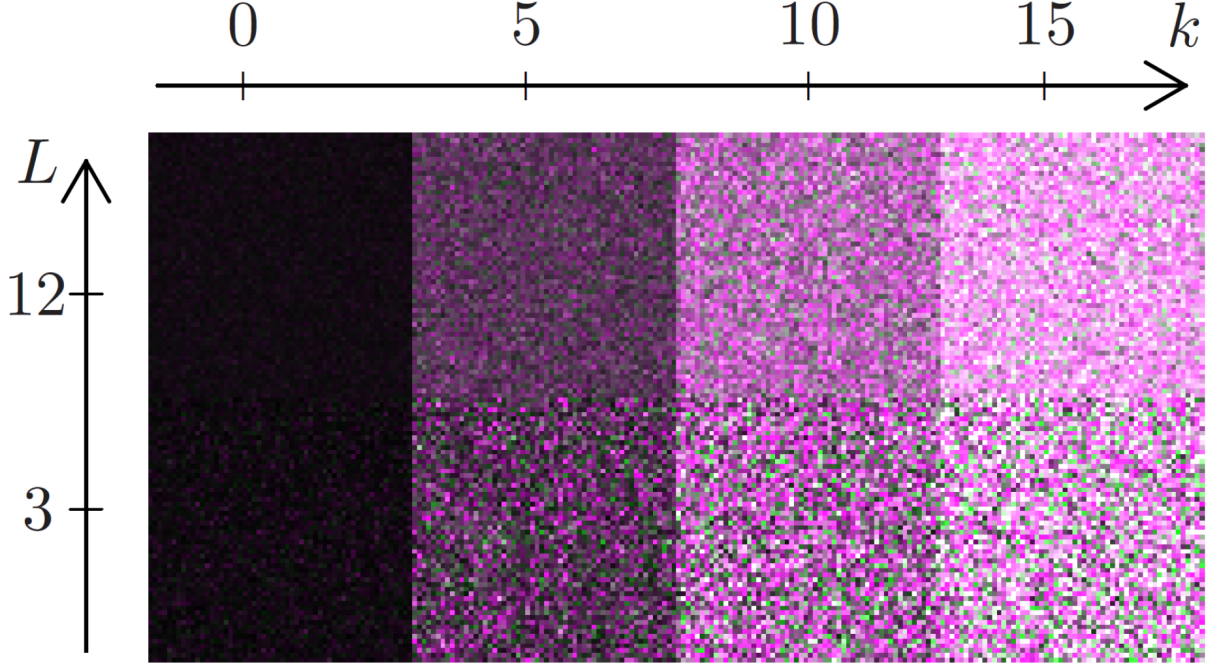
em que  $m = 3$ , os elementos  $B_{11}$ ,  $B_{22}$  e  $B_{33}$  representam os canais HH, HV e VV, respectivamente. Como a matriz  $\Sigma$  é hermitiana, apenas os elementos do triângulo superior e da diagonal principal são apresentados. Essa matriz foi considerada por Frery, Jacobo-Berlles, Gambini & Mejail (2010) para representar áreas de florestas. Seu valor é resultado da média amostral dos pixels de uma região selecionada. Note que, para um valor de  $L$  fixado, quando o  $k$  aumenta, o coeficiente cresce e a “potência”, o brilho, aumenta. Por outro lado, o aumento de  $L$ , dado um  $k$  fixo, diminui o efeito do ruído sobre a imagem.

### 3.1.3 Estimação dos parâmetros usuais em PolSAR

Devido às suas propriedades assintóticas (tais como consistência e convergência *em distribuição* a uma lei gaussiana), os estimadores de MV foram empregados nos estudos com imagens reais e sintéticas desta tese. A seguir, apresenta-se uma discussão em torno das propriedades da função de log-verossimilhança da lei Wishart complexa escalonada e da estimação por MV para parâmetros os  $\Sigma$  e  $L$ .

Seja  $Y$  uma matriz aleatória que segue uma distribuição Wishart complexa escalonada com vetor de parâmetros  $\theta = [L, \text{vec}(\Sigma)^t]^t$  e  $\dim(\theta) = 1 + m^2$ , em que  $\text{vec}(\cdot)$  é operador de vetorização e  $\dim(\cdot)$  é o tamanho de um vetor. A função de log-verossimilhança de  $Y$  é dada por

$$\begin{aligned} \ell(\theta) = & mL \log L + (L - m) \log |Y| - L \log |\Sigma| \\ & - \frac{m(m-1)}{2} \log \pi - \sum_{k=0}^{m-1} \log \Gamma(L - k) - L \text{tr}(\Sigma^{-1}Y). \end{aligned} \quad (3.10)$$



**Figura 3.4** Imagens PolSAR (segundo a codificação de Pauli, tal como discutida no Apêndice A.1.1) simuladas conforme uma lei Wishart complexa escalonada com parâmetros  $L = \{3, 12\}$  e  $\Sigma = (1 + k)\mathbf{B}$  tal que  $k = \{0, 5, 10, 15\}$ .

Baseando-se no resultado  $\partial \ell(\boldsymbol{\theta}) / \partial \text{vec}(\Sigma) = \text{vec}(\partial \ell(\boldsymbol{\theta}) / \partial \Sigma)$  proposto por Hjørungnes & Gesbert (2007), as funções escore em termos de  $L$  e  $\text{vec}(\Sigma)$  são dadas, respectivamente, por

$$\ell_L = m(\log L + 1) + \log |\mathbf{Y}| - \log |\Sigma| - \psi_m^{(0)}(L) - \text{tr}(\Sigma^{-1}\mathbf{Y})$$

e

$$\ell_{\text{vec}(\Sigma)} = L \text{vec}(\Sigma^{-1}\mathbf{Z}\Sigma^{-1} - \Sigma^{-1}),$$

em que  $\psi_m^{(0)}(\cdot)$  representa o termo de ordem zero da função poligama multivariada de  $v$ -ésima ordem dada por  $\psi_m^{(v)}(L) = \sum_{i=0}^{m-1} \psi^{(v)}(L - i)$ ,  $\psi^{(v)}(\cdot)$  é a função poligamma ordinária dada por  $\psi^{(v)}(L) = \partial^{v+1} \log \Gamma(L) / \partial L^{v+1}$ , para  $v \geq 0$  (neste caso,  $\psi^{(0)}(\cdot)$  é conhecida como a função digamma). Estes resultados estão demonstrados no Apêndice C. A fim de complementar o estudo sobre as propriedades da função de log-verossimilhança da Wishart, o seguinte parágrafo apresenta uma discussão sobre a obtenção de três quantidades importantes: as matrizes de informação de Fisher observada ou Hessiana ( $\mathcal{J}(\boldsymbol{\theta})$ ) e esperada ( $\mathcal{K}(\boldsymbol{\theta})$ ) e o limite inferior de Cramér-Rao (LCR) viesado<sup>1</sup>  $\mathcal{C}(\boldsymbol{\theta})$ .

<sup>1</sup> Esta notação foi considerada por Anfinssen, Doulgeris & Eltoft (2009) para indicar a inversa da matriz informação de Fisher esperada associada a um estimador MV viesado como uma boa proposta de avaliar a performance dos estimadores. Este fato é justificável pelos seguintes argumentos.



Baseando-se nos resultados do Apêndice C, as expressões analíticas para  $\mathcal{J}(\boldsymbol{\theta})$ ,  $\mathcal{K}(\boldsymbol{\theta})$  e  $\mathcal{C}(\boldsymbol{\theta})$  são dadas por

$$\mathcal{J}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{m}{L} - \psi_m^{(1)}(L) & \text{vec}(\boldsymbol{\Sigma}^{-1}\mathbf{Y}\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1})^t \\ \text{vec}(\boldsymbol{\Sigma}^{-1}\mathbf{Y}\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1})^* & \mathcal{J}_{\Sigma\Sigma} \end{bmatrix},$$

$$\mathcal{K}(\boldsymbol{\theta}) = \mathbb{E}(-\mathcal{J}(\boldsymbol{\theta})) = \begin{bmatrix} \psi_m^{(1)}(L) - \frac{m}{L} & \text{vec}(\mathbf{0}_m)^t \\ \text{vec}(\mathbf{0}_m) & L\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1} \end{bmatrix}$$

e

$$\mathcal{C}(\boldsymbol{\theta}) = \mathcal{K}(\boldsymbol{\theta})^{-1} = \begin{bmatrix} [\psi_m^{(1)}(L) - \frac{m}{L}]^{-1} & \text{vec}(\mathbf{0}_m)^t \\ \text{vec}(\mathbf{0}_m) & L^{-1}\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma} \end{bmatrix},$$

em que  $\mathbf{0}_m$  é a matriz nula de ordem  $m$ ,

$$\mathcal{J}_{\Sigma\Sigma} = L(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1} \mathbf{Z} \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \mathbf{Z})$$

e  $\otimes$  representa o produto de Kronecker.

Anfinsen, Doulgeris & Eltoft (2009) derivaram o LCR viesado para a lei Wishart complexa original. Naquela abordagem, encontrou-se que os parâmetros desta distribuição são não ortogonais. Baseando-se sobre  $\mathcal{K}(\boldsymbol{\theta})$  da versão escalonada, nós concluímos que

✎  $L$  e  $\boldsymbol{\Sigma}$  se tornam ortogonais, isto é, tem matriz informação de Fisher bloco-diagonal (Para maiores detalhes, consultar Cox & Reid 1987) baseado no seguinte escalonamento: multiplicação da distribuição Wishart complexa original por um fator  $L^{-1}$ . Entre outros benefícios, este escalonamento torna as equações de verossimilhança separáveis, como mostrado a seguir. Este resultado será mais discutido e inserido no contexto da Teoria da Informação na Seção 5.5.

Considere  $\{\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N\}$  uma amostra iid de tamanho  $N$  obtida de  $\mathbf{Y} \sim \mathcal{W}(\boldsymbol{\Sigma}, L)$ .

É conhecido que LCR fornece uma limite inferior para a matriz de covariância do estimador  $\hat{\boldsymbol{\theta}}$ . No caso em que  $\boldsymbol{\theta}$  é um vetor de parâmetros complexos, Van den Bos (1994) propôs a seguinte expressão para o LCR:

$$\text{Cov}(\hat{\boldsymbol{\theta}}) \succeq \mathcal{K}^{-1}(\boldsymbol{\theta}) + \frac{\partial B(\boldsymbol{\theta})}{\partial \theta^t} \mathcal{K}^{-1}(\boldsymbol{\theta}) + \mathcal{K}^{-1}(\boldsymbol{\theta}) \left[ \frac{\partial B(\boldsymbol{\theta})}{\partial \theta^t} \right]^H + \frac{\partial B(\boldsymbol{\theta})}{\partial \theta^t} \mathcal{K}^{-1}(\boldsymbol{\theta}) \left[ \frac{\partial B(\boldsymbol{\theta})}{\partial \theta^t} \right]^H,$$

em que  $B(\cdot)$  é o viés do estimador e  $\mathbf{A} \succeq \mathbf{0}$  representa que a matriz  $\mathbf{A}$  é positiva semi definida. Note que o LCR se reduz a  $\mathcal{K}^{-1}(\boldsymbol{\theta})$  para estimadores não viesados. Como o estimador MV para o número de looks é viesado,  $\mathcal{K}^{-1}(\boldsymbol{\theta})$  não representa o LCR, mas pode ser entendida como uma boa forma de avaliar a performance de  $\hat{L}$ , sendo nominada por *LCR viesado*.

Uma vez que  $N^{-1} \sum_{k=1}^N \nabla \ell_k(\hat{\theta}) = \mathbf{0}_{1+m^2}$ , nós temos que

$$\hat{\Sigma} = \frac{1}{N} \sum_{k=1}^N \mathbf{Y}_k = \bar{\mathbf{Y}},$$

e

$$m \log \hat{L} + \frac{1}{N} \sum_{k=1}^N \log |\mathbf{Y}_k| - \log |\bar{\mathbf{Y}}| - \psi_m^{(0)}(\hat{L}) = 0. \quad (3.11)$$

Assim, o estimador para  $\Sigma$  é a média amostral, enquanto que  $\hat{L}$  é dado pela solução da equação não linear (3.11). O método iterativo de Newton-Raphson (Gentle 2002) será utilizado para obter as estimativas MV nos experimentos desta tese (Um exemplo de estimação na presença de correlação é dado por Vorobyov, Gershman & Wong 2005).



**Figura 3.5** Imagem E-SAR (canal HH) com regiões selecionadas.

✍ Ainda no contexto de estimação dos parâmetros da distribuição Wishart complexa escalonada, esta teste avança na proposta de dois métodos de estimação melhorados

para o parâmetro  $L$  na Seção 5.2. Os novos estimadores apresentam um viés de ordem  $O(N^{-2})$  e  $O(N^{-3/2})$  ( $O(\cdot)$  é a notação de Landau para representar ordem de magnitude) superando o relativo a versão não corrigida,  $O(N^{-1})$  (Cribari-Neto & Cordeiro 1996). Para pequenas amostras recomenda-se o uso de estimadores melhorados.

A fim de contextualizar este processo de estimação, vamos considerar uma aplicação a dados reais. A Figura 3.5 apresenta uma imagem SAR polarimétrica com seis áreas selecionadas obtidas utilizando a banda C-(5.3 GHz). Adotando um número nominal de looks  $L = 3$  em seu processamento, esta imagem foi obtida pelo sensor E-SAR (cujas características serão comentadas na Seção 5) visando analisar as regiões de Munique - Alemanha. As áreas selecionadas exibem dois tipos de cenários: (i) homogêneos (áreas de pasto, denominada por “pasture- $i$ ” para  $i=1,2,3$ ) e (ii) heterogêneos (florestas, “forest- $i$ ” para  $i=1,2,3$ ). A Tabela 3.1 mostra os resultados das estimativas MV bem como os tamanhos de amostra envolvidos. Neste caso, as diferenças dos tamanhos de amostras envolvidos não redundam em estimativas pontuais muito diferentes das apresentadas, em virtude do estimador MV se assintoticamente não viciado.

**Tabela 3.1** Estimativas dos parâmetros

| Regiões   | $\hat{L}$ | $ \hat{\Sigma} $ | N (# pixels) |
|-----------|-----------|------------------|--------------|
| pasture-1 | 1.005     | 784050.800       | 2106         |
| pasture-2 | 1.121     | 143.437          | 2352         |
| pasture-3 | 1.000     | 190.130          | 3340         |
| forest-1  | 1.003     | 3854056.000      | 2870         |
| forest-2  | 2.480     | 10752870.000     | 1020         |
| forest-3  | 1.050     | 14568.150        | 2496         |

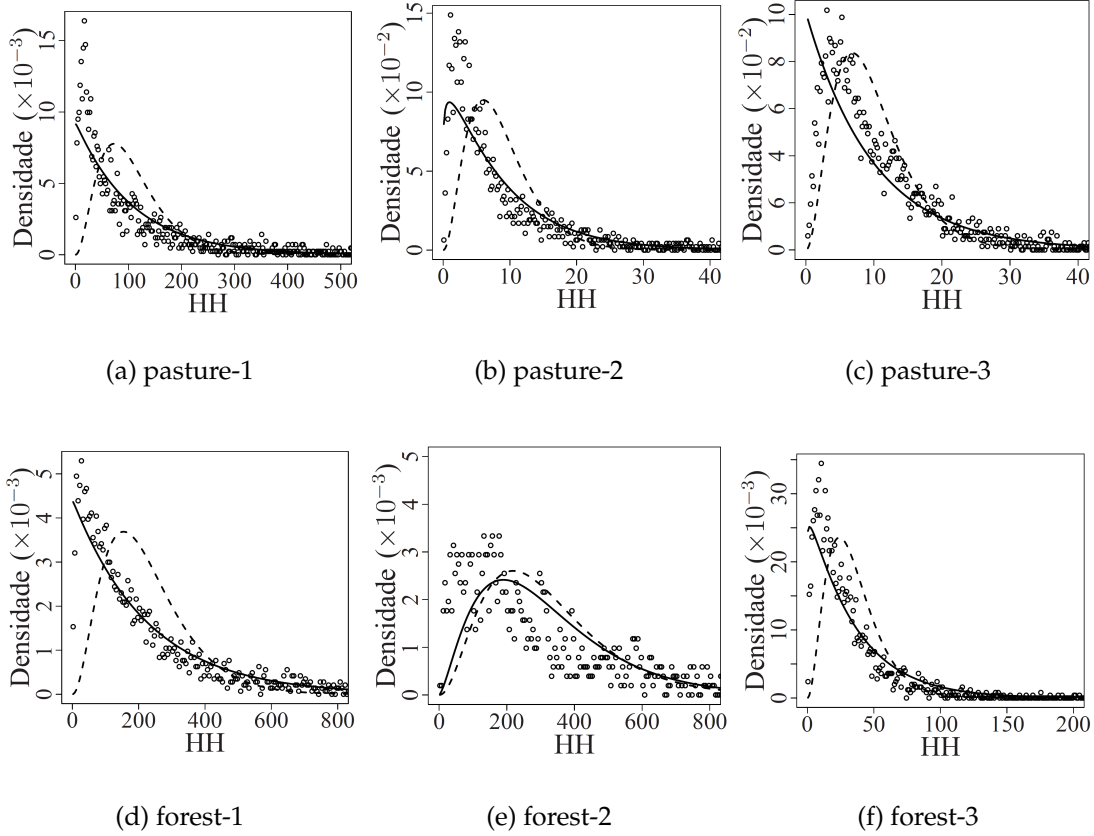
Como esperado da discussão apresentada por Anfinsen, Doulgeris & Eltoft (2009), as estimativas para  $L$  foram menores do que o número de looks pré-especificado  $L = 3$  e, para este caso, até menor do que o número de polarizações. Este fato sugere que o modelo estatístico  $\mathcal{W}(\Sigma, 3)$  é inadequado. Adicionalmente, a região “forest-2” apresentou maior estimativa para o número de looks em virtude de ser a menos contaminada de todas. Neste caso, Anfinsen, Eltoft & Doulgeris (2009) indicaram o relaxamento da restrição  $L \geq m$ , como uma possível solução. No modelo resultante, o parâmetro  $L$  é associado à forma da distribuição ao invés de ao número de looks. A distribuição Wishart escalonada resultante foi denominada como *Wishart relaxada*. A partir deste momento, a distribuição Wishart complexa escalonada  $\mathcal{W}_{\mathcal{R}}(\Sigma, L)$ , tal que  $L > 0$ , representará sua versão relaxada. Neste caso, vale-se a seguinte nota: Em algumas situações na prática (frequentemente em regiões muito texturadas e com número nominal de looks próximo de  $m$ ) o número de looks estimado pode ser menor do que a ordem da matriz,  $m$ . Para o tratamento destes casos, nós utilizaremos o ferramental derivado da Wishart considerando as seguintes substituições nas expressões resultan-

tes: as funções  $\Gamma(A)$  e  $\psi^{(v)}(A)$  assumirão os valores 1 e 0, respectivamente, quando  $A \leq 0$ . Estas atribuições preservam a positividade da forma da densidade.

Como complemento à discussão das estimativas, consideremos a visualização dos histogramas dos dados reais juntamente com as curvas das densidades ajustadas. Para tal fim, Hagedorn, Smith, Bones, Millane & Pairman (2006) provaram que a distribuição Wishart complexa escalonada se reduz à distribuição gamma com densidade dada por

$$f_{Z_i}(Z'_i; L/\sigma_i^2, L) = \frac{L^L Z_i'^{L-1}}{\sigma_i^{2L} \Gamma(L)} \exp(-L Z'_i / \sigma_i^2), \quad (3.12)$$

para  $i \in \{HH, HV, VV\}$ , em que  $\sigma_i^2$  é a entrada  $(i, i)$  de  $\Sigma$  e  $Z'_i$  é a entrada  $(i, i)$  da matriz  $Z$ . A Figura 3.6 exibe os ajustes das distribuições Wishart original,  $\mathcal{W}(\hat{\Sigma}, 3)$ , e relaxada,  $\mathcal{W}_{\mathcal{R}}(\hat{\Sigma}, \hat{L})$ , aos dados reais.



**Figura 3.6** Densidades empíricas e ajustadas para as distribuições Wishart relaxada (sólida) e original (tracejada longa) em função da intensidade do canal HH.

A fim de quantificar a comparação dos ajustes das distribuições  $\mathcal{W}(\hat{\Sigma}, 3)$  e  $\mathcal{W}_{\mathcal{R}}(\hat{\Sigma}, \hat{L})$ , nós obtivemos os valores do critério de informação de Akaike (*Akaike Information Criterion* - AIC). Esta medida de bondade de ajuste indica o melhor modelo como aquele

que apresenta o menor valor do AIC (Seghouane & Amari 2007). A Tabela 3.2 apresenta os AICs bem como a razão entre eles para cada região considerada. Em todos os

**Tabela 3.2** AICs para o canal HH das distribuições Wihsart relaxada e original sobre dados reais

| Regiões   | AIC                         |               | $\frac{AIC(\mathcal{W}_{\mathcal{R}})}{AIC(\mathcal{W})}$ |
|-----------|-----------------------------|---------------|-----------------------------------------------------------|
|           | $\mathcal{W}_{\mathcal{R}}$ | $\mathcal{W}$ |                                                           |
| pasture-1 | 23881.52                    | 27454.97      | 0,87                                                      |
| pasture-2 | 14885.62                    | 17345.38      | 0,86                                                      |
| pasture-3 | 21998.75                    | 22112.37      | 0,99                                                      |
| forest-1  | 36823.14                    | 40464.83      | 0,91                                                      |
| forest-2  | 14108.18                    | 14481.67      | 0,97                                                      |
| forest-3  | 22761.28                    | 25630.94      | 0,89                                                      |

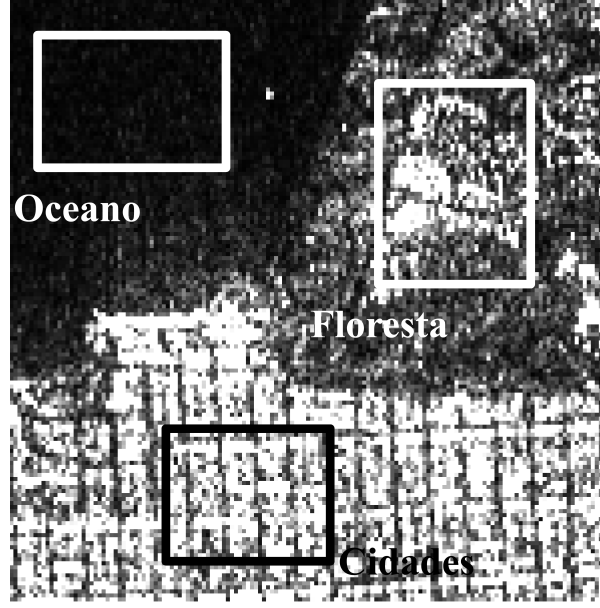
casos, a distribuição  $\mathcal{W}_{\mathcal{R}}$  apresentou o melhor desempenho.

### 3.2 Modelos multilook SAR bivariados

Teoricamente, as componentes co-polarizadas (seguindo uma mesma orientação linear, HH e VV) são correlatas, enquanto que as cros-polarizadas (por exemplo, HV e HH) tem correlação zero (Lee, Hoppel, Mango & Miller 1994). Na prática, tem-se que a correlação é mais intensa entre co-polarizações do que em outros casos. A fim de ilustrar este evento, consideremos uma área com  $150 \times 150$  pixels da imagem AIRSAR de São Francisco, como descrita no Apêndice A. A Figura 3.7 exibe esta área com três regiões selecionadas, representando oceano, floresta e cidades. A Tabela 3.3 apresenta os tamanhos das amostras utilizadas, # pixels (número de pixels), e os módulos dos valores médios das correlações co-polarizadas HH-VV,  $Z_{HH-VV}$ , e cros-polarizadas “HH-HV” e “HV-VV”,  $Z_{HH-HV}$  e  $Z_{HV-VV}$ . Os valores mais acentuadas estão formatados em negrito na tabela. Como esperado, os módulos das correlações entre co-polarizadas assumiram os maiores valores em todas as regiões. Em termos gerais, as correlações são mais intensas no cenário mais texturado, região de cidades. Esta rápida análise descritiva justifica o fato da maioria dos estudos bivariados estarem focados nas componentes co-polarizadas, HH e VV. (Para um estudo mais detalhado, consultar Ainsworth, Schuler & Lee 2008).

A seguinte discussão objetiva definir e comentar três principais resultados derivados por Lee, Hoppel, Mango & Miller (1994): funções densidade de probabilidade para (i) diferença de fases, (ii) magnitude do produto complexo e (iii) razões intensidade e amplitude entre os componentes HH e VV do vetor de retroespalhamento (3.2).

Primeiramente, considere a definição do coeficiente de correlação complexo ( $\rho_c$ ), conceito sobre o qual estarão baseados os principais resultados desta seção. Sejam  $S_i$



**Figura 3.7** Imagem AIRSAR de São Francisco (canal HH) com três áreas selecionadas.

**Tabela 3.3** Módulos das correlações médias entre os canais de polarizações

| Regiões  | Módulo das correlações médias ( $\times 10^{-4}$ ) |                          |                          | # pixels |
|----------|----------------------------------------------------|--------------------------|--------------------------|----------|
|          | $ \overline{Z_{HH-HV}} $                           | $ \overline{Z_{HH-VV}} $ | $ \overline{Z_{HV-VV}} $ |          |
| Oceano   | 9.89599                                            | <b>117.25436</b>         | 18.62410                 | 1632     |
| Floresta | 149.13040                                          | <b>771.21780</b>         | 33.01252                 | 1989     |
| Cidades  | 1291.7835                                          | <b>1307.4982</b>         | 655.9329                 | 1428     |

e  $S_j$  dois elementos do vetor de retroespalhamento (3.2). O parâmetro  $\rho_c$  é definido como

$$\rho_c = \frac{E(S_i S_j^*)}{\sqrt{E(|S_i|^2) E(|S_j|^2)}} = |\rho_c| \exp(i\theta), \quad (3.13)$$

em que  $\theta$  representa o ângulo de incidência,  $|\rho_c|$  representa a amplitude do coeficiente de correlação e  $i$  é a unidade imaginária. Além disso, obtida diretamente da equivalência  $\rho_I^{(L)} = |\rho_c|^2$ , a versão  $L$ -looks da intensidade do coeficiente de correlação ( $\rho_I^{(L)}$ ) é dada por

$$\rho_I^{(L)} = \frac{E((S_{ii} - \bar{S}_{ii})(S_{jj} - \bar{S}_{jj}))}{\sqrt{E((S_{ii} - \bar{S}_{ii})^2) E((S_{jj} - \bar{S}_{jj})^2)}}.$$

A seguir, definem-se as principais densidades dependentes dos parâmetros  $L$  e  $\rho_c$  para dados SAR bivariados.

### 3.2.1 Densidade para a diferença de fase $\psi$

Considere que temos a finalidade de descrever probabilisticamente a diferença de fase para um dos dois componentes HH ou VV do vetor de retroespalhamento (3.2). Para processamentos com número de looks 1 e  $L$ , as diferenças de fase são definidas, respectivamente, por:

$$\psi_1 = \text{Arg}(S_i S_j^*)$$

e

$$\psi_L = \text{Arg}\left[\frac{1}{L} \sum_{k=1}^L S_i(k) S_j^*(k)\right].$$

A quantidade  $\psi_L$  representa um argumento de um elemento fora da diagonal principal da matriz (3.6).

Buscando modelos para descrever a diferença de fase a partir dos parâmetros  $L$  e  $\rho_c$ , a submatriz  $\{\mathbf{Y}\}_{i,j=1,2}$  da equação (3.6) pode ser reparametrizada por

$$\mathbf{Y} = \begin{bmatrix} Y_{11} & \alpha \exp(\mathbf{i}\psi) \\ \alpha \exp(-\mathbf{i}\psi) & Y_{22} \end{bmatrix}$$

e sua matriz de covariância por

$$\mathbf{\Sigma} = \begin{bmatrix} \Sigma_{11} & \sqrt{\Sigma_{11}\Sigma_{22}}|\rho_c| \exp(\mathbf{i}\theta) \\ \sqrt{\Sigma_{11}\Sigma_{22}}|\rho_c| \exp(-\mathbf{i}\theta) & \Sigma_{22} \end{bmatrix}, \quad (3.14)$$

em que  $\Re[Y_{12}] + \mathbf{i}\Im[Y_{12}] = \alpha \exp(\mathbf{i}\psi)$  e  $\Sigma_{ii} = E(|Y_{ii}|^2)$ .

A partir deste ponto, (i) reparametrizando as variáveis  $Y_{11}$ ,  $Y_{12}$  e  $\alpha$  em termos de elementos da matriz  $\mathbf{\Sigma}$ ,

$$B_1 = \frac{Y_{11}}{\Sigma_{11}}, \quad B_2 = \frac{Y_{22}}{\Sigma_{22}} \quad \text{e} \quad \eta = \frac{\alpha}{\sqrt{\Sigma_{11}\Sigma_{22}}}, \quad (3.15)$$

e (ii) substituindo o vetor  $[Y_{11}, Y_{22}, \Re[Y_{12}], \Im[Y_{12}]]$  por  $[B_1, B_2, \eta, \psi]$ , Lee, Miller & Hopel (1994) derivaram que a equação (3.7) se reduz a

$$f_{\mathbf{Y}}(B'_1, B'_2, \eta', \psi'; \theta, \rho_c, L) = \frac{(B'_1 B'_2 - \eta'^2)^{L-2} \eta'}{\pi(1 - |\rho_c|^2)^2 \Gamma(L) \Gamma(L-1)} \times \exp\left(-\frac{B'_1 + B'_2 - 2\eta'|\rho_c| \cos(\psi' - \theta)}{(1 - |\rho_c|^2)}\right). \quad (3.16)$$

Finalmente, integrando esta fórmula em termos das variáveis  $B'_1$ ,  $B'_2$  e  $\eta'$ , a densidade marginal da diferença de fase  $\psi$  é dada por

$$f_{\psi_L}(\psi'; \beta, \rho_c, L) = \frac{\Gamma(L + 1/2)(1 - |\rho_c|^2)^L \beta}{2\sqrt{\pi}\Gamma(L)(1 - \beta^2)^{L+1/2}} + \frac{(1 - |\rho_c|^2)^L}{2\pi} {}_2F_1(L, 1; 1/2; \beta^2),$$

em que  $\beta = |\rho_c| \cos(\psi' - \theta)$ , para  $-\pi < \psi' < \pi$ , e  ${}_2F_1(\cdot)$  representa a função hipergéométrica.

✎ A Seção 5.3 faz uma contribuição propondo estatísticas dependentes do coeficiente de correlação complexo e do número de equivalente de looks para a reparametrização de Lee, Hoppel, Mango & Miller. Naquela seção, o contexto será o de distinguir áreas PolSAR a partir da informação de seus coeficientes de correlação complexos.

### 3.2.2 Densidade para a magnitude do produto $|S_i S_j^*|$

Adicionalmente à subseção anterior, é importante estudar a magnitude do produto entre  $S_i$  e  $S_j$ . Uma normalização usual para este termo é dada por

$$\epsilon = \frac{|L^{-1} \sum_{k=1}^L S_i(k) S_j^*(k)|}{\sqrt{E(|S_i|^2) E(|S_j|^2)}} = \frac{g}{h},$$

em que  $g$  é o produto multilook  $|S_i S_j^*|$ . Integrando a expressão (3.16) em termos das variáveis  $B_1$ ,  $B_2$  e  $\psi$ , a densidade para  $\epsilon$  é derivada,

$$f_\epsilon(\epsilon'; \rho_c, L) = \frac{4L^{L+1} \epsilon'^L}{\Gamma(L)(1 - |\rho_c|^2)} I_0\left(\frac{2|\rho_c|L\epsilon'}{1 - |\rho_c|^2}\right) K_{L-1}\left(\frac{2L\epsilon'}{1 - |\rho_c|^2}\right),$$

em que  $I_A(\cdot)$  e  $K_B(\cdot)$  são as funções modificadas de Bessel do primeiro e segundo tipo com ordens  $A$  e  $B$  cujas expressões podem ser obtidas por Abramowitz & Stegun (1964), respectivamente. Particularmente, Hu (2005) apresentou propriedades analíticas interessantes para  $K_B(\cdot)$  em termos de processamento de imagem PolSAR. Para tal fim, eles utilizaram a seguinte expressão:

$$K_\alpha(V_1) = \frac{1}{2} \int_0^\infty V^{\alpha-1} \exp\left(-\frac{V_1}{2}(V + V^{-1})\right) dV, \quad (3.17)$$

com  $V_1$  e  $V$  sendo variáveis reais positivas. Finalmente, a densidade de  $g$  é obtida por  $f_g(g'; \cdot) = f_\epsilon(g'/h; \cdot)$ .

### 3.2.3 Densidade conjunta para o par de intensidades $(|S_i|^2, |S_j|^2)$

A derivação da distribuição conjunta é importante em dois sentidos: (i) quando dados polarimétricos incompletos são considerados (como, por exemplo, em dados do processador sinótico AIRSAR JPL, vide Lee, Hoppel, Mango & Miller 1994) e, em um fim



teórico, (ii) quando se deseja obter distribuições marginais (a exemplo de Hagedorn et al. 2006, Dharmawansa & McKay 2009). Nesta seção, consideramos a tarefa de obter a distribuição conjunta para as quantidades

$$R_1 = \frac{1}{L} \sum_{k=1}^L |S_i(k)|^2 = \frac{B_1 C_{11}}{L} \text{ e } R_2 = \frac{1}{L} \sum_{k=1}^L |S_j(k)|^2 = \frac{B_2 C_{22}}{L}. \quad (3.18)$$

Com esse objetivo, integrando (3.16) com respeito às variáveis  $\eta'$  e  $\psi'$ , obtemos

$$f_{B_1, B_2}(B'_1, B'_2; \rho_c, L) = \frac{(B'_1 B'_2)^{(L-1)/2} \exp\left(-\frac{B'_1 + B'_2}{1 - |\rho_c|^2}\right)}{\Gamma(L)(1 - |\rho_c|^2)|\rho_c|^{L-1}} I_{L-1}\left(2\sqrt{B'_1 B'_2} \frac{|\rho_c|}{1 - |\rho_c|^2}\right).$$

Aplicando (3.18) na expressão acima temos que

$$\begin{aligned} f_{R_1, R_2}(R'_1, R'_2; \rho_c, L) &= \frac{(R'_1 R'_2)^{(L-1)/2} L^{L+1} \exp\left(-\frac{L(R'_1/\Sigma_{11} + R'_2/\Sigma_{22})}{1 - |\rho_c|^2}\right)}{(\Sigma_{11} \Sigma_{22})^{(L+1)/2} \Gamma(L)(1 - |\rho_c|^2)|\rho_c|^{L-1}} \\ &\quad \times I_{L-1}\left(2L \sqrt{\frac{R'_1 R'_2}{\Sigma_{11} \Sigma_{22}}} \frac{|\rho_c|}{1 - |\rho_c|^2}\right). \end{aligned}$$

### 3.2.4 Densidades para as razões entre intensidades ( $|S_i|^2/|S_j|^2$ ) e amplitudes ( $|S_i|/|S_j|$ )

Considere a seguinte razão normalizada das quantidades estócasticas  $B_1$  e  $B_2$  dadas em (3.15):

$$\mu = B_1/B_2 = \frac{\sum_{k=1}^L |S_i(k)|^2 / \Sigma_{11}}{\sum_{k=1}^L |S_j(k)|^2 / \Sigma_{22}}.$$

A seguinte discussão considera a derivação da densidade de  $\tau\mu$ , em que  $\tau = \Sigma_{11}/\Sigma_{22}$ . Este termo tem um papel importante em discriminação nos contextos de retornos polarimétricos.

Baseando-se nos resultados (i)  $f_\mu(\mu') = \int_0^\infty B_2 f_{B_1, B_2}(\mu' B_2, B_2) dB_2$  e, da identidade de integração de Prudnikov, Brychkov & Maichev (1986, p. 303, eq. 21), (iii)

$$\begin{aligned} \int_0^\infty B_2^L \exp\left(-\frac{B_2(1+\mu)}{1 - |\rho_c|^2}\right) I_{L-1}\left(2B_2 \sqrt{\mu} \frac{|\rho_c|}{1 - |\rho_c|^2}\right) dB_2 &= \\ \left(\frac{1+\mu}{1 - |\rho_c|^2}\right)^{-2L} \left(\frac{\sqrt{\mu} |\rho_c|}{1 - |\rho_c|^2}\right)^{L-1} \frac{\Gamma(2L)}{\Gamma(L)} {}_2F_1\left(L, (2L+1)/2; L; \frac{4|\rho_c|^2 \mu}{(1+\mu)^2}\right), \end{aligned}$$

Lee, Hoppel, Mango & Miller (1994) provaram que as densidades para intensidade e

amplitude multilook de  $\mu$  são dadas, respectivamente, por

$$f_{\mu}(\mu'; \rho_c, L) = \frac{\Gamma(2L)(1 - |\rho_c|^2)^L(1 + \mu')\mu'^{L-1}}{\Gamma(L)\Gamma(L)[(1 + \mu')^2 - |\rho_c|^2\mu']^{(2L+1)/2}}$$

e (definindo  $v = \sqrt{\mu}$ )

$$f_v(v'; \rho_c, L) = \frac{2\Gamma(2L)(1 - |\rho_c|^2)^L(1 + v'^2)v'^{2L-1}}{\Gamma(L)\Gamma(L)[(1 + v'^2)^2 - 4|\rho_c|^2v'^2]^{(2L+1)/2}}.$$

Assim, as densidades relativas aos termos  $w = \tau\mu$  e  $z = \sqrt{w}$  são dadas por  $f_w(w'; \rho_c, L) = \tau^{-1}f_{\mu}(w'/\tau; \rho_c, L)$  e  $f_z(z'; \rho_c, L) = 2z'f_w(z'^2; \rho_c, L)$ .

### 3.3 Modelos multiplicativos multilook PolSAR

Seja  $\mathbf{Y}$  uma matriz aleatória seguindo uma distribuição Wishart complexa escalonada,  $\mathbf{y}$  um vetor aleatório com distribuição gaussiana complexa com o vetor de médias nulo e  $X$  uma variável aleatória escalar positiva. Além disso, considere que a variável  $X$  é independente de ambas as variáveis  $\mathbf{Y}$  e  $\mathbf{y}$ . A *modelagem multiplicativa* para dados PolSAR assume que cada pixel segue uma das seguintes definições:

- Para dados multilook: A variável retorno é modelada segundo  $\mathbf{Z} = \mathbf{Y} \times X$ .
- Para dados com número de looks 1: O pulso retornado é descrito por  $\mathbf{z} = \mathbf{y} \times \sqrt{X}$ . Essa abordagem é conhecida como modelo para um *vetor aleatório esfericamente invariante* (Spherically Invariant Random Vector - SIRV) (Bombrun, Vassile, Gay & Totir 2011), que foi introduzido por Yao (1973). Neste caso, como o processo multilook representa o quadrado do modelo para 1-look em termos escalares, então a parte escalar desta modelagem é representada por  $\sqrt{X}$ .

Neste caso, as variáveis  $\mathbf{Y}$  e  $\mathbf{y}$  são destinadas a descrever o ruído speckle em cenários  $L$ -looks e 1-look, respectivamente; enquanto que  $X$  modela o comportamento do relevo pelo retroespalhamento dos pulsos emitidos (na língua inglesa, esta propriedade é sintetizada no termo *backscatter*). É possível notar ainda a seguinte relação entre os retornos de  $L$ -looks e 1-look: Considere um vetor com  $L$  vetores SIRV com tamanhos  $m$ ,  $\mathbf{v} = [z(1) z(2) \cdots z(L)]^t$ . Assim, o seguinte resultado é decorrente do produto  $\mathbf{v}^H \mathbf{v}$ :

$$\mathbf{v}^H \mathbf{v} = \frac{1}{L} \sum_{\ell=1}^L \mathbf{z}(\ell) \cdot \mathbf{z}(\ell)^H = \frac{X}{L} \sum_{\ell=1}^L \mathbf{y}(\ell) \cdot \mathbf{y}(\ell)^H = X \cdot \mathbf{Y} = \mathbf{Z},$$

em que  $\mathbf{z}(\ell)$  e  $\mathbf{y}(\ell)$  representam o  $\ell$ -ésimo retorno do vetor  $\mathbf{v}$  e seu modelo associado ao ruído, respectivamente. Adicionalmente, tem-se que a modelagem multiplicativa se está fortemente relacionada com a formação física de uma imagem PolSAR através da equação de radar dada pela equação (2.1) na página 19.

### 3.3.1 Ruído PolSAR

Sistemas PolSAR são afetados pelo ruído speckle. Diferentemente do sistema com apenas uma polarização, este ruído afeta tanto as intensidades (elementos das diagonais principais das matrizes que representam os pixels) como suas covariâncias (López-Martínez & Fabregas 2003, López-Martínez et al. 2005). Várias técnicas têm sido propostas com o fim de reduzir o ruído em imagens PolSAR. López-Martínez & Fabregas (2003) consideram que estas técnicas podem ser divididas em dois grupos: Um primeiro trabalha combinando canais de forma a ter uma imagem livre de ruído (Novak & Burl 1990, Lopes & Sery 1997). Estas técnicas mantêm a resolução espacial, mas perdem as informações polarimétricas. Uma segunda abordagem (sobre a qual se baseia este trabalho) busca tratar o ruído mantendo as informações dos canais e de suas covariâncias (López-Martínez & Pottier 2007). Neste sentido, esta tese considera que as variáveis  $\mathbf{Y}$  e  $\mathbf{y}$  são bem representadas pelas distribuições Wishart escalonada e complexas gaussiana (tal como discussão da Seção 3.1) para cenários  $L$ -looks e 1-look, respectivamente.

### 3.3.2 Retroespalhamento (backscatter) PolSAR

O backscatter carrega toda a informação relevante de uma cena mapeada; isto é, ele considera propriedades físicas de uma área sob estudo, tais como a umidade e relevo do terreno. O “número de elementos difusores” em uma célula de resolução pode ser modelado por uma distribuição de Poisson cuja média (número esperado de scatters) é geralmente modelada por um dentre os seguintes elementos no *sistema de distribuições de Pearson* (Delignon & Pieczynski 2002):

- (a) backscatter *gama* distribuído,  $X \sim \Gamma(\mathcal{L}, \mathcal{L}\mu^{-1})$ , com função densidade dada por

$$f_X(X'; \mu, \mathcal{L}) = \frac{\mathcal{L}}{\mu} \frac{1}{\Gamma(\mathcal{L})} \left( \frac{\mathcal{L}}{\mu} X' \right)^{\mathcal{L}-1} \exp\left(-\frac{\mathcal{L}}{\mu} X'\right) \mathbb{I}_{(0,\infty)}(X'), \quad (3.19)$$

em que  $\mu > 0$  e  $\mathcal{L} > 0$  são parâmetros de escala e forma, respectivamente, e  $\mathbb{I}_A(\cdot)$  representa a função indicadora no suporte  $A$ .

- (b) backscatter com distribuição *gama inversa*,  $X \sim \Gamma^{-1}(-\alpha, (-\alpha - 1)\mu)$  cuja densidade é

$$f_X(X'; \mu, \alpha) = \frac{1}{\Gamma(-\alpha)} \frac{1}{(-\alpha - 1)\mu} \left[ \frac{(-\alpha - 1)\mu}{X'} \right]^{\alpha+1} \exp\left(-\frac{(-\alpha - 1)\mu}{X'}\right) \mathbb{I}_{(0,\infty)}(X'), \quad (3.20)$$

em que  $\mu > 0$  (escala) e  $\alpha < 0$  (forma).

- (c) backscatter *beta do primeiro tipo* distribuído (Bombrun, Anfinen & Harant 2011),  $X \sim \text{beta}(\mathcal{L}, -\alpha - \mathcal{L}, (-\alpha - 1)\mu\mathcal{L}^{-1})$ , com densidade

$$f_X(X'; \mu, \mathcal{L}, w) = \frac{\mathcal{L}}{w\mu} \frac{\Gamma(w)}{\Gamma(\mathcal{L})\Gamma(w - \mathcal{L})} \left[ \frac{\mathcal{L}X'}{w\mu} \right]^{\mathcal{L}-1} \left[ 1 - \frac{\mathcal{L}X'}{w\mu} \right]^{w-\mathcal{L}-1} \mathbb{I}_{(0, \frac{w\mu}{\mathcal{L}})}(X'), \quad (3.21)$$

em que  $w > \mathcal{L}$  e os parâmetros  $\alpha$  e  $\mathcal{L}$  representam formas, enquanto que  $\mu$  está vinculado a escala. Essa distribuição é, também, conhecida como *Pearson do tipo I* (Delignon & Pieczynski 2002).

- (d) backscatter com distribuição *beta inversa*,  $X \sim \text{beta}^{-1}(-w + 1, w - \mathcal{L}, \mu\mathcal{L}w^{-1})$ , com densidade dada por

$$f_X(X'; \mu, \mathcal{L}, w) = \frac{w}{\mathcal{L}\mu} \frac{\Gamma(w)}{\Gamma(\mathcal{L})\Gamma(w - \mathcal{L})} \left[ \frac{wX'}{\mathcal{L}\mu} \right]^{-w} \left[ \frac{wX'}{\mathcal{L}\mu} - 1 \right]^{w-\mathcal{L}-1} \mathbb{I}_{(\frac{\mathcal{L}\mu}{w}, \infty)}(X'), \quad (3.22)$$

em que  $w \geq \mathcal{L}$ ,  $w$  e  $\mathcal{L}$  representam formas, enquanto  $\mu$  é um parâmetro de escala. Neste caso, ela também é chamada de *Pearson do tipo IV* (Delignon & Pieczynski 2002).

Com o objetivo de generalizar os modelos estatísticos para descrever o backscatter, dois modelos têm sido considerados na literatura de PolSAR. Tais distribuições são discutidas a seguir.

Freitas et al. (2005) propuseram uma variável aleatória de média unitária obedecendo a distribuição gaussiana inversa generalizada (GIG) como um modelo abrangente para o backscatter. Por questões de generalidade analítica, esta tese considera a mesma distribuição sem a restrição de ter uma média unitária. Assim, é possível representar, a partir desta e de outras distribuições para o backscatter com média diferente de um, tanto (i) modelos inclusos na abordagem multiplicativa polarimétrica (que necessitam ter média unitária) quanto (ii) distribuições para um canal de polarização que não supõem esta restrição. Assim, a versão da GIG considerada,  $X \sim \mathcal{N}^{-1}(\mu, \varsigma, w)$ , tem densidade dada por

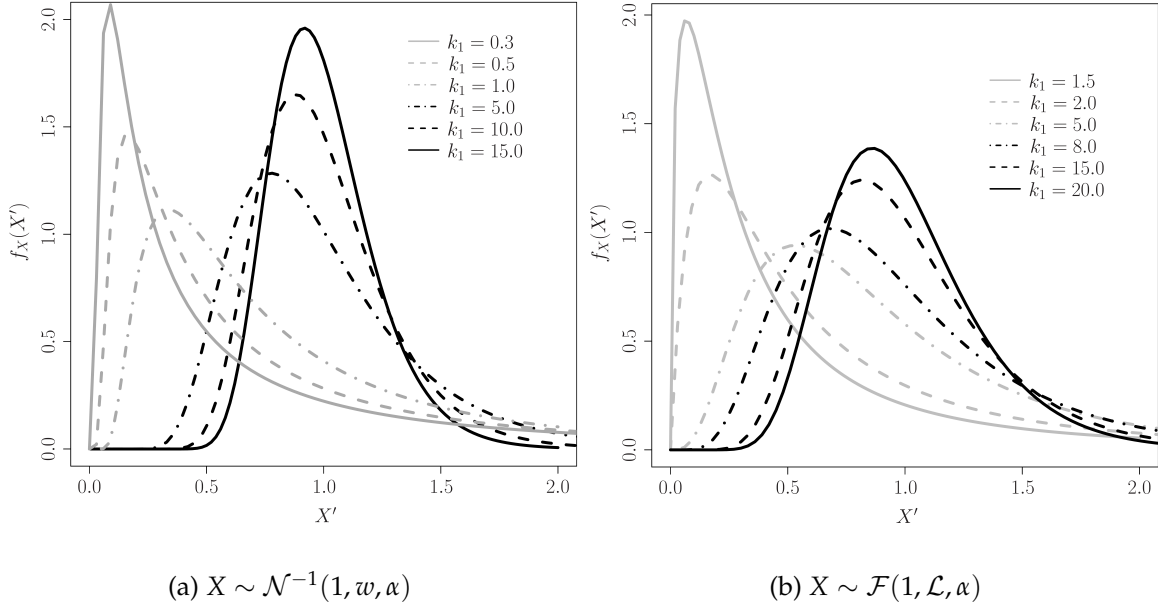
$$f_X(X'; \mu, \varsigma, w) = \frac{X'^{\varsigma-1} r_{\varsigma, w}^{\varsigma}}{2\mu^{\varsigma} K_{\varsigma}(w)} \exp\left(-\frac{w}{2} \left( \frac{\mu}{r_{\varsigma, w} X'} + \frac{r_{\varsigma, w} X'}{\mu} \right)\right) \mathbb{I}_{(0, \infty)}(X'), \quad (3.23)$$

em que  $r_{\varsigma, w} = K_{\varsigma+1}(w)/K_{\varsigma}(w)$  e  $-\infty < \varsigma, w < \infty$  (tal que  $\varsigma \notin [-1, 0)$ ) são parâmetros de forma e  $\mu$  controla a escala. Esta distribuição é uma generalização da distribuição gaussiana inversa (GI) cuja parametrização utilizada por Frery, Jacobo-Berlles, Gambini & Mejail (2010) (como modelo para o backscatter na proposta de uma nova distribuição polarimétrica) é dada por

$$f_X(X'; w, \mu) = \sqrt{\frac{w\mu}{2\pi X'^3}} \exp\left(-\frac{w}{2} \frac{(X' - \mu)^2}{X'\mu}\right) \mathbb{I}_{(0, \infty)}(X'),$$



fluência parecida nos comportamentos das densidades, embora com valores distintos. Na Figura 3.9(a), considerando a restrição  $w = -\alpha = k_1$ , quando  $k_1 \uparrow 1$  as densidades  $\mathcal{N}^{-1}(1, w, \alpha)$  se tornam mais achatadas (caracterizando eventos de maior variância), enquanto que para  $k_1 > 1$  e  $k_1 \uparrow \infty$  tais densidades tendem a ser mais pontiagudas. Para a distribuição  $\mathcal{F}(1, \mathcal{L}, \alpha)$  com  $w = -\alpha = k_1$ , as densidades mais achatadas são obtidas quando  $k_1 \uparrow 5$ , enquanto que as densidades que representam eventos de menores variâncias são obtidas quando, para  $k_1 \geq 8$ ,  $k_1$  aumenta.



**Figura 3.9** Curvas para os modelos backscatter mais gerais,  $\mathcal{F}$  e GIG, assumindo médias unitárias e parâmetros de escala e forma  $-\alpha = \mathcal{L} = w \in \{0.3, 0.5, 1, 5, 8, 10, 15, 20\}$ .

### 3.3.3 Distribuições dos Retornos PolSAR

A seguir nós apresentaremos os principais modelos polarimétricos para descrever o pulso retornado.

Considerando o modelo  $\mathbf{Z} = \mathbf{Y} \times X$ , Freitas et al. (2005) bem como Harant, Bombrun, Gay, Fallourd, Trouvã & Tupin (2009) desenvolveram novas propostas e estudos sobre distribuições polarimétricas. Estas abordagens trabalham aplicando os modelos para o ruído speckle e para o backscatter dentro da seguinte *convolução de Mellin* (Epstein 1948),

$$f_{\mathbf{Z}}(\mathbf{Z}') = \int_0^\infty X'^{-m^2} f_{\mathbf{Y}}(\mathbf{Z}' / X') f_X(X') dX', \quad (3.25)$$

em que  $m$  é o número de canais de polarizações. Bombrun, Vasile, Gay, Pascal &

Ovarlez (2010) também mostraram evidências que essa expressão é degenerada em

$$f_z(z') = \int_0^\infty X'^{-m} f_y(z' / X') f_X(X') dX', \quad (3.26)$$

quando cenários 1-look são considerados (Para maiores detalhes, consultar Lefebvre 2007). Ambas as abordagens,  $L$ -looks ou 1-look, podem ser justificadas pelos seguintes argumentos:

$$\begin{aligned} f_Z(\mathbf{Z}') &= \int_0^\infty f_{Z,X}(\mathbf{Z}', X') dX' = \int_0^\infty f_{Z|X}(\mathbf{Z}' | X') f_X(X') dX' \\ &= \int_0^\infty f_{Y \cdot X'}(\mathbf{Z}') f_X(X') dX'. \end{aligned} \quad (3.27)$$

Seguindo uma ideia similar, para modelos 1-look, temos que

$$f_z(z') = \int_0^\infty f_{y \cdot X'}(z') f_X(X') dX'. \quad (3.28)$$

Conforme resultados sobre *jacobiano de transformações matriciais* apresentados por Mathai (1997), temos que  $f_{Y \cdot X'}(\mathbf{Z}') = |\mathbf{J}_{Y \cdot X' \rightarrow Y}| f_Y(\mathbf{Z}' / X')$  e  $f_{y \cdot X'}(z') = |\mathbf{J}_{y \cdot X' \rightarrow y}| f_y(z' / X')$ , em que  $|\mathbf{J}_{Y \cdot X' \rightarrow Y}| = X'^{-m^2}$  e  $|\mathbf{J}_{y \cdot X' \rightarrow y}| = X'^{-m}$ . Portanto (3.27) e (3.28) resultam em (3.25) e (3.26), respectivamente.

Dentro dos argumentos acima e checando analiticamente as expressões das densidades polarimétricas disponíveis na literatura, entendemos que uma nova e elegante metodologia de representar os modelos polarimétricos  $L$ -looks e 1-look é definida a seguir:

- Para dados multilook:

$$f_Z(\mathbf{Z}'; \boldsymbol{\theta}, L) = \frac{|\mathbf{Z}'|^{L-m}}{\Gamma_m(L)} h_{\mathcal{M}}(\boldsymbol{\theta}, L) g_{\mathcal{M}}(\text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}'); \boldsymbol{\theta}, L), \quad (3.29)$$

em que  $\boldsymbol{\theta}$  representa o vetor de parâmetros adicionais ao número de looks,  $L$ . A Tabela 3.5 lista as várias funções  $h_{\mathcal{M}}(\cdot, \cdot)$  e  $g_{\mathcal{M}}(\cdot; \cdot, \cdot)$  relacionadas às suas distribuição polarimétricas. Esta expressão será utilizada no Capítulo 4 e no Apêndice E para derivar ferramentas da Teoria da Informação (distâncias e entropias) da forma mais abrangente possível. Desta forma, as distribuições particulares, mesmo quando não analiticamente tratáveis, poderão ser colocadas num formato apropriado a derivar aproximações e limites inferiores (como, por exemplo, o limite inferior para a distância de Kullback-Leibler da distribuição  $\mathcal{G}_m^0$  derivado no Apêndice E) das medidas.

- Para 1-look:

$$f_z(z'; \boldsymbol{\theta}, 1) = \frac{1}{\pi^m} h_{\mathcal{M}}(\boldsymbol{\theta}, 1) g_{\mathcal{M}}(z'^H \boldsymbol{\Sigma}^{-1} z'; \boldsymbol{\theta}, 1). \quad (3.30)$$

Neste caso, as expressões são obtidas também pela Tabela 3.5, considerando  $L = 1$  e substituindo o termo “ $\text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}')$ ” por “ $z'^H \boldsymbol{\Sigma}^{-1} z'$ ” em (3.30).

A seguir, as composições sobre cada distribuição da Tabela 3.5 bem como de suas versões 1-look são consideradas. Nós procuramos dividir as distribuições polarimétricas em três classes: *Simples* para as distribuições com dois e três parâmetros, *Moderadas* para aquelas com quatro parâmetros e dois casos particulares e as *Amplas* com quatro parâmetros e mais do que dois casos particulares.

### 3.3.3.1 Classes simples:

Inicialmente, consideremos três distribuições para o backscatter: a distribuição de Dirac, isto é, uma variável degenerada tal que  $\Pr(X = 1) = 1$ ; a distribuição gamma, com densidade (3.19) dado  $\mu = 1$ ; e a lei gama inversa, de densidade equação (3.20) com  $\mu = 1$ . Assuma também  $\mathcal{W}(\Sigma, L)$  e  $\mathcal{N}_m^c(\Sigma, 0)$  como modelos para o ruído speckle. Frery et al. (2007) apresentaram evidências que sugeriam como modelos polarimétricos a convolução dos três modelos para o backscatter com a distribuição Wishart complexa escalonada para diferentes graus de textura. Estes modelos juntamente com suas versões 1-look são comentados a seguir.

**Cenários homogêneos:** Estas situações geralmente são representadas por regiões de oceanos e pastos. As versões  $L$ -looks e 1-look de modelos para dados homogêneos são bem representadas pelas convoluções

$$\begin{cases} \mathcal{W}(\Sigma, L) = \delta(t - 1) \hat{*} \mathcal{W}(\Sigma, L), \\ \mathcal{N}_m^c(\Sigma, 0) = \delta(t - 1) \hat{*} \mathcal{N}_m^c(\Sigma, 0), \end{cases}$$

em que  $\delta(\cdot)$  representa a função *delta de Dirac* e  $\hat{*}$  representa a convolução de Mellin. Bombrun, Vasile, Gay & Totir (2011) utilizaram  $\hat{*}$  para representar a convolução de Mellin entre duas distribuição de probabilidade.

Neste caso, sugere-se que os retornos multilook e 1-look sejam modelados pelas próprias distribuições do ruído speckle. Hagedorn et al. (2006) mostraram evidências de que a distribuição Wishart complexa escalonada se degenera numa distribuição gamma (conforme discussão na Seção 3.1). Em termos 1-look, não é difícil verificar que a gaussiana complexa se degenera em um escalonamento da gaussiana complexa padrão,  $m = 1$ . Neste caso ( $m = 1$ ), a expressão  $\mathbf{y}'^H \Sigma^{-1} \mathbf{y}'$  se reduz a  $|Y'_i|^2 / \Sigma_{ii}$ , para  $i = 1, 2, \dots, m$ , em que  $\Sigma_{ii}$  é a entrada  $(i, i)$  da matriz  $\Sigma$  e  $Y'_i$  representa o  $i$ -ésimo elemento do vetor complexo  $\mathbf{y}'$ . Assim, tomando  $\mu_i = \Sigma_{ii}$ , a densidade (3.5) se reduz à seguinte densidade:

$$f_{Y_i}(Y'_i; \mu_i) = \frac{1}{(\pi \mu_i)} \exp(-|Y'_i|^2 / \mu_i).$$

Finalmente, é possível provar que a amplitude  $|Y_i|$  tem uma distribuição de *Rayleigh*, enquanto a intensidade  $|Y_i|^2$  tem uma distribuição *exponencial*. Para os demais casos univariados só iremos considerar modelos para intensidade.



**Cenários heterogêneos:** Estas situações representam regiões de florestas e têm brilhos mais intensos do que as homogêneas. Os valores dos determinantes das intensidades, definidos como *variância generalizada* por Goodman (1963a), nestas regiões são maiores do que em cenários homogêneos (Frery, Nascimento & Cintra 2011). As leis polarimétricas decorrem das convoluções

$$\begin{cases} \mathcal{K}_m^{(L)}(\mathbf{\Sigma}, \mathcal{L}, L) = \Gamma(\mathcal{L}, \mathcal{L}) \hat{*} \mathcal{W}(\mathbf{\Sigma}, L), \\ \mathcal{K}_m^{(1)}(\mathbf{\Sigma}, \mathcal{L}) = \Gamma(\mathcal{L}, \mathcal{L}) \hat{*} \mathcal{N}_m^c(\mathbf{\Sigma}, 0), \end{cases}$$

em que o sobre escrito  $[\cdot]^{(L)}$  se referirá ao caso  $L$ -looks. Esta lei é conhecida como  $\mathcal{K}_m^{(L)}$  multivariada e foi discutida inicialmente por Lee, Schuler, Lang & Ranson (1994). Quando  $m = 1$ , esta lei multivariada se reduz ao caso univariado multilook  $\mathcal{K}_I$  (Yueh, Kong, Jao, Shin & Novak 1989, Lee, Schuler, Lang & Ranson 1994) com densidade dada por

$$f_{\mathcal{K}_I}(Z'; \mathcal{L}, L) = \frac{2L\mathcal{L}}{\Gamma(\mathcal{L})\Gamma(L)} (\mathcal{L}LZ')^{\frac{\mathcal{L}+L}{2}-1} K_{\mathcal{L}-L}(2\sqrt{L\mathcal{L}Z'}) \mathbb{I}_{(0,\infty)}(Z'),$$

em que  $\mathcal{L}, L > 0$ .

**Cenários extremamente heterogêneos:** Este caso considera regiões urbanas. Estas áreas produzem o fenômeno “double bounce”, tal como discutido na Seção 2.2. Freitas et al. (2005) sugeriram como modelos apropriados as convoluções

$$\begin{cases} \mathcal{G}_m^{0(L)}(\mathbf{\Sigma}, \alpha, L) = \Gamma^{-1}(-\alpha, (-\alpha - 1)) \hat{*} \mathcal{W}(\mathbf{\Sigma}, L), \\ \mathcal{G}_m^{0(1)}(\mathbf{\Sigma}, \alpha) = \Gamma^{-1}(-\alpha, (-\alpha - 1)) \hat{*} \mathcal{N}_m^c(\mathbf{\Sigma}, 0). \end{cases}$$

Esta distribuição foi introduzida por Freitas et al. (2005) e estende a distribuição  $\mathcal{G}_I^0$  proposta por Frery et al. (1997) cuja densidade é dada por

$$f_{\mathcal{G}_I^0}(Z'; \alpha, L) = \frac{L^L \Gamma(L - \alpha)}{(-\alpha - 1)^\alpha \Gamma(-\alpha) \Gamma(L)} Z'^{L-1} [(-\alpha - 1) + LZ']^{\alpha-L} \mathbb{I}(0, \infty)(Z').$$

Adicionalmente, a distribuição  $\mathcal{G}_m^H$  proposta por Frery, Jacobo-Berlles, Gambini & Mejail (2010) também tem sido sugerida para modelar cenários extremamente heterogêneos. As versões  $L$ -looks e 1-look desta distribuição são resultantes de

$$\begin{cases} \mathcal{G}_m^{H(L)}(\mathbf{\Sigma}, w, L) = GI(1, w) \hat{*} \mathcal{W}(\mathbf{\Sigma}, L), \\ \mathcal{G}_m^{H(1)}(\mathbf{\Sigma}, w) = GI(1, w) \hat{*} \mathcal{N}_m^c(\mathbf{\Sigma}, 0). \end{cases}$$

Esta distribuição se reduz a  $\mathcal{G}_I^H$  com densidade dada por

$$f_{\mathcal{G}_I^H}(Z'; w, L) = \frac{L^L \exp(w)}{\Gamma(L)} \sqrt{\frac{2w}{\pi}} Z'^{L-1} \left( \frac{w}{w + 2LZ'} \right)^{L/2+1/4} K_{L+1/2}(\sqrt{w(w + 2LZ')}) \mathbb{I}(0, \infty)(Z').$$

As expressões para as densidades destas distribuições para  $L$ -looks estão escritas na Tabela 3.5. As expressões para os casos 1-look são facilmente obtidas considerando  $L = 1$  e substituindo o termo " $\text{tr}(\Sigma^{-1}Z')$ " por " $z'^H \Sigma^{-1} z'$ " conforme a expressão (3.30).

A Figura 3.10 apresenta os ajustes das principais distribuições para imagem SAR:  $\Gamma$ ,  $\mathcal{K}$  e  $\mathcal{G}^0$ . Em todos os casos foram utilizados recortes da imagem de São Francisco discutida no Apêndice A.1.1. Nota-se que a distribuição  $\mathcal{G}_I^0$  descreve bem dados SAR em diferentes situações, funcionando como um tipo de modelagem universal. Esta característica foi estudada primeiramente por Mejail, Frery, Jacobo-Berlles & Bustos (2001) e, depois, por Mejail, Jacobo-Berlles, Frery & Bustos (2003). Como o objetivo de investigar com mais detalhes os ajustes, a Tabela 3.4 apresenta os valores do AIC para as diferentes situações consideradas. Os resultados mostram evidências de que a distribuição  $\mathcal{K}_I$  apresentou os melhores desempenhos nas regiões de florestas e cidades, entretanto a distribuição  $\mathcal{G}_I^0$  se mostrou mais robusta a variação de texturas.

**Tabela 3.4** Módulos das correlações médias entre os canais de polarizações

| Regiões  | Distribuições |                 | (AIC)             |
|----------|---------------|-----------------|-------------------|
|          | $\Gamma$      | $\mathcal{K}_I$ | $\mathcal{G}_I^0$ |
| Oceano   | -56757.46     | -53790.35       | -57531.42         |
| Floresta | -22859.16     | -24941.16       | -24824.82         |
| Cidades  | 2095.55       | -2911.127       | -2561.849         |

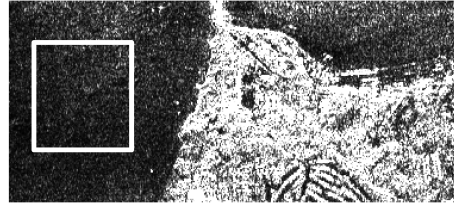
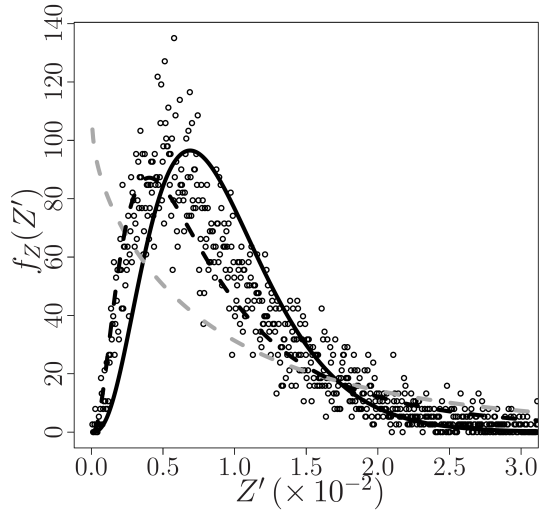
✎ A Seção 5.1 apresenta um avanço em termos da distribuição  $\mathcal{G}_I^0$ , (i) avaliando as estimativas de seus parâmetros por MV na presença de eventos de reflexão de canto, (ii) propondo testes de hipótese baseados em teoria da informação e (iii) comparando estes testes com o teste não paramétrico de Kolmogorov-Smirnov.

### 3.3.3.2 Classes moderadas:

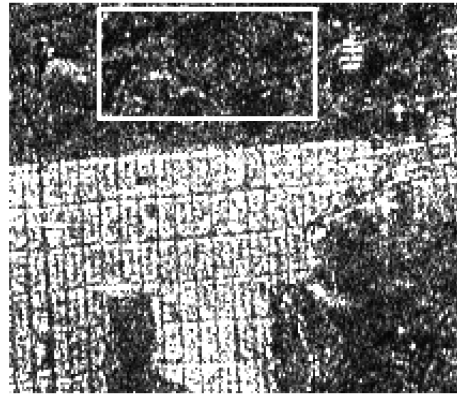
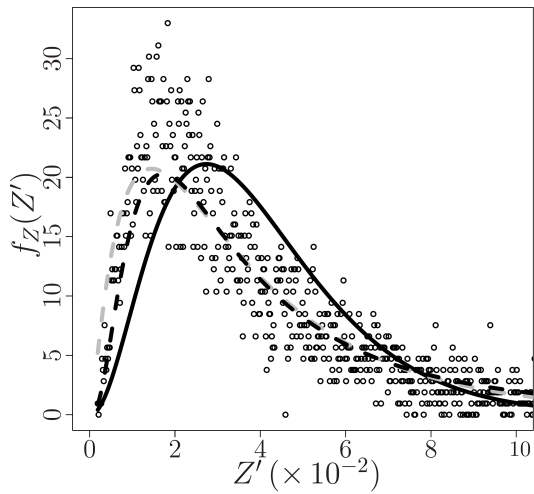
Adicionalmente, Bombrun, Anfinson & Harant (2011) consideraram as distribuições beta e beta inversa, (3.22) e (3.21), como modelos para o backscatter e as variáveis para o speckle como discutidas como no item anterior. Neste caso, as seguintes distribuições foram derivadas

$$\begin{cases} W_m^{(L)}(\Sigma, \mathcal{L}, w, L) = \text{beta}(\mathcal{L}, w - \mathcal{L}, w\mu\mathcal{L}^{-1}) \quad \hat{*} \quad \mathcal{W}(\Sigma, L), \\ M_m^{(L)}(\Sigma, \mathcal{L}, w, L) = \text{beta}^{-1}(-w + 1, w - \mathcal{L}, \mu\mathcal{L}w^{-1}) \quad \hat{*} \quad \mathcal{W}(\Sigma, L). \end{cases}$$

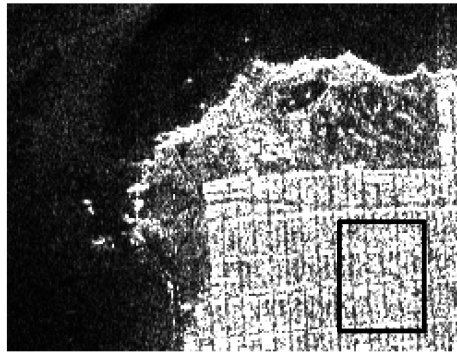
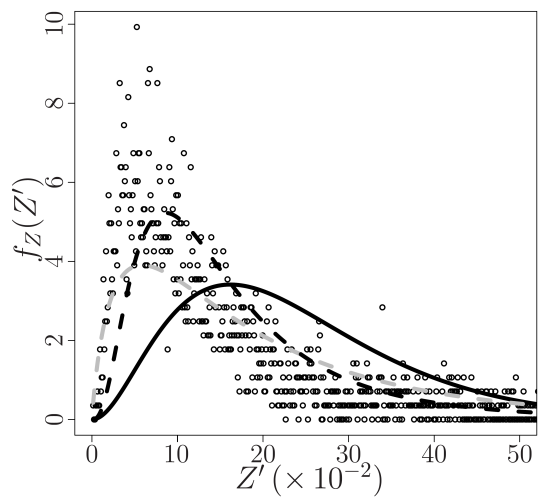
Suas expressões estão definidas na Tabela 3.5.



(a) Ajuste sobre região de oceano.



(b) Ajuste sobre região de florestas.



(c) Ajuste sobre região de cidades.

**Figura 3.10** Ajustes dos modelos clássicos para o canal HH:  $\Gamma$  (linha sólida preta),  $\mathcal{K}_I$  (tracejada cinza) e  $\mathcal{G}_I^0$  (tracejada preta).

## 3.3.3.3 Classes amplas:

Considerando as duas distribuições mais gerais para o backscatter, com densidades (3.24) e (3.23), combinadas aos modelos para o ruído  $\mathbf{Y} \sim \mathcal{W}(\mathbf{\Sigma}, L)$  e  $\mathbf{y} \sim \mathcal{N}_m^c(\mathbf{\Sigma}, 0)$ , os seguintes modelos polarimétricos podem ser derivados:

(a) a classe KummerU (Bombrun & Beaulieu 2008) definida pelas convoluções

$$\begin{cases} U_m^{(L)}(\mathbf{\Sigma}, \mathcal{L}, \alpha, L) = \mathcal{G}^0(\alpha, (-\alpha - 1), \mathcal{L}) \hat{*} \mathcal{W}(\mathbf{\Sigma}, L), \\ U_m^{(1)}(\mathbf{\Sigma}, \mathcal{L}, \alpha) = \mathcal{G}^0(\alpha, (-\alpha - 1), \mathcal{L}) \hat{*} \mathcal{N}_m^c(\mathbf{\Sigma}, 0). \end{cases}$$

(b) a classe  $\mathcal{G}_m$  (Freitas et al. 2005),

$$\begin{cases} \mathcal{G}_m^{(L)}(\mathbf{\Sigma}, \alpha, w, L) = \mathcal{N}^{-1}(1, \alpha, w) \hat{*} \mathcal{W}(\mathbf{\Sigma}, L), \\ \mathcal{G}_m^{(1)}(\mathbf{\Sigma}, \alpha, w) = \mathcal{N}^{-1}(1, \alpha, w) \hat{*} \mathcal{N}^c(\mathbf{\Sigma}, 0). \end{cases}$$

Suas expressões estão elencadas na Tabela 3.5 para  $L$ -looks e, para a 1-look, elas podem ser obtidas pela Fórmula (3.30). Estas expressões são dependentes de funções especiais: a classe  $\mathcal{G}_m$  depende da função modificada de Bessel definida em (3.17), enquanto que a KummerU depende da função  $U(a, b, c)$  definida por

$$U(a, b, c) = \Gamma(a)^{-1} \int_0^\infty \exp(-zt) t^{\alpha-1} (1+t)^{b-a-1} dt,$$

em que  $\Re[a], \Re[z] > 0$ , tal que  $\Re[\cdot]$  representa a parte real. A seguir, nós comentamos alguns trabalhos envolvendo estas classes de distribuições.

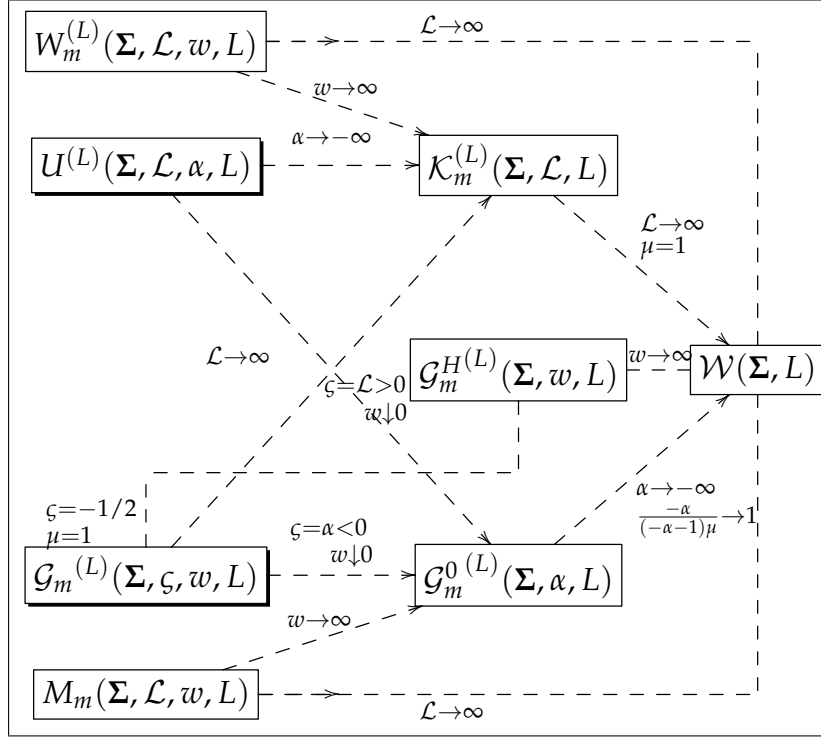
Quanto à classe KummerU, Bombrun, Vasile, Gay & Totir (2011) trabalharam recentemente com sua versão 1-look para modelar dados PolSAR heterogêneos. A fim de quantificar a eficiência da distribuição  $U_m^{(1)}$ , o contexto de segmentação hierárquica baseada na maximização de sua log-verossimilhança foi empregado. Sua versão multilook também foi derivada e aplicada ao contexto de segmentação hierárquica por Bombrun & Beaulieu (2008).

A classe  $\mathcal{G}_m$  foi tratada matematicamente e aplicada a dados PolSAR por Freitas et al. (2005). O estudo com sua versão SIRV ainda é um ponto em aberto na literatura.

Adicionalmente, sob algumas restrições ilustradas no diagrama a seguir, vê-se que ambas as classes KummerU (Bombrun, Anfinson & Harant 2011) e  $\mathcal{G}_m$  (Freitas et al. 2005) convergem para as distribuições particulares  $\mathcal{K}_m^{(L)}(\mathbf{\Sigma}, \mathcal{L}, L)$ ,  $\mathcal{G}_m^{(L)}(\mathbf{\Sigma}, \alpha, L)$  e  $\mathcal{W}(\mathbf{\Sigma}, L)$ . Além disso, a distribuição  $\mathcal{G}_m^{(L)}$  converge para o caso  $\mathcal{G}_m^{H(L)}$ . A Figura 3.11 apresenta um diagrama com as relações dos modelos polarimétricos considerados. Este diagrama foi organizado de forma que as distribuições das colunas da esquerda para a direita apresentam quatro, três e dois parâmetros, respectivamente.

Tabela 3.5 Fatores para formação das densidades polarimétricas

| Classe   | Modelo: $\mathcal{M}(\boldsymbol{\theta})$                  | $h_{\mathcal{M}}(\boldsymbol{\theta}, L)$                                                                                                                                             | $g_{\mathcal{M}}(\text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}'); \boldsymbol{\theta}, L)$                                                                                                                                                                                                                                                                                               |
|----------|-------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Ampla    | $\mathcal{G}_m^{(L)}(\boldsymbol{\Sigma}, \varsigma, w, L)$ | $\frac{L^{Lm} r_{\varsigma, w}^{\varsigma}}{ \boldsymbol{\Sigma} ^L K_{\varsigma}(w)}$                                                                                                | $\left( \frac{2L \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}') + \frac{w}{r_{\varsigma, w}}}{w r_{\varsigma, w}} \right) K_{\varsigma-Lm} \left( \sqrt{w r_{\varsigma, w} \left( 2L \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}') + \frac{w}{r_{\varsigma, w}} \right)} \right)$                                                                                                     |
|          | $U_m^{(L)}(\boldsymbol{\Sigma}, \mathcal{L}, \alpha, L)$    | $\frac{L^{Lm}}{ \boldsymbol{\Sigma} ^L} \frac{\Gamma(\mathcal{L}-\alpha) \Gamma(Lm-\alpha)}{\Gamma(\mathcal{L}) \Gamma(-\alpha)} \left[ \frac{\mathcal{L}}{(-\alpha-1)} \right]^{Lm}$ | $U \left( Lm - \alpha, 1 + Lm - \mathcal{L}, \frac{L \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}') \mathcal{L}}{(-\alpha-1)} \right),$<br>em que $U(\cdot, \cdot, \cdot)$ é a função de Kummer                                                                                                                                                                                         |
| Moderada | $W_m^{(L)}(\boldsymbol{\Sigma}, \mathcal{L}, w, L)$         | $\frac{1}{ \boldsymbol{\Sigma} ^L} \left( \frac{L\mathcal{L}}{w} \right)^{\frac{\mathcal{L}+Lm-1}{2}} \frac{\Gamma(w)}{\Gamma(\mathcal{L})}$                                          | $\text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}')^{\frac{\mathcal{L}+Lm-1}{2}} \exp \left( -\frac{L\mathcal{L}}{2w} \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}') \right) W_{\frac{w-Lm-2\mathcal{L}}{2}, \frac{Lm-\mathcal{L}}{2}} \left( -\frac{L\mathcal{L}}{w} \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}') \right),$<br>em que $W_{\mu, \delta}(\cdot)$ é a função de Whittaker |
|          | $M_m^{(L)}(\boldsymbol{\Sigma}, \mathcal{L}, w, L)$         | $\frac{1}{ \boldsymbol{\Sigma} ^L} \frac{\Gamma(w) \Gamma(Lm+\mathcal{L})}{\Gamma(\mathcal{L}) \Gamma(Lm+w)} \left( \frac{Lw}{\mathcal{L}} \right)^{\frac{Lm-w}{2}}$                  | $[-\text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}')]^{\frac{-w-Lm}{2}} \exp \left( -\frac{Lw}{2\mathcal{L}} \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}') \right) M_{\frac{w-Lm-2\mathcal{L}}{2}, \frac{Lm-\mathcal{L}}{2}} \left( -\frac{Lw}{\mathcal{L}} \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}') \right),$<br>em que $M_{\mu, \delta}(\cdot)$ é a função de Whittaker M       |
| Simples  | $\mathcal{K}_m^{(L)}(\boldsymbol{\Sigma}, \mathcal{L}, L)$  | $\frac{2(L\mathcal{L})^{\frac{\mathcal{L}+Lm}{2}}}{ \boldsymbol{\Sigma} ^L \Gamma(\mathcal{L})}$                                                                                      | $\text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}')^{\frac{\mathcal{L}-Lm}{2}} K_{\mathcal{L}-Lm} \left( 2\sqrt{L\mathcal{L} \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}')} \right)$                                                                                                                                                                                                      |
|          | $\mathcal{G}_m^{0(L)}(\boldsymbol{\Sigma}, \alpha, L)$      | $\frac{L^{Lm} \Gamma(Lm-\alpha)}{ \boldsymbol{\Sigma} ^L \Gamma(-\alpha) (-\alpha-1)^{\alpha}}$                                                                                       | $[L \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}') + (-\alpha-1)]^{\alpha-Lm}$                                                                                                                                                                                                                                                                                                          |
|          | $\mathcal{G}_m^{H(L)}(\boldsymbol{\Sigma}, w, L)$           | $\frac{\sqrt{2/L} L^{mL} \exp(w) w^{3L+1}}{\pi^3  \boldsymbol{\Sigma} ^L}$                                                                                                            | $\frac{K_{3\mathcal{L}-1/2} \left( \sqrt{w(2L \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}') + w)} \right)}{[w(2L \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}') + w)]^{3/2L+1/4}}$                                                                                                                                                                                                    |
|          | $\mathcal{W}(\boldsymbol{\Sigma}, L)$                       | $\frac{L^{mL}}{ \boldsymbol{\Sigma} ^L}$                                                                                                                                              | $\exp(-L \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}'))$                                                                                                                                                                                                                                                                                                                               |



**Figura 3.11** Diagrama de modelos para o retorno.

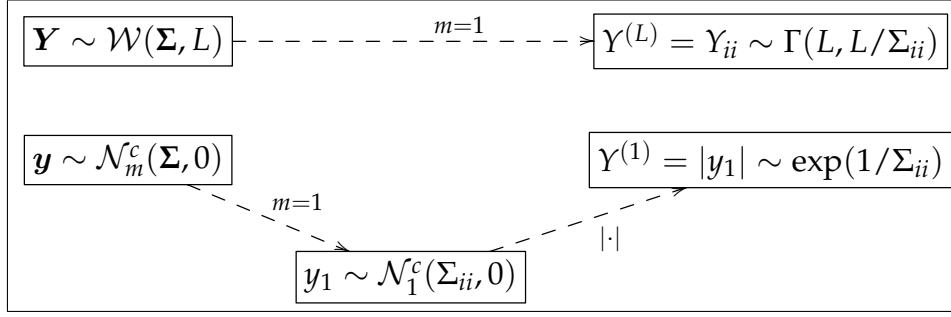
### 3.3.4 Retornos SAR

Neste caso, a modelagem multiplicativa considera que cada pixel é representado por uma variável aleatória  $Z^{(L)}$ , que é o produto de duas variáveis aleatórias independentes e positivas,  $X$  e  $Y^{(L)}$ . Assim, assumindo que  $f_X(X'; \theta_1)$  e  $f_{Y^{(L)}}(Y'; \theta_2)$  são as densidades associadas às variáveis  $X$  e  $Y^{(L)}$ , a densidade de  $Z^{(L)} = X \cdot Y^{(L)}$  pode ser obtida pelo seguinte gerador (Frery et al. 1997):

$$f_{Z^{(L)}}(Z'; \theta_1, \theta_2) = \int_0^\infty X'^{-1} f_{Y^{(L)}}(Z'/X'; \theta_2) f_X(X'; \theta_1) dX'. \quad (3.31)$$

Com respeito à modelagem do ruído, como discutido na Seção 3.3.3.1, as distribuições  $\mathcal{W}$  e  $\mathcal{N}_m^c$  se degeneram nas seguintes distribuições, quando  $m = 1$  e  $\Sigma_{ii}$  a entrada  $(i, i)$  da matriz  $\Sigma$ :

Desta forma, definindo  $\mu = \Sigma_{ii}$ , temos que os ruídos speckle para  $L$ -looks e 1-look são descritos por  $Y^{(L)} \sim \Gamma(L, L/\mu)$  e  $Y^{(1)} \sim \Gamma(1, 1/\mu) = \exp(1/\mu)$ , respectivamente, quando o objetivo é modelar a *intensidade* de um retorno SAR. Para modelar a amplitude do retorno, basta-se considerar a raiz quadrada destas distribuições cujas densidades são obtidas trivialmente pela transformação  $f_{\sqrt{Y^{(L)}}}(y) = 2yf_{Y^{(L)}}(y^2)$  (Para maiores detalhes, consultar Frery et al. 1997).



**Figura 3.12** Diagrama de modelos para o ruído speckle.

Considere como modelo para o backscatter da intensidade (i) a densidade da gaussiana inversa generalizada definida por Frery et al. (1997) e (ii) outras seis como dadas na Seção 3.3.2:  $\Gamma(\alpha, \lambda)$  com  $(\alpha, \lambda) = (\mathcal{L}, \mathcal{L}/\mu)$ ,  $\Gamma(-\alpha, \gamma)$  com  $\gamma = (-\alpha - 1)\mu$ ,  $\text{beta}(p, q, \beta)$  com  $(p, q, \beta) = (\mathcal{L}, -\alpha - \mathcal{L}, (-\alpha - 1)\mu/\mathcal{L})$  e  $\mathcal{F}(p, q, \beta)$  com  $(p, \alpha, \beta) = (\mathcal{L}, -\alpha, (-\alpha - 1)\mu/\mathcal{L})$ . Assim, aplicando estas densidades juntamente com a distribuição  $\Gamma(L, L)$  para o ruído na equação (3.31), tem-se as densidades para dados PolSAR monopolarizados.

Similarmente ao caso polarimétrico, é possível notar que as famílias para um canal de polarização podem ser colocadas na seguinte degeneração da formula equação (3.29) para  $m = 1$ :

$$f_Z(Z'; \boldsymbol{\theta}, L) = \frac{Z^{L-1}}{\Gamma(L)} h_{\mathcal{M}}(\boldsymbol{\theta}, L) g(Z'; \boldsymbol{\theta}, L), \quad (3.32)$$

As funções  $h_{\mathcal{M}}$  e  $g$  para gerar as respectivas densidades estão dadas na Tabela 3.6.

Outras abordagens de modelagem para dados monopolarizados, tais como modelos não paramétricos e modelos paramétricos baseados no teorema central do limite, em distribuições empíricas e nos modelos de misturas entre gamas e gaussianas, foram discutidas por Gao (2010). Esta tese desconsidera estas abordagens, uma vez que seu foco principal é a discussão de modelos polarimétricos.

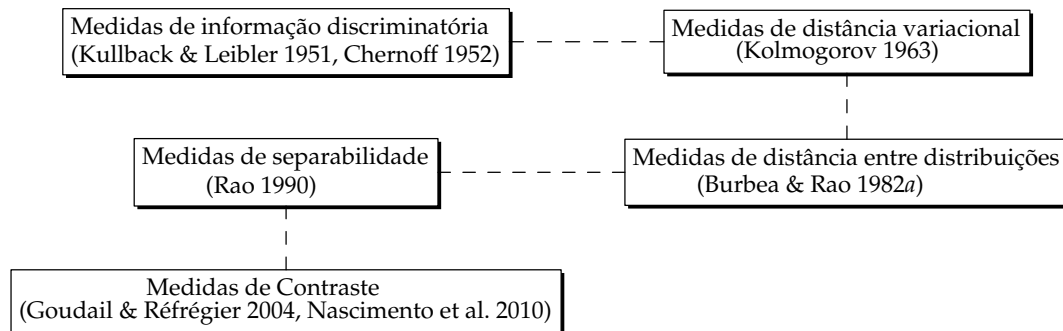
**Tabela 3.6** Fatores para formação das densidades para um canal

| Classe   | $\mathcal{M}(\boldsymbol{\theta})$          | $h_{\mathcal{M}}(\boldsymbol{\theta}, L)$                                  | $g_{\mathcal{M}}(Z'; \boldsymbol{\theta}, L)$                                                                              |
|----------|---------------------------------------------|----------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|
| Ampla    | $\mathcal{G}_I(\alpha, \gamma, \lambda, L)$ | $\frac{L^L(\lambda/\gamma)^{\alpha/2}}{K_{\alpha}(2\sqrt{\lambda\gamma})}$ | $(\frac{\gamma+LZ'}{\lambda})^{(\alpha-L)/2} K_{\alpha-L}(2\sqrt{\lambda(\gamma+LZ')})$                                    |
|          | $U_I(p, q, \beta, L)$                       | $(\frac{L}{\beta})^L \frac{\Gamma(L+q)}{B(p, q)}$                          | $U_{p+q, 1-L+p}(\frac{LZ'}{\beta})$                                                                                        |
| Moderada | $W_I(p, q, \beta, L)$                       | $\frac{\Gamma(q)}{B(p, q)} (\frac{L}{\beta})^L$                            | $(\frac{LZ'}{\beta})^{\frac{p-L-1}{2}} \exp(-\frac{Lx}{2\beta}) W_{-\frac{p-2q-L+1}{2}, \frac{L-p}{2}}(\frac{LZ'}{\beta})$ |
| Simples  | $\mathcal{K}_I(\alpha, \mathcal{L}, L)$     | $\frac{2(\mathcal{L}L)^L}{\Gamma(\alpha)\Gamma(L)}$                        | $(L\mathcal{L}Z')^{\frac{\alpha-L}{2}} K_{\alpha-L}(2\sqrt{L\mathcal{L}Z'})$                                               |
|          | $\mathcal{G}_I^0(\alpha, \gamma, L)$        | $\frac{L^L\Gamma(L-\alpha)}{\gamma^{\alpha}\Gamma(-\alpha)}$               | $[LZ' + \gamma]^{\alpha-L}$                                                                                                |
|          | $\mathcal{G}_I^H(w, \eta, L)$               | $L^L \sqrt{\frac{2w\eta}{\pi}} \exp(w)$                                    | $(\frac{w}{\eta(w\eta+2LZ')})^{\frac{L}{2}+\frac{1}{4}} K_{L+\frac{1}{2}}(\sqrt{\frac{w}{\eta}}(w\eta+2LZ'))$              |
|          | $\Gamma(L, L/\mu)$                          | $(\frac{L}{\mu})^L$                                                        | $\exp(-\frac{LZ'}{\mu})$                                                                                                   |



## Teoria Estatística da Informação

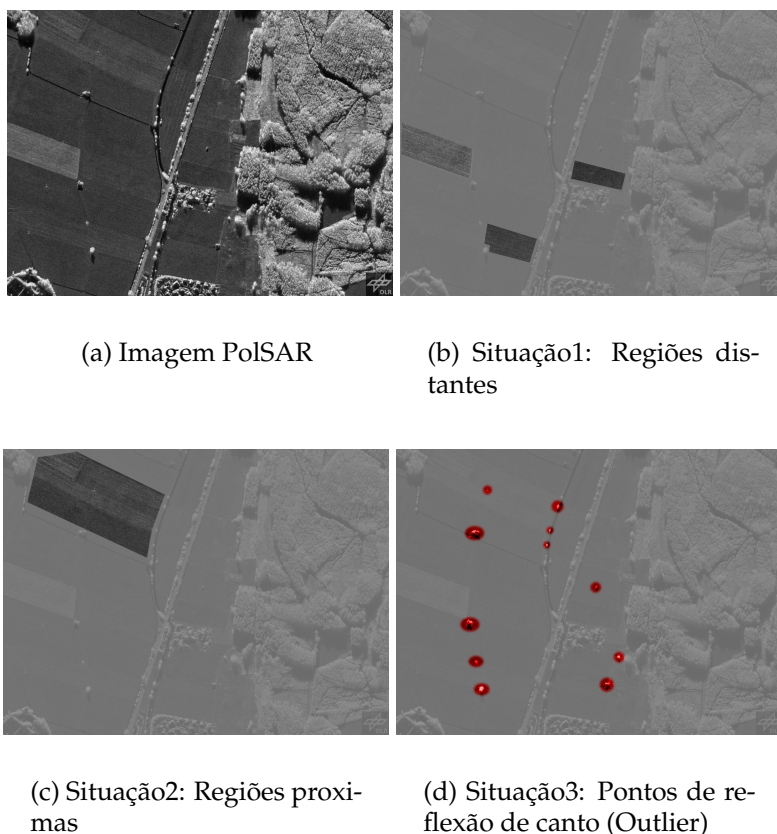
Desde que Mahalanobis (1936) introduziu o conceito de distância entre duas distribuições de probabilidade, várias medidas na literatura de Estatística têm sido propostas com o fim de responder a pergunta: *‘quão distinguíveis são duas distribuições de probabilidade?’* Tais medidas têm assumido vários nomes no decurso do tempo:



Muitos dos testes de hipótese usuais, tais como aqueles baseados na razão de verossimilhança, na distância qui-quadrado e nas estatísticas de escore e Wald, podem ser definidos como casos particulares dentro de classes de distâncias estocásticas apropriadas (Ullah 1996, Broniatowski & Keziou 2009). A formalização destes conceitos é o objetivo desta seção, a saber o estudo da *Teoria Estatística da Informação* (Blahut 1972). Adicionalmente, esta linha será aplicada ao tratamento de imagens PolSAR.

Em anos recentes, o interesse em se adaptar ferramentas da Teoria Estatística da Informação ao processamento destas imagens tem crescido (Morio, Réfrégier, Goudail, Dubois-Fernandez & Dupuis 2008, Goudail & Réfrégier 2004). Como motivação, a Figura 4 apresenta três situações em que se faz necessária a utilização de distâncias estocásticas. A Figura 4.1(a) representa uma imagem PolSAR de Oberpfaffenhofen (Alemanha) obtida pelo sensor E-SAR operando em cinco bandas de frequências, X -, C-, S-, L- e P- (para maiores detalhes, consultar o Apêndice A). As Figuras 4.1(b) e 4.1(c) exibem situações nas quais queremos ou quantificar contraste entre regiões distintas ou em uma mesma região, respectivamente; isto é, obter estimativas para o poder (ou probabilidade de detecção) e para o tamanho de um teste (ou, taxa de alarme falso) para a hipótese de que duas regiões são, a priori, similares. Adicionalmente, o problema de identificar pixels aberrantes e/ou influentes de um conjunto é representado pela Figura 4.1(d). Este problema é importante pois tais dados, apesar de influencia-

rem significativamente na análise de imagens reais, carregam informação sobre fenômenos importantes (tais como a reflexão de canto e o efeito *double bounce*, discutidos na Seção 2.2) e, portanto, não devem ser descartadas do processamento.



**Figura 4.1** Utilização de distâncias estocástica em processamento de imagem.

Desta forma, esta seção traz duas contribuições metodológicas: (i) uma discussão da teoria da informação em um sentido estocástico e (ii) sua adaptação ao domínio de dados PolSAR (tal como discutido no **Capítulo 3**).

Parece razoável começarmos pela conceituação do termo *Teoria da Informação*. A Teoria de Informação é um ramo relativamente novo da matemática, datando do final dos anos de 1940. De forma brevíssima, pode-se dizer que seu campo de atuação está relacionado com a sistematização matemática da transmissão de informações entre canais de comunicação. Os primeiros estudos direcionados a esse objetivo foram realizados por Nyquist nos anos de 1924 e 1928 e por Hartley (1928). Em 1948, Shannon estendeu os trabalhos de Nyquist e Hartley num notável artigo (Shannon 1948), que considerou, dentre outras coisas, o efeito do ruído em canais de comunicação. Os conceitos chave da teoria de Shannon recaem sobre duas medidas: *informação* e *entropia*. Tais conceitos têm um impacto considerável em diversas áreas, tais como Codificação e Compressão de dados (Donoho, Vetterli, DeVore & Daubechies 1998), e Telecomuni-

cações e Inferência Estatística (Blatt & Hero 2007).

A partir destes conceitos se introduziu uma ferramenta chamada de medida de entropia relativa (de fato, uma divergência), tendo como pioneiro o trabalho de Kullback & Leibler (1951). Esta última medida objetivava mensurar o erro de se escolher uma variável quando outra seria mais ajustável às informações. Com a primeira versão em 1959, Kullback escreveu o livro “Information Theory and Statistics” (Kullback 1978), o qual foi precursor da Teoria Estatística da Informação.

Posteriormente aos primeiros conceitos de entropias e divergências, começou-se uma investida na proposta de classes destas medidas. A seguir, nós apresentamos uma discussão das principais classes de entropias e divergências.

Seja  $X$  uma variável aleatória assumindo valores em um espaço amostral  $\mathcal{X}$ . Suponha ainda que  $F$  é a sua função de distribuição acumulada definida sobre o vetor de parâmetros desconhecidos  $\theta$ . Seja  $(\mathcal{X}, \beta_{\mathcal{X}}, P_{\theta})_{\theta \in \Theta}$  o espaço de probabilidade associado a variável  $X$ , obdecendo à seguinte notação:

- (a)  $\beta_{\mathcal{X}}$ :  $\sigma$ -álgebra dos subconjuntos de Borel  $A \subset \mathcal{X}$ , em que  $A$  é um subconjunto de Borel;
- (b)  $\{P_{\theta}\}_{\theta \in \Theta}$  é uma família de distribuições de probabilidade definida sobre o espaço mensurável  $(\mathcal{X}, \beta_{\mathcal{X}})$ , em que  $\Theta$  é um subconjunto de  $\mathbb{R}^m$  para  $m \geq 1$ .

Nós assumimos que a distribuição de probabilidade  $P_{\theta}$  é absolutamente contínua com respeito a  $\sigma$ -álgebra  $\mu$  sobre  $(\mathcal{X}, \beta_{\mathcal{X}})$ . Adicionalmente, assumiremos que o suporte de uma distribuição discreta é dado por  $S_{\mathcal{X}}$ .

Por simplicidade, vamos considerar  $\mu$  ou como uma medida de Lebesgue ou como uma medida de contagem (isto é, existe um conjunto finito ou contável  $S_{\mathcal{X}}$  tal que  $\Pr_{\theta}(\mathcal{X} - S_{\mathcal{X}}) = 0$ ). Assim, temos que

$$f(x; \theta) = \frac{dP_{\theta}}{d\mu}(x) = \begin{cases} f(x; \theta), & \text{se } \mu \text{ é uma medida de Lebesgue,} \\ \Pr_{\theta}(X = x), & \text{se } \mu \text{ é uma medida de contagem.} \end{cases}$$

Burbea & Rao (1982b) introduziram o conceito do funcional “ $\phi$ -entropia” dado por

$$\begin{aligned} H_{\phi}(X) &\equiv H_{\phi}(P_{\theta}) \equiv H_{\phi}(\theta) = \int_{\mathcal{X}} \phi(f(x; \theta)) d\mu(x) \\ &= \begin{cases} \int_{\mathcal{X}} \phi(f(x; \theta)) dx, & \text{se } \mu \text{ é uma medida de Lebesgue,} \\ \sum_{x \in S_{\mathcal{X}}} \phi(\Pr_{\theta}(X = x)), & \text{se } \mu \text{ é uma medida de contagem.} \end{cases} \end{aligned} \quad (4.1)$$

Burbea & Rao (1982a) ainda propuseram outra medida baseada na matriz Hessiana do funcional  $\phi$ -entropia. Salicrú et al. (1993) estabeleceram uma classe de medidas de

entropias mais geral do que a dada pela equação (4.1), a fim de considerar a inclusão de outras entropias ausentes na classe  $\phi$ -entropia. Esta medida é dada por

$$\boxed{H_\phi^h(X) \equiv H_\phi^h(P_\theta) \equiv H_\phi^h(\theta) = h(H_\phi(\theta)),} \quad (4.2)$$

em que  $h: \mathbb{R} \rightarrow \mathbb{R}$  é crescente e  $\phi: [0, \infty) \rightarrow \mathbb{R}$  é uma função côncava ou  $h: \mathbb{R} \rightarrow \mathbb{R}$  é decrescente e  $\phi: [0, \infty) \rightarrow \mathbb{R}$  é convexa.

Já para o caso de divergências, a seguinte discussão se faz necessária. Considere duas variáveis aleatórias  $X \sim f(x; \theta_1) = dP_{\theta_1}/d\mu(x)$  e  $Y \sim g(x; \theta_2) = dP_{\theta_2}/d\mu(x)$  com suporte comum:  $\mathcal{X}$ , para  $\mu$  como uma medida de Lebesgue, e  $S_{\mathcal{X}}$ , quando  $\mu$  é uma medida de contagem. As classes de divergências apresentadas posteriormente podem ser representadas pela seguinte medida:

$$\begin{aligned} D(X, Y) &\equiv D(P_{\theta_1}, P_{\theta_2}) \equiv D(\theta_1, \theta_2) = \int_{\mathcal{X}} \varphi(f(x; \theta_1), g(x; \theta_2)) d\mu(x) \\ &= \begin{cases} \int_{\mathcal{X}} \varphi(f(x; \theta_1), g(x; \theta_2)) dx, & \text{se } \mu \text{ é uma medida de Lebesgue,} \\ \sum_{x \in S_{\mathcal{X}}} \varphi(f(x; \theta_1), g(x; \theta_2)), & \text{se } \mu \text{ é uma medida de contagem,} \end{cases} \end{aligned} \quad (4.3)$$

em que  $\varphi(x, y)$  é uma função real das variáveis  $x, y > 0$ . Neste caso,  $\varphi(\cdot, \cdot)$  representa qualquer função definida em  $(x, y) \in [0, \infty) \times [0, \infty)$  tal que  $\varphi(0, 0) = 0$ ,  $\varphi(0, y) = \lim_{x \downarrow 0} \varphi(x, y)$  e  $\varphi(x, 0) = \lim_{y \downarrow 0} \varphi(x, y)$ .

A seguir, são listadas as três principais delas. Para estes casos, temos que  $\phi: (0, \infty) \rightarrow (-\infty, \infty)$  é convexa, estendida por continuidade na forma  $[0, \infty] \rightarrow (-\infty, \infty]$  pela satisfação de  $\phi(0) = \lim_{y \downarrow 0} \phi(y)$  e  $\phi(\infty) = \lim_{y \rightarrow \infty} \phi(y)$ . Adicionalmente,  $\phi(y)$  é duas vezes diferenciável sobre  $(0, \infty)$  com derivadas contínuas  $\phi'(y)$  e  $\phi''(y)$ , satisfazendo as condições (i)  $\phi(1) = 0$  e (ii)  $\phi''(1) > 0$ . Cada classe com suas respectivas funções  $\varphi(\cdot, \cdot)$  são apresentadas abaixo.

**Definição 2 (Divergência de Bregman: Bregman (1967)).** Considere o integrando da equação (4.3) como

$$\varphi(x, y) = \phi(x) - \phi(y) - \phi'(y)(x - y),$$

em que  $\phi'(x) = \partial\phi(x)/\partial x$ . Decorrentemente, considerando  $\dot{\varphi}(x, y) = \partial\varphi(x, y)/\partial x$  e  $\ddot{\varphi}(x, y) = \partial^2\varphi(x, y)/\partial x^2$ , tem-se as propriedades:

$$(i) \varphi(y, y) = 0, (ii) \dot{\varphi}(x, y) = \phi'(x) - \phi'(y) \text{ e } \dot{\varphi}(y, y) = 0, (iii) \ddot{\varphi}(x, y) = \phi''(x) \text{ e } \ddot{\varphi}(y, y) = \phi''(y) > 0,$$

em que  $\phi''(x) = \partial^2\phi(x)/\partial x^2$ . Aplicando  $\varphi(x, y)$  na equação (4.3), a seguinte classe é obtida:

$$\boxed{B_\phi(X, Y) = \int_{\mathcal{X}} [\phi(f(x; \theta_1)) - \phi(g(x; \theta_2))] - \phi'(g(x; \theta_2))[f(x; \theta_1) - g(x; \theta_2)] d\mu(x).}$$

**Definição 3 (Divergência de Burbea-Rao:** Burbea & Rao (1982c)). *Especifique, agora, o integrando de (4.3) como*

$$\varphi(x, y) = \frac{\phi(x) + \phi(y)}{2} - \phi\left(\frac{x+y}{2}\right).$$

*Desta definição, segue-se*

$$(i) \varphi(y, y) = 0, (ii) \dot{\varphi}(x, y) = \frac{1}{2}[\phi'(x) - \phi'\left(\frac{y+x}{2}\right)] \text{ e } \dot{\varphi}(y, y) = 0 \text{ e } (iii) \ddot{\varphi}(x, y) = \frac{1}{2}[\phi''(x) - \frac{1}{2}\phi''\left(\frac{y+x}{2}\right)] \text{ e } \ddot{\varphi}(y, y) = \frac{1}{4}\phi''(y).$$

*Assim, tem-se a classe segundo Burbea-Rao*

$$R_\phi(X, Y) \equiv R_\phi(\theta_1, \theta_2) = H_\phi\left(\frac{P_{\theta_1} + P_{\theta_2}}{2}\right) - \frac{H_\phi(P_{\theta_1}) + H_\phi(P_{\theta_2})}{2}.$$

*Baseando-se na extensão para  $\phi$ -entropia na equação (4.2), esta expressão é estendida em*

$$R_\phi^h(\theta_1, \theta_2) = H_\phi^h\left(\frac{P_{\theta_1} + P_{\theta_2}}{2}\right) - \frac{H_\phi^h(P_{\theta_1}) + H_\phi^h(P_{\theta_2})}{2}.$$

**Definição 4 ( $\phi$ -divergência ou divergência de Csiszár:** Csiszár (1967) e Ali & Silvey (1996)). *Finalmente, assuma a seguinte especificação para o integrando de equação (4.3):*

$$\varphi(x, y) = y\phi\left(\frac{x}{y}\right).$$

*Consequentemente,*

$$(i) \varphi(y, y) = 0, (ii) \dot{\varphi}(x, y) = \phi'(x/y) \text{ e } \dot{\varphi}(y, y) = \phi'(1) \text{ e } (iii) \ddot{\varphi}(x, y) = y^{-1}\phi''(x/y) \text{ e } \ddot{\varphi}(y, y) = y^{-1}\phi''(1) > 0.$$

*Assim, a classe de Csiszár é dada por*

$$D_\phi(X, Y) \equiv D_\phi(\theta_1, \theta_2) = \int_{\mathcal{X}} g(x; \theta_2) \phi\left(\frac{f(x; \theta_1)}{g(x; \theta_2)}\right) d\mu(x). \quad (4.4)$$

Para fins estatísticos, a importante medida proposta por Cressie & Read (1984),

$$I_\lambda(\theta_1, \theta_2) = \frac{1}{\lambda(\lambda+1)} \left[ \int \frac{f(x; \theta_1)^{\lambda+1}}{f(x; \theta_2)^\lambda} dx - 1 \right]$$

para  $\lambda \in \mathbb{R} \setminus \{0, 1\}$ , pode ser obtida aplicando

$$\phi_\lambda(x) = \frac{1}{\lambda(\lambda+1)} [x^{\lambda+1} - x - \lambda(x-1)],$$

na equação (4.4). Esta divergência obedece aos limites  $\lim_{\lambda \rightarrow 0} \phi_\lambda(x) = x \log x + 1$  e  $\lim_{\lambda \rightarrow 1} \phi_\lambda(x) = -\log x + x - 1$ , os quais redundam nas divergências de Kullback-Leibler

$$D_{\text{KL}}(\theta_1, \theta_2) = \int f(x; \theta_1) \log \frac{f(x; \theta_1)}{f(x; \theta_2)} dx \text{ e } D_{\text{KL}}(\theta_2, \theta_1) = \int f(x; \theta_2) \log \frac{f(x; \theta_2)}{f(x; \theta_1)} dx,$$

respectivamente. Entretanto, esta classe tem a desvantagem de não abordar medidas importantes de divergências, tais como Bhattacharyya (Kailath 1967), Rényi (Rényi 1961) e Sharma and Mittal (Liese & Vajda 2006). Como solução a este problema, Méndez, Morales, Pardo & Salicrú (1995) introduziram a classe de divergências  $(h, \phi)$ :

$$D_\phi^h(X, Y) \equiv D_\phi^h(P_{\theta_1}, P_{\theta_2}) \equiv D_\phi^h(\theta_1, \theta_2) = h(D_\phi(\theta_1, \theta_2)).$$

A seguir, as classes de entropias e divergências  $(h, \phi)$  serão detalhadas considerando funções densidade com suporte dado pelo conjunto das matrizes Hermitianas positivas definidas. Além disso, casos particulares clássicos serão considerados bem como algumas relações de suas expressões com tópicos da Inferência Estatística.

Este capítulo está estruturado nas seguintes seções. A classe de entropias proposta por Salicrú et al. (1993) adaptada à modelagem polarimétrica será apresentada na Seção 4.1. Esta seção, além de discutir formas analíticas para expressões de entropias, considera seus comportamentos assintóticos com vistas à proposição de testes de hipótese e intervalos de confiança nos contextos de dados polarimétricos. A Seção 4.2 complementa esta discussão com a classe de divergências  $(h, \phi)$ . Similarmente à seção anterior, os comportamentos assintóticos destas medidas e métodos para testar hipótese são discutidos. Em ambas as seções são propostos métodos gerais de se estudar informações polarimétricas por quaisquer uma das distribuições elencadas na Tabela 3.5. Uma proposta geral para entropia de Shannon em modelos PolSAR é apresentada na Seção 4.1. E a divergência de Kullback-Leibler para dados PolSAR é discutida na Seção 4.2.

### 4.1 A classe de entropias $(h, \phi)$

O conceito de “diversidade” é importante e está relacionado ao grau de variedade: o número de diferentes caminhos de avaliar uma característica com relação ao total de elementos com cada valor da característica. Esta quantidade também pode ser entendida como o grau de heterogeneidade de elementos relativo a uma dada característica. Uma vez que as medidas de entropias satisfazem todos requisitos de uma medida de diversidade (Rao 1984), nós podemos definir a entropia como uma medida de diversidade; isto é, como uma medida de heterogeneidade.

Como refinamentos de técnicas clássicas e soluções alternativas a problemas analíticos em aberto, métodos baseados em Teoria da Informação têm sido propostos em Inferência Estatística (Blatt & Hero 2007), Compressão de Dados (Donoho et al. 1998) e Processamento de Imagem (Morio, Réfrégier, Goudail, Dubois Fernandez & Dupuis 2009). Particularmente, o conceito de entropia tem recebido considerável atenção nestas propostas. Em Mecânica Estatística, esta medida é relacionada à noção de desordem (Kullback 1978, Rosso, De-Micco, Larrondo, Martín & Plastino 2010). Similarmente ao conceito de variância, esta medida está associada à variedade ou dispersão de um sistema (outras aplicações podem ser encontradas em Ullah 1996, Cover & Thomas 1991). O cálculo desta medida está baseado na configuração probabilística de um sistema dada a priori e redundante em um parâmetro que captura um tipo de variabilidade. Assim, tal medida se trata de um mapeamento de um espaço paramétrico em um parâmetro escalar real que mede variabilidade. Nesta seção, nós trabalhamos com estas medidas em dois passos: (i) derivação de expressões analíticas para entropias em modelos polarimétricos e (ii) consideração de teorias assintóticas baseadas no estimador MV das medidas derivadas.

Para satisfação do passo (ii), Salicrú et al. (1993) propuseram a classe de entropias  $(h, \phi)$ , a qual generaliza a proposta inicial de Burbea & Rao (1982b) apresentada na página 2. A seguir, nós reescrevemos a expressão da equação (4.2), dada na página 58, para  $\mu$  como uma medida de Lebesgue no espaço amostral do cone das matrizes positivas definidas Hermitianas,  $\mathcal{A}$ .

Seja  $\mathbf{Z}$  uma matriz aleatória com função densidade de probabilidade dada por  $f_{\mathbf{Z}}(\mathbf{Z}'; \boldsymbol{\theta})$  suportada pelo conjunto  $\mathcal{A}$ . A classe de entropias  $(h, \phi)$  relativa a  $\mathbf{Z}$  é dada por

$$H_{\phi}^h(\boldsymbol{\theta}) = h\left(\int_{\mathcal{A}} \phi(f_{\mathbf{Z}}(\mathbf{Z}'; \boldsymbol{\theta})) d\mathbf{Z}'\right),$$

em que  $h: \mathbb{R} \rightarrow \mathbb{R}$  é uma função crescente e  $\phi: [0, \infty) \rightarrow \mathbb{R}$  é concava, ou  $h$  decrescente

e  $\phi$  é convexa. Neste caso, o elemento diferencial  $d\mathbf{Z}'$  é dado por

$$d\mathbf{Z}' = \prod_{i=1}^m dZ_{ii} \prod_{\substack{i,j=1 \\ i < j}}^m d\Re\{Z_{ij}\} d\Im\{Z_{ij}\},$$

em que  $Z_{ij}$  representa a entrada  $(i, j)$  da matriz  $\mathbf{Z}'$ , e  $\Re$  e  $\Im$  representam as partes real e imaginária, respectivamente (Goodman 1963b). A Tabela 4.1 mostra as especificações de algumas entropias, dentre as quais três serão usadas na Seção 5.5: a saber, Shannon, Rényi, e Tsallis.

Com o objetivo de motivar o uso de entropia, vamos considerar seu cálculo para a família exponencial (FE). Uma variável aleatória  $\mathbf{g} \sim E(\boldsymbol{\theta})$  pertence a família exponencial multiparamétrica (com  $p$  parâmetros) se e somente se obedece a seguinte forma canônica:

$$f_{\mathbf{g}}(\mathbf{x}; \boldsymbol{\theta}) = \exp[\langle t(\mathbf{x}), c(\boldsymbol{\theta}) \rangle - F(\boldsymbol{\theta}) + k(\mathbf{x})] \quad (4.5)$$

( $\langle x, y \rangle = x^t y$  representa o produto interno), em que  $\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^d$ ,

- $t(\mathbf{x}) = [t_1(\mathbf{x}) \ t_2(\mathbf{x}) \ \cdots \ t_p(\mathbf{x})]^t: \mathcal{X} \rightarrow \mathbb{R}^p$  é uma estatística conjuntamente suficiente (Ferguson 1967),
- $k(\mathbf{x}): \mathcal{X} \rightarrow \mathbb{R}$ ,
- $c(\boldsymbol{\theta}) = [c_1(\boldsymbol{\theta}) \ c_2(\boldsymbol{\theta}) \ \cdots \ c_p(\boldsymbol{\theta})]: \Theta \rightarrow \mathbb{R}^p$ ,  $\boldsymbol{\theta} \in \Theta$ ,  $\Theta$  representa o espaço paramétrico
- e a constante log-normalizadora  $F(\boldsymbol{\theta}) = \log(\int_{\mathcal{X}} \exp[\langle t(\mathbf{x}), c(\boldsymbol{\theta}) \rangle + k(\mathbf{x})] d\mathbf{x})$  tem as quatro primeiras derivadas contínuas relativas ao vetor de parâmetros  $\boldsymbol{\theta}$ .

Este modelo é importante pois incorpora várias distribuições utilizadas em modelagem de dados, tais com gaussiana (uni e multivariada), gama, Poisson e binomial.

Neste caso, a medida de entropia de Shannon é dada pela seguinte expressão:

$$H_S(\mathbf{g}) = H_S(\boldsymbol{\theta}) = E(-\log f_{\mathbf{g}}(\mathbf{g}; \boldsymbol{\theta})) = F(\boldsymbol{\theta}) - \langle E(t(\mathbf{g})), c(\boldsymbol{\theta}) \rangle - E(k(\mathbf{g})).$$

Do Teorema 3.4.2 de Casella & Berger (2002, p. 112), temos que

$$E\left(\sum_{i=1}^m \frac{\partial c_i(\boldsymbol{\theta})}{\partial \theta_j} t_i(\mathbf{g})\right) = \frac{\partial F(\boldsymbol{\theta})}{\partial \theta_j}.$$

Então, temos que

$$\langle E(t(\mathbf{g})), c(\boldsymbol{\theta}) \rangle = \left\langle \left[ \frac{\partial c_i(\boldsymbol{\theta})}{\partial \theta_j} \right]_{i,j}^{-1} \frac{\partial F(\boldsymbol{\theta})}{\partial \theta_j}, c(\boldsymbol{\theta}) \right\rangle.$$



**Tabela 4.1**  $(h, \phi)$ -entropia e funções relacionadas

| $(h, \phi)$ -entropy                                                                                                                                          | $h(y)$                                                          | $\phi(x)$                                                 |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------|-----------------------------------------------------------|
| Shannon                                                                                                                                                       | $y$                                                             | $-x \ln x$                                                |
| $H_S(\theta) = E(-\log f_X)$ (Shannon 1948)                                                                                                                   |                                                                 |                                                           |
| Tsallis (ordem $\beta \in (0, \infty) \setminus \{1\}$ )                                                                                                      | $y$                                                             | $\frac{x^\beta - x}{1 - \beta}$                           |
| $H_T(\theta) = [1 - \beta]^{-1} [E(f_X^{\beta-1}) - 1]$ (Havrda & Chalvat 1967)                                                                               |                                                                 |                                                           |
| Rényi (ordem $\beta \in (0, \infty) \setminus \{1\}$ )                                                                                                        | $\frac{\log y}{1 - \beta}$                                      | $x^\beta$                                                 |
| $H_R(\theta) = [1 - \beta]^{-1} \log E(f_X^{\beta-1})$ (Rényi 1961)                                                                                           |                                                                 |                                                           |
| Varma ( $m - 1 < r < m, m \geq 1$ )                                                                                                                           | $(m - r)^{-1} \log y$                                           | $x^{r-m+1}$                                               |
| $H_V(\theta) = [m - r]^{-1} \log E(f_X^{r-m})$ (Varma 1966)                                                                                                   |                                                                 |                                                           |
| Arimoto (1971) ( $t \neq 1, t > 0$ )                                                                                                                          | $(t - 1)^{-1} (y^t - 1)$                                        | $x^{1/t}$                                                 |
| $H_A(\theta) = [t - 1]^{-1} [E^t(f_X^{1/t-1}) - 1]$ (Arimoto 1971)                                                                                            |                                                                 |                                                           |
| Sharma and Mittal ( $s \neq 1, s > 0$ )                                                                                                                       | $\frac{\exp[(s-1)y] - 1}{(1-s)}$                                | $x \log x$                                                |
| $H_{SM}(\theta) = [1 - s]^{-1} \{\exp[(s - 1) E(\log f_X)] - 1\}$ (Sharma & Mittal 1977)                                                                      |                                                                 |                                                           |
| Ferreri ( $\lambda > 0$ )                                                                                                                                     | $(1 + \frac{1}{\lambda}) \log(1 + \lambda) - \frac{y}{\lambda}$ | $(1 + \lambda x) \log(1 + \lambda x)$                     |
| $H_F(\theta) = (1 + \frac{1}{\lambda}) \log(1 + \lambda) - E\left(\frac{\log(1 + \lambda f_X)}{\lambda f_X}\right) + E(\log(1 + \lambda f_X))$ (Ferreri 1980) |                                                                 |                                                           |
| Kapur (1972) ( $s \neq 1$ )                                                                                                                                   | $y$                                                             | $\frac{x^s + (1-x)^s - 1}{(1-s)}$                         |
| $H_K(\theta) = [s - 1]^{-1} [E(f_X^{s-1}) + E\left(\frac{f_X^{s-1} - 1}{f_X}\right)]$ (Kapur 1972)                                                            |                                                                 |                                                           |
| Burbea                                                                                                                                                        | $y$                                                             | $\frac{x^s - (1+x)^s + 1 + (s-1)^{-1} (2^s - 2)x}{(s-2)}$ |
| $H_B(\theta) = [s - 2]^{-1} [E(f_X^{s-1}) - E\left(\frac{f_X^{s-1} - 1}{f_X}\right) + \frac{(2^s - 2)}{(s-1)}]$ (Burbea 1984)                                 |                                                                 |                                                           |

Finalmente, a entropia de Shannon para a FE é dada por

$$H_S(\boldsymbol{\theta}) = F(\boldsymbol{\theta}) - \left\langle \left[ \frac{\partial c_i(\boldsymbol{\theta})}{\partial \theta_j} \right]_{i,j}^{-1} \frac{\partial F(\boldsymbol{\theta})}{\partial \theta_j}, c(\boldsymbol{\theta}) \right\rangle - E(k(\mathbf{g})).$$

✎ Para o caso de retornos polarimétricos,  $\mathbf{Z} = \mathbf{Y} \cdot X$ , baseando-se na fórmula (3.29), temos a seguinte expressão para a entropia de Shannon:

$$H_S(\mathbf{Z}) = (m - L)[E(\log |\mathbf{Y}|) + m E(\log |X|)] - \log \frac{h_{\mathcal{M}}(\boldsymbol{\theta}, L)}{\Gamma_m(L)} - E(\log g(\text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}; \boldsymbol{\theta}, L))).$$

Como proposto por Anfinsen, Doulgeris & Eltoft (2009),

$$E(\log |\mathbf{Y}|) = \log |\boldsymbol{\Sigma}| + \psi_m^{(0)}(L) - m \log L$$

e, além disso, tem-se que  $E(\log |X|)$  é a esperança de uma das distribuições elencadas na Seção 3.3.2. Assim, para quaisquer distribuições polarimétricas, o único elemento intratável é

$$E \log g = E(\log g(\text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}; \boldsymbol{\theta}, L))).$$

Entretanto, em situações práticas em que temos uma sequência iid de matrizes  $\underline{\mathbf{Z}} = \{\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_N\}$  representando regiões (com  $N$  frequentemente grande), parece razoável, pela (i) lei fraca dos grandes números e pela (ii) consistência dos estimadores MV, propor que este termo seja substituído por

$$N^{-1} \sum_{i=1}^N \log g(\text{tr}(\tilde{\boldsymbol{\Sigma}}^{-1} \mathbf{Z}_i; \hat{\boldsymbol{\theta}}, \hat{L})),$$

em que  $\tilde{\boldsymbol{\Sigma}}$  representa a média amostral das  $N$  observações de  $\underline{\mathbf{Z}}$  e  $\hat{\boldsymbol{\theta}}$  e  $\hat{L}$  representam estimativas por MV.

✎ Para duas distribuições polarimétricas particulares, pode-se calcular os momentos  $E \log g$ :

- Para distribuição  $\mathcal{W}$ :  $E \log g = -mL$ . Neste caso, tanto a derivação para entropia quanto uma abordagem de teste assintótico para esta expressão será dada na Seção 5.5. Além disso, expressões para as entropias de Rényi e Tsallis são apresentadas nesta seção. Finalmente, testes assintóticos serão propostos e comparados àquele baseado na entropia de Shannon.
- Para distribuição  $\mathcal{G}_m^0$ :

$$E \log g = (\alpha - mL)[\psi^{(0)}(mL - \alpha) + \log(-\alpha - 1) - \psi^{(0)}(-\alpha)]$$

O tratamento algébrico deste resultado é dado no Apêndice D.

O seguinte resultado, derivado por Pardo et al. (1997), prepara o caminho para a proposta de métodos de Inferência Estatística Assintótica baseada em medidas de entropia.

**Lema 1.** *Seja o vetor  $p$ -variado  $\hat{\theta} = [\hat{\theta}_1 \ \hat{\theta}_2 \ \dots \ \hat{\theta}_p]^t$  o estimador de MV para  $\theta = [\theta_1 \ \theta_2 \ \dots \ \theta_p]^t$  baseado sobre uma amostra aleatória iid de tamanho  $N$  sob o modelo  $f(\mathbf{Z}'; \theta)$ . Então,*

$$\sqrt{N}[H_h^\phi(\hat{\theta}) - H_h^\phi(\theta)] \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \sigma_H^2(\theta)),$$

em que  $\mathcal{N}(\mu, \sigma^2)$  é a distribuição normal com média  $\mu$  e variância  $\sigma^2$ , ' $\xrightarrow{\mathcal{D}}$ ' representa a convergência em distribuição,

$$\sigma_H^2(\theta) = \delta^t \mathcal{K}(\theta)^{-1} \delta, \quad (4.6)$$

$\mathcal{K}(\theta) = E(-\partial^2 \ln f_Z(\mathbf{Z}; \theta) / \partial \theta^2)$  é matrix informação de Fisher esperada, e  $\delta = [\delta_1 \ \delta_2 \ \dots \ \delta_p]^t$  tal que  $\delta_i = \partial H_h^\phi(\theta) / \partial \theta_i$ , para  $i = 1, 2, \dots, p$ .

A seguir, nós introduziremos a metodologia de teste de hipótese e intervalos de confiança baseados em medidas de entropia.

#### 4.1.1 Teste de hipóteses

Seja  $\mathbf{Z}_i$  uma matriz aleatória com função densidade de probabilidade definida sobre  $\mathcal{A}$  e vetor de parâmetros  $\theta_i$ , para  $i = 1, 2, \dots, r$ , em que  $r$  é o número de populações sob estudo. Considere que nós estamos interessados em testar a hipótese  $\mathcal{H}_0: H_h^\phi(\theta_1) = H_h^\phi(\theta_2) = \dots = H_h^\phi(\theta_r) = v$  contra  $\mathcal{H}_1: H_h^\phi(\theta_i) \neq H_h^\phi(\theta_j)$  para algum  $i$  e  $j$ . Em outras palavras, evidência estatística é procurada para avaliar se, no mínimo, uma das  $r$  regiões de uma imagem PolSAR tem entropia diferente quando comparada com as restantes.

Seja  $\hat{\theta}_i$  o estimador de MV para  $\theta_i$  baseado sobre uma amostra iid de tamanho  $N_i$  sob  $\mathbf{Z}_i$ , para  $i = 1, 2, \dots, r$ . Do Lema 1, nós temos que

$$\frac{\sqrt{N_i}(H_h^\phi(\hat{\theta}_i) - v)}{\sigma_H(\hat{\theta}_i)} \xrightarrow[N_i \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, 1)$$

para  $i = 1, 2, \dots, r$ . Portanto,

$$\sum_{i=1}^r \frac{N_i(H_h^\phi(\hat{\theta}_i) - v)^2}{\sigma_H^2(\hat{\theta}_i)} \xrightarrow[N_i \rightarrow \infty]{\mathcal{D}} \chi_r^2. \quad (4.7)$$

Uma vez que  $v$  é, na prática, uma constante desconhecida, nós modificamos esta estatística. Considerando uma aplicação do teorema de Cochran (1934), o seguinte argumento é válido:

$$\sum_{i=1}^r \frac{N_i (H_h^\phi(\hat{\theta}_i) - v)^2}{\sigma_H^2(\hat{\theta}_i)} = \sum_{i=1}^r \frac{N_i (H_h^\phi(\hat{\theta}_i) - \bar{v})^2}{\sigma_H^2(\hat{\theta}_i)} + \sum_{i=1}^r \frac{N_i (\bar{v} - v)^2}{\sigma_H^2(\hat{\theta}_i)}, \quad (4.8)$$

em que

$$\bar{v} = \left[ \sum_{i=1}^r \frac{N_i}{\sigma_H^2(\hat{\theta}_i)} \right]^{-1} \sum_{i=1}^r \frac{N_i H_h^\phi(\hat{\theta}_i)}{\sigma_H^2(\hat{\theta}_i)}.$$

Salicrú et al. (1993) mostraram que a segunda soma do lado direito da equação (4.8) é distribuída como uma qui-quadrado com grau de liberdade um. Como o lado esquerdo de equação (4.8) possui uma distribuição qui-quadrado com  $r$  grau de liberdade (como mostrado na equação (4.7)), nós concluímos que:

$$\sum_{i=1}^r \frac{N_i (H_h^\phi(\hat{\theta}_i) - \bar{v})^2}{\sigma_H^2(\hat{\theta}_i)} \xrightarrow[N_i \rightarrow \infty]{\mathcal{D}} \chi_{r-1}^2.$$

Em particular, considere a seguinte estatística de teste:

$$S_\phi^h(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_r) = \sum_{i=1}^r \frac{N_i (H_h^\phi(\hat{\theta}_i) - \bar{v})^2}{\sigma_H^2(\hat{\theta}_i)}. \quad (4.9)$$

Os resultados acima demonstram a seguinte proposição.

**Proposição 1** (Salicrú et al. (1993)). *Seja  $N_i$ ,  $i = 1, 2, \dots, r$ , um tamanho de amostra suficientemente grande. Se  $S_\phi^h(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_r) = s$ , então a hipótese nula  $\mathcal{H}_0$  pode ser rejeitada no nível  $\alpha$  se  $\Pr(\chi_{r-1}^2 > s) \leq \alpha$ .*

#### 4.1.2 Intervalo de confiança

Seja  $\hat{\theta}$  o estimador de MV de  $\theta$  para um tamanho de amostra suficientemente grande  $N$ . Um intervalo de confiança aproximado para  $H_h^\phi(\theta)$  no nível nominal  $\alpha$  é dado por

$$H_h^\phi(\hat{\theta}) \pm z_{\alpha/2} \sqrt{\frac{\sigma_H^2(\hat{\theta})}{N}}, \quad (4.10)$$

em que  $z_{\alpha/2}$  é o  $\alpha/2$  quantil da distribuição normal padrão.

Considere agora  $\hat{\theta}_1$  e  $\hat{\theta}_2$  como estimativas de MV baseadas sob amostras suficientemente grandes  $N_1$  e  $N_2$ , respectivamente. Um intervalo de confiança aproximado para  $H_h^\phi(\theta_1) - H_h^\phi(\theta_2)$  é dado por

$$[H_h^\phi(\hat{\theta}_1) - H_h^\phi(\hat{\theta}_2)] \pm z_{\alpha/2} \sqrt{\frac{\sigma_H^2(\hat{\theta}_1)}{N_1} + \frac{\sigma_H^2(\hat{\theta}_2)}{N_2}}.$$

- ✎ Esta tese avança quanto à derivação de entropias e de metodologias baseada nestas medidas para realização de testes de homogeneidade e construção de intervalos de confiança em imagens PolSAR. Em particular, a Seção 5.5 e Apêndice E comportam os resultados neste sentido.

## 4.2 A classe de divergências $(h, \phi)$

Nesta seção, objetiva-se construir o conceito das distâncias estocásticas. Estas medidas são fortemente aplicadas a técnicas de processamento de imagem, tal como detecção de mudança. Como exemplo, Inglada & Mercier (2007) propuseram uma nova medida de similaridade baseada na distância de Kullback-Leibler. Neste caso, as densidades são aproximadas pela expansão de *Edgeworth* e aplicadas como argumentos da distância de Kullback-Leibler. Adicionalmente, Inglada & Mercier consideraram um estudo sobre detecção de mudança em imagens SAR monopolarizadas. Em termos polarimétricos, Morio et al. (2009) propuseram uma importante aplicação de medidas de contraste em detecção de mudanças. Neste caso, a distância de Bhattacharyya e a entropia de Shannon baseadas na distribuição gaussiana circular são utilizadas para descrever dados polarimétricos.

A seguir, nós aderimos à convenção de que uma “divergência” é alguma função não negativa entre duas medidas de probabilidade que obedece a propriedade de Definitividade, apresentada na página 70. Se a função é também simétrica, ela é chamada de “distância”. Finalmente, nós entendemos “métrica” como uma distância que satisfaz a desigualdade triangular (Capítulos 1 e 14 de Deza & Deza 2009).

Uma imagem pode ser entendida como um conjunto de regiões, as quais têm seus pixels descritos por uma variável aleatória que segue uma determinada distribuição. Baseando-se neste fato, Nascimento et al. (2010) propuseram medidas de contraste baseadas em divergências estocásticas para quantificar diferenças entre áreas de imagens SAR. Tais medidas estão incorporadas a uma classe que foi submetida a um tratamento analítico sistemático inicialmente por Csiszár (1967) e, posteriormente, estendido por Salicrú et al. (1994). Como resultado final, a classe de divergências  $(h, \phi)$  foi proposta por Salicrú et al. A seguir, esta classe será particularizada ao contexto de informações PolSAR.

Considere  $\mathbf{X}$  e  $\mathbf{Y}$  como duas matrizes aleatórias definidas sobre um mesmo suporte  $\mathbf{Z}' \in U \subset \mathcal{A}$  e equipadas com as densidades  $f_{\mathbf{X}}(\mathbf{Z}'; \theta_1)$  e  $f_{\mathbf{Y}}(\mathbf{Z}'; \theta_2)$ , respectivamente, em que  $\theta_1$  e  $\theta_2$  são vetores de parâmetros.

A classe de divergências  $(h, \phi)$  entre  $f_{\mathbf{X}}$  e  $f_{\mathbf{Y}}$  é definida por

$$D_{\phi}^h(\mathbf{X}, \mathbf{Y}) = h\left(\mathbb{E}_{\mathbf{Y}} \phi\left(\frac{f_{\mathbf{X}}(\mathbf{Y}; \theta_1)}{f_{\mathbf{Y}}(\mathbf{Y}; \theta_2)}\right)\right) = h\left(\int_U \phi\left(\frac{f_{\mathbf{X}}(\mathbf{Z}'; \theta_1)}{f_{\mathbf{Y}}(\mathbf{Z}'; \theta_2)}\right) f_{\mathbf{Y}}(\mathbf{Z}'; \theta_2) d\mathbf{Z}'\right) \quad (4.11)$$

em que  $h: (0, \infty) \rightarrow [0, \infty)$  é uma função estritamente crescente e  $\phi: (0, \infty) \rightarrow [0, \infty)$  é uma função convexa tal que  $0 \phi(0/0) = 0$ ,  $0 \phi(x/0) = \lim_{x \rightarrow \infty} \phi(x)/x$  e  $\phi(1) =$

$h(0) = 0$  e  $\phi''(1) > 0$ . Neste caso, o elemento diferencial  $d\mathbf{Z}'$  é dado por

$$d\mathbf{Z}' = \prod_{i=1}^m dZ'_{ii} \underbrace{\prod_{i,j=1}^m}_{i < j} d\Re\{Z'_{ij}\} d\Im\{Z'_{ij}\},$$

em que  $Z'_{ij}$  é a entrada  $(i, j)$  da matriz  $\mathbf{Z}'$  (a exemplo de Goodman 1963b). Esta expressão pode ser particularizada para variáveis aleatórias escalares. Considere duas variáveis aleatórias  $X$  e  $Y$  sobre o suporte  $Z' \in U \subset \mathbb{R}$  e com densidades  $f_X(Z'; \theta_1)$  e  $f_Y(Z'; \theta_2)$ , respectivamente. Assim, a expressão dada na equação (4.11) se degenera naturalmente na seguinte expressão

$$D_{\phi}^h(X, Y) = h\left(\int_U \phi\left(\frac{f_X(Z'; \theta_1)}{f_Y(Z'; \theta_2)}\right) f_Y(Z'; \theta_2) dZ'\right).$$

Divergências bem conhecidas são obtidas pelas escolhas adequadas de funções  $h$  e  $\phi$ . Nesta tese, nós analisamos as seguintes medidas: as divergências (i)  $\chi^2$  (Taneja 2006), (ii) Kullback-Leibler (Seghouane & Amari 2007) e (iii) Rényi (Abbasnejad & Arghami 2011), as distâncias segundo (iv) Bhattacharyya (Kailath 1967) e (v) Hellinger (Nascimento et al. 2010), a divergência (vi) aritmética-geométrica (Taneja 2006) e as distâncias (vii) triangular e (viii) média-harmônica (Taneja 2006).

Muito embora aptas a comparar variáveis aleatórias, as medidas de divergência frequentemente não satisfazem a desigualdade triangular e não são métricas, como avaliado em Burbea & Rao (1982c). Similarmente aos estudos apresentados por Nascimento et al. (2010) e por Goudail & Réfrégier (2004), nós nos referiremos as divergências simetrizadas como “distâncias” nesta tese. Seguindo este princípio, as divergências Kullback-Leibler, Rényi e aritmética-geométrica foram submetidas a um processo de simetriação,

$$d_{\phi}^h(\mathbf{X}, \mathbf{Y}) = \frac{D_{\phi}^h(\mathbf{X}, \mathbf{Y}) + D_{\phi}^h(\mathbf{Y}, \mathbf{X})}{2},$$

com o fim de serem obtidas distâncias (a exemplo de Seghouane & Amari 2007). Portanto, esta tese considera distância como se segue:

**Definição 5 (Distância estocástica).** Uma função  $d_{\phi}^h: U \times U \rightarrow \mathbb{R}$  é uma medida de distância se, para todo  $\mathbf{X}, \mathbf{Y} \in U \subset \mathcal{A}$ , as seguintes propriedades se verificam:

1.  $d_{\phi}^h(\mathbf{X}, \mathbf{Y}) \geq 0$  (Não-Negatividade).
2.  $d_{\phi}^h(\mathbf{X}, \mathbf{Y}) = d_{\phi}^h(\mathbf{Y}, \mathbf{X})$  (Simetria).
3.  $d_{\phi}^h(\mathbf{X}, \mathbf{Y}) = 0 \Leftrightarrow \mathbf{X} = \mathbf{Y}$  (Definitividade).

Note que, diferentemente da definição de distância usual, as medidas de distâncias estocásticas não necessariamente satisfazem a desigualdade triangular (Para maiores detalhes, consultar Burbea & Rao 1982c). A Tabela 4.2 apresenta a seleção das funções  $h$  e  $\phi$ , as quais geram cada uma das distâncias listadas a seguir.

Subsequentemente, apresentaremos as expressões para as distâncias  $(h, \phi)$  que serão discutidas neste trabalho bem como algumas aplicações ao contexto de Inferência Estatística. Por simplicidade, nós suprimimos a dependência explícita sobre a variável  $Z'$  e sobre o suporte  $U$ .

(i) A distância  $\chi^2$ :

$$d_{\chi^2}(\mathbf{X}, \mathbf{Y}) = \frac{1}{2}[D_{\chi^2}(\mathbf{X}, \mathbf{Y}) + D_{\chi^2}(\mathbf{Y}, \mathbf{X})] = \frac{1}{2} \int \frac{(f_{\mathbf{X}} - f_{\mathbf{Y}})^2}{2f_{\mathbf{X}}} + \int \frac{(f_{\mathbf{X}} - f_{\mathbf{Y}})^2}{2f_{\mathbf{Y}}}.$$

A divergência  $D_{\chi^2}$  tem sido aplicada em muitos contextos. Por exemplo, Broniatowski & Keziou (2009) propuseram um eficiente teste de hipótese combinando-a com técnicas de dualidade.

(ii) A distância de Kullback-Leibler:

$$\begin{aligned} d_{\text{KL}}(\mathbf{X}, \mathbf{Y}) &= \frac{1}{2}[D_{\text{KL}}(\mathbf{X}, \mathbf{Y}) + D_{\text{KL}}(\mathbf{Y}, \mathbf{X})] = \frac{1}{2} \left[ \int f_{\mathbf{X}} \log \frac{f_{\mathbf{X}}}{f_{\mathbf{Y}}} + \int f_{\mathbf{Y}} \log \frac{f_{\mathbf{Y}}}{f_{\mathbf{X}}} \right] \\ &= \frac{1}{2} \int (f_{\mathbf{X}} - f_{\mathbf{Y}}) \log \frac{f_{\mathbf{X}}}{f_{\mathbf{Y}}}. \end{aligned} \quad (4.12)$$

Eguchi & Copas (2006) apresentaram resultados mostrando que a divergência  $D_{\text{KL}}$  tem uma estreita relação com o lema de Neyman-Pearson. Adicionalmente, Seghouane & Amari (2007) utilizaram a simetrização desta medida de divergência como uma forma de correção para o critério de informação de Akaike (medida discutida na página 35), medida descritiva para avaliar a adequação de modelos estatísticos.

(iii) A distância de Rényi de ordem  $\beta$ :

$$\widetilde{d}_{\text{R}}^{\beta}(\mathbf{X}, \mathbf{Y}) = \frac{1}{2}[D_{\text{R}}^{\beta}(\mathbf{X}, \mathbf{Y}) + D_{\text{R}}^{\beta}(\mathbf{Y}, \mathbf{X})] = \frac{\log \int f_{\mathbf{X}}^{\beta} f_{\mathbf{Y}}^{1-\beta} + \log \int f_{\mathbf{X}}^{1-\beta} f_{\mathbf{Y}}^{\beta}}{2(\beta - 1)},$$

em que  $0 < \beta < 1$ . Andai (2009) utilizou a divergência  $D_{\text{R}}^{\beta}$  para estudar características geométricas com respeito a leis de probabilidades. Pela desigualdade de Fejér apresentada por Neuman (1990), nós temos que

$$d_{\text{R}}^{\beta}(\mathbf{X}, \mathbf{Y}) \triangleq \frac{1}{\beta - 1} \log \frac{\int f_{\mathbf{X}}^{\beta} f_{\mathbf{Y}}^{1-\beta} + \int f_{\mathbf{X}}^{1-\beta} f_{\mathbf{Y}}^{\beta}}{2} \leq \widetilde{d}_{\text{R}}^{\beta}(\mathbf{X}, \mathbf{Y}).$$



Note que expressão  $d_R^\beta$  pode ser colocada na estrutura  $(h, \phi)$ , diferentemente de  $\widetilde{d_R^\beta}$ . Assim, nós selecionamos  $d_R^\beta$  para as análises de testes de hipótese. Entretanto, esta última expressão redonda em formas analíticas mais complicadas e, por consequência, mais fadadas a instabilidades numéricas.

**(iv) A distância de Bhattacharyya:**

$$d_B(\mathbf{X}, \mathbf{Y}) = -\log \int \sqrt{f_X f_Y}.$$

Goudail et al. (2004) mostraram que essa distância é uma ferramenta mais precisa do que a distância de Kullback-Leibler quanto a definição de contraste em algoritmos para processamento de imagens PolSAR.

**(v) A distância de Hellinger:**

$$d_H(\mathbf{X}, \mathbf{Y}) = 1 - \int \sqrt{f_X f_Y}.$$

Giet & Lubrano (2008) empregaram métodos de estimação baseados na minimização de  $d_H$  em contextos de equações diferenciais estocásticas. Adicionalmente, resultados das propriedades assintóticas destes estimadores foram considerados.

**(vi) A distância aritmética-geométrica:**

$$\begin{aligned} d_{AG}(\mathbf{X}, \mathbf{Y}) &= \frac{D_{AG}(\mathbf{X}, \mathbf{Y}) + D_{AG}(\mathbf{Y}, \mathbf{X})}{2} = \int (f_X + f_Y) \log \left( \frac{f_Y + f_X}{2f_X} \right) \\ &+ \int (f_X + f_Y) \log \left( \frac{f_Y + f_X}{2f_Y} \right) = \frac{1}{2} \int (f_X + f_Y) \log \left( \frac{f_Y + f_X}{2\sqrt{f_Y f_X}} \right). \end{aligned}$$

**(vii) A distância triangular:**

$$d_T(\mathbf{X}, \mathbf{Y}) = \int \frac{(f_X - f_Y)^2}{f_X + f_Y}.$$

**(viii) A distância média-harmônica:**

$$d_{MH}(\mathbf{X}, \mathbf{Y}) = -\log \left( \int \frac{2f_X f_Y}{f_X + f_Y} \right) = -\log \left( 1 - \frac{d_T(\mathbf{X}, \mathbf{Y})}{2} \right).$$

Dentre as distâncias consideradas, a distância de Hellinger é a única limitada, apresentando valores no intervalo  $[0, 1]$ . Este fato é interessante pois tal medida termina funcionando como um percentual de discriminação. As demais distâncias variam nos reais não negativos.

Ao considerar a distância entre as distribuições de uma mesma família, apenas seus parâmetros são relevantes. Neste caso, os vetores de parâmetros  $\theta_1$  e  $\theta_2$  substituem os

**Tabela 4.2** Distâncias tipo  $(h, \phi)$  e funções  $h$  e  $\phi$  relacionadas

| $(h, \phi)$ -distância | $h(y)$                                                               | $\phi(x)$                                                            |
|------------------------|----------------------------------------------------------------------|----------------------------------------------------------------------|
| $\chi^2$               | $y/4$                                                                | $(x-1)^2(x+1)/x$                                                     |
| Kullback-Leibler       | $y/2$                                                                | $(x-1)\log(x)$                                                       |
| Rényi (order $\beta$ ) | $\frac{1}{\beta-1} \log((\beta-1)y+1), 0 \leq y < \frac{1}{1-\beta}$ | $\frac{x^{1-\beta}+x^\beta-\beta(x-1)-2}{2(\beta-1)}, 0 < \beta < 1$ |
| Bhattacharyya          | $-\log(-y+1), 0 \leq y < 1$                                          | $-\sqrt{x} + \frac{x+1}{2}$                                          |
| Hellinger              | $y/2, 0 \leq y < 2$                                                  | $(\sqrt{x}-1)^2$                                                     |
| aritmética-geométrica  | $y$                                                                  | $(\frac{x+1}{2}) \log(\frac{x+1}{2x}) + (\frac{x-1}{2})$             |
| triangular             | $y, 0 \leq y < 2$                                                    | $\frac{(x-1)^2}{x+1}$                                                |
| média-harmônica        | $-\log(-\frac{y}{2}+1), 0 \leq y < 2$                                | $\frac{(x-1)^2}{x+1}$                                                |

símbolos para variáveis aleatórias  $\mathbf{X}$  e  $\mathbf{Y}$  como os argumentos das medidas de divergências e de distâncias. Esta notação está em concordância com aquela apresentada por Salicrú et al. (1994) e será adotada no decurso desta tese.

A fim de motivar o uso de divergência, considere seu estudo no contexto de distribuições na família exponencial. Seja  $\mathbf{x} \sim E(\boldsymbol{\theta}_1)$  e  $\mathbf{y} \sim E(\boldsymbol{\theta}_2)$ , a seguinte derivação aborda o cálculo de  $d_{\text{KL}}(\mathbf{x}, \mathbf{y})$ . Aplicando a equação (4.5) na equação (4.12), temos o seguinte resultado:

$$\begin{aligned}
 D_{\text{KL}}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) &= \int_{\mathcal{X}} \log \frac{\exp\{c(\boldsymbol{\theta}_1)t(\mathbf{x}) - F(\boldsymbol{\theta}_1) + k(\mathbf{x})\}}{\exp\{c(\boldsymbol{\theta}_2)t(\mathbf{x}) - F(\boldsymbol{\theta}_2) + k(\mathbf{x})\}} f_g(\mathbf{x}; \boldsymbol{\theta}_1) d\mathbf{x} \\
 &= E(t(\mathbf{x}))^t [c(\boldsymbol{\theta}_1) - c(\boldsymbol{\theta}_2)] - [F(\boldsymbol{\theta}_1) - F(\boldsymbol{\theta}_2)], \tag{4.13}
 \end{aligned}$$

em que  $E(t(\mathbf{x})) = [\partial c_i(\boldsymbol{\theta}_1) / \partial \theta_j]_{i,j}^{-1} \partial F(\boldsymbol{\theta}_1) / \partial \theta_j$ . Assim,

$$\begin{aligned}
 d_{\text{KL}}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) &= \{E(t(\mathbf{x})) - E(t(\mathbf{y}))\}^t [c(\boldsymbol{\theta}_1) - c(\boldsymbol{\theta}_2)] \\
 &= \left\langle \left[ \frac{\partial c_i(\boldsymbol{\theta}_1)}{\partial \theta_{1j}} \right]_{i,j}^{-1} \frac{\partial F(\boldsymbol{\theta}_1)}{\partial \theta_{1j}} - \left[ \frac{\partial c_i(\boldsymbol{\theta}_2)}{\partial \theta_{2j}} \right]_{i,j}^{-1} \frac{\partial F(\boldsymbol{\theta}_2)}{\partial \theta_{2j}}, c(\boldsymbol{\theta}_1) - c(\boldsymbol{\theta}_2) \right\rangle.
 \end{aligned}$$

◻ Similarmente à subseção sobre entropias, consideremos a derivação da distância de Kulback-Leibler para as distribuições polarimétricas discutidas na Seção 3.3: Considere duas matrizes  $\mathbf{Z}_1$  e  $\mathbf{Z}_2$  aleatórias com densidade (conforme (3.29)):

$$f_{\mathbf{Z}_i}(\mathbf{Z}'; \boldsymbol{\theta}_i, L_i) = \frac{|\mathbf{Z}'|^{L_i-m}}{\Gamma_m(L)} h_{\mathcal{M}}(\boldsymbol{\theta}_i, L_i) g(\text{tr}(\boldsymbol{\Sigma}_i^{-1} \mathbf{Z}'); \boldsymbol{\theta}_i, L_i)$$

para  $i = 1, 2$ . Algumas manipulações algébricas nos levam ao seguinte resultado para

a distância de Kullback-Leibler:

$$\begin{aligned} d_{\text{KL}}([\theta_1, L_1], [\theta_2, L_2]) &= (L_1 - L_2)[E(\log |\mathbf{Y}_1|) - E(\log |\mathbf{Y}_2|) + E(\log X_1) - E(\log X_2)] \\ &\quad + \sum_{i,j=1}^2 \text{Elog} g_{ij}, \\ &= (L_1 - L_2) \left[ \log \frac{|\Sigma_1|}{|\Sigma_2|} + \psi_m^{(0)}(L_1) - \psi_m^{(0)}(L_2) - m \log \frac{L_1}{L_2} \right] \\ &\quad + [E(\log X_1) - E(\log X_2)] + \sum_{i,j=1}^2 \text{Elog} g_{ij} \end{aligned}$$

em que

$$\text{Elog} g_{ij} = E(\log g(\text{tr}(\Sigma_i^{-1} \mathbf{Z}_j; \theta_i, L_i))).$$

Maiores detalhes são dados no Apêndice E. Neste caso, em situações práticas com duas amostras suficientemente grandes  $\underline{\mathbf{Z}}_1 = \{\mathbf{Z}_{11}, \mathbf{Z}_{12}, \dots, \mathbf{Z}_{1N_1}\}$  e  $\underline{\mathbf{Z}}_2 = \{\mathbf{Z}_{21}, \mathbf{Z}_{22}, \dots, \mathbf{Z}_{2N_2}\}$ , este termo pode ser representado por

$$\widetilde{\text{Elog} g_{ij}} = N_i^{-1} \sum_{k=1}^{N_i} \log g(\text{tr}(\tilde{\Sigma}_j^{-1} \hat{\mathbf{Z}}_{ik}; \hat{\theta}_j, \hat{L}_j)).$$

em que  $\hat{\mathbf{Z}}_{ik}$  é a  $k$ -ésima matriz da amostra  $i$ ,  $\tilde{\Sigma}_j$  é a média amostral relativa ao vetor matricial  $\underline{\mathbf{Z}}_j$  e  $\hat{\theta}_j$  e  $\hat{L}_j$  representam os estimadores de MV sob a amostra  $\underline{\mathbf{Z}}_j$ .

✎ Para a distribuição Wishart complexa, este termo é dado por:

$$\widetilde{\text{Elog} g_{ij}} = \begin{cases} -mL_i, & \text{se } i = j \\ -L_i \text{tr}(\Sigma_i^{-1} \Sigma_j), & \text{se } i \neq j. \end{cases}$$

A partir deste resultado, a distância de Kullback-Leibler pode ser derivada. O Apêndice E apresenta a derivação de várias distâncias entre elementos da família Wishart complexa, incluindo  $d_{\text{KL}}$ , e a Seção 5.3 apresenta aplicações desta expressão bem como de outras distâncias aos contextos de processamento de imagens PolSAR.

Várias propriedades de convergência da classe  $(h, \phi)$  de divergências foram discutidas por Salicrú et al. (1994). Uma vez asseguradas as condições de regularidade dadas em Salicrú et al. (1994, p. 380), se os vetores de parâmetros  $\theta_1$  e  $\theta_2$  são iguais, então, com  $N_1, N_2 \rightarrow \infty$ , a quantidade

$$\frac{2N_1N_2}{N_1 + N_2} \frac{D_\phi^h(\hat{\theta}_1, \hat{\theta}_2)}{h'(0)\phi''(1)}$$

é assintoticamente  $\chi_p^2$  com  $p$  graus de liberdade. Nesta variável,  $\hat{\theta}_1 = (\hat{\theta}_{11}, \dots, \hat{\theta}_{1p})$  e  $\hat{\theta}_2 = (\hat{\theta}_{21}, \dots, \hat{\theta}_{2p})$  são estimadores de MV para  $\theta_1$  e  $\theta_2$  baseado sobre amostras aleatórias de tamanhos  $N_1$  e  $N_2$  (Salicrú et al. 1994), respectivamente.

Assim, considerando a definição de distância em termos das funções  $h$  e  $\phi$  bem como aplicando os resultados acima sobre convergência em distribuição das medidas  $(h, \phi)$  a uma lei  $\chi_p^2$  (Salicrú et al. 1994), o seguinte lema é provado.

**Lema 2.** *Considere válidas as condições de regularidade apresentadas por Salicrú et al. (1994). Se  $\frac{N_1}{N_1+N_2} \xrightarrow{N_1, N_2 \rightarrow \infty} \lambda \in (0, 1)$  e  $\theta_1 = \theta_2$ , então*

$$\frac{2N_1N_2}{N_1+N_2} \frac{d_\phi^h(\hat{\theta}_1, \hat{\theta}_2)}{h'(0)\phi''(1)} \xrightarrow[N_1, N_2 \rightarrow \infty]{\mathcal{D}} \chi_p^2.$$

Em particular, a seguinte estatística de teste é considerada:

$$S_\phi^h(\hat{\theta}_1, \hat{\theta}_2) = \frac{2N_1N_2v}{N_1+N_2} d_\phi^h(\hat{\theta}_1, \hat{\theta}_2),$$

em que  $v = [h'(0)\phi''(1)]^{-1}$  é uma constante que depende das funções escolhidas. Baseado sobre o Lema 2, o teste de hipóteses dependente da classe  $(h, \phi)$  para avaliar a hipótese nula  $\theta_1 = \theta_2$  pode ser derivado na seguinte proposição.

**Proposição 2.** *Supondo que  $N_1, N_2 \rightarrow \infty$  e  $S_\phi^h(\hat{\theta}_1, \hat{\theta}_2) = s$ , então a hipótese nula  $\theta_1 = \theta_2$  é rejeitada no nível  $\eta_1$  se  $\Pr(\chi_p^2 > s) \leq \eta_1$ .*

✎ A Seção 5.3 e o Apêndice E agrupam as principais contribuições desta tese com respeito a distâncias estocásticas e seus comportamentos assintóticos. Além disso, a Seção 5.4 aplica as ferramentas derivadas em detecção de borda.

Como considerações finais, os seguintes comentários merecem ser citados quanto à análise de imagem PolSAR. As Proposições 1 e 2 oferecem métodos de se refutar estatisticamente a hipótese que duas amostras distintas sejam classificadas como advindas de uma mesma região. Tal avaliação é relevante em diversos sentidos; por exemplo, quando escolhemos amostras de treinamento em processos de aprendizagem supervisionada ou em técnicas baseadas em medidas de similaridade, tal como segmentação.

# Estudos Desenvolvidos

Este capítulo faz uso dos princípios e derivações apresentados até o momento para propor contribuições práticas e teóricas no processamento e modelagem de imagens PolSAR. Em outras palavras, nós aprofundamos as contribuições brevemente comentadas na Seção 1.2.

Inicialmente, na Seção 5.1, uma imagem SAR é encarada como um “quebra-cabeça de pixels” em que cada peça é representada por uma variável aleatória escalar, distribuída conforme uma lei  $\mathcal{G}_I^0$ . A razoabilidade desta afirmação é justificável porque este modelo estocástico é considerado como uma distribuição universal (para detalhes, consultar Mejail et al. 2003); isto é, adequado a diferentes tipos de regiões de uma imagem SAR. Desta forma, recursos de Inferência Estatística, tais como estimação por MV, teste de hipóteses e medidas descritivas das estimativas obtidas, tentam mapear comportamentos e comparações entre as ferramentas desenvolvidas dentro de uma imagem SAR. Ainda, os resultados são submetidos a um modelo de contaminação espelho do fenômeno de reflexão de canto, discutido na Seção 2.2.

Como um segundo estudo, a Seção 5.2 apresenta duas formas melhoradas de estimação por MV para o número de equivalente de looks. Baseando-se na expressão geral de Cox & Snell (1968) para o viés de segunda ordem dos estimadores de MV, nós derivamos uma expressão para este termo no contexto de matrizes aleatórias modeladas pela distribuição Wishart complexa escalonada. Além disso, uma expressão para a verossimilhança perfilada modificada segundo Barndorff-Nielsen (1994) também é derivada para tais matrizes. A partir destas ferramentas, duas metodologias de estimação por MV melhorada são propostas.

Um terceiro estudo é apresentado na Seção 5.3. Neste caso, a idéia é entender os pixels como realizações de matrizes pertencentes a família de distribuições Wishart complexa escalonada. Inicialmente, Goodman (1963b) sugeriu a distribuição Wishart complexa como um modelo razoável a descrever dados Hermitianos positivos definidos. Conradsen et al. (2003) e Goudail & Réfrégier (2004) propuseram metodologias de teste de hipótese e algoritmos de processamento de imagens para dados PolSAR Wishart distribuídos. Subsequentemente, pesquisadores na área de sistemas PolSAR entenderam um escalonamento desta distribuição como uma proposta de modelo mais razoável. Nesta perspectiva, a Seção 5.3 propõe ferramentas estocásticas baseadas na Teoria da Informação (sob o funcional discutido na Seção 4.2) para o contexto

da distribuição Wishart complexa escalonada.

Posteriormente, a Seção 5.4 aplica o ferramental derivado na Seção 5.3 ao contexto de detecção de borda em imagens PolSAR. Seguindo (i) a metodologia proposta por Gambini, Mejail, Jacobo-Berlles & Frery (2006) baseada em curvas *B-spline* e (ii) o algoritmo *k*-médias proposto por Cao et al. (2007), as ferramentas derivadas são entendidas como quatro novos detectores de bordas e comparadas a uma abordagem usual.

Na Seção 5.5, um quinto estudo considera a proposta de testes de homogeneidade em imagens PolSAR via a utilização de entropias estocásticas. Esta metodologia é uma especificação da proposta de Salicrú et al. (1993) para a distribuição Wishart complexa escalonada e as entropias de Shannon e Rényi.

Finalmente, a Seção 5.6 discute a comparação de três testes de homogeneidade envolvendo ferramentas da Teoria da Informação com a abordagem baseada na razão de verossimilhança proposta por Conradsen et al. (2003). O desempenho dos testes é avaliada tendo como critério de comparação seus poderes e o tamanhos estimados assim como outras medidas descritivas.

## 5.1 Testes paramétricos e não paramétrico para dados speckle

Sistemas do tipo SAR têm sido sugeridos como métodos eficientes de sensoria-mento remoto. Pelo fato de serem obtidas por um processo de iluminação coerente, as imagens SAR são contaminadas pelo ruído speckle. Este ruído é bem descrito pela abordagem de modelo multiplicativo, que implica na eficiente e bem conhecida famí-lia de distribuições  $\mathcal{G}_I^0$  (vide Tabela 3.6 da página 50).

Vários métodos de estimação para seus parâmetros,  $\alpha$  (parâmetro associado à ru-gosidade) e  $\gamma$  (ao brilho), têm sido propostos, incluindo processos de redução de viés (Cribari-Neto, Frery & Silva 2002, Silva, Cribari-Neto & Frery 2008, Vasconcellos, Frery & Silva 2005), técnicas de robustez (Allende, Frery, Galbiati & Pizarro 2006, Bus-tos, Lucini & Frery 2002) e algoritmos utilizando pequenas amostras (Frery, Cribari-Neto & Souza 2004). Nesta tese, visando checar os resultados à luz de suas proprie-dades assintóticas apresentadas por Casella & Berger (2002), a estimação por MV dos parâmetros  $\alpha$  e  $\gamma$  é empregada.

Baseando-se em uma amostra aleatória de tamanho  $N$ ,  $z = (Z_1, Z_2, \dots, Z_N)$ , a função de verossimilhança relativa à distribuição  $\mathcal{G}_I^0(\alpha, \gamma, L)$  é dada por

$$\mathcal{L}(\alpha, \gamma; z') = \left( \frac{L^L \Gamma(L - \alpha)}{\gamma^\alpha \Gamma(-\alpha) \Gamma(L)} \right)^N \prod_{i=1}^N Z_i'^{L-1} (\gamma + LZ_i')^{\alpha-L}.$$

Assim, os estimadores para  $\alpha$  e  $\gamma$ , a saber  $\hat{\alpha}$  e  $\hat{\gamma}$ , respectivamente, são solução do sis-tema de equações não lineares:

$$\begin{cases} \psi^0(L - \hat{\alpha}) - \psi^0(-\hat{\alpha}) - \log(\hat{\gamma}) + \frac{1}{N} \sum_{i=1}^N \log(\hat{\gamma} + LZ_i') = 0, \\ -\frac{\hat{\alpha}}{\hat{\gamma}} + \frac{\hat{\alpha}-L}{N} \sum_{i=1}^N (\hat{\gamma} + LZ_i')^{-1} = 0. \end{cases}$$

Entretanto, o sistema de equações acima não possui solução em forma fechada, então o método de otimização numérica proposto por Broyden-Fletcher-Goldfarb-Shanno (BFGS), como discutido por Cribari-Neto & Zarkos (1999), é usado para estimação de parâmetros neste estudo.

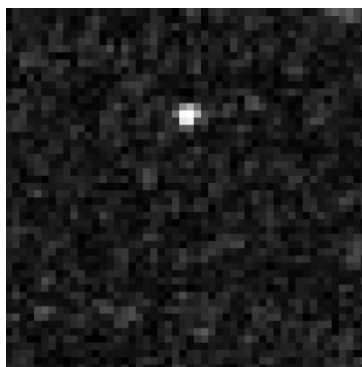
O entendimento do sistema de imageamento e a identificação dos elementos de uma cena (região de uma imagem SAR) são objetivos importantes neste campo. Em particular, medidas de contraste confiáveis são procuradas, a fim de discriminar for-mas em dados do tipo speckle. Para este fim, nós

apresentamos testes de hipóteses baseados sobre distâncias estocásticas propostos por Nascimento et al. (2010) e sobre a distância de Kolmogorov-Smirnov, tal como discutido por Beauchemin, Thomson & Edwards (1998), no tratamento com dados  $\mathcal{G}_I^0$ . Este trabalho pode ser visto como um complemento de (Nascimento et al. 2010) nos seguintes termos:

- (i) um estudo mais apurado de estimação MV é considerado, envolvendo cenários puros e contaminados (seguindo um modelo de contaminação paramétrico) de elementos da família  $\mathcal{G}_I^0$ ;
- (ii) adicionalmente, para ambos os cenários, os teste de hipóteses  $(h, \phi)$  são quantificados (adotando como critérios os tamanhos e poderes do teste) e comparados ao teste não paramétrico de Kolmogorov-Smirnov.

As distâncias  $(h, \phi)$  trabalhadas foram: Kullback-Leibler, triangular, Bhattacharyya e aritmética-geométrica. Os testes correspondentes a estas distâncias serão denominados por  $S_{KL}$ ,  $S_T$ ,  $S_B$  e  $S_{AG}$ , respectivamente, enquanto que o de Kolmogorov-Smirnov será rotulado por  $S_{KS}$ . Suas performances são quantificadas e comparadas de acordo com seus tamanhos e poderes do teste.

Na prática, é bastante difícil garantir que todas as observações em uma amostra venham de variáveis com a mesma distribuição. Ao coletar amostras por inspeção visual, frequentemente um conjunto de diferentes tipos de contaminação é experimentado. Um tipo de contaminação usual na obtenção de imagem SAR está relacionado ao fenômeno *double bounce*<sup>1</sup>. O efeito de *double bounce* é característico em áreas de florestas inundadas e de cidades. A presença de um pixel simples sofrendo de *double bounce* deve causar perda de precisão nas estimativas para os parâmetros e em estatísticas de teste, salvo se elas exibirem algum tipo de robustez à presença de *double bounce*. A Figura 5.1 exemplifica o efeito do fenômeno *double bounce* sobre um campo homogêneo.



**Figura 5.1** Efeito de uma reflexão de canto sobre um cenário homogêneo.

Objetivando avaliar a robustez dos testes considerados, nós supomos a presença de contaminação em um padrão paramétrico: os dados em uma amostra vem ou da distribuição  $\mathcal{G}_I^0$  (sob a hipótese nula) com probabilidade  $1 - \epsilon$ , ou vem de uma versão escalonada da mesma distribuição com probabilidade  $\epsilon$ . O escalonamento foi escolhido

<sup>1</sup>Alguns pixels exibem alto valor de retorno (pois a região envia de volta muito da energia recebida), causado tanto por efeitos naturais bem como pela ação do homem (para mais informações, consultar Frery, Nascimento & Cintra 2010b).



arbitráriamente, entretanto está baseado na forma do seguinte estudo prévio (Frery, Nascimento & Cintra 2010b). A distribuição escalonada é tal que seu valor médio é cem vezes maior do que o valor médio da distribuição original. Um estudo futuro considerará a substituição desta distribuição escalonada por outros modelos, tal como a distribuição de Ricean (Fannjiang, Solna & Yan 2009). Desejando escrever este procedimento em um quadro probabilístico, represente o evento “presença de outlier” pela letra  $A$  com  $\Pr(A) = \epsilon$  e defina  $A^c$  como seu complemento. Assim, pela regra de probabilidade total, a função distribuição acumulada do retorno é expressada por

$$\begin{aligned} F_Z(Z') &= \Pr(\{Z \leq Z'\}) = \Pr(\{Z \leq Z'\} \cap A) + \Pr(\{Z \leq Z'\} \cap A^c) \\ &= \Pr(A^c) \Pr(Z \leq Z' \mid A^c) + \Pr(A) \Pr(Z \leq Z' \mid A) \\ &= (1 - \epsilon) F_{\mathcal{G}^0(\alpha, \gamma, L)}(Z') + \epsilon F_{\mathcal{G}^0(\alpha, 100\gamma, L)}(Z'). \end{aligned}$$

Desta forma, usando experimentos de Monte Carlo,

✎ um estudo de robustez para diferentes graus de contaminação também foi considerado. Conforme Allende et al. (2006) e Bustos et al. (2002), nós escolhemos avaliar as situações sem contaminação ( $\epsilon = 0$ ) e com leves graus de contaminação ( $\epsilon \in \{10^{-4}, 5 \times 10^{-3}\}$ ).

Buscando satisfazer aos objetivos destacados acima, o manuscrito

### “Parametric and Nonparametric Tests for Speckled Imagery”

será anexado a esta seção. Apesar dos ensaios teóricos deste trabalho serem relacionados com as Seções 3.3.4 e 4.2, este artigo tem uma notação própria e pode ser analisado de forma independente da tese. Este fato evita que o leitor esteja retornando constantemente aos conceitos anteriores. A seguir, elencamos suas principais contribuições.

#### 5.1.1 Contribuições

##### **Linha teórica:**

Considerando  $\mathbf{X} \sim \mathcal{G}_I^0(\boldsymbol{\theta}_1) : \boldsymbol{\theta}_1 = (\alpha_1, \gamma_1, L_1)$  e  $\mathbf{Y} \sim \mathcal{G}_I^0(\boldsymbol{\theta}_2) : \boldsymbol{\theta}_2 = (\alpha_2, \gamma_2, L_2)$ , nós derivamos as seguintes expressões:

(i) A distância de Kullback-Leibler:

$$d_{\text{KL}}(\theta_1, \theta_2) = \sum_{i=1}^2 (-1)^i k_i d \int_0^\infty \frac{x^{a_i}}{(b_i + L_i x)^{c_i}} dx + \sum_{i=1}^2 (-1)^i k_i d \int_0^\infty \frac{\log(b_i + L_i x) x^{a_i}}{(b_i + L_i x)^{c_i}} dx +$$

$$\sum_{i=1}^2 (-1)^i k_i (L_2 - L_1) \int_0^\infty \frac{\log(x) x^{a_i}}{(b_i + L_i x)^{c_i}} dx - k_1 d \int_0^\infty \frac{x_1^a \log(b_2 + L_2 x)}{(b_1 + L_1 x)^{c_1}} dx +$$

$$k_2 d \int_0^\infty \frac{x^{a_2} \log(b_1 + L_1 x)}{(b_2 + L_2 x)^{c_2}} dx,$$

em que  $k_i = \frac{L_i^{L_i} \Gamma(L_i - \alpha_i)}{(\frac{\gamma_i}{2})^{\alpha_i} \Gamma(\alpha_i) \Gamma(L_i)}$ ,  $a_i = L_i - 1$ ,  $b_i = \frac{\gamma_i}{2}$ ,  $c_i = |\alpha_i| + L_i$  para  $i = 1, 2$  e  $d = \log(k_2/k_1)$ .

(ii) A distância de Rényi de ordem  $\beta$ :

$$d_{\text{R}}^\beta(\theta_1, \theta_2) = \frac{1}{\beta - 1} \left[ \log(h_1 \int_0^\infty x^{g_1} (b_1 + L_1 x)^{e_1} (b_2 + L_2 x)^{e_2} dx) \right.$$

$$\left. + \log(h_2 \int_0^\infty x^{g_2} (b_1 + L_1 x)^{m_1} (b_2 + L_2 x)^{m_2} dx) \right],$$

em que  $h_1 = k_2^\beta k_1^{1-\beta}$ ,  $h_2 = k_1^\beta k_2^{1-\beta}$ ,  $g_1 = (L_1 - 1)(1 - \beta) + (L_2 - 1)\beta$ ,  $g_2 = (L_1 - 1)\beta + (L_2 - 1)(1 - \beta)$ ,  $e_1 = (\alpha_1 - L_1)(1 - \beta)$ ,  $e_2 = (\alpha_2 - L_2)\beta$ ,  $m_1 = (\alpha_1 - L_1)\beta$  e  $m_2 = (\alpha_2 - L_2)(1 - \beta)$ , e  $0 < \beta < 1$ .

(iii) A distância Hellinger:

$$d_{\text{H}}(\theta_1, \theta_2) = 2(1 - \sqrt{k_1 k_2} \int_0^\infty x^{(a_1+a_2)/2} (b_1 + L_1 x)^{f_1} (b_2 + L_2 x)^{f_2} dx),$$

em que  $f_i = \frac{\alpha_i - L_i}{2}$ , para  $i = 1, 2$ .

(iv) A distância de Bhattacharyya:

$$d_{\text{B}}(\theta_1, \theta_2) = -\log(\sqrt{k_1 k_2} \int_0^\infty x^{(a_1+a_2)/2} (b_1 + L_1 x)^{f_1} (b_2 + L_2 x)^{f_2} dx).$$

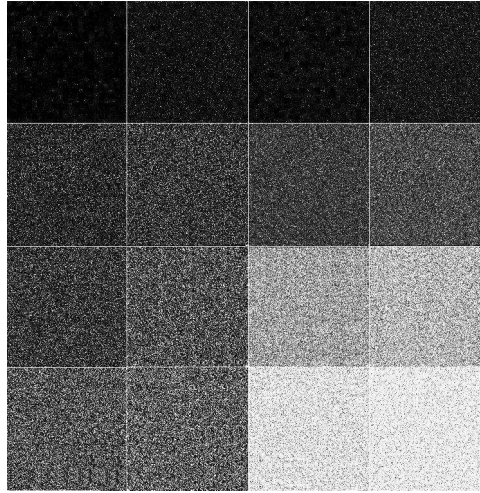
As integrais são apresentadas na Figura B.1 do Apêndice B. As outras distâncias listadas na Seção 4.2 são dependentes de integrais cujas formas não se mostraram harmônicas. Assim, foram omitidas da listagem acima. No processo de simulação via Monte Carlo presente nesta seção, as distâncias utilizadas foram obtidas numericamente pela função de integração numérica “QAGI” presente na linguagem de programação  $\text{\textcircled{O}x}$  (para detalhes, consultar Piessens, de Doncker-Kapenga, Uberhuber & Kahaner 1983).

**Linha aplicada:**

### 1) Discussão dos cenários de simulação:

Este estudo apresentou uma comparação entre os desempenhos de quatro testes paramétricos baseados em distâncias estocásticas e o teste de Kolmogorov-Smirnov. A avaliação foi feita com experimentos de Monte Carlo variando os parâmetros da distribuição  $\mathcal{G}_I^0$  e os níveis de contaminação, conforme seguinte discussão.

Dados do tipo speckle foram gerados conforme uma distribuição  $\mathcal{G}_I^0$ , visando avaliar estatísticas de teste paramétricas e uma não paramétrica em imagens de um único canal de polarização. As situações utilizadas consideraram três níveis de rugosidade ( $\alpha \in \{-1.5, -3, -5\}$  representando arbitrariamente cenários extremamente heterogêneos, heterogêneos e homogêneos), três valores médios da distribuição  $\mathcal{G}_I^0$  ( $\mu \in \{1, 2, 5\}$ , em que  $\mu = -\gamma/(1 + \alpha)$ ), os quais estão diretamente relacionados ao brilho da imagem, e dois valores para o número de looks ( $L \in \{1, 8\}$ ). A fim de ilustrar a interpretabilidade física dos parâmetros sobre uma imagem, considere a imagem sintética dada na Figura 5.2. Esta figura mostra dezesseis regiões  $\mathcal{G}_I^0$  distribuídas com  $256 \times 256$  pixels cada seguindo a seguinte configuração paramétrica:  $L = \{1, 8\}$ ,  $\alpha \in \{-15, -8, -5, -3\}$  e  $E(Z) = \mu \in \{1, 10\}$ . Nota-se que as imagens ficam mais brilhantes quando submetidas aos mais altos valores de  $\alpha$  e/ou  $\mu$ . Nós realizamos os testes ao



**Figura 5.2** Imagem sintética  $\mathcal{G}_I^0$  distribuída para  $(\mu, L) \in \{(1, 1), (1, 8), (10, 1), (10, 8)\}$  (da esquerda para direita) and  $\alpha \in \{-15, -8, -5, -3\}$  (de cima para baixo).

nível de significância 1%, utilizando janelas quadradas de  $7 \times 7$ ,  $9 \times 9$  e  $11 \times 11$  pixels; isto é, os tamanhos de amostra considerados foram  $N_X = N_Y = N \in \{49, 81, 121\}$ . Os tamanhos e poderes empíricos do teste foram utilizados como critérios para medir e comparar suas performances.

Experimentos de Monte Carlo (Rubinstein & Kroese 2008) em cinco diferentes cenários foram usados. Considerando duas regiões sobre uma imagem SAR especificadas pelos vetores de parâmetros  $(\alpha_1, \mu_1, L)$  e  $(\alpha_2, \mu_2, L)$ , os seguintes cenários foram

definidos: (i)  $\alpha_1 = \alpha_2, \mu_1 = \mu_2$ , (ii)  $\alpha_1 = \alpha_2, \mu_1 \neq \mu_2$ , (iii)  $\alpha_1 > \alpha_2, \mu_1 > \mu_2$ , (iv)  $\alpha_1 > \alpha_2, \mu_1 < \mu_2$  e (v)  $\alpha_1 \neq \alpha_2, \mu_1 = \mu_2$ . Quanto às suas interpretações práticas, a primeira situação corresponde às regiões de uma imagem com os mesmos brilhos e rugosidades. O cenário (ii) leva em consideração imagens com as mesmas rugosidades, mas com brilhos diferentes. Imagens com ambas as rugosidades e os brilhos distintos foram capturadas pelos cenários (iii) e (iv). Adicionalmente, o quinto cenário se atém às regiões de diferentes rugosidades e brilhos semelhantes.

## 2) Desempenho das estimativas MV sob cenários puros e contaminados:

Os resultados de simulação são baseados em 5500 replicações independentes e amostras foram consideradas válidas se  $\hat{\alpha} \in [10\alpha, \alpha/20]$ ; isto é, um tipo de censura foi aplicado, a exemplo de Nascimento et al. (2010). A Tabela 5.1 apresenta o viés estimado (B), o coeficiente de variação (CV%, em porcentagem) e o erro quadrático médio (EQM) dos estimadores de MV,  $\hat{\alpha}$  e  $\hat{\gamma}$ , de acordo com as várias condições de simulação e os diferentes graus de contaminação.

Primeiramente, consideramos a análise de viés sobre dados não contaminados. Nesta análise, apenas as seguintes situações apresentaram redução de viés com o aumento do tamanho de amostra: (a)  $\alpha = -1.5$ , (b)  $\alpha = -3, L = 8$  e (c)  $\alpha = -5, L = 8$ . Esta situação, aparentemente surpreendente, é frequente quando se lida com esta distribuição. Aqui, o conteúdo das informações é reduzido em situações homogêneas e heterogêneas e, portanto, as propriedades assintóticas dos estimadores MV não são verificadas nas amostras com os tamanhos considerados (problemas semelhantes foram discutidos por Cribari-Neto et al. 2002, Frery et al. 2004, Vasconcellos et al. 2005). Um aumento do número de looks pode compensar parcialmente esse comportamento. Esta perda do conteúdo da informação também é verificada pela dependência do viés ao número de looks. Espera-se que o primeiro seja reduzido aumentando o segundo, mas isto só é verificado quando  $\alpha = -1.5$ , ou seja, em áreas extremamente heterogêneas. Ao tratar dados heterogêneos ( $\alpha = -3$ ), esta redução só ocorre a partir de  $N = 81$  e com dados homogêneos ( $\alpha = -5$ ) em  $N = 121$ . Focando-se sobre a situação em que  $N = 121$  e  $L = 8$ , a qual é esperado revelar o comportamento mais preciso dos estimadores, o viés decresce quando  $\alpha$  é reduzido de  $-1.5$  a  $-5$ .

Considerando agora a influência da contaminação sobre o viés, em todos os casos a exceção de um (a saber,  $N = 49, L = 1, \epsilon = 10^{-4}$ ), o crescimento da primeira acarreta no aumento do segundo, para ambos os estimadores. Este aumento de viés varia aproximadamente de 34% ( $\alpha = -1.5, L = 1, \gamma = 0.5, \epsilon = 10^{-4}$ ) a 1000% ( $\alpha = -3, L = 1, \gamma = 2, \epsilon = 5 \times 10^{-3}$ ), em ambos os casos em que  $N = 49$ .

O coeficiente de variação é reduzido pelo aumento de  $L$  e/ou  $N$  e/ou  $\alpha$ , como esperado, em dados não contaminados. Na presença de contaminação, esses comportamentos são menores com respeito a  $\alpha$ .

O comportamento usual de redução de EQM com o aumento de  $L$  e/ou  $N$  e/ou  $\alpha$  é observado na ausência de contaminação. A presença de contaminação aumenta o

EQM, exceto em quatro das 324 situações consideradas, em que as flutuações aleatórias mascararam os resultados. Por exemplo, o EQM de  $\hat{\gamma}$  dobra na situação  $\alpha = -5$ ,  $\gamma = 20$ ,  $L = 1$ ,  $N = 49$  quando  $\epsilon = 10^{-4}$ , isto é, ao menor nível de contaminação. Se  $\alpha = -5$ ,  $\gamma = 20$ ,  $L = 8$ ,  $N = 121$  e  $\epsilon = 5 \times 10^{-3}$ , isto é, a mais forte contaminação, o EQM foi afetado por um fator de 13794. A forte influência da contaminação sobre o EQM alerta os pesquisadores de imagem SAR sobre a necessidade de amostras de qualidade quando se lida com este tipo de dados.

### 3) Performance dos testes em cenários puros e contaminados:

A Figura 5.3 mostra o tamanho empírico do teste, isto é, sob uma avaliação do cenário (i). Com a exceção de  $S_T$ , os níveis de significância estimados estão acima do nível nominal (mostrado pela linha tracejada). Os testes tendem a rejeitar mais frequentemente quando  $\alpha$  cresce, com pouco efeito do tamanho de amostra. Em geral,  $S_T$  é a medida mais eficiente considerando o tamanho do teste, uma vez que suas estimativas são em média mais próximas ao nível nominal;  $S_T$  é imediatamente seguida por  $S_{KS}$ . Como será visto posteriormente,  $S_{KS}$  surpreendentemente perde essa habilidade na presença de contaminação.

Goudail & Réfrégier (2004) apresentaram uma relação entre poder do teste e as distâncias de Kullback-Leibler e Bhattacharyya. Similarmente, nós ilustramos as estimativas dos poderes com respeito à medida de discrepância dada por

$$\bar{D} = (d_{KL} + d_T + d_B + d_{AG} + d_{KS}) / 5, \quad (5.1)$$

em que  $L$  é fixado, a distância de Kolmogorov-Smirnov,  $d_{KS}$ , é dada por

$$d_{KS} = \max_{-\infty < Z' < \infty} \left| L^L Z'^L \left\{ \frac{\Gamma(L - \hat{\alpha}_1)}{\hat{\gamma}_1^L \Gamma(-\hat{\alpha}_1)} {}_2F_1\left(L, L - \hat{\alpha}_1, L + 1, \frac{-LZ'}{\hat{\gamma}_1}\right) - \frac{\Gamma(L - \hat{\alpha}_2)}{\hat{\gamma}_2^L \Gamma(-\hat{\alpha}_2)} {}_2F_1\left(L, L - \hat{\alpha}_2, L + 1, \frac{-LZ'}{\hat{\gamma}_2}\right) \right\} \right|,$$

e, por questões de simplicidade, os argumentos  $(\hat{\theta}_1, \hat{\theta}_2)$  foram suprimidos. Note ainda que a medida (5.1) é imprecisa quando as diferentes distâncias assumem propriedades estatísticas significativamente distintas. Uma abordagem futura, visará substituir tal medida pela componente principal amostral associada ao maior autovalor da matriz de covariância amostral baseada nas distâncias. Por hora, na intenção de preservarmos a idéia utilizada no artigo em anexo, manteremos a medida (5.1).

As Tabelas 5.2 e 5.3 apresentam o poder empírico do teste nos cenários (ii) e (iii)-(v), respectivamente. Como um comportamento geral, o poder empírico do teste cresce com o número de looks ( $L$ ) e/ou tamanho de amostra ( $N$ ).

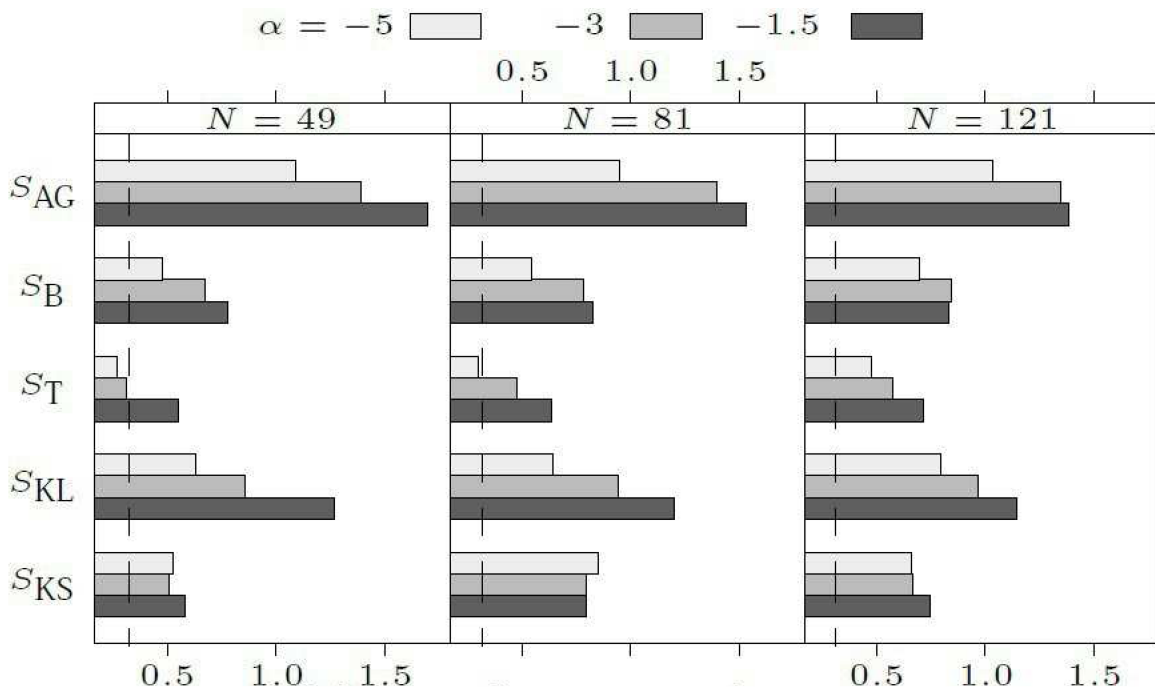
Considerando uma rugosidade fixada, a Tabela 5.2 fornece evidência que a mudança de brilho é mais influente com respeito ao poder em regiões homogêneas. A Figura 5.4 exhibe os resultados desta tabela com respeito a  $\bar{D}$  para diferentes tamanhos

Tabela 5.1 Estimativas para estimador de MV, viés (B), coeficiente de variação (CV%) e erro quadrático médio (EQM) em termos do grau de contaminação ( $\epsilon$ ) e dos parâmetros ( $\alpha, \gamma, L$ )

| $\alpha$  | $\gamma$ | $L$ | $\epsilon$         | $\hat{\alpha}$   |                  |                  | $\hat{\gamma}$     |                     |                     | $N = 121$           |                 |       |
|-----------|----------|-----|--------------------|------------------|------------------|------------------|--------------------|---------------------|---------------------|---------------------|-----------------|-------|
|           |          |     |                    | $N = 49$         | $N = 81$         | $N = 121$        | $N = 49$           | $N = 81$            | $N = 121$           |                     |                 |       |
| $-1.5$    | $0.5$    | $1$ | $0$                | $B$              | $CV\%$           | $EQM$            | $B$                | $CV\%$              | $EQM$               | $B$                 | $CV\%$          | $EQM$ |
|           |          |     |                    | 0.23 37.64 0.47  | 0.18 32.50 0.33  | 0.14 27.15 0.22  | 0.12 56.10 0.13    | 0.09 47.64 0.09     | 0.07 40.00 0.06     | 0.07 40.00 0.06     |                 |       |
|           |          |     |                    | 0.11 23.58 0.15  | 0.06 17.39 0.08  | 0.04 13.95 0.05  | 0.05 29.83 0.03    | 0.03 21.76 0.01     | 0.02 17.67 0.01     | 0.02 17.67 0.01     |                 |       |
|           |          |     |                    | 0.22 37.70 0.50  | 0.17 32.24 0.34  | 0.12 26.86 0.21  | 0.24 56.60 0.60    | 0.19 47.66 0.37     | 0.13 39.59 0.23     | 0.13 39.59 0.23     |                 |       |
|           |          |     |                    | 0.10 22.73 0.15  | 0.06 17.30 0.08  | 0.05 14.01 0.05  | 0.09 28.24 0.11    | 0.05 21.78 0.06     | 0.04 17.64 0.04     | 0.04 17.64 0.04     |                 |       |
|           |          |     |                    | 0.20 36.63 0.47  | 0.19 32.59 0.33  | 0.13 26.74 0.20  | 0.55 54.94 3.47    | 0.49 48.49 2.25     | 0.35 39.55 1.35     | 0.35 39.55 1.35     |                 |       |
|           |          |     |                    | 0.12 23.31 0.15  | 0.06 17.38 0.08  | 0.04 13.87 0.05  | 0.26 29.26 0.66    | 0.14 21.88 0.35     | 0.09 17.45 0.22     | 0.09 17.45 0.22     |                 |       |
|           |          |     |                    | 0.31 36.75 0.54  | 0.27 32.33 0.40  | 0.23 27.15 0.27  | 0.19 52.79 0.17    | 0.16 46.73 0.12     | 0.14 39.15 0.08     | 0.14 39.15 0.08     |                 |       |
|           |          |     |                    | 0.17 23.78 0.19  | 0.12 18.03 0.10  | 0.10 14.08 0.06  | 0.09 28.50 0.04    | 0.07 21.98 0.02     | 0.06 17.14 0.01     | 0.06 17.14 0.01     |                 |       |
|           |          |     |                    | 0.29 36.41 0.51  | 0.26 31.36 0.37  | 0.23 27.19 0.28  | 0.37 53.47 0.67    | 0.31 45.41 0.45     | 0.28 38.05 0.31     | 0.28 38.05 0.31     |                 |       |
| $10^{-4}$ | $1$      | $1$ | $10^{-3}$          | 0.17 23.51 0.18  | 0.13 17.45 0.10  | 0.10 14.16 0.06  | 0.18 28.72 0.15    | 0.14 20.98 0.08     | 0.12 17.53 0.05     | 0.12 17.53 0.05     | 0.12 17.53 0.05 |       |
|           |          |     |                    | 0.30 36.71 0.52  | 0.28 32.05 0.40  | 0.23 26.78 0.26  | 0.91 52.60 4.05    | 0.83 45.77 3.01     | 0.67 38.11 1.91     | 0.67 38.11 1.91     |                 |       |
|           |          |     |                    | 0.17 23.34 0.18  | 0.12 17.54 0.09  | 0.10 14.20 0.06  | 0.47 28.79 0.95    | 0.34 21.54 0.49     | 0.31 17.48 0.33     | 0.31 17.48 0.33     |                 |       |
|           |          |     |                    | 0.94 47.42 2.22  | 0.84 42.28 1.68  | 0.75 35.29 1.19  | 0.74 60.30 1.10    | 0.66 53.66 0.82     | 0.60 44.81 0.60     | 0.60 44.81 0.60     |                 |       |
|           |          |     |                    | 0.44 25.90 0.44  | 0.38 19.51 0.28  | 0.35 15.16 0.20  | 0.35 29.15 0.19    | 0.32 22.41 0.14     | 0.30 17.38 0.11     | 0.30 17.38 0.11     |                 |       |
|           |          |     |                    | 0.91 47.46 2.14  | 0.82 40.98 1.58  | 0.75 34.74 1.17  | 1.44 60.88 4.30    | 1.31 52.46 3.17     | 1.20 43.15 2.33     | 1.20 43.15 2.33     |                 |       |
|           |          |     |                    | 0.44 25.98 0.44  | 0.38 19.06 0.27  | 0.35 15.14 0.20  | 0.70 29.95 0.75    | 0.64 21.67 0.53     | 0.61 17.52 0.45     | 0.61 17.52 0.45     |                 |       |
|           |          |     |                    | 0.93 48.01 2.23  | 0.85 41.88 1.69  | 0.75 34.69 1.16  | 3.65 60.33 27.04   | 3.34 52.76 20.68    | 2.98 43.82 14.68    | 2.98 43.82 14.68    |                 |       |
|           |          |     |                    | 0.44 25.83 0.45  | 0.37 18.84 0.26  | 0.35 15.12 0.20  | 1.76 29.81 4.73    | 1.58 21.76 3.28     | 1.52 17.47 2.81     | 1.52 17.47 2.81     |                 |       |
|           |          |     |                    | 1.45 44.24 3.79  | 1.42 40.45 3.42  | 1.41 36.47 3.11  | 1.93 53.58 5.41    | 1.87 48.52 4.82     | 1.85 43.66 4.45     | 1.85 43.66 4.45     |                 |       |
| $10^{-3}$ | $1$      | $1$ | $5 \times 10^{-3}$ | 0.75 26.83 0.93  | 0.69 20.99 0.68  | 0.64 16.35 0.53  | 1.05 29.32 1.30    | 1.00 23.29 1.11     | 0.96 18.29 1.00     | 0.96 18.29 1.00     |                 |       |
|           |          |     |                    | 1.43 44.82 3.78  | 1.43 39.83 3.40  | 1.42 36.66 3.16  | 3.84 55.04 21.86   | 3.76 47.85 19.30    | 3.71 43.21 17.90    | 3.71 43.21 17.90    |                 |       |
|           |          |     |                    | 0.75 27.64 0.95  | 0.68 20.33 0.66  | 0.64 16.13 0.53  | 2.10 30.80 5.31    | 1.98 22.65 4.37     | 1.92 18.05 3.96     | 1.92 18.05 3.96     |                 |       |
|           |          |     |                    | 1.46 45.50 3.93  | 1.42 40.18 3.40  | 1.41 36.94 3.16  | 9.69 55.03 138.98  | 9.36 48.42 120.59   | 9.24 43.82 111.92   | 9.24 43.82 111.92   |                 |       |
|           |          |     |                    | 0.76 27.45 0.97  | 0.68 20.17 0.65  | 0.64 16.01 0.53  | 5.28 30.76 33.60   | 4.95 22.63 27.39    | 4.82 17.80 24.89    | 4.82 17.80 24.89    |                 |       |
|           |          |     |                    | 0.15 42.11 1.81  | 0.29 37.21 1.62  | 0.29 34.61 1.42  | 0.16 56.46 1.55    | 0.25 48.36 1.27     | 0.25 44.09 1.12     | 0.25 44.09 1.12     |                 |       |
|           |          |     |                    | 0.31 27.28 0.90  | 0.18 22.14 0.49  | 0.12 17.58 0.30  | 0.24 31.38 0.55    | 0.14 25.21 0.29     | 0.09 20.19 0.18     | 0.09 20.19 0.18     |                 |       |
|           |          |     |                    | 0.17 41.35 1.79  | 0.29 37.47 1.63  | 0.30 34.99 1.45  | 0.35 55.24 5.95    | 0.53 48.57 5.35     | 0.53 44.84 4.33     | 0.53 44.84 4.33     |                 |       |
|           |          |     |                    | 0.29 27.51 0.88  | 0.18 21.29 0.49  | 0.11 18.28 0.32  | 0.45 31.65 2.14    | 0.28 24.47 1.17     | 0.18 20.84 0.75     | 0.18 20.84 0.75     |                 |       |
|           |          |     |                    | 0.18 41.92 1.84  | 0.32 38.17 1.66  | 0.31 35.07 1.51  | 0.90 55.33 39.62   | 1.42 49.42 32.20    | 1.38 44.93 29.69    | 1.38 44.93 29.69    |                 |       |
| $10^{-2}$ | $1$      | $1$ | $10^{-2}$          | 0.28 27.79 0.88  | 0.16 21.15 0.51  | 0.12 17.93 0.29  | 1.09 31.91 13.19   | 0.65 24.30 7.71     | 0.48 20.61 4.36     | 0.48 20.61 4.36     |                 |       |
|           |          |     |                    | 0.32 41.92 2.03  | 0.43 37.48 1.84  | 0.46 34.04 1.60  | 0.33 54.06 1.69    | 0.42 48.17 1.54     | 0.44 43.13 1.31     | 0.44 43.13 1.31     |                 |       |
|           |          |     |                    | 0.40 27.47 1.03  | 0.29 21.30 0.57  | 0.23 17.48 0.37  | 0.35 31.15 0.66    | 0.26 24.49 0.37     | 0.22 19.85 0.24     | 0.22 19.85 0.24     |                 |       |
|           |          |     |                    | 0.33 40.13 1.90  | 0.47 37.01 1.87  | 0.47 33.76 1.59  | 0.75 52.75 6.85    | 0.92 47.45 6.29     | 0.91 43.08 5.30     | 0.91 43.08 5.30     |                 |       |
|           |          |     |                    | 0.38 27.10 0.99  | 0.29 21.72 0.59  | 0.23 17.18 0.36  | 0.68 30.78 2.53    | 0.53 24.77 1.53     | 0.43 19.70 0.94     | 0.43 19.70 0.94     |                 |       |
|           |          |     |                    | 0.31 41.12 1.95  | 0.40 36.90 1.74  | 0.48 33.94 1.63  | 1.68 54.01 42.65   | 2.05 47.10 36.42    | 2.31 42.67 32.91    | 2.31 42.67 32.91    |                 |       |
|           |          |     |                    | 0.39 27.33 1.01  | 0.30 21.61 0.60  | 0.23 17.03 0.36  | 1.68 30.98 15.93   | 1.39 24.67 9.82     | 1.09 19.26 5.74     | 1.09 19.26 5.74     |                 |       |
|           |          |     |                    | 0.84 35.35 2.54  | 1.14 31.11 2.95  | 1.30 28.34 3.18  | 1.22 44.44 3.55    | 1.49 38.04 3.98     | 1.65 34.06 4.28     | 1.65 34.06 4.28     |                 |       |
|           |          |     |                    | 1.07 25.91 2.24  | 1.03 21.99 1.84  | 0.97 18.74 1.50  | 1.27 28.36 2.49    | 1.24 24.24 2.15     | 1.18 20.27 1.82     | 1.18 20.27 1.82     |                 |       |
|           |          |     |                    | 0.93 34.41 2.68  | 1.16 30.90 3.01  | 1.34 27.64 3.22  | 2.70 43.67 15.85   | 3.04 37.61 16.23    | 3.35 33.24 17.21    | 3.35 33.24 17.21    |                 |       |
| $10^{-1}$ | $1$      | $1$ | $10^{-1}$          | 1.07 25.88 2.25  | 1.03 22.05 1.84  | 0.97 18.47 1.48  | 2.55 28.48 10.00   | 2.47 24.07 8.53     | 2.37 20.11 7.26     | 2.37 20.11 7.26     |                 |       |
|           |          |     |                    | 0.87 35.12 2.59  | 1.10 31.03 2.84  | 1.35 28.24 3.33  | 6.31 44.07 91.51   | 7.47 38.16 100.21   | 8.47 33.70 110.56   | 8.47 33.70 110.56   |                 |       |
|           |          |     |                    | 1.07 25.95 2.26  | 1.05 22.39 1.93  | 0.99 18.49 1.52  | 6.35 28.37 61.86   | 6.31 24.51 55.79    | 5.99 19.97 46.09    | 5.99 19.97 46.09    |                 |       |
|           |          |     |                    | 1.64 27.29 4.29  | 1.93 24.11 5.12  | 2.19 20.16 5.91  | 3.99 33.89 20.00   | 4.33 28.75 22.08    | 4.72 23.54 24.78    | 4.72 23.54 24.78    |                 |       |
|           |          |     |                    | 2.03 20.87 5.22  | 2.18 18.18 5.62  | 2.28 15.81 5.87  | 4.05 23.00 18.33   | 4.22 19.88 19.32    | 4.34 17.00 19.96    | 4.34 17.00 19.96    |                 |       |
|           |          |     |                    | 1.62 28.30 4.34  | 1.96 23.49 5.18  | 2.26 19.92 6.20  | 7.85 34.20 77.99   | 8.79 28.54 90.59    | 9.61 23.78 102.92   | 9.61 23.78 102.92   |                 |       |
|           |          |     |                    | 2.03 20.69 5.23  | 2.18 17.96 5.64  | 2.27 15.97 5.88  | 8.10 22.78 73.24   | 8.47 19.56 77.78    | 8.67 17.21 79.98    | 8.67 17.21 79.98    |                 |       |
|           |          |     |                    | 1.52 28.95 4.02  | 1.88 23.62 4.86  | 2.14 21.23 5.76  | 19.13 35.19 471.04 | 21.52 27.91 540.58  | 23.29 25.41 614.05  | 23.29 25.41 614.05  |                 |       |
|           |          |     |                    | 2.05 20.42 5.27  | 2.19 18.12 5.69  | 2.30 15.95 6.00  | 20.35 22.38 460.41 | 21.22 19.70 488.21  | 21.85 17.19 507.52  | 21.85 17.19 507.52  |                 |       |
|           |          |     |                    | 0.33 46.97 4.85  | 0.02 42.92 4.75  | 0.22 39.77 4.27  | 0.30 58.61 4.73    | 0.03 51.94 4.58     | 0.22 47.82 3.90     | 0.22 47.82 3.90     |                 |       |
| $10^0$    | $1$      | $1$ | $10^0$             | 0.52 30.38 3.01  | 0.35 28.62 2.04  | 0.13 26.71 1.28  | 0.47 33.62 2.42    | 0.33 31.08 1.61     | 0.13 28.51 1.02     | 0.13 28.51 1.02     |                 |       |
|           |          |     |                    | 0.30 46.86 5.11  | 0.06 42.50 4.75  | 0.32 40.23 4.23  | 0.48 57.82 19.89   | 0.05 51.74 17.98    | 0.64 48.68 15.97    | 0.64 48.68 15.97    |                 |       |
|           |          |     |                    | 0.51 30.85 3.14  | 0.28 29.04 2.08  | 0.11 27.77 1.28  | 0.92 34.18 10.23   | 0.52 31.38 6.64     | 0.21 29.41 4.06     | 0.21 29.41 4.06     |                 |       |
|           |          |     |                    | 0.25 47.26 4.99  | 0.00 43.57 4.74  | 0.18 39.53 4.13  | 1.09 58.27 122.69  | 0.05 53.04 110.26   | 0.89 47.03 98.05    | 0.89 47.03 98.05    |                 |       |
|           |          |     |                    | 0.52 30.61 2.99  | 0.31 28.25 2.04  | 0.13 27.63 1.30  | 2.28 33.64 60.17   | 1.43 30.69 39.68    | 0.63 29.41 25.36    | 0.63 29.41 25.36    |                 |       |
|           |          |     |                    | 0.14 46.39 5.11  | 0.20 42.56 4.93  | 0.39 37.67 4.26  | 0.02 57.34 5.21    | 0.27 51.54 4.93     | 0.45 44.61 4.15     | 0.45 44.61 4.15     |                 |       |
|           |          |     |                    | 0.70 30.21 3.44  | 0.57 25.28 2.31  | 0.46 21.34 1.57  | 0.68 33.27 2.90    | 0.57 27.83 1.95     | 0.47 23.43 1.32     | 0.47 23.43 1.32     |                 |       |
|           |          |     |                    | 0.83 38.48 5.72  | 1.38 33.55 6.48  | 1.76 30.81 7.45  | 3.05 46.15 35.33   | 4.26 39.08 41.14    | 4.94 35.14 45.13    | 4.94 35.14 45.13    |                 |       |
|           |          |     |                    | 1.80 26.35 6.47  | 1.89 23.43 6.20  | 1.92 20.74 5.76  | 4.68 28.27 34.79   | 4.88 25.32 34.44    | 4.91 21.99 32.12    | 4.91 21.99 32.12    |                 |       |
|           |          |     |                    | 0.77 38.63 5.57  | 1.32 33.96 6.36  | 1.71 30.66 7.16  | 7.69 47.46 231.81  | 10.21 39.33 245.49  | 12.31 35.57 283.43  | 12.31 35.57 283.43  |                 |       |
| $10^1$    | $1$      | $1$ | $10^1$             | 1.88 25.52 6.13  | 1.93 20.86 5.82  | 1.19 29.86 22.98 | 11.90 46.62 24.98  | 12.11 25.08 21.47   | 12.35 22.24 204.18  | 12.35 22.24 204.18  |                 |       |
|           |          |     |                    | 1.90 31.41 25.32 | 2.66 26.17 29.23 | 3.39 19.56 41.10 | 5.69 35.86 100.86  | 6.88 30.02 115.62   | 7.95 22.52 151.55   | 7.95 22.52 151.55   |                 |       |
|           |          |     |                    | 3.52 18.38 38.51 | 3.97 14.84 42.14 | 4.21 12.92 44.56 | 7.56 20.03 124.35  | 8.19 16.09 133.49   | 8.49 14.03 143.03   | 8.49 14.03 143.03   |                 |       |
|           |          |     |                    | 2.74 24.23 32.54 | 3.19 22.72 32.54 | 3.19 22.72 32.54 | 12.41 30.54 441.66 | 13.86 27.47 498.89  | 15.01 25.64 640.03  | 15.01 25.64 640.03  |                 |       |
|           |          |     |                    | 3.46 18.40 38.25 | 3.89 15.32 40.99 | 4.24 12.72 44.41 | 14.95 19.77 490.83 | 16.12 16.65 522.42  | 17.09 13.67 558.04  | 17.09 13.67 558.04  |                 |       |
|           |          |     |                    | 2.07 29.49 19.21 | 2.48 23.50 25.52 | 3.26 21.64 36.30 | 30.16 34.71 199.53 | 34.69 27.82 2571.16 | 38.22 23.54 3357.49 | 38.22 23.54 3357.49 |                 |       |
|           |          |     |                    | 3.45 19.05 37.62 | 3.88 15.45 41.52 | 4.19 13.21 44.84 | 37.37 20.75 307.61 | 40.32 16.61 3304.94 | 42.35 14.20 3523.39 | 42.35 14.20 3523.39 |                 |       |
|           |          |     |                    | 0.84 35.35 2.54  | 1.14 31.11 2.95  | 1.30 28.34 3.18  | 1.22 44.44 3.55    | 1.49 38.04 3.98     | 1.65 34.06 4.28     | 1.65 34.06 4.28     |                 |       |
|           |          |     |                    | 1.07 25.91 2.24  | 1.03 21.99 1.84  | 0.97 18.74 1.50  | 1.27 28.36 2.49    | 1.24 24.24 2.15     | 1.18 20.27 1.82     | 1.18 20.27 1.82     |                 |       |
|           |          |     |                    | 0.93 34.41 2.68  | 1.16 30.90 3.01  | 1.34 27.64 3.22  | 2.70 43.67 15.85   | 3.04 37.61 16.23    | 3.35 33.24 17.21    | 3.35 33.24 17.21    |                 |       |

de amostra. Neste caso, nós notamos que os testes paramétricos superam o teste de Kolmogorov-Smirnov.

Na Tabela 5.3, a desigualdade entre os respectivos poderes  $\text{Poder}(S_{KS}) \leq \text{Poder}(S_T) \leq \text{Poder}(S_B) \leq \text{Poder}(S_{KL}) \leq \text{Poder}(S_{AG})$  é satisfeita em todos os casos dos cenários (iii) e (iv). Para o cenário (v), este fato segue quando o número de looks é 8. A Figura 5.5 apresenta os poderes como uma função de  $\bar{D}$  para diferentes pares de rugosidade. O poder aumenta com o tamanho de amostra e com  $\bar{D}$ , como esperado. Esses resultados mostram que o teste baseado em  $S_{AG}$  supera todas as outras técnicas com respeito ao poder do teste, uma vez que ela é mais apta para discriminar amostras provenientes de diferentes distribuições.



**Figura 5.3** Valores para o tamanho empírico do teste médio no nível nominal 1% multiplicado por um fator de 10.

Baseando-se nos tamanhos e poderes empíricos do teste, o seguinte estudo de robustez foi limitado às melhores estatísticas, a saber  $S_T$  e  $S_{AG}$ , juntamente com  $S_{KS}$ , a qual é esperada ser robusta. Para essas medidas, a Figura 5.6 mostra o comportamento do tamanho do teste na presença de níveis fixados de contaminação.

Considerando a presença de contaminação, tamanhos de amostra maiores implicam que os melhores tamanhos empíricos do teste estão relacionados aos testes paramétricos. O teste baseado sobre  $S_T$  é o mais robusto para regiões homogêneas. Os tamanhos empíricos do teste relativos à classe  $(h, \phi)$  são maiores para  $\alpha = -1.5$  do que para aqueles observados quando  $\alpha \in \{-3, -5\}$ ; esse comportamento é mais pronunciado quando  $\epsilon$  cresce.

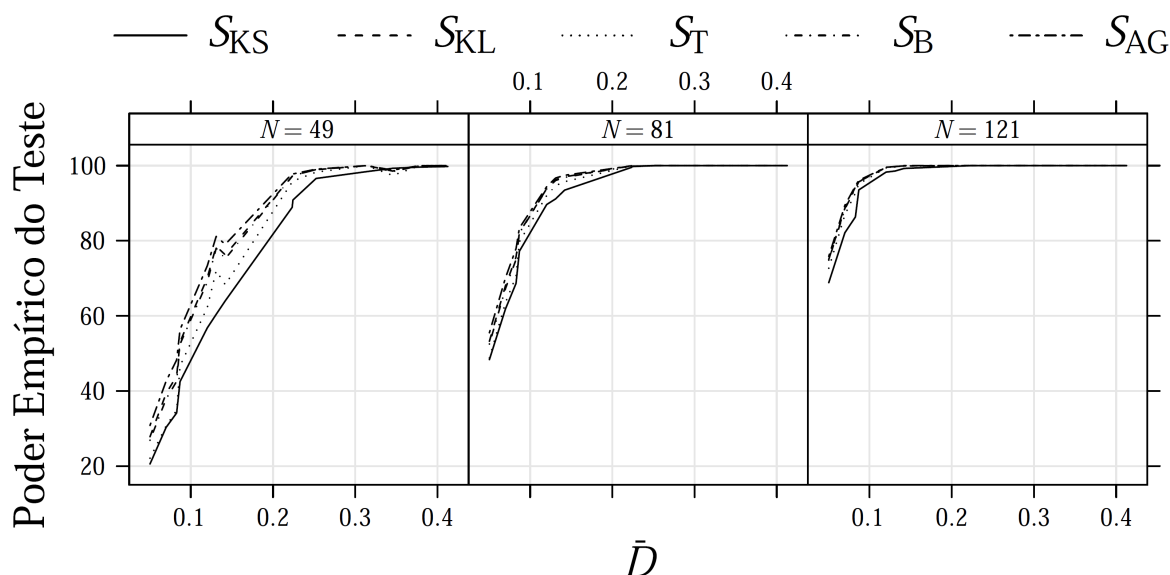
**Tabela 5.2** Taxas de rejeição dos testes relativos às distâncias tipo  $(h, \phi)$  e a de Kolmogorov-Smirnov para o cenário (ii)

| $\mu_2/\mu_1$ | $L$ | $N$ | $\alpha$ | $S_{KS}$ | $S_{KL}$ | $S_T$  | $S_B$  | $S_{AG}$ | $\alpha$ | $S_{KS}$ | $S_{KL}$ | $S_T$  | $S_B$  | $S_{AG}$ |
|---------------|-----|-----|----------|----------|----------|--------|--------|----------|----------|----------|----------|--------|--------|----------|
| 2.0           | 1   | 49  | -1.5     | 20.58    | 27.84    | 22.11  | 26.89  | 30.85    | -3.0     | 30.31    | 39.34    | 30.70  | 37.87  | 42.63    |
|               |     | 81  |          | 48.36    | 53.24    | 48.79  | 52.51  | 55.52    |          | 61.85    | 68.03    | 63.52  | 67.26  | 70.14    |
|               |     | 121 |          | 68.85    | 75.04    | 72.68  | 74.84  | 75.85    |          | 82.13    | 88.64    | 86.89  | 88.42  | 89.42    |
|               | 8   | 49  |          | 60.40    | 78.53    | 72.00  | 78.02  | 81.15    |          | 90.85    | 97.41    | 95.89  | 97.34  | 97.78    |
|               |     | 81  |          | 91.13    | 96.20    | 94.65  | 96.04  | 96.69    |          | 99.75    | 99.84    | 99.82  | 99.84  | 99.84    |
|               |     | 121 |          | 98.55    | 99.73    | 99.62  | 99.73  | 99.75    |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |
| 2.5           | 1   | 49  |          | 42.51    | 53.07    | 46.66  | 52.46  | 56.36    |          | 56.80    | 70.15    | 62.23  | 69.00  | 73.33    |
|               |     | 81  |          | 77.22    | 82.18    | 79.73  | 82.05  | 83.41    |          | 89.64    | 93.87    | 92.03  | 93.78  | 94.38    |
|               |     | 121 |          | 93.53    | 95.73    | 95.13  | 95.68  | 96.08    |          | 98.27    | 99.54    | 99.40  | 99.50  | 99.56    |
|               | 8   | 49  |          | 88.89    | 97.16    | 95.65  | 97.13  | 97.85    |          | 99.60    | 99.96    | 99.95  | 99.96  | 99.98    |
|               |     | 81  |          | 99.53    | 99.95    | 99.91  | 99.95  | 99.96    |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |
|               |     | 121 |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |
| 5.0           | 1   | 49  |          | 96.58    | 98.91    | 98.27  | 98.95  | 99.06    |          | 99.29    | 98.55    | 97.39  | 98.42  | 98.81    |
|               |     | 81  |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |
|               |     | 121 |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |
|               | 8   | 49  |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |
|               |     | 81  |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |
|               |     | 121 |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |



Tabela 5.3 Taxas de rejeição dos testes relativos às distâncias tipo  $(l, \phi)$  e a de Kolmogorov-Smirnov para os cenários (iii)-(v)

| $\alpha$               | Constantes |         | $\gamma$               | $\mu_1 > \mu_2$ |          |        |        |          | $\gamma$               | $\mu_1 = \mu_2$ |          |        |        |           | $\gamma$               | $\mu_1 < \mu_2$ |          |        |        |          |
|------------------------|------------|---------|------------------------|-----------------|----------|--------|--------|----------|------------------------|-----------------|----------|--------|--------|-----------|------------------------|-----------------|----------|--------|--------|----------|
| $(\alpha_1, \alpha_2)$ | $L$        | $N$     | $(\gamma_1, \gamma_2)$ | $S_{KS}$        | $S_{KL}$ | $S_T$  | $S_B$  | $S_{AG}$ | $(\gamma_1, \gamma_2)$ | $S_{KS}$        | $S_{KL}$ | $S_T$  | $S_B$  | $S_{AG}$  | $(\gamma_1, \gamma_2)$ | $S_{KS}$        | $S_{KL}$ | $S_T$  | $S_B$  | $S_{AG}$ |
| $(-1.5, -3)$           | 1          | 49      | (1,2)                  | 1.05            | 7.12     | 1.17   | 3.53   | 11.94    | (0.5,2)                | 13.45           | 16.15    | 11.64  | 15.14  | 18.56     | (0.5,4)                | 82.27           | 89.51    | 86.89  | 89.48  | 90.47    |
|                        |            | 81      |                        | 2.47            | 14.10    | 3.45   | 9.46   | 20.33    |                        | 31.80           | 35.39    | 30.38  | 34.11  | 38.25     |                        | 98.53           | 99.28    | 99.20  | 99.26  | 99.35    |
|                        |            | 121     |                        | 2.64            | 24.98    | 10.20  | 19.17  | 31.83    |                        | 51.13           | 58.36    | 54.75  | 57.68  | 60.30     |                        | 99.89           | 100.00   | 100.00 | 100.00 | 100.00   |
|                        | 8          | 49      |                        | 2.58            | 29.18    | 12.68  | 23.52  | 36.20    |                        | 62.35           | 92.41    | 87.65  | 91.71  | 94.06     |                        | 99.98           | 100.00   | 100.00 | 100.00 | 100.00   |
|                        |            | 81      |                        | 7.33            | 50.77    | 36.49  | 47.39  | 56.12    |                        | 93.31           | 99.51    | 99.16  | 99.45  | 99.53     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        |            | 121     |                        | 12.40           | 73.73    | 64.55  | 71.93  | 77.13    |                        | 99.22           | 99.85    | 99.84  | 99.84  | 99.85     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        | 1          | 49      | (2.5,2)                | 69.45           | 89.53    | 82.24  | 87.95  | 91.89    | (1,4)                  | 13.78           | 15.55    | 10.85  | 14.41  | 18.28     | (0.5,10)               | 99.95           | 100.00   | 100.00 | 100.00 | 100.00   |
|                        |            | 81      |                        | 96.51           | 99.24    | 98.39  | 99.17  | 99.45    |                        | 31.65           | 35.47    | 30.65  | 34.15  | 37.74     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        |            | 121     |                        | 99.62           | 100.00   | 99.98  | 100.00 | 100.00   |                        | 49.98           | 58.51    | 54.67  | 57.57  | 60.48     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        | 8          | 49      |                        | 99.49           | 100.00   | 100.00 | 100.00 | 100.00   |                        | 62.35           | 92.19    | 87.23  | 91.47  | 94.03     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        |            | 81      |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |                        | 92.93           | 99.55    | 99.29  | 99.49  | 99.62     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        |            | 121     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |                        | 99.47           | 99.98    | 99.96  | 99.98  | 100.00    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
| $(-1.5, -5)$           | 1          | 49      | (2.5,4)                | 5.29            | 18.77    | 6.70   | 13.06  | 24.88    | (2.5,10)               | 13.71           | 14.83    | 10.65  | 14.00  | 17.47     | (1,10)                 | 94.65           | 97.74    | 96.96  | 97.74  | 97.92    |
|                        |            | 81      |                        | 12.44           | 37.45    | 19.88  | 31.21  | 44.52    |                        | 32.27           | 33.52    | 28.79  | 32.34  | 36.47     |                        | 99.93           | 100.00   | 99.98  | 100.00 | 100.00   |
|                        |            | 121     |                        | 20.35           | 57.55    | 41.69  | 53.04  | 63.38    |                        | 50.87           | 58.23    | 54.08  | 57.37  | 60.09     |                        | 99.98           | 100.00   | 100.00 | 100.00 | 100.00   |
|                        | 8          | 49      |                        | 18.71           | 51.03    | 33.57  | 46.56  | 57.55    |                        | 62.78           | 92.53    | 87.72  | 91.97  | 94.00     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        |            | 81      |                        | 44.84           | 78.12    | 67.69  | 76.02  | 81.37    |                        | 92.11           | 99.58    | 99.25  | 99.56  | 99.69     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        |            | 121     |                        | 66.75           | 93.38    | 90.38  | 92.84  | 94.44    |                        | 99.27           | 100.00   | 100.00 | 100.00 | 100.00    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        | 1          | 49      | (1,4)                  | 0.65            | 11.15    | 0.85   | 4.07   | 19.42    | (0.5,4)                | 25.36           | 29.40    | 21.80  | 27.09  | 33.07     | (0.5,8)                | 91.91           | 95.52    | 94.08  | 95.55  | 95.96    |
|                        |            | 81      |                        | 1.33            | 25.82    | 3.85   | 13.12  | 35.38    |                        | 54.31           | 60.22    | 53.02  | 57.79  | 63.30     |                        | 99.76           | 99.92    | 99.92  | 99.92  | 99.92    |
|                        |            | 121     |                        | 1.31            | 44.16    | 15.45  | 32.35  | 53.23    |                        | 75.85           | 84.22    | 81.42  | 83.47  | 85.89     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        | 8          | 49      |                        | 3.02            | 74.90    | 55.05  | 69.6   | 80.11    |                        | 92.13           | 99.91    | 99.77  | 99.89  | 99.96     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        |            | 81      |                        | 11.13           | 93.72    | 88.57  | 92.61  | 94.8     |                        | 99.95           | 100.00   | 100.00 | 100.00 | 100.00    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        |            | 121     |                        | 20.38           | 97.87    | 97.43  | 97.76  | 98.03    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
| 1                      | 49         | (2.5,4) | 60.47                  | 90.83           | 78.27    | 87.63  | 93.55  | (1,8)    | 24.11                  | 29.42           | 21.33    | 26.57  | 33.43  | (0.5,200) | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          |
|                        | 81         |         | 91.87                  | 99.48           | 98.51    | 99.32  | 99.66  |          | 53.78                  | 59.83           | 53.71    | 58.03  | 62.63  |           | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          |
|                        | 121        |         | 98.84                  | 100.00          | 100.00   | 100.00 | 100.00 |          | 75.71                  | 84.85           | 81.74    | 84.20  | 86.11  |           | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          |
| $(-3, -5)$             | 8          | 49      |                        | 99.49           | 100.00   | 99.98  | 100.00 | 100.00   |                        | 92.13           | 99.89    | 99.79  | 99.89  | 99.91     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        |            | 81      |                        | 99.98           | 100.00   | 100.00 | 100.00 | 100.00   |                        | 99.80           | 100.00   | 100.00 | 100.00 | 100.00    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        |            | 121     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        | 1          | 49      | (2.5,8)                | 2.18            | 21.98    | 4.46   | 11.99  | 31.40    | (2.5,20)               | 23.85           | 27.70    | 21.51  | 25.43  | 31.72     | (1,20)                 | 97.49           | 98.86    | 98.67  | 98.83  | 99.02    |
|                        |            | 81      |                        | 7.02            | 45.56    | 15.82  | 33.52  | 55.15    |                        | 54.78           | 60.56    | 53.55  | 58.22  | 63.61     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        |            | 121     |                        | 10.25           | 68.37    | 42.64  | 59.95  | 75.50    |                        | 75.38           | 85.19    | 81.98  | 84.22  | 86.58     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        | 8          | 49      |                        | 13.67           | 79.52    | 58.63  | 74.64  | 84.73    |                        | 92.73           | 99.89    | 99.74  | 99.87  | 99.91     |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        |            | 81      |                        | 38.49           | 96.08    | 91.27  | 95.11  | 96.94    |                        | 99.89           | 100.00   | 100.00 | 100.00 | 100.00    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        |            | 121     |                        | 62.85           | 99.69    | 99.05  | 99.6   | 99.84    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|                        | 1          | 49      | (4,4)                  | 19.2            | 32.15    | 20.93  | 28.45  | 37.20    | (2,4)                  | 1.07            | 0.58     | 0.39   | 0.54   | 1.00      | (2,8)                  | 46.45           | 52.51    | 45.57  | 51.89  | 55.28    |
|                        |            | 81      |                        | 44.62           | 62.18    | 51.78  | 59.56  | 65.80    |                        | 1.96            | 0.98     | 0.58   | 0.82   | 1.49      |                        | 81.96           | 83.39    | 80.52  | 83.33  | 84.91    |
|                        |            | 121     |                        | 66.18           | 85.15    | 80.11  | 84.24  | 86.91    |                        | 2.13            | 1.52     | 0.89   | 1.12   | 2.56      |                        | 95.53           | 97.25    | 96.89  | 97.22  | 97.48    |
| 8                      | 49         |         | 81.25                  | 94.27           | 91.28    | 93.94  | 95.16  |          | 3.71                   | 12.14           | 6.36     | 10.70  | 15.60  |           | 99.62                  | 99.92           | 99.90    | 99.92  | 99.96  |          |
|                        | 81         |         | 98.49                  | 98.75           | 98.6     | 98.75  | 98.79  |          | 9.22                   | 26.17           | 19.21    | 24.48  | 29.42  |           | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          |
|                        | 121        |         | 99.98                  | 98.21           | 98.2     | 98.21  | 98.21  |          | 14.55                  | 42.30           | 36.07    | 41.31  | 45.10  |           | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          |
| 1                      | 49         | (10,4)  | 98.71                  | 99.81           | 99.61    | 99.77  | 99.81  | (4,8)    | 1.16                   | 0.54            | 0.38     | 0.50   | 1.15   | (2,20)    | 99.78                  | 100.00          | 100.00   | 100.00 | 100.00 |          |
|                        | 81         |         | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          | 1.69                   | 1.36            | 0.81     | 1.11   | 1.90   |           | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          |
|                        | 121        |         | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          | 2.53                   | 2.11            | 1.27     | 1.71   | 2.93   |           | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          |
| 8                      | 49         |         | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          | 3.15                   | 13.05           | 6.36     | 11.37  | 16.25  |           | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          |
|                        | 81         |         | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          | 8.56                   | 25.55           | 18.48    | 23.98  | 28.87  |           | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          |
|                        | 121        |         | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          | 12.98                  | 43.64           | 36.85    | 42.35  | 46.68  |           | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          |
| 1                      | 49         | (10,8)  | 46.93                  | 65.05           | 52.45    | 62.71  | 69.91  | (10,20)  | 0.91                   | 0.70            | 0.31     | 0.54   | 1.05   | (4,20)    | 74.84                  | 81.73           | 76.48    | 81.42  | 84.18  |          |
|                        | 81         |         | 81.65                  | 91.53           | 87.37    | 90.84  | 92.98  |          | 2.18                   | 1.30            | 0.96     | 1.06   | 2.11   |           | 96.71                  | 98.04           | 97.77    | 98.07  | 98.25  |          |
|                        | 121        |         | 94.95                  | 99.06           | 98.55    | 98.88  | 99.36  |          | 2.29                   | 2.19            | 1.25     | 1.58   | 3.10   |           | 99.71                  | 99.95           | 99.92    | 99.95  | 99.95  |          |
| 8                      | 49         |         | 99.22                  | 99.98           | 99.94    | 99.98  | 99.98  |          | 2.96                   | 12.71           | 6.67     | 11.31  | 16.28  |           | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          |
|                        | 81         |         | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          | 8.91                   | 26.29           | 19.33    | 24.93  | 29.79  |           | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          |
|                        | 121        |         | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          | 13.58                  | 42.24           | 36.07    | 41.11  | 44.92  |           | 100.00                 | 100.00          | 100.00   | 100.00 | 100.00 |          |



**Figura 5.4** Estimativas dos poderes dos teste na situação (ii) para medidas de contraste diferentes no nível nominal de 1%.

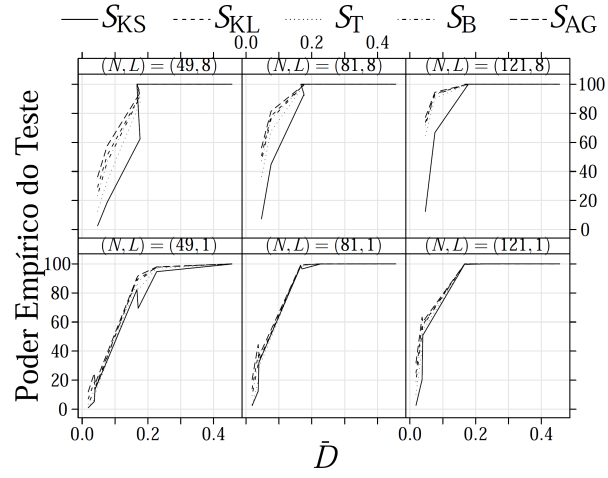
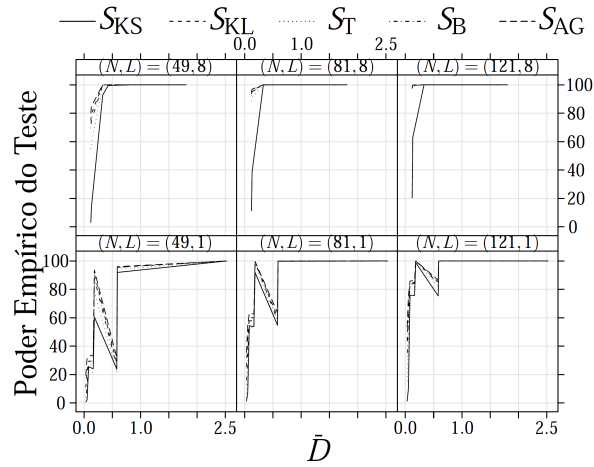
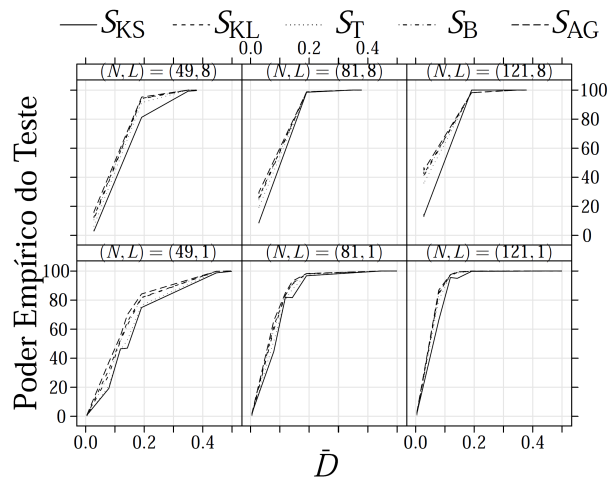
Inesperadamente, o teste de Kolmogorov-Smirnov tende a ser menos robusto do que os testes paramétricos quando a contaminação aumenta. A Figura 5.6 indica uma baixa variabilidade de  $S_{KS}$  baseada no tamanho do teste através das diferentes especificações da simulação. Isso sugere que a rugosidade e o brilho da imagem não afetam significativamente as estimativas para tamanhos de teste.

Outros gráficos e análises complementares podem ser encontrados no artigo em anexo.

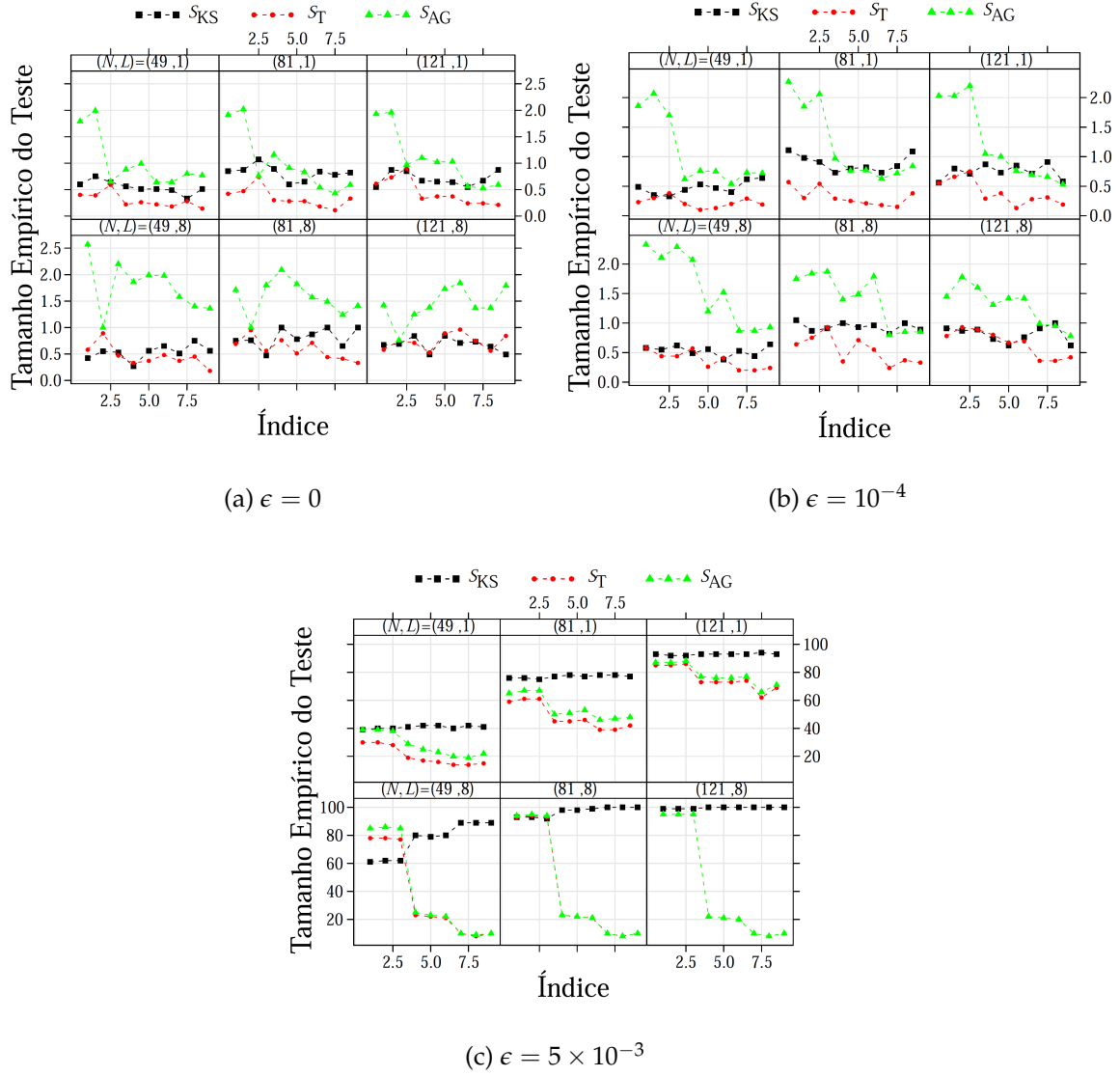
#### 4) Aplicação a dados reais tipo SAR:

A Figura 5.7 apresenta uma imagem SAR obtida pelo sensor E-SAR da cidade de Munique, Alemanha (como discutido no Apêndice A), utilizando um número de looks de 3.2. A área exibe três tipos de rugosidade para o relevo: (i) homogênea (correspondendo a pastos), (ii) heterogênea (floresta) e (iii) extremamente heterogênea (áreas urbanas). Amostras são selecionadas e submetidas a uma breve análise estatística descritiva. A Tabela 5.4 mostra as estimativas em cada uma das amostras, bem como seus tamanhos de amostra.

A Tabela 5.5 apresenta estimativas para o tamanho do teste nos níveis de significância de 1% e 10% com respeito as informações obtidas das regiões descritas na Tabela 5.4. Os resultados mostram que os testes baseados em  $S_{KS}$ ,  $S_{KL}$  e  $S_T$  têm excelentes resultados com respeito a esse critério e que a situação mais complicada está relacionada às áreas urbanas. Além disso, teste sob  $S_{AG}$  mostra ótimo tamanho empírico para o nível nominal de 1%, falindo para as situações urban-1 e urban-2 no nível 10%,

(a)  $\alpha_1 = -1.5$  e  $\alpha_2 = -3.0$ (b)  $\alpha_1 = -1.5$  e  $\alpha_2 = -5.0$ (c)  $\alpha_1 = -3.0$  e  $\alpha_2 = -5.0$ 

**Figura 5.5** Estimativas para os poderes dos testes nos cenários (iii)-(v) para medidas de contraste diferentes no nível nominal 1%.



**Figura 5.6** Estimativas para o tamanho do teste no nível nominal de 1% e a uma contaminação  $\epsilon$ . Para cada tamanho de amostra, todos os cenários considerados foram ordenados lexicograficamente:  $[i: (i; \alpha, \mu) = \{(1; -1.5, 1), \dots, (4; -3, 1), \dots, (7; -5, 1), \dots\}]$ . O índice  $i$  foi empregado como valor para abscissa.

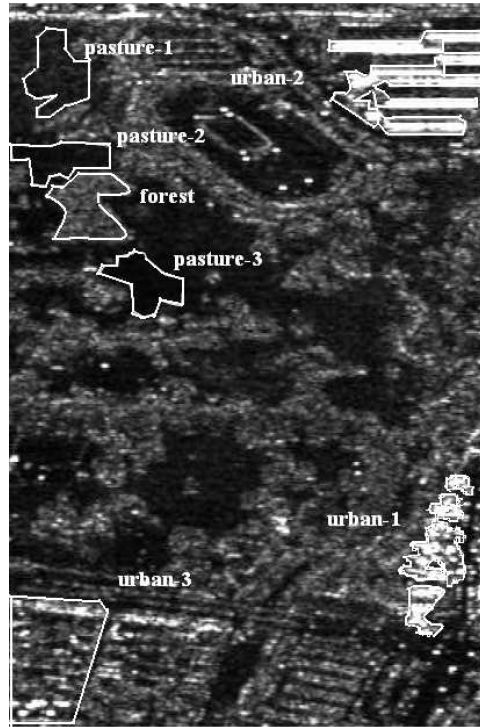


Figura 5.7 Imagem E-SAR e regiões selecionadas.

**Tabela 5.4** Estimativas dos parâmetros

| Regiões   | $\hat{\alpha}$ | $\hat{\gamma}$ | $\hat{\mu}$ | # pixels |
|-----------|----------------|----------------|-------------|----------|
| pasture-1 | -15.702        | 39259          | 2670.32     | 1235     |
| pasture-2 | -12.698        | 80320          | 6866.13     | 1216     |
| pasture-3 | -11.304        | 162292         | 15750.39    | 1602     |
| forest    | -9.339         | 661183         | 79288.04    | 1606     |
| urban-1   | -0.759         | 148413         | $\infty$    | 2005     |
| urban-2   | -0.388         | 110183         | $\infty$    | 3481     |
| urban-3   | -1.079         | 55583          | 703582.30   | 4657     |

enquanto que o tamanho  $S_B$  é inapropriado no caso urban-2. Assim, como  $S_{AG}$  apresentou as melhores taxas de discriminação no estudo de simulação, tem-se que a medida  $S_{AG}$  é a mais recomendável nos cenários: forest, pasture-1, pasture-2 e pasture-3.

**Tabela 5.5** Taxas de rejeição dos testes discutidos sob  $H_0: (\alpha_1, \gamma_1) = (\alpha_2, \gamma_2)$  nos níveis nominal 1% e 10%

| Regiões   | Nível nominal 1% |          |       |       |          | Nível Nominal 10% |          |       |       |          |
|-----------|------------------|----------|-------|-------|----------|-------------------|----------|-------|-------|----------|
|           | $S_{KS}$         | $S_{KL}$ | $S_T$ | $S_B$ | $S_{AG}$ | $S_{KS}$          | $S_{KL}$ | $S_T$ | $S_B$ | $S_{AG}$ |
| pasture-1 | 0.33             | 0.00     | 0.00  | 0.00  | 0.00     | 5.00              | 0.00     | 0.00  | 0.00  | 0.00     |
| pasture-2 | 0.36             | 0.00     | 0.00  | 0.00  | 0.00     | 6.16              | 0.00     | 0.00  | 0.00  | 0.00     |
| pasture-3 | 0.60             | 0.00     | 0.00  | 0.36  | 0.36     | 6.05              | 3.62     | 3.62  | 3.62  | 3.99     |
| forest    | 2.22             | 0.00     | 0.00  | 0.00  | 0.00     | 9.68              | 2.67     | 1.67  | 2.00  | 2.33     |
| urban-1   | 0.00             | 0.00     | 0.00  | 0.13  | 0.00     | 2.18              | 4.10     | 2.18  | 34.49 | 90.64    |
| urban-2   | 0.00             | 0.00     | 0.00  | 99.96 | 0.00     | 0.97              | 2.01     | 1.21  | 99.96 | 100.00   |
| urban-3   | 0.29             | 4.23     | 1.97  | 3.56  | 5.64     | 5.98              | 20.58    | 16.84 | 20.85 | 31.02    |

A Tabela 5.6 mostra o poder do teste nos níveis nominais 1% e 10%, isto é, a habilidade de rejeitar amostras de diferentes tipos. Os melhores desempenhos são observados em testes sob a classe  $(h, \phi)$ , principalmente quando confrontamos áreas cujas distribuições diferem mais significativamente. Os testes estatísticos  $S_T$  e  $S_{KL}$  apresentaram os piores resultados. O teste de Kolmogorov-Smirnov apresentou resultados inaceitáveis quando contrastamos as amostras urban-3 contra forest, e urban-1 contra urban-2. Em particular, a situação urban-3 comparada a forest contrasta duas regiões notavelmente distintas (como mostrado na Figura 5.7 e na Tabela 5.4). A Figura 5.8 apresentam as funções de distribuição empírica para as diferentes regiões da aplicação. Como esperado, as regiões mais heterogêneas possuem funções de distribuição empírica que crescem mais lentamente para o valor 1 a medida que o nível cinza cresce, justificando o retorno de mais energia por essas regiões. Além disso, o baixo desempenho do poder do teste para o caso forest contra urban-3 é explicado pela intersecção de suas funções de distribuição acumulada empírica.

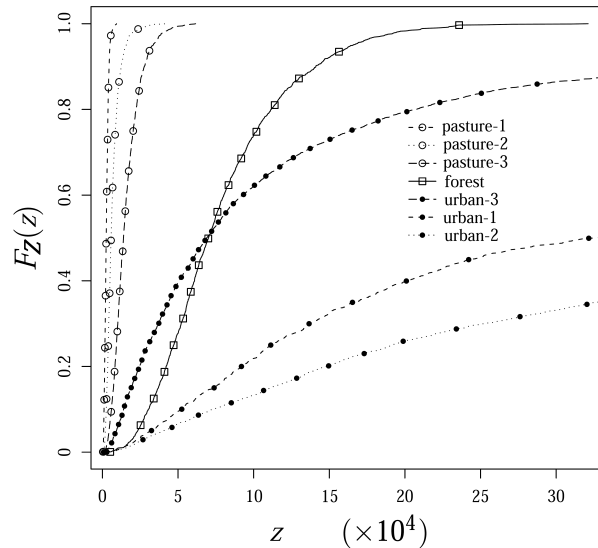
### 5.1.2 Considerações da Seção 5.1

Nós mostramos evidência numérica que o teste baseado na distância triangular,  $S_T$ , tem, em geral, menores tamanhos empírico do que o de Kolmogorov-Smirnov,  $S_{KS}$ . Usando uma modelo de contaminação paramétrico plausível, nós mostramos que testes paramétricos sob a classe  $(h, \phi)$  são mais eficientes do que aquele baseado sobre  $S_{KS}$  diante de contaminação. Este fato é um evento importante devido ao uso comum do teste de Kolmogorov-Smirnov ao tratar dados não gaussianos.

O teste baseado na medida medida  $S_{AG}$  apresentou a melhor desempenho com respeito à estimativa de seu poder. Entretanto, para um dado número de looks, nós podemos observar que a performance do poder do teste para as quatro medidas pa-

**Tabela 5.6** Taxas de rejeição dos testes discutidos sob  $H_1: (\alpha_1, \gamma_1) \neq (\alpha_2, \gamma_2)$  com  $\mu_1 = \mu_2$  nos níveis nominais 1% e 10%

| Regiões                      | 1% nominal level |          |        |        |          | 10% nominal level |          |        |        |          |
|------------------------------|------------------|----------|--------|--------|----------|-------------------|----------|--------|--------|----------|
|                              | $S_{KS}$         | $S_{KL}$ | $S_T$  | $S_B$  | $S_{AG}$ | $S_{KS}$          | $S_{KL}$ | $S_T$  | $S_B$  | $S_{AG}$ |
| pasture-1 $\times$ pasture-2 | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| pasture-1 $\times$ pasture-3 | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| pasture-2 $\times$ pasture-3 | 100.00           | 94.64    | 87.50  | 90.48  | 95.24    | 100.00            | 98.21    | 97.02  | 98.21  | 98.21    |
| forest $\times$ pasture-1    | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| forest $\times$ pasture-2    | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| forest $\times$ pasture-3    | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| urban-1 $\times$ pasture-1   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| urban-1 $\times$ pasture-2   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| urban-1 $\times$ pasture-3   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| urban-2 $\times$ pasture-1   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| urban-2 $\times$ pasture-2   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| urban-2 $\times$ pasture-3   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| urban-3 $\times$ pasture-1   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| urban-3 $\times$ pasture-2   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| urban-3 $\times$ pasture-3   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| urban-1 $\times$ forest      | 100.00           | 98.70    | 98.10  | 99.70  | 98.40    | 100.00            | 98.90    | 99.40  | 100.00 | 100.00   |
| urban-2 $\times$ forest      | 100.00           | 98.03    | 99.49  | 100.00 | 97.30    | 100.00            | 98.03    | 99.49  | 100.00 | 100.00   |
| urban-3 $\times$ forest      | 12.17            | 94.74    | 80.17  | 91.79  | 96.17    | 61.78             | 98.65    | 96.93  | 98.40  | 97.14    |
| urban-1 $\times$ urban-2     | 32.04            | 29.68    | 26.37  | 94.19  | 30.60    | 83.87             | 53.20    | 51.73  | 99.47  | 100.00   |
| urban-2 $\times$ urban-3     | 98.58            | 91.05    | 89.92  | 94.26  | 91.50    | 100.00            | 96.26    | 96.13  | 97.29  | 98.53    |
| urban-1 $\times$ urban-3     | 100.00           | 98.84    | 98.50  | 100.00 | 99.01    | 100.00            | 99.84    | 99.79  | 100.00 | 100.00   |

**Figura 5.8** Funções de distribuição acumulada empírica para diferentes regiões seleccionadas a partir de uma imagem SAR.

ramétricas foram rigorosamente as mesmas. Não obstante, eles se mostraram mais eficientes do que aquele baseado sobre a medida  $S_{KS}$ .

Essas conclusões sugerem o uso de  $S_T$  para processamento de dados SAR devido às estimativas para seus tamanhos serem mais próximas do nível nominal. Por outro lado, o teste baseado em  $S_{AG}$  é uma ferramenta mais apropriada quando o foco é o poder do teste.

Complementando este resultado, uma análise de dados reais ilustrou o caráter discriminatório destas medidas. Testes baseados sobre as medidas clássicas discutidas (a saber, as distâncias segundo Kullback-Leibler, Kolmogorov-Smirnov e Bhattacharyya) foram superados pelas distâncias  $d_T$  e  $d_{AG}$  e, portanto, devem ser substituídos pelos testes em função das distâncias triangular e aritmética geométrica.

### 5.1.3 Trabalhos futuros

Como complemento desta seção, objetiva-se

- ✎ Considerar técnicas de detecção de borda baseada em distâncias estocásticas e comparar com a metodologia proposta por Gambini et al. (2006),
- ✎ Analisar o efeito de correções de *bootstrap* sobre os tamanhos dos testes de homogeneidade propostos,
- ✎ Avaliar a performance assintótica das estatísticas definidas em outros tipos de estimadores (tais como momentos e outros baseados na transformada de Mellin).



---

## Anexo da Seção 5.1

### ARTIGO 1:

“Parametric and Nonparametric Tests for Speckled Imagery”

Published in *Pattern Analysis and Application*

---

# Parametric and nonparametric tests for speckled imagery

Renato J. Cintra · Alejandro C. Frery ·  
Abraão D. C. Nascimento

Received: 4 July 2010 / Accepted: 21 September 2011  
© Springer-Verlag London Limited 2011

**Abstract** Synthetic aperture radar (SAR) has a pivotal role as a remote imaging method. Obtained by means of coherent illumination, SAR images are contaminated with speckle noise. The statistical modeling of such contamination is well described according to the multiplicative model and its implied  $\mathcal{G}^0$  distribution. The understanding of SAR imagery and scene element identification is an important objective in the field. In particular, reliable image contrast tools are sought. Aiming the proposition of new tools for evaluating SAR image contrast, we investigated new methods based on stochastic divergence. We propose several divergence measures specifically tailored for  $\mathcal{G}^0$  distributed data. We also introduce a nonparametric approach based on the Kolmogorov–Smirnov distance for  $\mathcal{G}^0$  data. We devised and assessed tests based on such measures, and their performances were quantified according to their test sizes and powers. Using Monte Carlo simulation, we present a robustness analysis of test statistics and of maximum likelihood estimators for several

degrees of innovative contamination. It was identified that the proposed tests based on triangular and arithmetic-geometric measures outperformed the Kolmogorov–Smirnov methodology.

**Keywords** Robust statistics · Information theory · Nonparametric methods · Parametric inference

## 1 Introduction

Synthetic aperture radar (SAR), ultrasound-B, sonar, and laser images are obtained with coherent illumination and, consequently, their appearance is affected by a signal-dependent granular noise called speckle [41]. Due to deviations from classical properties of additivity and Gaussian distribution, speckled images require tailored processing and analysis techniques which could conveniently rely on the statistical properties of the data.

The ability to discriminate SAR image regions is an important tool for identifying and classifying distinct targets in the scene. Such capability has also been employed as a validation criterion for novel methodologies in SAR image change detection [26]. These facts have motivated the proposal of contrast measures based on the statistical properties of image data [14, 27, 39]. Therefore, an accurate modeling of speckled data is a major step for SAR image analysis [13]. Indeed, the probability distribution that describes the image data is a crucial information for any subsequent statistical analysis.

Among the proposals for SAR image modeling, the multiplicative model [13] has proven to be an important framework for the statistical processing and analysis of SAR data [3]. Assuming that the hypotheses suggested by the multiplicative model are valid, the  $\mathcal{G}^0$  distribution

---

R. J. Cintra (✉)  
Departamento de Estatística, Universidade Federal de  
Pernambuco, Cidade Universitária, 50740-540 Recife,  
Pernambuco, Brazil  
e-mail: rjdsc@stat.ufpe.org

A. C. Frery  
CPMAT & LCCV, Instituto de Computação,  
Universidade Federal de Alagoas, Maceió, Alagoas, Brazil  
e-mail: acfrery@pesquisador.cnpq.br

A. D. C. Nascimento  
Graduate Program in Statistics,  
Universidade Federal de Pernambuco,  
Recife, Pernambuco, Brazil  
e-mail: abraao.susej@gmail.com

arises as an efficient and effective probabilistic model for speckled data [15, 21, 36, 37, 46].

Hypothesis test methods have been considered as a venue to quantifying the image contrast between different regions in SAR imagery. Edge detection [9, 20], classification [24, 33], target identification [23], change detection [38] and oil spill identification [34] rely on parametric and nonparametric hypothesis tests. Although some nonparametric methods are computationally convenient, they seldom take explicitly the roughness information into account [9]. As a consequence, such simple approaches are unable to provide any information about that important parameter. Additionally, these methods cannot offer any known statistical property that could be utilized for the design of hypothesis testing procedures [20]. Thus, the proposal of hypothesis testing methods based on roughness-dependent contrast measures is a much sought tool for speckled data analysis.

In recent years, the interest in adapting information-theoretic tools to image processing has increased notably. In [39], the Bhattacharyya and (symmetrized) Kullback–Leibler distances were applied as scalar contrast measures for polarimetric and interferometric SAR imagery. The works [17, 40] proposed several parametric methods based on  $(h, \phi)$ -divergence measures for speckled data.

Often regarded as a classical benchmark tool, the Kolmogorov–Smirnov distance has been used for segmentation and labeling of SAR images [32]. This nonparametric test has found applications in polarimetric speckled data analysis [24]. Additionally, the Kolmogorov–Smirnov distance was also applied as a goodness-of-fit measure for SAR data modeling [8].

Whereas on the one hand a careful statistical modeling is fundamental, on the other hand nature is more complex than most models. Even sophisticated statistical models and tools may have their effectiveness diminished when data deviate from the assumed underlying hypotheses, even mildly. Robust statistics is a way of tackling this problem, and resistant techniques have been used with success in the SAR literature [1, 5, 18, 28].

This paper presents three contributions regarding the analysis of SAR imagery with stochastic measures of contrast. Firstly, it quantifies the ability of selected parametric methods based on the divergences measures by means of the size and the power of associated hypothesis tests using Monte Carlo. Such quantities were compared to the ones related to the nonparametric Kolmogorov–Smirnov test. It also presents an expression for the Kolmogorov–Smirnov distance between  $\mathcal{G}^0$  distributions. Secondly, it employs a contamination model for representing ‘innovative outliers’ [12] and assesses the behavior of maximum likelihood (ML) estimators under the  $\mathcal{G}^0$  model in the presence of

contamination. Finally, it submits hypothesis tests to different degrees of contamination, aiming to investigate the robustness of their sizes. Additionally, it presents an application to actual SAR data.

The remainder of the paper is organized as follows. Section 2 discusses the multiplicative model and the  $\mathcal{G}^0$  distribution. The hypothesis tests under consideration, namely those derived from  $(h, \phi)$ -divergences and the Kolmogorov–Smirnov distance, are derived in Sect. 3. Section 4 presents Monte Carlo experiments to assess ML estimation and hypothesis tests under situations which include “pure” and contaminated data. Actual SAR data are analyzed in Sect. 5. Section 6 concludes the paper.

## 2 The multiplicative model and the $\mathcal{G}^0$ distribution

Meaningful statistical processing of SAR imagery is always associated to a statistical model.

In [21], Gao provided a comprehensive exposition of existings models. In particular, three major categories of models were identified: (1) the multiplicative model [13, 16], (2) models based on the generalized central limit theorem [29], and (3) the empirical distribution model [2]. Other approaches include the joint distribution model [4], the mixed Gaussian model [11], mixtures of Gamma laws [49], and the correlation based model [30].

The multiplicative model has a phenomenological nature which is closely tied to the physics of the image formation [41]. Zhang [47] showed evidence that the distributions stemming from the product between independent random variables which describe the speckle noise and the backscatter outperform other models. In [21, p. 776], Gao includes the following note about the multiplicative model: The most attractive achievement among them is the statistical modeling on extremely heterogeneous region of SAR images proposed by Frery et al. [16] who [...] has introduced the original idea that for the purpose of statistical modeling, SAR images can be divided into homogeneous regions, heterogeneous regions and extremely heterogeneous regions, according to their contents.

Such assertion, related to  $\mathcal{G}^0$  distribution, is due to the following facts:

1. It has as many parameters as the  $K$  law, i.e., scale, roughness and number of looks, and they can also be interpreted, leading to scene understanding;
2. The roughness is expressive, in the sense that it is able to describe more targets than the  $K$  law [36];
3. Its density does not depend on special functions, as is the case of the  $K$  law [16];

4. Its cumulative distribution function can be obtained using its relationship with the Snedekor's  $F$  distribution [37].

We, therefore, admit the  $\mathcal{G}^0$  law as the model for the observed data, regardless the target and, without loss of generality, in its intensity format, i.e., quadratic detection.

Consider positive and independent random variables  $X$  and  $Y$  as models for the terrain backscatter and the speckle noise, respectively. The model for the return is  $Z = X \cdot Y$ .

The generalized inverse Gaussian distribution is proposed as a general model for the intensity backscatter. However, a tractable particular case is the reciprocal gamma law [13], whose density function is given by

$$f_X(x; \alpha, \gamma) = \frac{\gamma^{-\alpha}}{\Gamma(-\alpha)} x^{\alpha-1} \exp\left(-\frac{\gamma}{x}\right), \quad -\alpha, \gamma, x > 0. \quad (1)$$

In intensity format, the random variable  $Y$  obeys the gamma distribution with unitary mean [13], which has the following density

$$f_Y(y; L) = \frac{L^L}{\Gamma(L)} y^{L-1} \exp(-Ly), \quad y > 0, L \geq 1, \quad (2)$$

where  $L$  is the number of looks. Throughout this paper, the number of looks is assumed to be known and constant over the whole image. A detailed account of the until recently largely unexplored issue of estimating  $L$  is provided in [3].

Considering independent random variables with the densities presented in Eqs. (1) and (2), we obtain that the density of  $Z$  is expressed by

$$\begin{aligned} f_Z(z; \alpha, \gamma, L) &= \frac{L^L \Gamma(L - \alpha)}{\gamma^\alpha \Gamma(-\alpha) \Gamma(L)} z^{L-1} (\gamma + Lz)^{\alpha-L}, \\ &= \frac{\Gamma(L - \alpha)}{\Gamma(-\alpha) \Gamma(L)} \left(\frac{L}{\gamma}\right)^L \frac{z^{L-1}}{\left(1 + \frac{L}{\gamma} z\right)^{L-\alpha}}, \end{aligned} \quad (3)$$

where  $z > 0$ ,  $\alpha < 0$  is the roughness,  $\gamma > 0$  is the scale, and  $L \geq 1$ . We indicate this situation as  $Z \sim \mathcal{G}^0(\alpha, \gamma, L)$ .

As shown in [36, 37], this distribution can be used as an universal model for speckled data. The  $r$ th moment of  $Z$  is expressed by

$$E(Z^r) = \left(\frac{\gamma}{L}\right)^r \frac{\Gamma(-\alpha - r) \Gamma(L + r)}{\Gamma(-\alpha) \Gamma(L)}, \quad (4)$$

if  $-r > \alpha$ . Otherwise, it is infinite.

Reparametrizing (3) with  $M = -\alpha > 0$  and  $\gamma = M\mu$ , we obtain the density of the Fisher model [45] whose density is given by

$$f_Z(z; M, \mu, L) = \frac{\Gamma(L + M)}{\Gamma(M) \Gamma(L)} \left(\frac{L}{M\mu}\right)^L \frac{z^{L-1}}{\left(1 + \frac{L}{M\mu} z\right)^{L+M}}.$$

Using some Mellin transform properties for Fisher distribution, Galland et al. [19] derived the so-called log-cumulants  $w_i = E(Z^i \log Z)$ :

$$\begin{aligned} w_0 &= \log \frac{M\mu}{L} + \psi^0(L) - \psi^0(M), \\ w_1 &= \frac{M\mu}{M-1} \left[ \log \frac{M\mu}{L} + \psi^0(L+1) - \psi^0(M-1) \right], \\ w_2 &= \frac{M^2(L+1)}{(M-1)(M-2)L} \left[ \log \frac{M\mu}{L} + \psi^0(L+2) - \psi^0(M-2) \right], \end{aligned}$$

where  $\psi^0(\cdot)$  is the digamma function. As the Fisher distribution is a reparametrization of the  $\mathcal{G}^0$  distribution, the results in [19, 45] apply to the latter.

Figure 1 shows sixteen  $256 \times 256$  pixel regions of simulated speckled data according to the  $\mathcal{G}^0$  distribution for  $L = \{1, 8\}$ ,  $\alpha \in \{-15, -8, -5, -3\}$ , and  $E(Z) = \mu \in \{1, 10\}$ . The homogeneity depends on  $\alpha$ , the roughness parameter, while  $\mu$  controls the mean brightness. Figure 2a and b shows the density of  $\mathcal{G}^0$  distributions with varying  $(\alpha, L) \in \{(-1.5, 20), (-1.5, 8), (-1.5, 3), (-15, 20), (-15, 8), (-15, 3)\}$ , and  $\mu \in \{1, 2, 5\}$  for  $\mu = 1$  and  $(\alpha, L) = (-2.5, 3)$ , respectively. Increasing  $L$  or  $\mu$  flattens the densities, while small values of  $\alpha$  lead to images of low variability.

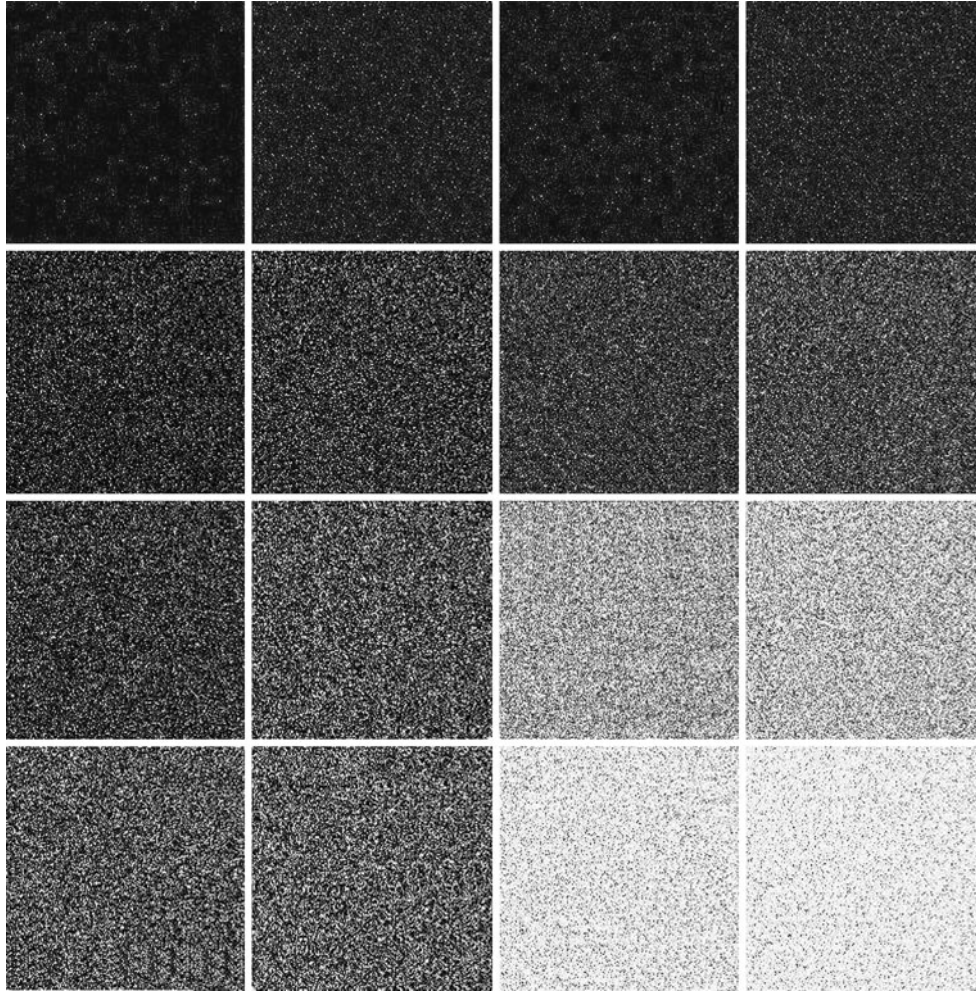
Parameters  $\alpha$  and  $\gamma$  can be estimated according to several methods, including bias-reduced procedures [6, 43, 46] and robust techniques (such as  $L$ - and  $M$ -estimators adapted to asymmetric distributions) [1, 5]. Small sample methods are particularly relevant in image processing and analysis, for instance, in the determination of the equivalent number of looks, in adaptive filters and for training classifiers. In such cases, samples are usually very small when compared to the full image [17], ranging from a single observation to, typically, less than a hundred.

ML estimation was adopted due to its optimal asymptotic properties [42]. Let  $Z_1, \dots, Z_n$  be  $n$  iid  $\mathcal{G}^0(\alpha, \gamma, L)$  distributed random variables. The ML estimator for  $(\alpha, \gamma)$ , for a given  $L$ , namely  $(\hat{\alpha}, \hat{\gamma})$ , is any solution of the following system of non-linear equations [40]:

$$\begin{aligned} \psi^0(-\hat{\alpha}) - \psi^0(L - \hat{\alpha}) - \log \hat{\gamma} + \frac{1}{n} \sum_{i=1}^n \log(\hat{\gamma} + LZ_i) &= 0, \\ -\frac{\hat{\alpha}}{\hat{\gamma}} + \frac{\hat{\alpha} - L}{n} \sum_{i=1}^n (\hat{\gamma} + LZ_i)^{-1} &= 0. \end{aligned}$$

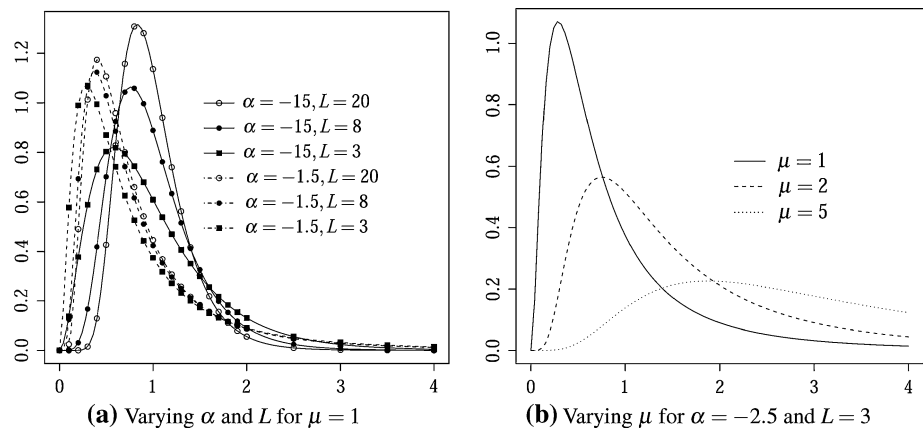
In general, the above system of equations does not possess a closed form solution and numerical optimization methods are required. We used the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method, which is regarded as an accurate technique [7]. Next section presents several contrast





**Fig. 1** Synthetic  $\mathcal{G}^0$  imagery for  $(\mu, L) \in \{(1, 1), (1, 8), (10, 1), (10, 8)\}$  (left to right) and  $\alpha \in \{-15, -8, -5, -3\}$  (top to bottom)

**Fig. 2**  $\mathcal{G}^0$  densities

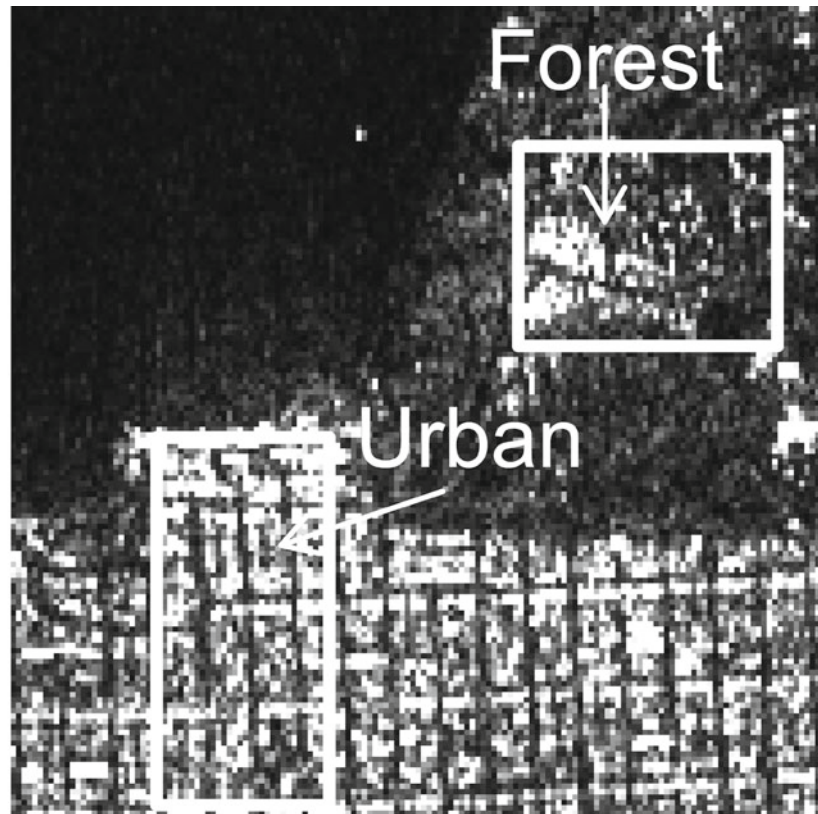


measures for  $\mathcal{G}^0$  distributed images based on the above estimators.

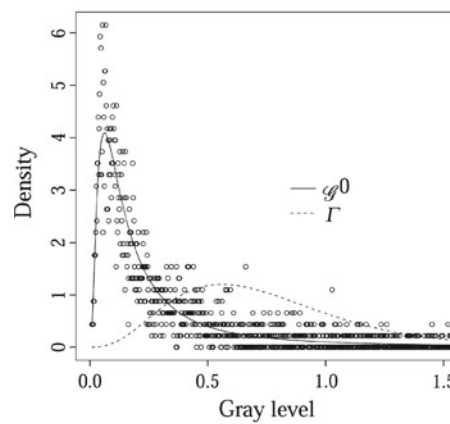
Figure 3a shows an image over San Francisco, CA, obtained by the AIRSAR sensor at band  $L$  with four (nominal) looks. Urban and forest regions were selected,

and Table 1 presents the sample sizes and ML estimates of  $\alpha$  and  $\gamma$ . Moreover, this table shows the sum of squared errors (SSE) between the histogram  $f_n$  and the fitted density  $\hat{f}_n$ , which is given by [48]

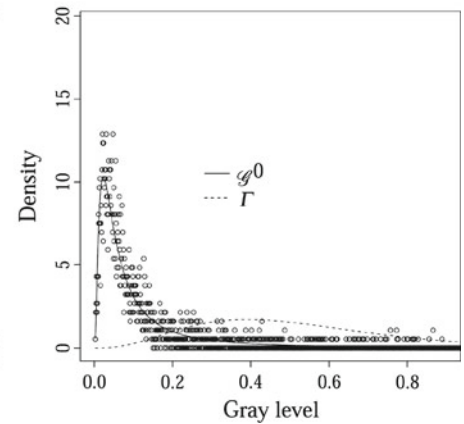
**Fig. 3** Real SAR image with selected regions and empirical and fitted densities for different urban and forest regions



(a) SAR image with selected regions



(b) Urban region



(c) Forest region

**Table 1** ML estimates and SSE

| Regions | $\bar{Z}$ | $\hat{\alpha}$ | $\hat{\gamma} (\times 10^{-2})$ | SSE             |          | # pixels |
|---------|-----------|----------------|---------------------------------|-----------------|----------|----------|
|         |           |                |                                 | $\mathcal{G}^0$ | $\Gamma$ |          |
| Urban   | 0.374     | -1.370         | 19.550                          | 0.062           | 0.318    | 2,277    |
| Forest  | 0.260     | -1.275         | 7.461                           | 0.108           | 0.389    | 1,813    |

$$SSE = \sum_{n=1}^{\# \text{pixels}} \frac{(\hat{f}_n - f_n)^2}{\# \text{pixels}}.$$

Two densities are considered: the basic gamma model [the scaled version of the density presented in (2)] and the  $\mathcal{G}^0$  distribution.

The values of SSE provide evidence that the  $\mathcal{G}^0$  distribution outperforms the gamma distribution for the considered regions. The strongest difference occurs in urban regions, i.e., in situation with intense roughness, as presented in Fig. 3b, c. As discussed in [36], among other references, the  $\mathcal{G}^0$  law includes the gamma distribution as a particular case and, therefore, it is able to describe homogeneous targets as pasture, crops, and sea.

### 3 Contrast measures for the $\mathcal{G}^0$ distribution

An image can be interpreted as a set of regions described by possibly different probability laws, and contrast analysis addresses the problem of quantifying how distinguishable two image regions are. Many statistical approaches have been developed to reach this goal, leading to the difficult problem of specifying accurate, expressive, and tractable metrics. In general terms, these methods can be either parametric or nonparametric.

In a previous work, we proposed a class of parametric measures based on stochastic distances. Advancing that study, among the existing nonparametric methods, we now include an approach based on the Kolmogorov–Smirnov test. This method is reported to possess good discriminatory properties [31] and has found applications in several instances [24, 32]. In the following, we briefly describe the parametric methods. Moreover, we adapt the Kolmogorov–Smirnov test to  $\mathcal{G}^0$  distributed data.

We assume that  $Z'$  and  $Z''$  are  $\mathcal{G}^0$  distributed random variables with cumulative distribution functions  $F_{Z'}$  and  $F_{Z''}$ , respectively. These distributions share the support  $\mathbb{R}_+$  and are defined over the same probability space, equipped with densities  $f_{Z'}(z; \theta_1)$  and  $f_{Z''}(z; \theta_2)$ , respectively, where  $\theta_1$  and  $\theta_2$  are parameter vectors.

Section 3.1 presents contrast measures derived from stochastic distances, while Sect. 3.2 discusses the use of the Kolmogorov–Smirnov distance for the  $\mathcal{G}^0$  model.

#### 3.1 Contrast measures based on divergence

Information theoretical tools known as divergence measures offer entropy-based methods for contrasting stochastic distributions. Divergence measures were submitted

to a systematic and comprehensive treatment and, as a result, the class of  $(h, \phi)$ -divergences was proposed [42].

The  $(h, \phi)$ -divergence between  $f_{Z'}$  and  $f_{Z''}$  with common support  $I$  is defined by

$$D_{\phi}^h(Z', Z'') = h \left( \int_I \phi \left( \frac{f_{Z'}(z; \theta_1)}{f_{Z''}(z; \theta_2)} \right) f_{Z''}(z; \theta_2) dz \right),$$

where  $\phi: (0, \infty) \rightarrow [0, \infty)$  is a convex function,  $h: (0, \infty) \rightarrow [0, \infty)$  is a strictly increasing function with  $h(0) = 0$ , and indeterminate forms are assigned value zero. Table 2 shows the  $h$  and  $\phi$  functions considered in this work.

These divergences are often asymmetric, and to solve this issue and to turn them into distances, a new measure  $d_{\phi}^h$  can be expressed as

$$d_{\phi}^h(Z', Z'') = \frac{D_{\phi}^h(Z', Z'') + D_{\phi}^h(Z'', Z')}{2}. \quad (5)$$

We will be interested in comparing samples from possibly the same distribution  $\mathcal{D}(\theta)$ , so instead of indexing (5) by the random variables, we will do it by the parameters  $\theta_1$  and  $\theta_2$ . Furthermore, denote  $\hat{\theta}_1$  and  $\hat{\theta}_2$  the maximum likelihood estimators of  $\theta_1$  and  $\theta_2$  based on independent samples of sizes  $m$  and  $n$ , respectively.

**Lemma 1** *Let the regularity conditions proposed in [p. 380, 42] hold. If  $\frac{m}{m+n} \rightarrow \lambda \in (0, 1)$  and  $\theta_1 = \theta_2$ , then*

$$\frac{2mn}{m+n} \frac{1}{h'(0)\phi''(1)} d_{\phi}^h(\hat{\theta}_1, \hat{\theta}_2) \xrightarrow[m, n \rightarrow \infty]{\mathcal{D}} \chi_M^2,$$

where “ $\xrightarrow{\mathcal{D}}$ ” denotes convergence in distribution.

Based on Lemma 1, the following statistical hypothesis test for the null hypothesis  $\theta_1 = \theta_2$  can be derived:

$$S_{\phi}^h(\hat{\theta}_1, \hat{\theta}_2) = \frac{2mnv}{m+n} d_{\phi}^h(\hat{\theta}_1, \hat{\theta}_2).$$

Table 2 shows the notation we for each test statistic, along with the values of  $v = [h'(0)\phi''(1)]^{-1}$ .

We are now in position to state the following result.

**Proposition 1** *Let  $m$  and  $n$  be large, and consider the outcome  $S_{\phi}^h(\hat{\theta}_1, \hat{\theta}_2) = s$ . The null hypothesis  $\theta_1 = \theta_2$  can be rejected at a nominal level  $\eta$  if  $\Pr(\chi_M^2 > s) \leq \eta$ .*

**Table 2**  $(h, \phi)$ -distances, statistics,  $h$  and  $\phi$  functions, and related quantity

| $(h, \phi)$ -distance             | Statistic | $h(y)$                      | $\phi(x)$                                                        | $v = [h'(0)\phi''(1)]^{-1}$ |
|-----------------------------------|-----------|-----------------------------|------------------------------------------------------------------|-----------------------------|
| Kullback–Leibler ( $d_{KL}$ )     | $S_{KL}$  | $y/2$                       | $(x-1) \log x$                                                   | 1                           |
| Triangular ( $d_T$ )              | $S_T$     | $y, 0 \leq y < 2$           | $\frac{(x-1)^2}{x+1}$                                            | 1                           |
| Bhattacharyya ( $d_B$ )           | $S_B$     | $-\log(-y+1), 0 \leq y < 1$ | $-\sqrt{x} + \frac{x+1}{2}$                                      | 4                           |
| Arithmetic-geometric ( $d_{AG}$ ) | $S_{AG}$  | $y$                         | $\frac{x+1}{2} \log \frac{x+1}{2x} + \left(\frac{x-1}{2}\right)$ | 4                           |



### 3.2 Kolmogorov–Smirnov contrast measure

The empirical function of the sample  $\mathbf{Z} = (Z_1, \dots, Z_n)$  of a random variable  $Z$  is given by  $\hat{F}_{\mathbf{Z}}(z) = n^{-1} \#\{j : Z_j \leq z\}$ . Empirical functions can be used to test the null hypothesis  $H_0: F_{Z'} = F_{Z''}$  against the alternative  $H_1: F_{Z'} \neq F_{Z''}$  with random samples  $\mathbf{Z}'$  and  $\mathbf{Z}''$  of sizes  $m$  and  $n$ , respectively, using the test statistic [44]

$$S_{KS}(\mathbf{Z}', \mathbf{Z}'') = \sqrt{\frac{mn}{m+n}} \underbrace{\max_{-\infty < z < \infty} |\hat{F}_{\mathbf{Z}'}(z) - \hat{F}_{\mathbf{Z}''}(z)|}_{\hat{d}_{KS}},$$

where  $\hat{d}_{KS}$  is the empirical estimate for the Kolmogorov–Smirnov distance  $d_{KS}$ . In this case, the analytic expression for  $d_{KS}$  under  $\mathcal{G}^0$  distributions with same number of looks is given by

$$d_{KS}(\mathbf{Z}', \mathbf{Z}'') = \max_{-\infty < z < \infty} \left| L^L z^L \left[ \frac{\Gamma(L - \alpha_1)}{\gamma_1^L \Gamma(-\alpha_1)} {}_2F_1 \left( L, L - \alpha_1, L + 1, \frac{-Lz}{\gamma_1} \right) - \frac{\Gamma(L - \alpha_2)}{\gamma_2^L \Gamma(-\alpha_2)} {}_2F_1 \left( L, L - \alpha_2, L + 1, \frac{-Lz}{\gamma_2} \right) \right] \right|,$$

where  ${}_2F_1$  is the hypergeometric function.

**Proposition 2** Assume  $S_{KS}(\mathbf{Z}', \mathbf{Z}'') = s$ . The null hypothesis  $F_{Z'} = F_{Z''}$  can be rejected at level  $\eta$  if  $\Pr[S_{KS}(\mathbf{Z}', \mathbf{Z}'') > \delta] = 1 - L(s) \leq \eta$ , where  $L(s) = 1 - 2 \sum_{k=1}^{\infty} (-1)^{k-1} \exp(-2k^2 s^2)$  and  $\delta$  is a critical value given in [44], provided that  $m$  and  $n$  are large.

In terms of image analysis, Propositions 1 and 2 offer methods to statistically refute the hypothesis that two samples obtained in different regions can be described by the same distribution. Such assessment is relevant, for instance, when choosing input training samples in supervised learning procedures, or when proposing similarity-based techniques such as segmentation and classification.

Since the aforementioned results are asymptotic and valid only when the underlying hypothesis holds, the following section shows a Monte Carlo approach for their assessment with finite samples and under contamination. Such assessment is needed since finite samples and outliers are unavoidable in practical situations.

## 4 Validation

This section presents a qualitative assessment of the test proposed for checking whether two samples are distributed according to the same law when finite size samples are used. The proposed tests are submitted to uncontaminated data in Sect. 4.1. Subsequently, a contamination model based on innovative outliers, inspired by corner reflectors,

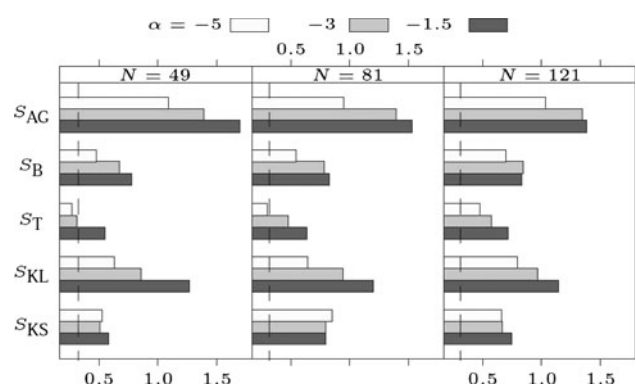
is considered in Sect. 4.2. The effect of contamination on the ML estimates of  $\mathcal{G}^0$  parameters is also assessed.

### 4.1 Analysis with simulated data

Speckled data simulated with the  $\mathcal{G}^0$  law were employed to assess the discussed parametric and nonparametric test statistics. The situations herein assessed were three levels of roughness ( $\alpha \in \{-1.5, -3, -5\}$ ), three mean return values ( $\mu \in \{1, 2, 5\}$ ), and two values of the number of looks ( $L \in \{1, 8\}$ ). The mean  $\mu = -\gamma/(1 + \alpha)$  is directly related to the image brightness. We performed the tests at the 1% significance level using square windows of  $7 \times 7$ ,  $9 \times 9$ , and  $11 \times 11$  pixels. Therefore, the considered sample sizes were  $N = m = n \in \{49, 81, 121\}$ .

Empirical size and power of the proposed tests were sought as a means to measure and compare their performances. Monte Carlo experiments under five different scenarios were used. Considering two image regions specified by parameter vectors  $(\alpha_1, \mu_1, L)$  and  $(\alpha_2, \mu_2, L)$ , the following scenarios were set: (i)  $\alpha_1 = \alpha_2$ ,  $\mu_1 = \mu_2$ , (ii)  $\alpha_1 = \alpha_2$ ,  $\mu_1 \neq \mu_2$ , (iii)  $\alpha_1 > \alpha_2$ ,  $\mu_1 > \mu_2$ , (iv)  $\alpha_1 > \alpha_2$ ,  $\mu_1 < \mu_2$ , and (v)  $\alpha_1 \neq \alpha_2$ ,  $\mu_1 = \mu_2$ . Whereas the first situation corresponds to image regions of same roughness and brightness, scenario (ii) accounts for images of same roughness but with different brightness. Images of both distinct roughness and brightness are captured by scenarios (iii) and (iv). Additionally, the fifth scenario considers regions of different roughness and equal brightness.

Figure 4 shows the mean empirical test sizes, i.e., under scenario (i). With the exception of  $S_T$ , significance levels are above the nominal value (shown as a dashed line). The tests tend to reject more often when  $\alpha$  increases, with little effect from sample sizes. In general,  $S_T$  is the most efficient measure regarding the test size, since its estimative is in mean the closest to nominal values;  $S_T$  is immediately followed by  $S_{KS}$ . As will be seen later,  $S_{KS}$  surprisingly loses this ability in the presence of contamination.



**Fig. 4** Mean of empirical test size scaled by 10 at 1% nominal level



**Table 3** Rejection rates of  $(h, \phi)$ -divergence and Kolmogorov–Smirnov tests for scenarios (ii)

| $\mu_2/\mu_1$ | $L$ | $N$ | $\alpha$ | $S_{KS}$ | $S_{KL}$ | $S_T$  | $S_B$  | $S_{AG}$ | $\alpha$ | $S_{KS}$ | $S_{KL}$ | $S_T$  | $S_B$  | $S_{AG}$ | $\alpha$ | $S_{KS}$ | $S_{KL}$ | $S_T$  | $S_B$  | $S_{AG}$ |
|---------------|-----|-----|----------|----------|----------|--------|--------|----------|----------|----------|----------|--------|--------|----------|----------|----------|----------|--------|--------|----------|
| 2.0           | 1   | 49  | -1.5     | 20.58    | 27.84    | 22.11  | 26.89  | 30.85    | -3.0     | 30.31    | 39.34    | 30.70  | 37.87  | 42.63    | -5.0     | 34.16    | 43.97    | 34.89  | 42.58  | 48.26    |
|               |     | 81  |          | 48.36    | 53.24    | 48.79  | 52.51  | 55.52    |          | 61.85    | 68.03    | 63.52  | 67.26  | 70.14    |          | 68.64    | 75.65    | 70.50  | 75.01  | 77.89    |
|               |     | 121 |          | 68.85    | 75.04    | 72.68  | 74.84  | 75.85    |          | 82.13    | 88.64    | 86.89  | 88.42  | 89.42    |          | 86.33    | 94.30    | 92.73  | 93.93  | 94.58    |
|               | 8   | 49  |          | 60.40    | 78.53    | 72.00  | 78.02  | 81.15    |          | 90.85    | 97.41    | 95.89  | 97.34  | 97.78    |          | 98.40    | 100.00   | 100.00 | 100.00 | 100.00   |
|               |     | 81  |          | 91.13    | 96.20    | 94.65  | 96.04  | 96.69    |          | 99.75    | 99.84    | 99.82  | 99.84  | 99.84    |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |
|               |     | 121 |          | 98.55    | 99.73    | 99.62  | 99.73  | 99.75    |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |
| 2.5           | 1   | 49  |          | 42.51    | 53.07    | 46.66  | 52.46  | 56.36    |          | 56.80    | 70.15    | 62.23  | 69.00  | 73.33    |          | 64.00    | 76.25    | 68.30  | 75.20  | 79.03    |
|               |     | 81  |          | 77.22    | 82.18    | 79.73  | 82.05  | 83.41    |          | 89.64    | 93.87    | 92.03  | 93.78  | 94.38    |          | 93.47    | 97.02    | 95.72  | 96.86  | 97.42    |
|               |     | 121 |          | 93.53    | 95.73    | 95.13  | 95.68  | 96.08    |          | 98.27    | 99.54    | 99.40  | 99.50  | 99.56    |          | 99.24    | 99.93    | 99.88  | 99.93  | 100.00   |
|               | 8   | 49  |          | 88.89    | 97.16    | 95.65  | 97.13  | 97.85    |          | 99.60    | 99.96    | 99.95  | 99.96  | 99.98    |          | 99.96    | 100.00   | 100.00 | 100.00 | 100.00   |
|               |     | 81  |          | 99.53    | 99.95    | 99.91  | 99.95  | 99.96    |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |
|               |     | 121 |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |
| 5.0           | 1   | 49  |          | 96.58    | 98.91    | 98.27  | 98.95  | 99.06    |          | 99.29    | 98.55    | 97.39  | 98.42  | 98.81    |          | 99.76    | 100.00   | 100.00 | 100.00 | 100.00   |
|               |     | 81  |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |
|               |     | 121 |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |
|               | 8   | 49  |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |
|               |     | 81  |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |
|               |     | 121 |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |          | 100.00   | 100.00   | 100.00 | 100.00 | 100.00   |

Goudail and Réfrégier [22] present graphically a relationship between test power and the Kullback and Bhattacharyya distances. Similarly, we illustrate the estimates of test powers with respect to the discrepancy measure given by

$$\bar{D} = (d_{KL} + d_T + d_B + d_{AG} + d_{KS})/5,$$

where, for the sake of simplicity, arguments  $(\hat{\theta}_1, \hat{\theta}_2)$  were suppressed.

Tables 3, 4, 5, and 6 present the test empirical power in scenarios (ii)–(v), respectively. As a general behavior, the empirical power increases with the number of looks ( $L$ ) and/or the sample size ( $N$ ).

Considering a fixed roughness, Table 3 provides evidence that the brightness change has more influence on the power in homogeneous regions. Figure 5 illustrates the results of this table with respect to  $\bar{D}$  for different sample sizes. In this case, we noticed that parametric tests outperformed the Kolmogorov–Smirnov test.

In Tables 4, 5, and 6, the inequality  $\text{Power}(S_{KS}) \leq \text{Power}(S_T) \leq \text{Power}(S_B) \leq \text{Power}(S_{KL}) \leq \text{Power}(S_{AG})$  is satisfied in every case of scenarios (iii) and (iv). For scenario (v), this fact holds when the number of looks is 8. Figure 6 presents the powers as functions of  $\bar{D}$  for different roughness pairs. The power increases with the sample size and with  $\bar{D}$ , as expected. These results show that the  $S_{AG}$  test outperforms other techniques with respect to the test power, since it is able to discriminate more often samples from different distributions.

#### 4.2 ML estimation in contaminated data

In practice, it is quite hard to guarantee that all observations in a sample come from random variables with the same distribution. When collecting samples by visual inspection, frequently the onset of some type of contamination is experienced.

One of the most common sources of contamination in SAR imagery is related to the *double bounce* phenomenon: a few pixels exhibit a very high return caused by either natural or man-made targets that send back most of the energy they receive. This is the case of flooded forests and urban areas. The presence of a single pixel suffering from double bounce may produce meaningless estimates and test statistics, unless they exhibit some robustness with respect to this kind of contamination. Figure 7 shows an example of the double bounce effect on an homogeneous field.

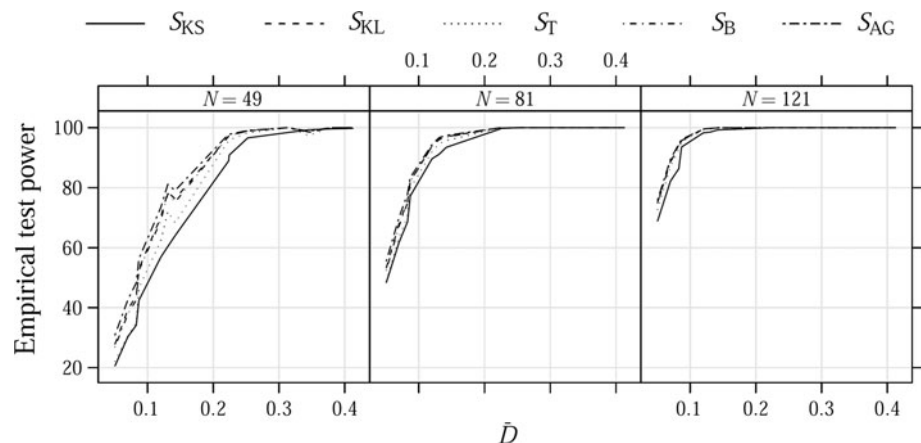
To assess the robustness of the considered tests, we assume contamination in a parametric way: the data in a sample come either from the hypothesized  $\mathcal{G}^0$  distribution with probability  $1 - \epsilon$ , or from a scaled version of the same law with probability  $\epsilon$ . The considered scaled distribution is such that its mean value is 100 times larger than the mean of the original distribution. Let  $A$  be the event “presence of outlier”, which has  $\Pr(A) = \epsilon$ , and  $A'$  its complement. By the total probability rule, the cumulative distribution function of the return is

$$\begin{aligned} F_Z(z) &= \Pr(Z \leq z) \\ &= \Pr(\{Z \leq z\} \cap A') + \Pr(\{Z \leq z\} \cap A) \\ &= \Pr(A') \Pr(Z \leq z | A') + \Pr(A) \Pr(Z \leq z | A) \\ &= (1 - \epsilon)F_{\mathcal{G}^0(\alpha, \gamma, L)}(z) + \epsilon F_{\mathcal{G}^0(\alpha, 100\gamma, L)}(z). \end{aligned}$$

Following [1, 5], we chose to assess the situations without contamination  $\epsilon = 0$  and with mild levels of contamination  $\epsilon \in \{10^{-4}, 5 \times 10^{-3}\}$ .

Simulation results are based on 5,500 independent replications, and instances were considered valid only if  $\hat{\alpha} \in [10\alpha, \alpha/20]$ , i.e., censoring was applied as in [15, 40]. Tables 7, 8, and 9 present the estimated bias (B), the coefficient of variation (CV, in percentage), and the mean squared error (MSE) of the ML estimators  $\hat{\alpha}$  and  $\hat{\gamma}$  for the roughness degrees  $-1.5$ ,  $-3$ , and  $-5$  according to several simulation conditions and contamination levels, respectively.

**Fig. 5** Empirical test power in situations (ii) for different contrast measure at 1% nominal level



**Table 4** Rejection rates of  $(h, \phi)$ -divergence and Kolmogorov–Smirnov tests for scenarios (iii)–(v) for roughness pair  $(\alpha_1, \alpha_2) = (-1.5, -3)$ 

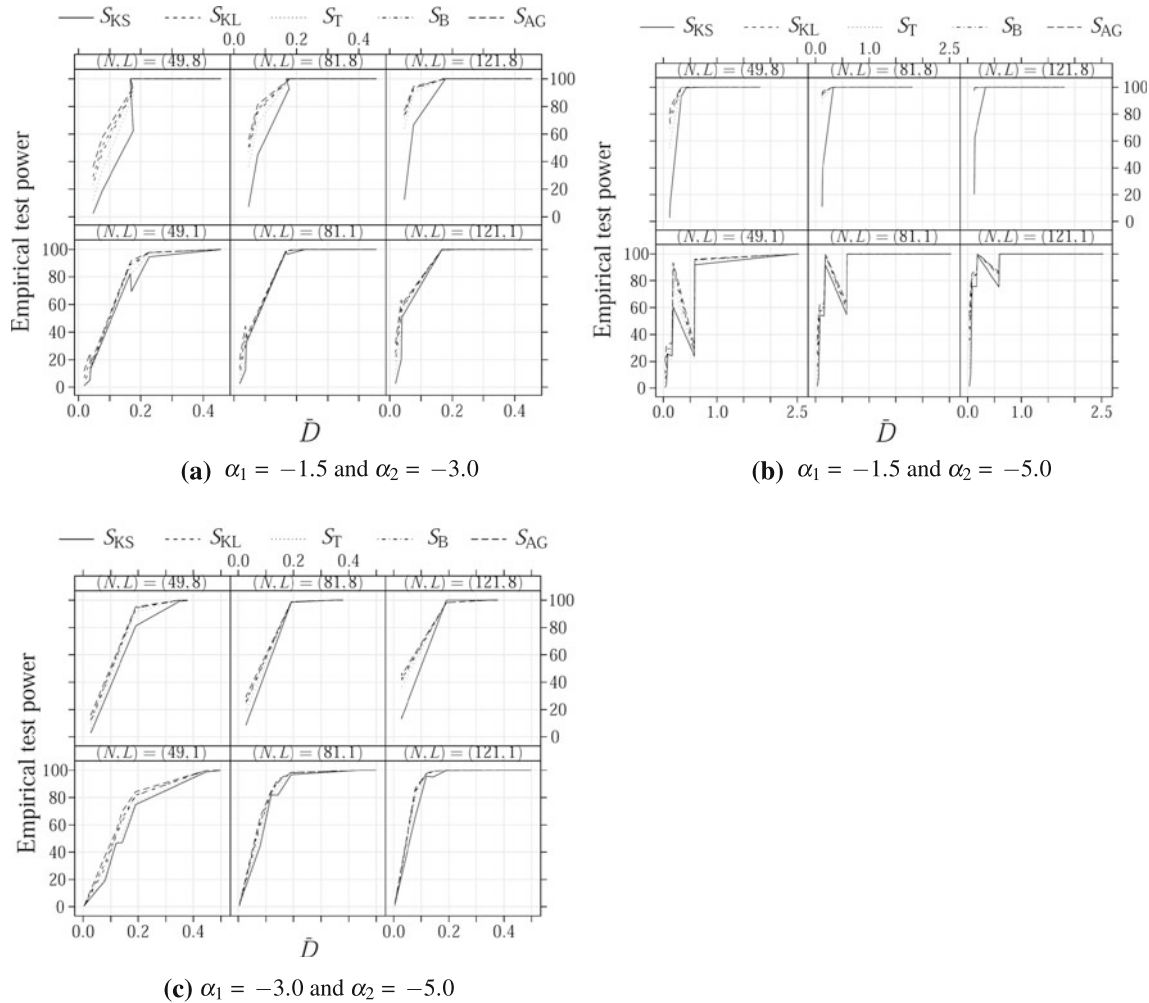
| Constants |     | $\gamma$               | $\mu_1 > \mu_2$ |          |        | $\gamma$ |          |                        | $\mu_1 = \mu_2$ |          |        | $\gamma$ |          |                        | $\mu_1 < \mu_2$ |          |        |        |          |
|-----------|-----|------------------------|-----------------|----------|--------|----------|----------|------------------------|-----------------|----------|--------|----------|----------|------------------------|-----------------|----------|--------|--------|----------|
| $L$       | $N$ | $(\gamma_1, \gamma_2)$ | $S_{KS}$        | $S_{KL}$ | $S_T$  | $S_B$    | $S_{AG}$ | $(\gamma_1, \gamma_2)$ | $S_{KS}$        | $S_{KL}$ | $S_T$  | $S_B$    | $S_{AG}$ | $(\gamma_1, \gamma_2)$ | $S_{KS}$        | $S_{KL}$ | $S_T$  | $S_B$  | $S_{AG}$ |
| 1         | 49  | (1, 2)                 | 1.05            | 7.12     | 1.17   | 3.53     | 11.94    | (0.5, 2)               | 13.45           | 16.15    | 11.64  | 15.14    | 18.56    | (0.5, 4)               | 82.27           | 89.51    | 86.89  | 89.48  | 90.47    |
|           | 81  |                        | 2.47            | 14.10    | 3.45   | 9.46     | 20.33    |                        | 31.80           | 35.39    | 30.38  | 34.11    | 38.25    |                        | 98.53           | 99.28    | 99.20  | 99.26  | 99.35    |
|           | 121 |                        | 2.64            | 24.98    | 10.20  | 19.17    | 31.83    |                        | 51.13           | 58.36    | 54.75  | 57.68    | 60.30    |                        | 99.89           | 100.00   | 100.00 | 100.00 | 100.00   |
| 8         | 49  |                        | 2.58            | 29.18    | 12.68  | 23.52    | 36.20    |                        | 62.35           | 92.41    | 87.65  | 91.71    | 94.06    |                        | 99.98           | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 81  |                        | 7.33            | 50.77    | 36.49  | 47.39    | 56.12    |                        | 93.31           | 99.51    | 99.16  | 99.45    | 99.53    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 121 |                        | 12.40           | 73.73    | 64.55  | 71.93    | 77.13    |                        | 99.22           | 99.85    | 99.84  | 99.84    | 99.85    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
| 1         | 49  | (2.5, 2)               | 69.45           | 89.53    | 82.24  | 87.95    | 91.89    | (1, 4)                 | 13.78           | 15.55    | 10.85  | 14.41    | 18.28    | (0.5, 10)              | 99.95           | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 81  |                        | 96.51           | 99.24    | 98.39  | 99.17    | 99.45    |                        | 31.65           | 35.47    | 30.65  | 34.15    | 37.74    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 121 |                        | 99.62           | 100.00   | 99.98  | 100.00   | 100.00   |                        | 49.98           | 58.51    | 54.67  | 57.57    | 60.48    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
| 8         | 49  |                        | 99.49           | 100.00   | 100.00 | 100.00   | 100.00   |                        | 62.35           | 92.19    | 87.23  | 91.47    | 94.03    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 81  |                        | 100.00          | 100.00   | 100.00 | 100.00   | 100.00   |                        | 92.93           | 99.55    | 99.29  | 99.49    | 99.62    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 121 |                        | 100.00          | 100.00   | 100.00 | 100.00   | 100.00   |                        | 99.47           | 99.98    | 99.96  | 99.98    | 100.00   |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
| 1         | 49  | (2.5, 4)               | 5.29            | 18.77    | 6.70   | 13.06    | 24.88    | (2.5, 10)              | 13.71           | 14.83    | 10.65  | 14.00    | 17.47    | (1, 10)                | 94.65           | 97.74    | 96.96  | 97.74  | 97.92    |
|           | 81  |                        | 12.44           | 37.45    | 19.88  | 31.21    | 44.52    |                        | 32.27           | 33.52    | 28.79  | 32.34    | 36.47    |                        | 99.93           | 100.00   | 99.98  | 100.00 | 100.00   |
|           | 121 |                        | 20.35           | 57.55    | 41.69  | 53.04    | 63.38    |                        | 50.87           | 58.23    | 54.08  | 57.37    | 60.09    |                        | 99.98           | 100.00   | 100.00 | 100.00 | 100.00   |
| 8         | 49  |                        | 18.71           | 51.03    | 33.57  | 46.56    | 57.55    |                        | 62.78           | 92.53    | 87.72  | 91.97    | 94.00    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 81  |                        | 44.84           | 78.12    | 67.69  | 76.02    | 81.37    |                        | 92.11           | 99.58    | 99.25  | 99.56    | 99.69    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 121 |                        | 66.75           | 93.38    | 90.38  | 92.84    | 94.44    |                        | 99.27           | 100.00   | 100.00 | 100.00   | 100.00   |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |

**Table 5** Rejection rates of  $(h, \phi)$ -divergence and Kolmogorov–Smirnov tests for scenarios (iii)–(v) for roughness pair  $(\alpha_1, \alpha_2) = (-1.5, -5)$ 

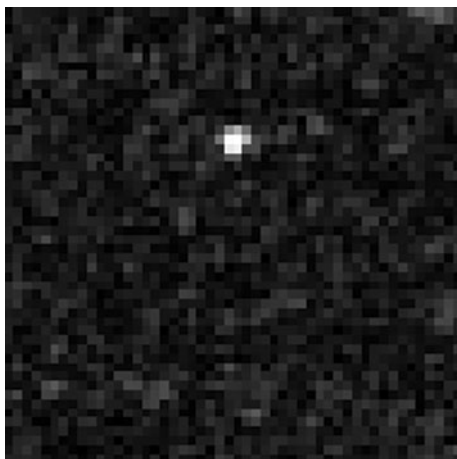
| Constants |     | $\gamma$               | $\mu_1 > \mu_2$ |          |        | $\gamma$ |          |                        | $\mu_1 = \mu_2$ |          |        | $\gamma$ |          |                        | $\mu_1 < \mu_2$ |          |        |        |          |
|-----------|-----|------------------------|-----------------|----------|--------|----------|----------|------------------------|-----------------|----------|--------|----------|----------|------------------------|-----------------|----------|--------|--------|----------|
| $L$       | $N$ | $(\gamma_1, \gamma_2)$ | $S_{KS}$        | $S_{KL}$ | $S_T$  | $S_B$    | $S_{AG}$ | $(\gamma_1, \gamma_2)$ | $S_{KS}$        | $S_{KL}$ | $S_T$  | $S_B$    | $S_{AG}$ | $(\gamma_1, \gamma_2)$ | $S_{KS}$        | $S_{KL}$ | $S_T$  | $S_B$  | $S_{AG}$ |
| 1         | 49  | (1, 4)                 | 0.65            | 11.15    | 0.85   | 4.07     | 19.42    | (0.5, 4)               | 25.36           | 29.40    | 21.80  | 27.09    | 33.07    | (0.5, 8)               | 91.91           | 95.52    | 94.08  | 95.55  | 95.96    |
|           | 81  |                        | 1.33            | 25.82    | 3.85   | 13.12    | 35.38    |                        | 54.31           | 60.22    | 53.02  | 57.79    | 63.30    |                        | 99.76           | 99.92    | 99.92  | 99.92  | 99.92    |
|           | 121 |                        | 1.31            | 44.16    | 15.45  | 32.35    | 53.23    |                        | 75.85           | 84.22    | 81.42  | 83.47    | 85.89    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
| 8         | 49  |                        | 3.02            | 74.90    | 55.05  | 69.6     | 80.11    |                        | 92.13           | 99.91    | 99.77  | 99.89    | 99.96    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 81  |                        | 11.13           | 93.72    | 88.57  | 92.61    | 94.8     |                        | 99.95           | 100.00   | 100.00 | 100.00   | 100.00   |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 121 |                        | 20.38           | 97.87    | 97.43  | 97.76    | 98.03    |                        | 100.00          | 100.00   | 100.00 | 100.00   | 100.00   |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
| 1         | 49  | (2.5, 4)               | 60.47           | 90.83    | 78.27  | 87.63    | 93.55    | (1, 8)                 | 24.11           | 29.42    | 21.33  | 26.57    | 33.43    | (0.5, 20)              | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 81  |                        | 91.87           | 99.48    | 98.51  | 99.32    | 99.66    |                        | 53.78           | 59.83    | 53.71  | 58.03    | 62.63    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 121 |                        | 98.84           | 100.00   | 100.00 | 100.00   | 100.00   |                        | 75.71           | 84.85    | 81.74  | 84.20    | 86.11    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
| 8         | 49  |                        | 99.49           | 100.00   | 99.98  | 100.00   | 100.00   |                        | 92.13           | 99.89    | 99.79  | 99.89    | 99.91    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 81  |                        | 99.98           | 100.00   | 100.00 | 100.00   | 100.00   |                        | 99.80           | 100.00   | 100.00 | 100.00   | 100.00   |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 121 |                        | 100.00          | 100.00   | 100.00 | 100.00   | 100.00   |                        | 100.00          | 100.00   | 100.00 | 100.00   | 100.00   |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
| 1         | 49  | (2.5, 8)               | 2.18            | 21.98    | 4.46   | 11.99    | 31.40    | (2.5, 20)              | 23.85           | 27.70    | 21.51  | 25.43    | 31.72    | (1, 20)                | 97.49           | 98.86    | 98.67  | 98.83  | 99.02    |
|           | 81  |                        | 7.02            | 45.56    | 15.82  | 33.52    | 55.15    |                        | 54.78           | 60.56    | 53.55  | 58.22    | 63.61    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 121 |                        | 10.25           | 68.37    | 42.64  | 59.95    | 75.50    |                        | 75.38           | 85.19    | 81.98  | 84.22    | 86.58    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
| 8         | 49  |                        | 13.67           | 79.52    | 58.63  | 74.64    | 84.73    |                        | 92.73           | 99.89    | 99.74  | 99.87    | 99.91    |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 81  |                        | 38.49           | 96.08    | 91.27  | 95.11    | 96.94    |                        | 99.89           | 100.00   | 100.00 | 100.00   | 100.00   |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |
|           | 121 |                        | 62.85           | 99.69    | 99.05  | 99.6     | 99.84    |                        | 100.00          | 100.00   | 100.00 | 100.00   | 100.00   |                        | 100.00          | 100.00   | 100.00 | 100.00 | 100.00   |

**Table 6** Rejection rates of  $(h, \phi)$ -divergence and Kolmogorov–Smirnov tests for scenarios (iii)–(v) for roughness pair  $(\alpha_1, \alpha_2) = (-3, -5)$ 

| Constants |     | $\gamma$               | $\mu_1 > \mu_2$ |          |        | $\mu_1 = \mu_2$ |          |                        | $\gamma$ | $\mu_1 < \mu_2$ |       |       |          |
|-----------|-----|------------------------|-----------------|----------|--------|-----------------|----------|------------------------|----------|-----------------|-------|-------|----------|
| $L$       | $N$ | $(\gamma_1, \gamma_2)$ | $S_{KS}$        | $S_{KL}$ | $S_T$  | $S_B$           | $S_{AG}$ | $(\gamma_1, \gamma_2)$ | $S_{KS}$ | $S_{KL}$        | $S_T$ | $S_B$ | $S_{AG}$ |
| 1         | 49  | (4, 4)                 | 19.2            | 32.15    | 20.93  | 28.45           | 37.20    | (2, 4)                 | 1.07     | 0.58            | 0.39  | 0.54  | 1.00     |
|           | 81  |                        | 44.62           | 62.18    | 51.78  | 59.56           | 65.80    |                        | 1.96     | 0.98            | 0.58  | 0.82  | 1.49     |
|           | 121 |                        | 66.18           | 85.15    | 80.11  | 84.24           | 86.91    |                        | 2.13     | 1.52            | 0.89  | 1.12  | 2.56     |
| 8         | 49  |                        | 81.25           | 94.27    | 91.28  | 93.94           | 95.16    |                        | 3.71     | 12.14           | 6.36  | 10.70 | 15.60    |
|           | 81  |                        | 98.49           | 98.75    | 98.6   | 98.75           | 98.79    |                        | 9.22     | 26.17           | 19.21 | 24.48 | 29.42    |
|           | 121 |                        | 99.98           | 98.21    | 98.2   | 98.21           | 98.21    |                        | 14.55    | 42.30           | 36.07 | 41.31 | 45.10    |
| 1         | 49  | (10, 4)                | 98.71           | 99.81    | 99.61  | 99.77           | 99.81    | (4, 8)                 | 1.16     | 0.54            | 0.38  | 0.50  | 1.15     |
|           | 81  |                        | 100.00          | 100.00   | 100.00 | 100.00          | 100.00   |                        | 1.69     | 1.36            | 0.81  | 1.11  | 1.90     |
|           | 121 |                        | 100.00          | 100.00   | 100.00 | 100.00          | 100.00   |                        | 2.53     | 2.11            | 1.27  | 1.71  | 2.93     |
| 8         | 49  |                        | 100.00          | 100.00   | 100.00 | 100.00          | 100.00   |                        | 3.15     | 13.05           | 6.36  | 11.37 | 16.25    |
|           | 81  |                        | 100.00          | 100.00   | 100.00 | 100.00          | 100.00   |                        | 8.56     | 25.55           | 18.48 | 23.98 | 28.87    |
|           | 121 |                        | 100.00          | 100.00   | 100.00 | 100.00          | 100.00   |                        | 12.98    | 43.64           | 36.85 | 42.35 | 46.68    |
| 1         | 49  | (10, 8)                | 46.93           | 65.05    | 52.45  | 62.71           | 69.91    | (10, 20)               | 0.91     | 0.70            | 0.31  | 0.54  | 1.05     |
|           | 81  |                        | 81.65           | 91.53    | 87.37  | 90.84           | 92.98    |                        | 2.18     | 1.30            | 0.96  | 1.06  | 2.11     |
|           | 121 |                        | 94.95           | 99.06    | 98.55  | 98.88           | 99.36    |                        | 2.29     | 2.19            | 1.25  | 1.58  | 3.10     |
| 8         | 49  |                        | 99.22           | 99.98    | 99.94  | 99.98           | 99.98    |                        | 2.96     | 12.71           | 6.67  | 11.31 | 16.28    |
|           | 81  |                        | 100.00          | 100.00   | 100.00 | 100.00          | 100.00   |                        | 8.91     | 26.29           | 19.33 | 24.93 | 29.79    |
|           | 121 |                        | 100.00          | 100.00   | 100.00 | 100.00          | 100.00   |                        | 13.58    | 42.24           | 36.07 | 41.11 | 44.92    |



**Fig. 6** Empirical test power in situations (iii)–(v) for different contrast measure at 1% nominal level



**Fig. 7** A corner reflector producing double bounce on a homogeneous background

Let us first consider the bias for uncontaminated data. The only situations for which the bias reduces with increasing sample size are: (a)  $\alpha = -1.5$ , (b)  $\alpha = -3$ ,

$L = 8$ , and (c)  $\alpha = -5$ ,  $L = 8$ . This seemingly surprising situation is frequent when dealing with the  $\mathcal{G}^0$  distribution. Indeed, the information content is reduced in homogeneous and heterogeneous situations; therefore, the good asymptotic properties of ML estimators are not verified in samples of the considered size [6, 15, 46]. An increase of the number of looks can partially compensate this behavior.

This loss of information content is also verified by the dependence of the bias on the number of looks. It is expected that the former is reduced increasing the latter, but this is only verified when  $\alpha = -1.5$ , i.e., in extremely heterogeneous areas. When dealing with heterogeneous data ( $\alpha = -3$ ), this reduction only takes place starting at  $N = 81$ , and with homogeneous data ( $\alpha = -5$ ) at  $N = 121$ . Again, the more homogeneous the sample is, the smaller its information content becomes under the  $\mathcal{G}^0$  model. Regarding the  $N = 121$  and  $L = 8$  situation, which is expected to reveal the asymptotic behavior of the estimators, the bias increases when  $\alpha$  is reduced from  $-1.5$  to  $-5$ .

**Table 7** Maximum likelihood estimates bias (B), coefficients of variation (CV%), and mean squared errors (MSE) for high roughness

| $\alpha$           | $\gamma$             | $L$ | $\epsilon$ | $\hat{\alpha}$ |       |      |      |       |      |          |       |      |      |       |        |
|--------------------|----------------------|-----|------------|----------------|-------|------|------|-------|------|----------|-------|------|------|-------|--------|
|                    |                      |     |            | $N = 49$       |       |      |      |       |      | $N = 81$ |       |      |      |       |        |
|                    |                      |     |            | B              | CV%   | MSE  | B    | CV%   | MSE  | B        | CV%   | MSE  | B    | CV%   | MSE    |
| -1.5               | 0.5                  | 1   | 0          | 0.23           | 37.64 | 0.47 | 0.18 | 32.50 | 0.33 | 0.14     | 27.15 | 0.22 | 0.12 | 56.10 | 0.13   |
|                    |                      |     |            | 0.11           | 23.58 | 0.15 | 0.06 | 17.39 | 0.08 | 0.04     | 13.95 | 0.05 | 0.05 | 29.83 | 0.03   |
|                    |                      |     |            | 0.22           | 37.70 | 0.50 | 0.17 | 32.24 | 0.34 | 0.12     | 26.86 | 0.21 | 0.24 | 56.60 | 0.60   |
|                    | 1                    | 1   | 0          | 0.10           | 22.73 | 0.15 | 0.06 | 17.30 | 0.08 | 0.05     | 14.01 | 0.05 | 0.09 | 28.24 | 0.11   |
|                    |                      |     |            | 0.20           | 36.63 | 0.47 | 0.19 | 32.59 | 0.33 | 0.13     | 26.74 | 0.20 | 0.55 | 54.94 | 3.47   |
|                    |                      |     |            | 0.12           | 23.31 | 0.15 | 0.06 | 17.38 | 0.08 | 0.04     | 13.87 | 0.05 | 0.26 | 29.26 | 0.66   |
|                    | 2.5                  | 1   | 0          | 0.31           | 36.75 | 0.54 | 0.27 | 32.33 | 0.40 | 0.23     | 27.15 | 0.27 | 0.19 | 52.79 | 0.17   |
|                    |                      |     |            | 0.17           | 23.78 | 0.19 | 0.12 | 18.03 | 0.10 | 0.10     | 14.08 | 0.06 | 0.09 | 28.50 | 0.04   |
|                    |                      |     |            | 0.29           | 36.41 | 0.51 | 0.26 | 31.36 | 0.37 | 0.23     | 27.19 | 0.28 | 0.37 | 53.47 | 0.67   |
|                    | 8                    | 1   | $10^{-4}$  | 0.17           | 23.51 | 0.18 | 0.13 | 17.45 | 0.10 | 0.10     | 14.16 | 0.06 | 0.18 | 28.72 | 0.15   |
|                    |                      |     |            | 0.30           | 36.71 | 0.52 | 0.28 | 32.05 | 0.40 | 0.23     | 26.78 | 0.26 | 0.91 | 52.60 | 4.05   |
|                    |                      |     |            | 0.17           | 23.34 | 0.18 | 0.12 | 17.54 | 0.09 | 0.10     | 14.20 | 0.06 | 0.47 | 28.79 | 0.95   |
| $5 \times 10^{-3}$ | 10 <sup>-3</sup>     | 1   | 0          | 0.94           | 47.42 | 2.22 | 0.84 | 42.28 | 1.68 | 0.75     | 35.29 | 1.19 | 0.74 | 60.30 | 1.10   |
|                    |                      |     |            | 0.44           | 25.90 | 0.44 | 0.38 | 19.51 | 0.28 | 0.35     | 15.16 | 0.20 | 0.35 | 29.15 | 0.19   |
|                    |                      |     |            | 0.91           | 47.46 | 2.14 | 0.82 | 40.98 | 1.58 | 0.75     | 34.74 | 1.17 | 1.44 | 60.88 | 4.30   |
|                    | 5 × 10 <sup>-3</sup> | 1   | 0          | 0.44           | 25.98 | 0.44 | 0.38 | 19.06 | 0.27 | 0.35     | 15.14 | 0.20 | 0.70 | 29.95 | 0.75   |
|                    |                      |     |            | 0.93           | 48.01 | 2.23 | 0.85 | 41.88 | 1.69 | 0.75     | 34.69 | 1.16 | 3.65 | 60.33 | 27.04  |
|                    |                      |     |            | 0.44           | 25.83 | 0.45 | 0.37 | 18.84 | 0.26 | 0.35     | 15.12 | 0.20 | 1.76 | 29.81 | 4.73   |
|                    | 5 × 10 <sup>-3</sup> | 1   | 0          | 1.45           | 44.24 | 3.79 | 1.42 | 40.45 | 3.42 | 1.41     | 36.47 | 3.11 | 1.93 | 53.58 | 5.41   |
|                    |                      |     |            | 0.75           | 26.83 | 0.93 | 0.69 | 20.99 | 0.68 | 0.64     | 16.35 | 0.53 | 1.05 | 29.32 | 1.30   |
|                    |                      |     |            | 1.43           | 44.82 | 3.78 | 1.43 | 39.83 | 3.40 | 1.42     | 36.66 | 3.16 | 3.84 | 55.04 | 21.86  |
|                    | 5 × 10 <sup>-3</sup> | 1   | 0          | 0.75           | 27.64 | 0.95 | 0.68 | 20.33 | 0.66 | 0.64     | 16.13 | 0.53 | 2.10 | 30.80 | 5.31   |
|                    |                      |     |            | 1.46           | 45.50 | 3.93 | 1.42 | 40.18 | 3.40 | 1.41     | 36.94 | 3.16 | 9.69 | 55.03 | 138.98 |
|                    |                      |     |            | 0.76           | 27.45 | 0.97 | 0.68 | 20.17 | 0.65 | 0.64     | 16.01 | 0.53 | 5.28 | 30.76 | 33.60  |
|                    | 5 × 10 <sup>-3</sup> | 1   | 0          | 0.09           | 47.64 | 0.09 | 0.09 | 47.64 | 0.09 | 0.09     | 47.64 | 0.09 | 0.09 | 47.64 | 0.09   |
|                    |                      |     |            | 0.01           | 21.76 | 0.01 | 0.03 | 21.76 | 0.01 | 0.03     | 21.76 | 0.01 | 0.03 | 21.76 | 0.01   |
|                    |                      |     |            | 0.13           | 39.59 | 0.23 | 0.13 | 39.59 | 0.23 | 0.13     | 39.59 | 0.23 | 0.13 | 39.59 | 0.23   |

**Table 8** Maximum likelihood estimates bias (B), coefficients of variation (CV%), and mean squared errors (MSE) for mid roughness

| $\alpha$ | $\gamma$ | $L$ | $\epsilon$         | $\hat{\alpha}$ |       |      |      |       |      |          |       |      |       |       |        | $\hat{\gamma}$ |       |        |       |       |        |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|----------|----------|-----|--------------------|----------------|-------|------|------|-------|------|----------|-------|------|-------|-------|--------|----------------|-------|--------|-------|-------|--------|----------|-----|-----|---|-----|-----|----------|-----|-----|---|-----|-----|-----------|--|--|--|--|--|
|          |          |     |                    | $N = 49$       |       |      |      |       |      | $N = 81$ |       |      |       |       |        | $N = 121$      |       |        |       |       |        | $N = 49$ |     |     |   |     |     | $N = 81$ |     |     |   |     |     | $N = 121$ |  |  |  |  |  |
|          |          |     |                    | B              | CV%   | MSE  | B    | CV%   | MSE  | B        | CV%   | MSE  | B     | CV%   | MSE    | B              | CV%   | MSE    | B     | CV%   | MSE    | B        | CV% | MSE | B | CV% | MSE | B        | CV% | MSE | B | CV% | MSE |           |  |  |  |  |  |
| -3.0     | 2.0      | 1   | 0                  | 0.15           | 42.11 | 1.81 | 0.29 | 37.21 | 1.62 | 0.29     | 34.61 | 1.42 | 0.16  | 56.46 | 1.55   | 0.25           | 48.36 | 1.27   | 0.25  | 44.09 | 1.12   |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          | 8   |                    | 0.31           | 27.28 | 0.90 | 0.18 | 22.14 | 0.49 | 0.12     | 17.58 | 0.30 | 0.24  | 31.38 | 0.55   | 0.14           | 25.21 | 0.29   | 0.09  | 20.19 | 0.18   |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          | 4.0      | 1   |                    | 0.17           | 41.35 | 1.79 | 0.29 | 37.47 | 1.63 | 0.30     | 34.99 | 1.45 | 0.35  | 55.24 | 5.95   | 0.53           | 48.57 | 5.35   | 0.53  | 44.84 | 4.33   |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          | 8   |                    | 0.29           | 27.51 | 0.88 | 0.18 | 21.29 | 0.49 | 0.11     | 18.28 | 0.32 | 0.45  | 31.65 | 2.14   | 0.28           | 24.47 | 1.17   | 0.18  | 20.84 | 0.75   |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          | 10.0     | 1   |                    | 0.18           | 41.92 | 1.84 | 0.32 | 38.17 | 1.66 | 0.31     | 35.07 | 1.51 | 0.90  | 55.33 | 39.62  | 1.42           | 49.42 | 32.20  | 1.38  | 44.93 | 29.69  |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          | 8   |                    | 0.28           | 27.79 | 0.88 | 0.16 | 21.15 | 0.51 | 0.12     | 17.93 | 0.29 | 1.09  | 31.91 | 13.19  | 0.65           | 24.30 | 7.71   | 0.48  | 20.61 | 4.36   |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     | $10^{-4}$          | 0.32           | 41.92 | 2.03 | 0.43 | 37.48 | 1.84 | 0.46     | 34.04 | 1.60 | 0.33  | 54.06 | 1.69   | 0.42           | 48.17 | 1.54   | 0.44  | 43.13 | 1.31   |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     |                    | 0.40           | 27.47 | 1.03 | 0.29 | 21.30 | 0.57 | 0.23     | 17.48 | 0.37 | 0.35  | 31.15 | 0.66   | 0.26           | 24.49 | 0.37   | 0.22  | 19.85 | 0.24   |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     |                    | 0.33           | 40.13 | 1.90 | 0.47 | 37.01 | 1.87 | 0.47     | 33.76 | 1.59 | 0.75  | 52.75 | 6.85   | 0.92           | 47.45 | 6.29   | 0.91  | 43.08 | 5.30   |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     |                    | 0.38           | 27.10 | 0.99 | 0.29 | 21.72 | 0.59 | 0.23     | 17.18 | 0.36 | 0.68  | 30.78 | 2.53   | 0.53           | 24.77 | 1.53   | 0.43  | 19.70 | 0.94   |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     |                    | 0.31           | 41.12 | 1.95 | 0.40 | 36.90 | 1.74 | 0.48     | 33.94 | 1.63 | 1.68  | 54.01 | 42.65  | 2.05           | 47.10 | 36.42  | 2.31  | 42.67 | 32.91  |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     | $10^{-3}$          | 0.39           | 27.33 | 1.01 | 0.30 | 21.61 | 0.60 | 0.23     | 17.03 | 0.36 | 1.68  | 30.98 | 15.93  | 1.39           | 24.67 | 9.82   | 1.09  | 19.26 | 5.74   |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     |                    | 0.84           | 35.35 | 2.54 | 1.14 | 31.11 | 2.95 | 1.30     | 28.34 | 3.18 | 1.22  | 44.44 | 3.55   | 1.49           | 38.04 | 3.98   | 1.65  | 34.06 | 4.28   |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     |                    | 1.07           | 25.91 | 2.24 | 1.03 | 21.99 | 1.84 | 0.97     | 18.74 | 1.50 | 1.27  | 28.36 | 2.49   | 1.24           | 24.24 | 2.15   | 1.18  | 20.27 | 1.82   |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     |                    | 0.93           | 34.41 | 2.68 | 1.16 | 30.90 | 3.01 | 1.34     | 27.64 | 3.22 | 2.70  | 43.67 | 15.85  | 3.04           | 37.61 | 16.23  | 3.35  | 33.24 | 17.21  |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     |                    | 1.07           | 25.88 | 2.25 | 1.03 | 22.05 | 1.84 | 0.97     | 18.47 | 1.48 | 2.55  | 28.48 | 10.00  | 2.47           | 24.07 | 8.53   | 2.37  | 20.11 | 7.26   |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     |                    | 0.87           | 35.12 | 2.59 | 1.10 | 31.03 | 2.84 | 1.35     | 28.24 | 3.33 | 6.31  | 44.07 | 91.51  | 7.47           | 38.16 | 100.21 | 8.47  | 33.70 | 110.56 |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     |                    | 1.07           | 25.95 | 2.26 | 1.05 | 22.39 | 1.93 | 0.99     | 18.49 | 1.52 | 6.35  | 28.37 | 61.86  | 6.31           | 24.51 | 55.79  | 5.99  | 19.97 | 46.09  |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     | $5 \times 10^{-3}$ | 1.64           | 27.29 | 4.29 | 1.93 | 24.11 | 5.12 | 2.19     | 20.16 | 5.91 | 3.99  | 33.89 | 20.00  | 4.33           | 28.75 | 22.08  | 4.72  | 23.54 | 24.78  |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     |                    | 2.03           | 20.87 | 5.22 | 2.18 | 18.18 | 5.62 | 2.28     | 15.81 | 5.87 | 4.05  | 23.00 | 18.33  | 4.22           | 19.88 | 19.32  | 4.34  | 17.00 | 19.96  |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     |                    | 1.62           | 28.30 | 4.34 | 1.96 | 23.49 | 5.18 | 2.26     | 19.92 | 6.20 | 7.85  | 34.20 | 77.99  | 8.79           | 28.54 | 90.59  | 9.61  | 23.78 | 102.92 |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     |                    | 2.03           | 20.69 | 5.23 | 2.18 | 17.96 | 5.64 | 2.27     | 15.97 | 5.88 | 8.10  | 22.78 | 73.24  | 8.47           | 19.56 | 77.78  | 8.67  | 17.21 | 79.98  |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     |                    | 1.52           | 28.95 | 4.02 | 1.88 | 23.62 | 4.86 | 2.14     | 21.23 | 5.76 | 19.13 | 35.19 | 471.04 | 21.52          | 27.91 | 540.58 | 23.29 | 25.41 | 614.05 |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |
|          |          |     |                    | 2.05           | 20.42 | 5.27 | 2.19 | 18.12 | 5.69 | 2.30     | 15.95 | 6.00 | 20.35 | 22.38 | 460.41 | 21.22          | 19.70 | 488.21 | 21.85 | 17.19 | 507.52 |          |     |     |   |     |     |          |     |     |   |     |     |           |  |  |  |  |  |



**Table 9** Maximum likelihood estimates bias (B), coefficients of variation (CV%), and mean squared errors (MSE) for low roughness

| $\alpha$ | $\gamma$ | $L$ | $\epsilon$         | $\hat{\gamma}$ |       |       |      |          |       |      |       |           |       |       |         |
|----------|----------|-----|--------------------|----------------|-------|-------|------|----------|-------|------|-------|-----------|-------|-------|---------|
|          |          |     |                    | $N = 49$       |       |       |      | $N = 81$ |       |      |       | $N = 121$ |       |       |         |
|          |          |     |                    | B              | CV%   | MSE   | B    | CV%      | MSE   | B    | CV%   | MSE       | B     | CV%   | MSE     |
| -5.0     | 4.0      | 1   | 0                  | 0.33           | 46.97 | 4.85  | 0.02 | 42.92    | 4.75  | 0.22 | 39.77 | 4.27      | 0.30  | 58.61 | 4.73    |
|          |          |     |                    | 0.52           | 30.38 | 3.01  | 0.35 | 28.62    | 2.04  | 0.13 | 26.71 | 1.28      | 0.47  | 33.62 | 2.42    |
|          |          |     |                    | 0.30           | 46.86 | 5.11  | 0.06 | 42.50    | 4.75  | 0.32 | 40.23 | 4.23      | 0.48  | 57.82 | 19.89   |
|          |          |     |                    | 0.51           | 30.85 | 3.14  | 0.28 | 29.04    | 2.08  | 0.11 | 27.77 | 1.28      | 0.92  | 34.18 | 10.23   |
|          | 8.0      | 1   |                    | 0.25           | 47.26 | 4.99  | 0.00 | 43.57    | 4.74  | 0.18 | 39.53 | 4.13      | 1.09  | 58.27 | 122.69  |
|          |          |     |                    | 0.52           | 30.61 | 2.99  | 0.31 | 28.25    | 2.04  | 0.13 | 27.63 | 1.30      | 2.28  | 33.64 | 60.17   |
|          |          |     |                    | 0.14           | 46.39 | 5.11  | 0.20 | 42.56    | 4.93  | 0.39 | 37.67 | 4.26      | 0.02  | 57.34 | 5.21    |
|          |          |     |                    | 0.70           | 30.21 | 3.44  | 0.57 | 25.28    | 2.31  | 0.46 | 21.34 | 1.57      | 0.68  | 33.27 | 2.90    |
|          |          | 8   | $10^{-4}$          | 0.14           | 46.19 | 5.06  | 0.21 | 41.15    | 4.65  | 0.49 | 38.28 | 4.66      | 0.05  | 56.87 | 20.43   |
|          |          |     |                    | 0.68           | 30.30 | 3.43  | 0.58 | 25.85    | 2.41  | 0.44 | 20.82 | 1.48      | 1.33  | 33.63 | 11.63   |
|          |          |     |                    | 0.19           | 46.55 | 5.04  | 0.17 | 41.88    | 4.71  | 0.48 | 38.78 | 4.70      | 0.37  | 57.22 | 126.31  |
|          |          |     |                    | 0.64           | 29.89 | 3.26  | 0.55 | 25.60    | 2.32  | 0.45 | 20.83 | 1.49      | 3.16  | 32.88 | 67.97   |
|          |          |     | $10^{-3}$          | 0.76           | 38.72 | 5.54  | 1.38 | 34.95    | 6.87  | 1.73 | 30.71 | 7.29      | 1.48  | 46.67 | 8.71    |
|          |          |     |                    | 1.84           | 26.58 | 6.68  | 1.93 | 23.19    | 6.30  | 1.93 | 20.90 | 5.81      | 2.40  | 28.81 | 9.15    |
|          |          |     |                    | 0.83           | 38.48 | 5.72  | 1.38 | 33.55    | 6.48  | 1.76 | 30.81 | 7.45      | 3.05  | 46.15 | 35.33   |
|          |          |     |                    | 1.80           | 26.35 | 6.47  | 1.89 | 23.43    | 6.20  | 1.92 | 20.74 | 5.76      | 4.68  | 28.27 | 34.79   |
|          |          |     |                    | 0.77           | 38.63 | 5.57  | 1.32 | 33.96    | 6.36  | 1.71 | 30.66 | 7.16      | 7.69  | 47.46 | 231.81  |
|          |          |     |                    | 1.84           | 26.76 | 6.72  | 1.88 | 23.52    | 6.13  | 1.93 | 20.86 | 5.82      | 11.90 | 28.62 | 224.98  |
|          |          |     |                    | 1.90           | 31.41 | 25.32 | 2.66 | 26.17    | 29.23 | 3.39 | 19.56 | 41.10     | 5.69  | 35.86 | 100.86  |
|          |          |     |                    | 3.52           | 18.38 | 38.51 | 3.97 | 14.84    | 42.14 | 4.21 | 12.92 | 44.56     | 7.56  | 20.03 | 124.35  |
|          |          |     | $5 \times 10^{-3}$ | 2.23           | 29.07 | 28.14 | 2.74 | 24.23    | 32.54 | 3.19 | 22.56 | 43.42     | 12.41 | 33.54 | 441.66  |
|          |          |     |                    | 3.46           | 18.40 | 38.25 | 3.89 | 15.32    | 40.99 | 4.24 | 12.72 | 44.41     | 14.95 | 19.77 | 490.83  |
|          |          |     |                    | 2.07           | 29.49 | 19.21 | 2.64 | 23.50    | 25.52 | 3.26 | 21.64 | 36.30     | 30.16 | 34.71 | 1959.53 |
|          |          |     |                    | 3.45           | 19.05 | 37.62 | 3.88 | 15.45    | 41.52 | 4.19 | 13.21 | 44.84     | 37.37 | 20.75 | 3047.61 |

Regarding the influence of contamination on the bias, in all but a single case (namely  $N = 49$ ,  $L = 1$ ,  $\epsilon = 10^{-4}$ ), increasing the former increases the latter for both estimators. This increase in the bias ranges from, approximately, 34% ( $\alpha = -1.5$ ,  $L = 1$ ,  $\gamma = 0.5$ ,  $\epsilon = 10^{-4}$ ) to 1,000% ( $\alpha = -3$ ,  $L = 1$ ,  $\gamma = 2$ ,  $\epsilon = 5 \times 10^{-3}$ ), in both cases  $N = 49$ .

The coefficient of variation is reduced by increasing  $L$  and/or  $N$  and/or  $\alpha$ , as expected, in uncontaminated data. In the presence of contamination, this behavior is lost with respect to  $\alpha$ .

The usual behavior of reduced MSE with increasing  $L$  and/or  $N$  and/or  $\alpha$  is observed in the absence of contamination. The presence of contamination consistently increases the MSE, except in four out of 324 situations, where random fluctuations masked the results. For instance, the MSE of  $\hat{\gamma}$  doubled in the  $\alpha = -5$ ,  $\gamma = 20$ ,  $L = 1$ ,  $N = 49$  situation when  $\epsilon = 10^{-4}$ , i.e., the mildest level of contamination. If

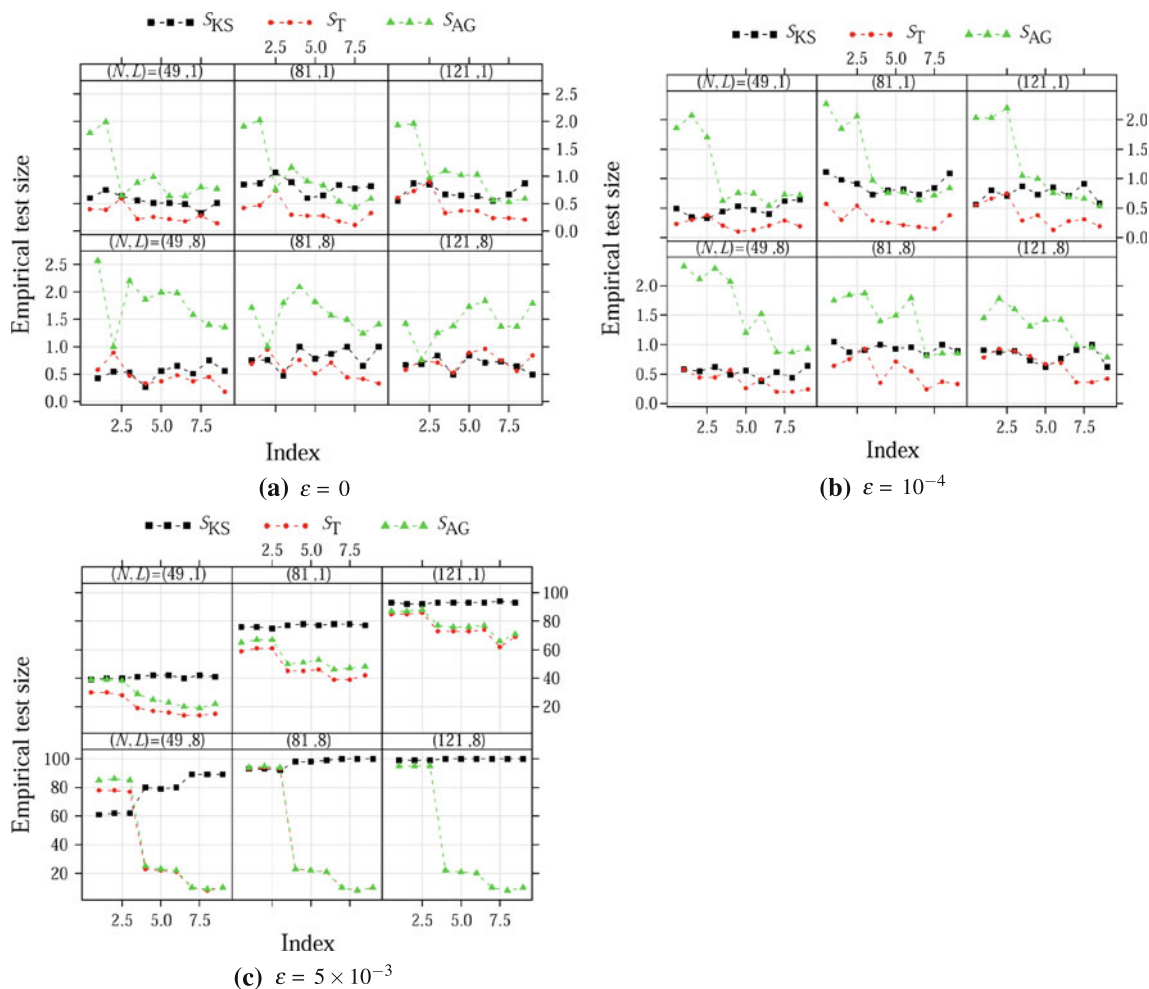
$\alpha = -5$ ,  $\gamma = 20$ ,  $L = 8$ ,  $N = 121$ , and  $\epsilon = 5 \times 10^{-3}$ , i.e., the strongest contamination, the MSE was affected by a factor of 13,794.

This strong influence of contamination on the MSE warns users about the need of quality samples and of robust estimators when dealing with this type of data.

#### 4.3 Test sizes in contaminated data

Based on the above results, the forthcoming robustness study is limited to the best statistics, namely  $S_T$  and  $S_{AG}$ , along with  $S_{KS}$ , which is expected to be robust. For these measures, Fig. 8 shows the behavior of the test size in the presence of fixed levels  $\epsilon$  of contamination.

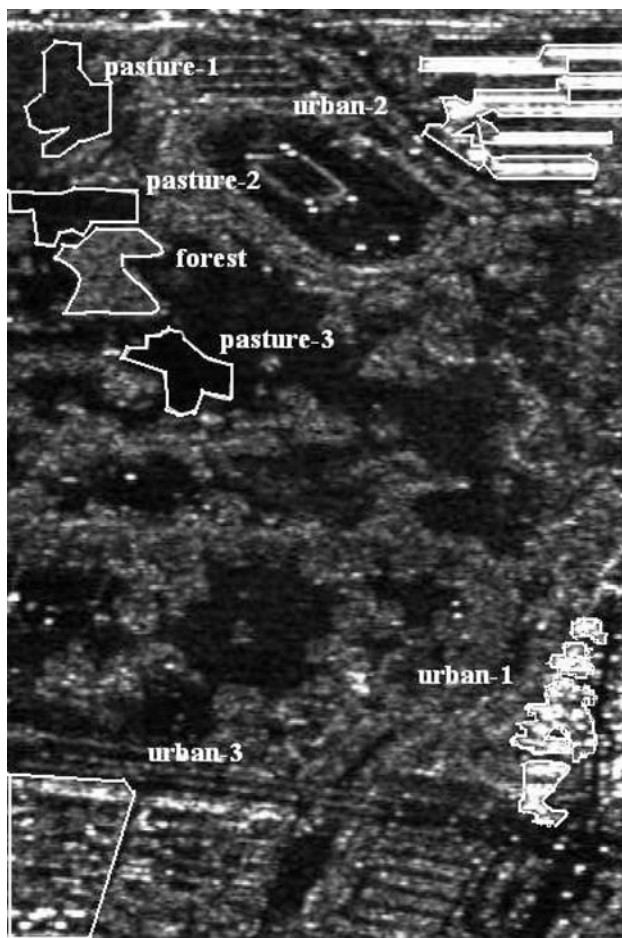
Under contamination, larger windows imply better empirical test sizes for the parametric tests. Tests based on  $S_T$  are the most robust for homogeneous regions. The empirical test sizes of  $(h, \phi)$ -distance based tests for



**Fig. 8** Test sizes at 1% nominal level and  $\epsilon$ . For each sample size, all considered scenarios were lexicographically ordered:  $[i: (i; \alpha, \mu) = \{(1; -1.5, 1), \dots, (4; -3, 1), \dots, (7; -5, 1), \dots\}]$ . Index  $i$  was employed as abscissa value

$\alpha = -1.5$  are greater than the ones observed when  $\alpha \in \{-3, -5\}$ ; this behavior is more pronounced as  $\epsilon$  increases.

Unexpectedly, the Kolmogorov–Smirnov test tends to be less robust than parametric tests as contamination increases. Figure 8 indicates a low variability of the  $S_{KS}$  based test size across different simulation specifications. This suggests that image roughness and brightness do not significantly affect these particular estimated test sizes.



**Fig. 9** E-SAR image and selected regions

## 5 Results

In the following, we present an application of the proposed tools to real data. The performance of the considered hypothesis tests is quantified by means of their empirical sizes and powers. Based on Fig. 9, samples of same and different classes are contrasted.

The image was obtained by the E-SAR sensor over surroundings of Munich, Germany [25], and its estimated number of looks is 3.2. The area exhibits three distinct types of target roughness: (1) homogeneous (corresponding to pasture), (2) heterogeneous (forest), and (3) extremely heterogeneous (urban areas). Samples were selected and submitted to statistical analysis, after parameter estimation by maximum likelihood methods.

Table 10 shows the  $\hat{\alpha}$ ,  $\hat{\gamma}$ , and  $\hat{\mu}$  estimates (whose interpretability was discussed in Sect. 4.1), along with the sample size (# pixels). Values of  $\hat{\mu}$  for urban regions are infinite in accordance to the theoretical background as seen in (4). The table also presents the number of disjoint  $7 \times 7$  pixels blocks (# parts) which were used to estimate the test sizes and powers.

The estimated roughness and brightness parameters satisfied the following inequalities:  $\alpha_{\text{pasture-1}} < \alpha_{\text{pasture-2}} < \alpha_{\text{pasture-3}} < \alpha_{\text{forest}} < \alpha_{\text{urban-3}} < \alpha_{\text{urban-1}} < \alpha_{\text{urban-2}}$  and  $\mu_{\text{pasture-1}} < \mu_{\text{pasture-2}} < \mu_{\text{pasture-3}} < \mu_{\text{forest}} < \mu_{\text{urban-3}}$ . Those inequalities confirm the interpretability of the parameters, and show how hard it is to discern different samples from the same class.

Table 11 presents the rejection rates at significance levels 1 and 10% of samples from the same area, i.e., the test size or Type I error. Such nominal levels were considered useful in practical applications (cf. [43]). These results show that  $S_{KS}$ ,  $S_{KL}$ , and  $S_T$  tests have excellent performance with respect to this criterion, and that the hardest situation is related to urban areas. Furthermore, the  $S_{AG}$  test shows optimal size test for 1% nominal level, failing on urban-1 and urban-2 situations at 10%, while the  $S_B$  test size is inappropriate in case urban-2.

Table 12 shows the test powers at 1 and 10% nominal levels, i.e., the inability to reject samples from different types. The best performances are observed in  $(h, \phi)$ -divergence

**Table 10** Parameter estimates

| Regions   | $\hat{\alpha}$ | $\hat{\gamma}$ | $\hat{\mu}$ | # pixels | # parts |
|-----------|----------------|----------------|-------------|----------|---------|
| Pasture-1 | −15.702        | 39,259         | 2,670.32    | 1,235    | 25      |
| Pasture-2 | −12.698        | 80,320         | 6,866.13    | 1,216    | 24      |
| Pasture-3 | −11.304        | 162,292        | 15,750.39   | 1,602    | 32      |
| Forest    | −9.339         | 661,183        | 79,288.04   | 1,606    | 32      |
| Urban-1   | −0.759         | 148,413        | $\infty$    | 2,005    | 40      |
| Urban-2   | −0.388         | 110,183        | $\infty$    | 3,481    | 71      |
| Urban-3   | −1.079         | 55,583         | 703,582.30  | 4,657    | 95      |

**Table 11** Rejection rates of  $(h, \phi)$ -divergence tests under  $H_1: (\alpha_1, \gamma_1) \neq (\alpha_2, \gamma_2)$ , with  $\mu_1 = \mu_2$  at 1 and 10 % nominal levels

| Regions   | 1% Nominal level |          |       |       |          | 10% Nominal level |          |       |       |          |
|-----------|------------------|----------|-------|-------|----------|-------------------|----------|-------|-------|----------|
|           | $S_{KS}$         | $S_{KL}$ | $S_T$ | $S_B$ | $S_{AG}$ | $S_{KS}$          | $S_{KL}$ | $S_T$ | $S_B$ | $S_{AG}$ |
| Pasture-1 | 0.33             | 0.00     | 0.00  | 0.00  | 0.00     | 5.00              | 0.00     | 0.00  | 0.00  | 0.00     |
| Pasture-2 | 0.36             | 0.00     | 0.00  | 0.00  | 0.00     | 6.16              | 0.00     | 0.00  | 0.00  | 0.00     |
| Pasture-3 | 0.60             | 0.00     | 0.00  | 0.36  | 0.36     | 6.05              | 3.62     | 3.62  | 3.62  | 3.99     |
| Forest    | 2.22             | 0.00     | 0.00  | 0.00  | 0.00     | 9.68              | 2.67     | 1.67  | 2.00  | 2.33     |
| Urban-1   | 0.00             | 0.00     | 0.00  | 0.13  | 0.00     | 2.18              | 4.10     | 2.18  | 34.49 | 90.64    |
| Urban-2   | 0.00             | 0.00     | 0.00  | 99.96 | 0.00     | 0.97              | 2.01     | 1.21  | 99.96 | 100.00   |
| Urban-3   | 0.29             | 4.23     | 1.97  | 3.56  | 5.64     | 5.98              | 20.58    | 16.84 | 20.85 | 31.02    |

**Table 12** Rejection rates of  $(h, \phi)$ -divergence tests under  $H_1: (\alpha_1, \gamma_1) \neq (\alpha_2, \gamma_2)$ , with  $\mu_1 = \mu_2$  at 1 and 10% nominal levels

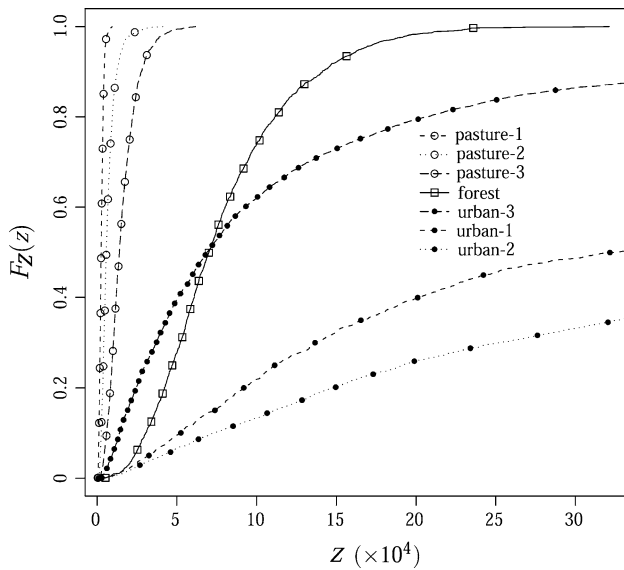
| Regions                      | 1% Nominal level |          |        |        |          | 10% Nominal level |          |        |        |          |
|------------------------------|------------------|----------|--------|--------|----------|-------------------|----------|--------|--------|----------|
|                              | $S_{KS}$         | $S_{KL}$ | $S_T$  | $S_B$  | $S_{AG}$ | $S_{KS}$          | $S_{KL}$ | $S_T$  | $S_B$  | $S_{AG}$ |
| Pasture-1 $\times$ pasture-2 | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| Pasture-1 $\times$ pasture-3 | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| Pasture-2 $\times$ pasture-3 | 100.00           | 94.64    | 87.50  | 90.48  | 95.24    | 100.00            | 98.21    | 97.02  | 98.21  | 98.21    |
| Forest $\times$ pasture-1    | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| Forest $\times$ pasture-2    | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| Forest $\times$ pasture-3    | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| Urban-1 $\times$ pasture-1   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| Urban-1 $\times$ pasture-2   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| Urban-1 $\times$ pasture-3   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| Urban-2 $\times$ pasture-1   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| Urban-2 $\times$ pasture-2   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| Urban-2 $\times$ pasture-3   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| Urban-3 $\times$ pasture-1   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| Urban-3 $\times$ pasture-2   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| Urban-3 $\times$ pasture-3   | 100.00           | 100.00   | 100.00 | 100.00 | 100.00   | 100.00            | 100.00   | 100.00 | 100.00 | 100.00   |
| Urban-1 $\times$ forest      | 100.00           | 98.70    | 98.10  | 99.70  | 98.40    | 100.00            | 98.90    | 99.40  | 100.00 | 100.00   |
| Urban-2 $\times$ forest      | 100.00           | 98.03    | 99.49  | 100.00 | 97.30    | 100.00            | 98.03    | 99.49  | 100.00 | 100.00   |
| Urban-3 $\times$ forest      | 12.17            | 94.74    | 80.17  | 91.79  | 96.17    | 61.78             | 98.65    | 96.93  | 98.40  | 97.14    |
| Urban-1 $\times$ urban-2     | 32.04            | 29.68    | 26.37  | 94.19  | 30.60    | 83.87             | 53.20    | 51.73  | 99.47  | 100.00   |
| Urban-2 $\times$ urban-3     | 98.58            | 91.05    | 89.92  | 94.26  | 91.50    | 100.00            | 96.26    | 96.13  | 97.29  | 98.53    |
| Urban-1 $\times$ urban-3     | 100.00           | 98.84    | 98.50  | 100.00 | 99.01    | 100.00            | 99.84    | 99.79  | 100.00 | 100.00   |

tests, mainly when confronting areas whose distributions differ more significantly; among them,  $S_T$  and  $S_{KL}$  are the weakest ones. The worst result was observed in the urban-1 versus urban-2 situation whose samples are, in fact, very close and hard to discern. The Kolmogorov–Smirnov test produces unacceptable results when contrasting urban-3 versus forest, and urban-1 versus urban-2 samples. In particular, the situation urban-3 versus forest contrasts two noticeably distinct regions, as shown in Fig. 9 and Table 10. Figure 10 presents the empirical distribution functions of the areas considered. It illustrates the close connection of the  $S_{KS}$  efficiency to these functions, regardless the texture

differences, which makes the  $S_{KS}$  test an inadequate tool for discrimination in speckled imagery.

## 6 Conclusions

This paper presented a comparison among the performance of four parametric tests based on stochastic distances and the Kolmogorov–Smirnov test. We also presented compact formulas for the Kolmogorov–Smirnov contrast measure, and a study of the robustness of the maximum likelihood estimator under the  $\mathcal{G}^0$  law. The assessment was made with



**Fig. 10** Empirical cumulative distribution function for different regions in SAR data for sample values

Monte Carlo experiments varying the parameters of the  $\mathcal{G}^0$  distribution, regarded as a universal model for speckled data, and the level of innovative contamination. In doing so, we provided new results regarding both the accuracy and the robustness of the measures under examination.

We show numerical evidence that the test based on the triangular distance  $S_T$  has, in general, smaller empirical Type I error than the test based on the Kolmogorov–Smirnov distance  $S_{KS}$ . Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data increase, the test based on  $S_{KS}$  has its performance diminished with respect to the results obtained from the  $(h, \phi)$ -distance methods. This is an important warning due to the widespread use of the Kolmogorov–Smirnov test when dealing with non-normal data.

The  $S_{AG}$  measure presented the best performance with respect to the test power. However, for a given number of looks, we observed that the test power performance of the four proposed parametric measures was roughly the same, and that they were more accurate than the test related to the  $S_{KS}$  measure.

These conclusions suggest the use of the  $S_T$  for SAR data processing due to its resulting accurate test empirical size. On the other hand, the  $S_{AG}$  test is a more appropriate tool whenever the focus is on the test power.

We illustrated the discriminatory capability of these measures using real data. Tests based on the classical measures here tested, namely the Kullback–Leibler, Kolmogorov–Smirnov, and Bhattacharyya distances, were outperformed by the other distances and should not be employed.

**Acknowledgments** Authors are grateful to CNPq and FACEPE for funding this research.

## Computational information

Simulations of the  $(h, \phi)$ -distance based parametric tests were performed in Ox programming environment [10]; function QAGI was employed for numerical integration. The Kolmogorov–Smirnov test was coded in R and function `ks.test` was called [35]. The computation time of the triangular and arithmetic-geometric distances took typically less than one and four millisecond, respectively, when performed in a Pentium processor at 3.20 GHz. We used the George Marsaglia’s multiply-with-carry with 52 bits pseudorandom number generator, which has an approximate period of  $2^{8222}$ .

## References

1. Allende H, Frery AC, Galbiati J, Pizarro L (2006) M-estimators with asymmetric influence functions: the GA0 distribution case. *J Stat Comput Simul* 76(11):941–956
2. Andai A (2009) On the geometry of generalized Gaussian distributions. *J Multivar Anal* 100(4):777–793
3. Anfinson SN, Doulgeris AP, Eltoft T (2009) Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery. *IEEE Trans Geosci Remote Sens* 47(11): 3795–3809
4. Blacknell D, Blake AP, Oliver CJ (1995) High resolution SAR clutter texture analysis and simulation. In: *SPIE conference synthetic aperture radar and passive microwave sensing*, vol 2584. Paris, France, pp 101–108
5. Bustos OH, Lucini MM, Frery AC (2002) M-estimators of roughness and scale for GA0-modelled SAR imagery. *EURASIP J Appl Signal Process* 2002(1): 105–114
6. Cribari-Neto F, Frery AC, Silva MF (2002) Improved estimation of clutter properties in speckled imagery. *Comput Stat Data Anal* 40(4):801–824
7. Cribari-Neto F, Zarkos SG (1999) R: yet another econometric programming environment. *J Appl Econom* 14:319–329
8. Daba JS, Bell MR (1994) Statistical distributions of partially developed speckle based on a small number of constant scatterers with random phase. In: *Geoscience and remote sensing symposium IGARSS '94*, vol 4, pp 2338–2341
9. Donohue KD, Rahmati M, Hassebrook LG, Gopalakrishnan P (1993) Parametric and nonparametric edge detection for speckle degraded images. *Opt Eng* 32(8):1935–1946
10. Doornik JA (1999) *Object-oriented matrix programming using Ox*, 3 edn.
11. Doulgeris AP, Eltoft T (2010) Scale mixture of Gaussian modelling of polarimetric SAR data. *EURASIP J Adv Signal Process* 2010(874592)
12. Fox AJ (1972) Outliers in time series. *J Royal Stat Soc Ser B (Methodological)* 34(3):350–363
13. Freitas CC, Frery AC, Correia AH (2005) The polarimetric  $G$  distribution for SAR data analysis. *Environmetrics* 16(1):13–31
14. Frery AC, Correia AH, Freitas CC (2007) Classifying multifrequency fully polarimetric imagery with multiple sources of

- statistical evidence and contextual information. *IEEE Trans Geosci Remote Sens* 45:3098–3109
15. Frery AC, Cribari-Neto F, Souza MO (2004) Analysis of minute features in speckled imagery with maximum likelihood estimation. *EURASIP J Appl Signal Process* 2004(16):2476–2491
  16. Frery AC, Muller HJ, Yanasse CCF, Sant'Anna SJS (1997) A model for extremely heterogeneous clutter. *IEEE Trans Geosci Remote Sens* 35(3): 648–659
  17. Frery AC, Nascimento ADC, Cintra RJ (2010) Contrast in speckled imagery with stochastic distances. In: International conference on image processing (ICIP), Hong Kong, pp 26–29
  18. Frery AC, Sant'Anna SJS, Mascarenhas NDA, Bustos OH (1997) Robust inference techniques for speckle noise reduction in 1-look amplitude SAR images. *Appl Signal Process* 4:61–76
  19. Galland F, Nicolas JM, Sportouche H, Roche M, Tupin F, Réfrégier P (2009) Unsupervised synthetic aperture radar image segmentation using Fisher distributions. *IEEE Trans Geosci Remote Sens* 47(8): 2966–2972
  20. Gambini J, Mejail M, Jacobo-Berlles J, Frery AC (2008) Accuracy of edge detection methods with local information in speckled imagery. *Stat Comput* 18(1):15–26
  21. Gao G (2010) Statistical modeling of SAR images: a survey. *Sensors* 10:775–795
  22. Goudail F, Réfrégier P (2004) Contrast definition for optical coherent polarimetric images. *IEEE Trans Pattern Anal Mach Intell* 26(7):947–951
  23. Gudnason J, Cui J, Brookes M (2009) HRR automatic target recognition from superresolution scattering center features. *IEEE Trans Aerosp Electron Syst* 45(4): 1512–1524
  24. Hoekman DH, Quiñones MJ (2000) Land cover type and biomass classification using AirSAR data for evaluation of monitoring scenarios in the Colombian Amazon. *IEEE Trans Geosci Remote Sens* 38: 685–696
  25. Horn R (1996) The DLR airborne SAR project E-SAR. In: Geoscience and remote sensing symposium, vol 3. IEEE Press, New Jersey, pp 1624–1628
  26. Inglada J, Mercier G (2007) A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis. *IEEE Trans Geosci Remote Sens* 45(5):1432–1445
  27. Karoui I, Fablet R, Boucher JM, Pieczynski W, Augustin JM (2008) Fusion of textural statistics using a similarity measure: Application to texture recognition and segmentation. *Pattern Anal Appl* 11(3–4):425–434
  28. Kersten PR, Lee JS, Ainsworth TL (2005) Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and EM clustering. *IEEE Trans Geosci Remote Sens* 43(3): 519–527
  29. Kuruoglu EE, Zerubia J (2004) Modeling SAR images with a generalization of the Rayleigh distribution. *IEEE Trans Image Process* 13(4):527–533
  30. Lee JS, Hoppel KW, Mango SA, Miller AR (1994) Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery. *IEEE Trans Geosci Remote Sens* 32(5): 1017–1028
  31. Lim DH, Ju Jang S (2002) Comparison of two-sample tests for edge detection in noisy images. *J Royal Stat Soc Ser D (the statistician)* 51(1):21–30
  32. Maillard P, Clausi DA (2005) Comparing classification metrics for labeling segmented remote sensing images. In: Computer and robot vision proceedings of the 2nd Canadian conference on computer and robot vision (CRV 2005), pp 421–428
  33. Manolova A, Guérin-Dugué A (2008) Classification of dissimilarity data with a new flexible Mahalanobis-like metric. *Pattern Anal Appl* 11(3–4):337–351
  34. Marghany M, Hashim M (2010) Texture entropy algorithm for automatic detection of oil spill from RADARSAT-1 SAR data. *Int J Phys Sci* 5(9): 1475–1480
  35. Marsaglia G, Tsang WW, Wang J (2003) Evaluating Kolmogorov's distribution. *J Stat Softw* 8(18):1–4
  36. Mejail ME, Frery AC, Jacobo-Berlles J, Bustos OH (2001) Approximation of distributions for SAR images: Proposal, evaluation and practical consequences. *Lat Am Appl Res* 31:83–92
  37. Mejail ME, Jacobo-Berlles J, Frery AC, Bustos OH (2003) Classification of SAR images using a general and tractable multiplicative model. *Int J Remote Sens* 24(18): 3565–3582
  38. Mercier G, Moser G, Serpico S (2008) Conditional copulas for change detection in heterogeneous remote sensing images. *IEEE Trans Geosci Remote Sens* 46(5):1428–1441
  39. Morio J, Réfrégier P, Goudail F, Dubois-Fernandez PC, Dupuis X (2008) Information theory-based approach for contrast analysis in polarimetric and/or interferometric SAR images. *IEEE Trans Geosci Remote Sens* 46(8):2185–2196
  40. Nascimento ADC, Cintra RJ, Frery AC (2010) Hypothesis testing in speckled data with stochastic distances. *IEEE Trans Geosci Remote Sens* 48(1): 373–385
  41. Oliver C, Quegan S (1998) Understanding synthetic aperture radar images. Artech House
  42. Salicrú M, Menéndez ML, Pardo L, Morales D (1994) On the applications of divergence type measures in testing statistical hypothesis. *J Multivar Anal* 51:372–391
  43. Silva M, Cribari-Neto F, Frery AC (2008) Improved likelihood inference for the roughness parameter of the GA0 distribution. *Environmetrics* 19(4):347–368
  44. Smirnov NV (1933) Estimate of deviation between empirical distribution functions in two independent. *Mosc Univ Math Bull* 2(2):3–16
  45. Tison C, Nicolas JM, Tupin F, Maitre H (2004) A new statistical model for Markovian classification of urban areas in high-resolution SAR images. *IEEE Trans Geosci Remote Sens* 42(10): 2046–2057
  46. Vasconcellos KLP, Frery AC, Silva LB (2005) Improving estimation in speckled imagery. *Comput Stat* 20(3):503–519
  47. Zhang Q (2005) Research on detection methods of vehicle targets from sar images based on statistical model. Master's thesis, National University of Defence Technology, Hunan, China
  48. Zhang YD, Wu LN, Wei G (2009) A new classifier for polarimetric SAR. *Prog Electromagn Res* 94:83–104
  49. Ziou D, Bouguila N, Allili MS, El-Zaart A (2009) Finite gamma mixture modeling using minimum message length inference: application to SAR image analysis. *Int J Remote Sens* 30: 771–792



## 5.2 Verossimilhança perfilada modificada e correção por viés na distribuição Wishart complexa escalonada

Inerente ao seu processo de formação, imagens PolSAR são fortemente afetadas pelo ruído speckle multidimensional, como discutido por López-Martínez et al. (2005). No tocante ao tratamento deste efeito, um importante parâmetro no contexto de imagens PolSAR é o número de equivalente de looks ( $L$ ). Na Seção 3.1.3, por exemplo, a Figura 3.4 ilustra que quanto maior é o valor de  $L$  menor é o efeito do ruído sobre imagens PolSAR sintéticas. Desta forma, propor estimadores melhorados para  $L$  é uma atividade importante no processamento de imagens PolSAR. Esta seção se dedica a propor estimadores MV melhorados por dois diferentes métodos.

O estudo das propriedades dos estimadores MV em amostras finitas tem sido revisitado por vários pesquisadores de teoria assintótica. É conhecido que tais estimadores apresentam vieses de ordem  $O(N^{-1})$ , em que  $N$  é o tamanho de amostra e  $O(\cdot)$  é a notação de Landau para representar ordem. Na prática, este termo é desconsiderado pois seu valor é desprezível se comparado com o erro padrão cuja ordem é  $O(N^{-1/2})$ . Entretanto, tais vieses podem ser substanciais em situações em que se consideram tamanhos de amostras pequenos ou moderados e/ou a informação de Fisher é reduzida. Assim, expressões analíticas para o viés do estimador de MV têm sido procuradas para obter estimadores mais precisos em amostras pequenas.

Cox & Snell (1968) propuseram uma fórmula geral para o viés de segunda ordem do estimador MV. A partir deste trabalho, várias pesquisas têm aplicado esta fórmula no contexto de várias distribuições como discutido por Cordeiro (2011). Vasconcellos et al. (2005) consideraram esta expressão para propor estimadores MV corrigidos para o parâmetro de rugosidade,  $\alpha$ , da distribuição  $\mathcal{G}_I^0$  (vide Tabela 3.6), que é utilizada para modelar dados SAR monopolarizados.

O problema de propor estimadores MV mais precisos também pode ser tratado utilizando verossimilhança perfilada modificada (para detalhes, consultar Severini 2000). Neste caso, um estimador melhorado pode ser obtido pela maximização da verossimilhança perfilada modificada. Silva et al. (2008) propuseram métodos de estimação pontual e teste de hipótese sobre o parâmetro de rugosidade da distribuição  $\mathcal{G}_I^0$ , baseando-se em funções de verossimilhança perfilada modificada.

Considere, então, a suposição que uma imagem é bem representada pela distribuição Wishart complexa escalonada (vide, Seção 3.1). Esta distribuição é definida por dois parâmetros: a matriz de covariância e o parâmetro  $L$ . A Seção 3.1 apresenta uma discussão que conclui que esta distribuição está relacionada com a física do processamento *multilook*. Um tratamento matemático deste fato foi apresentado por Goodman (1963b). Partindo desta suposição, Anfinson, Doulgeris & Eltoft (2009) apresentaram uma metodologia da estimação de  $L$  por MV bem como por outros métodos de estima-

ção baseados em momentos dependentes de operações do traço envolvendo a matriz de covariância. Para aquele caso, considerou-se o modelo Wishart complexo não escalonado. Diferentemente, esta seção se baseia no escalonamento deste modelo.

A fim de corrigir a abordagem de estimação MV de  $L$  em amostras pequenas ou moderadas, esta seção

- ✎ deriva a expressão para o viés de segunda ordem do estimador MV pela fórmula proposta por Cox & Snell (1968),
- ✎ propõe uma expressão para a verossimilhança perfilada modificada de Barndorff-Nielsen (1994),
- ✎ como consequência direta dos itens anteriores, deriva dois novos estimadores com vieses de ordens  $O(N^{-2})$  e  $O(N^{-3/2})$ , baseando-se no primeiro e no segundo itens, respectivamente,
- ✎ apresenta um estudo de simulação de Monte Carlo para mensurar a performance dos estimadores, tendo como critérios de avaliação (i) o erro quadrático médio, (ii) o coeficiente de variação e (iii) estimativas para o vieses,
- ✎ aplica a metodologia desenvolvida a dados PolSAR da imagem AIRSAR de São Francisco.

Uma discussão para satisfazer estes objetivos é anexada à esta seção na forma do trabalho

### **“Modified Profile Likelihood and Bias Correction in Complex Wishart Distribution”.**

Similarmente ao anexo anterior, este trabalho pode ser lido de forma independente da tese. A seguir, as principais contribuições deste trabalho são apresentadas. Vale-se, anteriormente, a seguinte nota: este trabalho não foi publicado e está na sua versão mais bruta, tratando-se apenas de uma proposta inicial para discussão.

#### **5.2.1 Contribuições**

##### **Linha teórica:**

##### **1) Expressão para o viés de segunda ordem do estimador MV do parâmetro $L$ :**

Seja  $\{Z_1, Z_2, \dots, Z_N\}$  uma amostra iid de  $Z \sim \mathcal{W}(\theta)$  tal que  $\theta = [L, \text{vec}(\Sigma)^t]^t$ .



Primeiramente, nós propomos a seguinte expressão como o viés de segunda ordem para o estimador MV de  $L$ :

$$B(\hat{L}) = -\frac{1}{2N} \frac{\frac{m^2}{2L} + \sum_{i=0}^{m-1} \psi^{(2)}(L-i)}{\left[\sum_{i=0}^{m-1} \psi^{(0)}(L-i) - \frac{m}{L}\right]^2} + \frac{m^2}{2NL \left[\sum_{i=0}^{m-1} \psi^{(0)}(L-i) - \frac{m}{L}\right]}.$$

Desta forma, um estimador corrigido pode ser dado por  $\tilde{L} = \hat{L} - \hat{B}(\hat{L})$ , em que a estimativa de  $\hat{L}$  é obtida da equação (3.11). Os detalhes matemáticos desta proposta são discutidos no Apêndice C.

## 2) Estimação melhorada para $L$ baseada na verossimilhança perfilada modificada:

Substituindo a matriz de covariância por seu estimador MV, a função de log-verossimilhança (3.10) pode ser reescrita como

$$\begin{aligned} \ell_p(L) = & mNL(\log L - 1) + (L - m) \sum_{i=1}^N \log |Z_k| - NL \log |\bar{Z}| - N \frac{m(m-1)}{2} \log \pi \\ & - N \sum_{k=0}^{m-1} \log \Gamma(L - k). \end{aligned}$$

Esta função representa a verossimilhança perfilada. Assim, a escore perfilada é dada por

$$U_p(L) = mL \log L - N \sum_{k=0}^{m-1} \psi^{(0)}[\Gamma(L - k)] + N \log |\bar{Z}| - \sum_{i=1}^N \log |Z_k|.$$

Seguindo a definição de Barndorff-Nielsen (1994), sua verossimilhança perfilada modificada e sua função escore relativa são dadas, respectivamente, por

$$\ell_{\text{BN}}(L) = \ell_p(L) - \frac{m^2}{2} (\log N + \log L) - \frac{1}{2} \log |\bar{Z}^{-1} \otimes \bar{Z}^{-1}|$$

e

$$U_{\text{BN}}(L) = U_p(L) - \frac{m^2}{2L}. \quad (5.2)$$

Estas derivações se tornaram tratáveis pois  $L$  e  $\text{vec}(\Sigma)$  são ortogonais e, portanto, a expressão de Barndorff-Nielsen (1994), que depende de derivadas em termos de estimadores, reduz-se a expressão proposta por Cox & Reid (1992). Uma discussão mais robusta, com as definições necessárias, é apresentada no arquivo em anexo. O estimador baseado na log-verossimilhança  $\ell_{\text{BN}}(L)$  é mais preciso, de ordem  $O(N^{-3/2})$ , do que aquele baseado na log-verossimilhança original.

**Linha aplicada:**

Os métodos de estimação para o número de looks foram comparados a outros três estimadores propostos por Anfinssen, Doulgeris & Eltoft (2009) por meio de experimentos de Monte Carlo:

1. Estimadores de momentos do traço:  $\hat{L}_{MM1}$  e  $\hat{L}_{MM2}$ , em que

$$\bar{L}_{MM1} = \frac{\text{tr}(\bar{\mathbf{Z}} \bar{\mathbf{Z}})}{\text{tr}(\mathbf{Z})^2 - \text{tr}(\bar{\mathbf{Z}})^2}, \quad \bar{L}_{MM2} = \frac{\text{tr}(\bar{\mathbf{Z}})^2}{\text{tr}(\mathbf{Z}\mathbf{Z}) - \text{tr}(\bar{\mathbf{Z}} \bar{\mathbf{Z}})},$$

$$\overline{\text{tr}(\mathbf{Z})^2} = N^{-1} \sum_{i=1}^N \text{tr}(\mathbf{Z}_i)^2 \text{ e } \overline{\text{tr}(\mathbf{Z}\mathbf{Z})} = N^{-1} \sum_{i=1}^N \text{tr}(\mathbf{Z}_i \mathbf{Z}_i).$$

2. Estimador MV padrão para Wishart escalonada ( $\hat{L}_{ML}$ ) proposto por Anfinssen, Doulgeris & Eltoft (2009),
3. Estimador MV corrigido via verossimilhança perfilada modificada de Barndorff-Nielsen (1994) ( $\hat{L}_{BN}$ ) e
4. Estimador MV corrigido via expressão de Cox & Snell (1968) para o viés de segunda ordem ( $\tilde{L}_{ML}$ ).

Tomando como critérios de comparação três medidas estatísticas [o erro quadrático médio ( $\text{MSE}_{\hat{L}}$ ),

$$\text{MSE}_{\hat{L}} = \sum_{i=1}^{5500} \frac{(\hat{L}_i - L)^2}{5500},$$

o coeficiente de variação ( $\text{CV}_{\hat{L}}$ ) e o viés  $B(\hat{L})$ ], a performance dos estimadores foram quantificadas. A Tabela 5.7 apresenta os resultados para dados Wishart distribuídos conforme a seguinte estrutura de simulação:

**Replicações:** Como proposto por Nascimento et al. (2010), nós utilizamos 5500 cenários de Monte Carlo.

**Parâmetros não estimáveis:** Foram utilizados os tamanhos de amostra  $N \in \{49, 81, 121\}$ .

**Parâmetros estimáveis** Considerou-se os números de looks  $L \in \{4, 6, 8, 12\}$  e a matriz de covariância dada por

$$\Sigma = \begin{bmatrix} 962892 & 19171 - 3579i & -154638 + 191388i \\ & 56707 & -5798 + 16812i \\ & & 472251 \end{bmatrix}$$

(proposta por Frery, Jacobo-Berlles, Gambini & Mejail 2010). Neste caso, utilizou-se apenas uma matriz de covariância, por se tratar de um estudo inicial.



**Figura 5.9** Imagem AIRSAR de São Francisco com regiões selecionadas.

A Tabela 5.7 apresenta as estimativas dos estimadores considerados e suas medidas de comparação estimadas. Os resultados mostraram evidências que os estimadores corrigidos superaram os propostos na literatura: as estimativas  $\hat{L}_{BN}$  e  $\tilde{L}_{ML}$  foram menos viesadas e apresentaram tanto  $MSE_{\hat{L}}$  como  $CV_{\hat{L}}$  menores do que os propostos na literatura.

Finalmente, uma aplicação a dados PolSAR reais da cidade de São Francisco foi realizada. Esta imagem foi obtida pelo sensor AIRSAR operando com um número nominal de looks quatro (como discutido no Apêndice A). A Figura 5.9 apresenta uma amostra  $150 \times 150$  desta imagem com três áreas selecionadas: uma envolvendo cenários de oceano referenciada por “Ocean”, outra situações de florestas nominada por “Forest” e uma terceira considerando cidades, “Urban”. A Tabela 5.8 mostra as estimativas para dados reais, indicando que os estimadores corrigidos apresentam uma melhor performance em cenários homogêneos, isto é, cenários mais adequados ao uso da distribuição Wishart complexa escalonada.

### 5.2.2 Considerações da Seção 5.2

Nesta seção, nós apresentamos dois métodos de estimação melhorada por MV para o número de looks da distribuição Wishart complexa escalonada. Estas propostas são significantes pela importância deste parâmetro em tecnologia PolSAR.

**Tabela 5.7** Estimativas médias e suas medidas de comparabilidade em dados sintéticos

|                 |           | $L = 4$     |                  |                   | $L = 6$     |                  |                   | $L = 8$     |                  |                   | $L = 12$    |                  |                   |
|-----------------|-----------|-------------|------------------|-------------------|-------------|------------------|-------------------|-------------|------------------|-------------------|-------------|------------------|-------------------|
| $N$             | $\hat{L}$ | $\tilde{L}$ | $CV_{\tilde{L}}$ | $MSE_{\tilde{L}}$ | $\tilde{L}$ | $CV_{\tilde{L}}$ | $MSE_{\tilde{L}}$ | $\tilde{L}$ | $CV_{\tilde{L}}$ | $MSE_{\tilde{L}}$ | $\tilde{L}$ | $CV_{\tilde{L}}$ | $MSE_{\tilde{L}}$ |
| $\hat{L}_{MM1}$ | 9         | 6.278       | 0.676            | 23.174            | 9.344       | 0.689            | 52.605            | 12.478      | 0.698            | 95.978            | 18.119      | 0.683            | 190.432           |
|                 |           | 4.957       | 0.291            | 2.993             | 7.401       | 0.285            | 6.399             | 9.849       | 0.294            | 11.800            | 14.606      | 0.292            | 24.977            |
|                 |           | 4.339       | 0.126            | 0.414             | 6.663       | 0.145            | 1.373             | 8.967       | 0.153            | 2.810             | 13.538      | 0.158            | 6.963             |
|                 |           | 4.090       | 0.118            | 0.243             | 6.150       | 0.140            | 0.760             | 8.197       | 0.148            | 1.518             | 12.259      | 0.155            | 3.700             |
|                 |           | 3.998       | 0.118            | 0.221             | 6.000       | 0.139            | 0.695             | 7.989       | 0.148            | 1.398             | 11.937      | 0.155            | 3.435             |
| 49              |           | 4.333       | 0.223            | 1.045             | 6.452       | 0.219            | 2.201             | 8.601       | 0.216            | 3.809             | 12.847      | 0.215            | 8.363             |
|                 |           | 4.165       | 0.124            | 0.294             | 6.235       | 0.122            | 0.633             | 8.313       | 0.120            | 1.093             | 12.435      | 0.119            | 2.388             |
|                 |           | 4.055       | 0.049            | 0.042             | 6.110       | 0.057            | 0.133             | 8.157       | 0.059            | 0.258             | 12.269      | 0.063            | 0.661             |
|                 |           | 4.014       | 0.048            | 0.037             | 6.026       | 0.057            | 0.117             | 8.031       | 0.059            | 0.225             | 12.059      | 0.062            | 0.568             |
|                 |           | 4.000       | 0.048            | 0.037             | 6.002       | 0.056            | 0.115             | 7.998       | 0.059            | 0.222             | 12.007      | 0.062            | 0.559             |
| 121             |           | 4.131       | 0.140            | 0.350             | 6.182       | 0.136            | 0.738             | 8.212       | 0.134            | 1.261             | 12.310      | 0.134            | 2.820             |
|                 |           | 4.063       | 0.080            | 0.108             | 6.097       | 0.077            | 0.233             | 8.113       | 0.077            | 0.403             | 12.164      | 0.076            | 0.871             |
|                 |           | 4.023       | 0.031            | 0.016             | 6.041       | 0.036            | 0.048             | 8.064       | 0.038            | 0.096             | 12.100      | 0.039            | 0.237             |
|                 |           | 4.006       | 0.031            | 0.015             | 6.008       | 0.035            | 0.045             | 8.014       | 0.038            | 0.091             | 12.016      | 0.039            | 0.223             |
|                 |           | 4.001       | 0.031            | 0.015             | 5.998       | 0.035            | 0.045             | 8.001       | 0.038            | 0.090             | 11.995      | 0.039            | 0.222             |

**Tabela 5.8** Estimativas médias e suas medidas de comparabilidade em dados reais

|         |     | $\hat{L}_{ML}$ |                  |                   | $\tilde{L}_{ML}$ |                  |                   | $\hat{L}_{BN}$ |                  |                   |
|---------|-----|----------------|------------------|-------------------|------------------|------------------|-------------------|----------------|------------------|-------------------|
| Regiões | $N$ | $\tilde{L}$    | $CV_{\tilde{L}}$ | $MSE_{\tilde{L}}$ | $\tilde{L}$      | $CV_{\tilde{L}}$ | $MSE_{\tilde{L}}$ | $\tilde{L}$    | $CV_{\tilde{L}}$ | $MSE_{\tilde{L}}$ |
| Ocean   | 9   | 4.520          | 0.153            | 0.747             | 3.776            | 0.145            | 0.350             | 3.995          | 0.294            | 1.376             |
|         | 36  | 4.074          | 0.193            | 0.623             | 3.932            | 0.193            | 0.582             | 3.979          | 0.216            | 0.739             |
|         | 121 | 4.140          | 0.029            | 0.034             | 4.099            | 0.029            | 0.024             | 4.123          | 0.029            | 0.029             |
|         | 144 | 4.133          | 0.026            | 0.029             | 4.099            | 0.025            | 0.021             | 4.118          | 0.026            | 0.025             |
| Forest  |     | 3.293          | 0.095            | 0.599             | 3.069            | 0.087            | 0.938             | 3.159          | 0.092            | 0.792             |
|         |     | 3.119          | 0.043            | 0.794             | 3.075            | 0.042            | 0.872             | 3.089          | 0.049            | 0.852             |
|         |     | 2.809          | 0.281            | 2.041             | 2.798            | 0.281            | 2.065             | 2.818          | 0.273            | 1.987             |
|         |     | 3.082          | 0.017            | 0.845             | 3.072            | 0.017            | 0.864             | 3.075          | 0.017            | 0.858             |
| Urban   |     | 2.738          | 0.336            | 2.439             | 2.492            | 0.394            | 3.239             | 2.678          | 0.313            | 2.451             |
|         |     | 2.567          | 0.256            | 2.486             | 2.529            | 0.260            | 2.594             | 2.454          | 0.316            | 2.989             |
|         |     | 2.218          | 0.399            | 3.959             | 2.209            | 0.401            | 3.990             | 2.177          | 0.417            | 4.146             |
|         |     | 2.447          | 0.265            | 2.833             | 2.439            | 0.265            | 2.857             | 2.426          | 0.275            | 2.923             |

Primeiramente, apresentamos uma expressão para o viés de segunda ordem do estimador MV para o parâmetro  $L$ . Esta derivação foi realizada baseando-se na expressão proposta por Cox & Snell (1968). Além disso, nós propomos uma fórmula para a verossimilhança perfilada modificada segundo Barndorff-Nielsen (1994) da distribuição Wishart complexa escalonada.

Por meio de experimentos de Monte Carlo, os métodos derivados foram comparados aos métodos propostos por Anfinsen, Doulgeris & Eltoft (2009). Considerando como critérios de comparação (i) o erro quadrático médio, (ii) o coeficiente de variação e (iii) o viés das estimativas obtidas, os resultados mostram evidências que as estimativas corrigidas superam os demais métodos propostos por Anfinsen, Doulgeris & Eltoft. Finalmente uma aplicação a dados PolSAR da imagem AIRSAR de São Francisco sugere o uso das correções em cenários homogêneos, nos quais a distribuição Wishart se ajusta melhor. Este resultado é útil, especialmente, em áreas de processamento de imagem que requerem a estimação do parâmetro  $L$  em pequenas amostras, tais como técnicas de filtragem.

### 5.2.3 Trabalhos futuros

Baseando-se sobre os resultados teóricos desenvolvidos, esta seção deixa os seguintes pontos científicos:

- ✎ Estudar estimação por MV dos parâmetros de rugosidade das distribuições polarimétricas apresentadas na Tabela 4.1 da página 51,
- ✎ Avaliar o comportamento das estimativas na presença de contaminação; em termos práticos, diante de eventos de reflexão de canto em imagens PolSAR.

## Anexo da Seção 5.2

### ARTIGO 2:

“Modified Profile Likelihood and Bias Correction in Complex  
Wishart Distribution”

Em processo de revisão com meus orientadores

---

# Modified Profile Likelihood and Bias Correction in Complex Wishart Distribution

Abraão D. C. Nascimento, *Student Member*, Alejandro C. Frery, *Member*, and Renato J. Cintra, *Senior Member*

**Abstract**—This paper proposes improved methods for the maximum likelihood (ML) estimation of the equivalent number of looks  $L$ . This parameter has a meaningful interpretation in the context of polarimetric synthetic aperture radar (PolSAR) images. Due to the presence of coherent illumination in their processing, PolSAR systems generate images which present a granular noise called speckle. As a potential solution for reducing such interference, the parameter  $L$  controls the signal-noise ratio. Thus, the proposal of efficient estimation methodologies for  $L$  has been sought. To that end, we consider firstly that a PolSAR image is well described by the scaled complex Wishart distribution. In recent years, Anfinson *et al.* [IEEE Transactions on Geoscience and Remote Sensing, vol. 47, no. 11, pp. 3795–3809, 2009.] derived and analyzed estimation methods based on the ML and on trace statistical moments for obtaining the parameter  $L$  of the unscaled version of such probability law. This paper generalizes that approach. We present the second-order bias expression proposed by Cox-Snell [Journal of the Royal Statistical Society. Series B (Methodological), vol. 30, no. 2, pp. 248–275, 1968.] for the ML estimator of this parameter. Moreover, the formula of the profile likelihood modified by Barndorff-Nielsen [Journal of the Royal Statistical Society. Series B (Methodological), vol. 56, no. 1, pp. 125–140, 1994.] in terms of  $L$  is proposed. Such derivations are degenerated in two new ML estimators for the parameter  $L$ , which are compared to the estimators proposed by Anfinson *et al.*. The performance of these estimators is assessed by means of Monte Carlo experiments, adopting three statistical measures as comparison criterion: the mean square error, the bias, and the coefficient of variation. Equivalently to the simulation study, an application to actual PolSAR data concludes that the proposed estimators outperform all the others in homogeneous scenarios.

## I. INTRODUCTION

Multilook polarimetric synthetic aperture radar (PolSAR) is a technology which has proven to be an effective tool for geophysical remote sensing. This particular system operates by sending electromagnetic pulses of several polarizations toward a target; then the returning echoes are captured and registered. This process is characterized by the use of coherent illumination. Thus, the obtained images are affected by granular noise called ‘speckle’. Therefore, a specialized modeling is required for processing and analyzing such images.

In this context, the scaled complex Wishart distribution has been suggested as a successful approach for modeling backscatter PolSAR data. This distribution is equipped with two parameters: the covariance matrix and the number of looks ( $L$ ). The first of them is directly related to the data structure, which has as support the set of positive-definite Hermitian matrices. On the other hand, parameter  $L$  is defined as the mean number of statistically independent samples represented by each pixel. This parameter  $L$  has the crucial role of controlling the signal-noise ratio, diminishing the speckle noise effect on PolSAR images when its value is increased. Thus, many papers have addressed the proposal of efficient estimation procedures for the parameter  $L$  [1], [2].

The study of maximum likelihood (ML) estimator properties in finite sample has been revisited by several researches in asymptotic theory. Such estimators are generally biased with respect to the true parameter value. Indeed, the ML estimator bias has order  $O(N^{-1})$ , where  $N$  is the sample size and  $O(\cdot)$  is Landau notation to represent order. In practice, this quantity is not commonly considered because its value is negligible when compared to the standard error, which has order  $O(N^{-1/2})$ . However, such biases can be substantial in situations where small or moderate sample sizes are used. Thus, analytic expressions for ML estimator bias have been sought in order to obtain a more accurate estimator, for finite sample sizes.

Considering a set of independent random variables, Cox-Snell [3] proposed a general formula for the second-

This work was supported by CNPq, Fapeal and FACEPE, Brazil.

A. C. Frery is with the Instituto de Computação, Universidade Federal de Alagoas, BR 104 Norte km 97, 57072-970, Maceió, AL, Brazil, email: acfrery@gmail.com

A. D. C. Nascimento is with the Departamento de Estatística, Universidade Federal de Pernambuco, Cidade Universitária, 50740-540, Recife, PE, Brazil, e-mail: abraao.susej@gmail.com

R. J. Cintra is with the Departamento de Estatística, Universidade Federal de Pernambuco, Cidade Universitária, 50740-540, Recife, PE, Brazil, e-mail: rjdcsc@de.ufpe.br

order bias of the ML estimator. Many papers have considered this formula in the context of several distributions [4]. Vasconcellos *et al.* [5] applied it to the univariate  $\mathcal{G}^0$  distribution, which is often used for modeling synthetic aperture radar (SAR) images with one polarization channel. In recent years, Pianto and Cribari-Neto [6] presented a study concerning the bias correction for the roughness parameter of the  $\mathcal{G}^0$  distribution. This could be achieved by combining the methodology of [7] with bootstrapping [8].

The problem of proposing more accurate ML estimators can also be treated in the context of the modified profile likelihood [9, chapter 9]. In this case, an improved parameter estimation can be achieved by maximizing an adjusted profiled likelihood function [10]. Barndorff-Nielsen proposed a successful adjustment for profile likelihood [11]. The parameters of several distributions were given statistical inference techniques benefited by this methodology. The Birnbaum-Saunders distribution [12] and elliptical structural models [13] are illustrative examples. This approach has also been introduced to models which have their parameters linked to SAR data. Silva *et al.* [14] proposed point estimation and hypothesis test methods for the roughness parameter of the  $\mathcal{G}^0$  distribution based on modified profile likelihood.

Anfinssen *et al.* [1] proposed three methods for estimating the number of looks. Their method was favorably compared with classic methods based on fractional moment and the coefficient of variation. Assuming that PolSAR images are well modeled by the complex Wishart distribution, the ML estimator was found to excel when compared to other classic techniques. In this article, such approach is generalized. The second-order bias expression for the ML estimator of the parameter  $L$  is derived. Such a result implies an improved estimator for it.

Additionally, the expression for the profile likelihood modified by Barndorff-Nielsen [11] is derived. In this case, working on the scaled complex Wishart law, the number of looks is considered as the interest parameter and the covariance matrix elements as a vector of nuisance parameters. Thus, two corrected ML estimators for the parameter  $L$  are proposed and their performance are compared to the others indexed by estimators derived by Anfinssen *et al.*. Such study is assessed by means of Monte Carlo experiments to quantify the results in several possible scenarios. To that end, the following measures are considered as comparison criterion: the mean square error, the bias, and the coefficient of variation. In order to confirm the simulation study results, the discussed methodology is also applied to the actual PolSAR data.

The remainder of this paper is structured as follows. The employed probability modeling is introduced in Section II. Section III discusses statistical methodologies for obtaining corrected ML estimators. Additionally, such methods are applied to the scaled complex Wishart distribution. Subsequently, the derived estimators are analyzed and compared to others proposals in Section IV. Finally, the main conclusions are organized in Section V.

## II. COMPLEX WISHART DISTRIBUTION

The polarimetric coherent information is characterized by a vector with dimension four for each pixel of an image:  $S_{VV}$ ,  $S_{VH}$ ,  $S_{HV}$ , and  $S_{HH}$ , where  $S_{ij}$  represents the backscattered signal at the linear polarization  $i$ th reception and  $j$ th transmission (for  $i, j = H, V$ ). Under the reciprocity theorem conditions [15], the scattering matrix is often simplified into a three-component vector, i.e.,  $S_{HV} = S_{VH}$ . Thus, we have a scattering vector

$$[S_{VV} \ S_{VH} \ S_{HH}]^t,$$

where the superscript  $(\cdot)^t$  indicates the vector transposition. This three-component representation is explained and exemplified in [16].

Firstly, this paper considers systems with  $m$  polarizations, i.e., complex random vector termed by

$$\mathbf{s} = [S_1 \ S_2 \ \cdots \ S_m]^t.$$

Such a vector has been successfully modeled by a multivariate complex Gaussian distribution with zero-mean, as discussed in [17]. The single-look PolSAR data representation is essentially based on these conditions.

Multilook polarimetric speckled data obtained from scattered signal is commonly represented by means of a Hermitian positive definite matrix, defined by

$$\mathbf{Z} = \frac{1}{L} \sum_{i=1}^L \mathbf{s}_i \mathbf{s}_i^H,$$

where superscript  $(\cdot)^H$  represents the complex conjugate transpose of a vector,  $\mathbf{s}_i$ ,  $i = 1, 2, \dots, L$ , are scattering vectors, and  $L$  is the number of looks. The estimated covariance matrix  $\mathbf{Z}$  is termed sample covariance matrix [1] and follows a scaled complex Wishart distribution. Having  $\Sigma$  and  $L$  as parameters, the scaled complex Wishart distribution is associated to the following probability density function:

$$f_{\mathbf{Z}}(\mathbf{Z}'; \Sigma, L) = \frac{L^m |\mathbf{Z}'|^{L-m}}{|\Sigma|^L \Gamma_m(L)} \exp[-L \operatorname{tr}(\Sigma^{-1} \mathbf{Z}')], \quad (1)$$

where  $\Gamma_m(L) = \pi^{m(m-1)/2} \prod_{i=0}^{m-1} \Gamma(L-i)$  for  $L \geq m$ ,  $\Gamma(\cdot)$  is the gamma function,  $\operatorname{tr}(\cdot)$  represents the trace



operator,  $|\cdot|$  denotes the determinant operator,  $\Sigma$  is the covariance matrix of  $\mathbf{Z}$  given by

$$\Sigma = E(\mathbf{y}\mathbf{y}^*) = \begin{bmatrix} E(S_1 S_1^*) & E(S_1 S_2^*) & \cdots & E(S_1 S_m^*) \\ E(S_2 S_1^*) & E(S_2 S_2^*) & \cdots & E(S_2 S_m^*) \\ \vdots & \vdots & \ddots & \vdots \\ E(S_m S_1^*) & E(S_m S_2^*) & \cdots & E(S_m S_m^*) \end{bmatrix},$$

and  $(\cdot)^*$  means the conjugate of a complex number. This distribution is denoted by  $\mathbf{Z} \sim \mathcal{W}(\Sigma, L)$  and it satisfies  $E(\mathbf{Z}) = \Sigma$  [1], where  $E(\cdot)$  indicates the statistical expectation operator.

### III. CORRECTED ML ESTIMATION METHODS

Many articles have addressed the study of ML estimator properties in finite sample, so much in SAR data modeling [5], [6], [14] as well as for other distributions [4], [12], [13]. This section describes briefly two methodologies for the obtaining of improved ML estimators: (i) one based on the ML estimator second-order bias according to the Cox-Snell general formula [3] and (ii) the ML estimator based on modified profile likelihood proposed by Barndorff-Nielsen [11]. To that end, both the methods assume the following stochastic context.

Let  $\{\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_N\}$  be a random sample from a random positive-definite Hermitian matrix  $\mathbf{Z}$  equipped with density  $f_{\mathbf{Z}}(\mathbf{Z}'; \boldsymbol{\theta})$ , where  $\dim(\boldsymbol{\theta}) = p$  and  $\dim(\cdot)$  represents the vector dimension.

#### A. Corrected ML estimator by its second-order bias

This section approaches the methodology for obtaining the corrected ML estimator based on the second-order bias expression. In this method, it is required the derivation from expected values of second and third order derivatives of the total log-likelihood function based on  $N$  observations,

$$\ell(\boldsymbol{\theta}) = \sum_{k=1}^N \log f_{\mathbf{Z}}(\mathbf{Z}_k; \boldsymbol{\theta}).$$

To the following, it is introduced a common notation to researches in the asymptotic theory [4].

Let  $U_{\theta_i}$  (for  $i = 1, 2, \dots, p$ ) be the derivatives of  $\ell(\boldsymbol{\theta})$  at the  $i$ th element of vector  $\boldsymbol{\theta}$ , termed by  $\theta_i$ . We also define the higher order derivatives by

$$U_{\theta_i \theta_j} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \quad \text{and} \quad U_{\theta_i \theta_j \theta_k} = \frac{\partial^3 \ell(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j \partial \theta_k},$$

for  $i, j, k = 1, 2, \dots, p$ . Based on these terms, the following expressions for cumulants of  $\ell(\boldsymbol{\theta})$  are defined:

$\kappa_{\theta_i \theta_j} = E(U_{\theta_i \theta_j})$ ,  $\kappa_{\theta_i, \theta_j} = E(U_{\theta_i} U_{\theta_j})$ ,  $\kappa_{\theta_i \theta_j \theta_k} = E(U_{\theta_i \theta_j \theta_k})$ ,  $\kappa_{\theta_i, \theta_j \theta_k} = E(U_{\theta_i} U_{\theta_j \theta_k})$ , and so on. Moreover, the derivatives of cumulants over the parameters are represented by

$$\kappa_{\theta_i \theta_j}^{(\theta_k)} = \frac{\partial \kappa_{\theta_i \theta_j}}{\partial \theta_k} \quad \text{and} \quad \kappa_{\theta_i \theta_j \theta_k}^{(\theta_h)} = \frac{\partial \kappa_{\theta_i \theta_j \theta_k}}{\partial \theta_h},$$

for  $i, j, k, h = 1, 2, \dots, p$ . Notice that the element  $\kappa_{\theta_i, \theta_j} = -\kappa_{\theta_i \theta_j}$  (such equation is known as one of Bartlett identities [18]) denotes an entry of the Fisher information matrix  $\mathcal{K}(\boldsymbol{\theta})$  for the vector of parameters  $\boldsymbol{\theta}$ . Thus,  $\kappa_{\theta_i, \theta_j} = -\kappa_{\theta_i \theta_j}$  represents an element of  $\mathcal{K}(\boldsymbol{\theta})^{-1}$ .

Considering the above definitions, Cox-Snell [3] proposed a general formula for the ML estimator second-order bias given by

$$B(\hat{\theta}_i) = \sum_{r,s,t} \kappa_{\theta_i \theta_r} \kappa_{\theta_s \theta_t} \left( \kappa_{\theta_r \theta_s}^{(\theta_t)} - \frac{1}{2} \kappa_{\theta_r \theta_s \theta_t} \right), \quad (2)$$

where  $\hat{\theta}_i$  is the ML estimator relative to  $i$ th term of  $\boldsymbol{\theta}$ , for  $i = 1, 2, \dots, p$ . Therefore, the corrected estimator is defined by

$$\tilde{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}} - \hat{B}(\hat{\boldsymbol{\theta}}), \quad (3)$$

where  $\hat{B}(\hat{\boldsymbol{\theta}})$  indicates the bias  $B(\hat{\boldsymbol{\theta}})$  evaluated at  $\hat{\boldsymbol{\theta}}$ . It is known that the bias of  $\tilde{\boldsymbol{\theta}}$  has order of  $N^{-2}$  (i.e.,  $E(\tilde{\boldsymbol{\theta}}) = \boldsymbol{\theta} + O(N^{-2})$ ) while that  $\hat{\boldsymbol{\theta}}$  presents a bias with order of  $N^{-1}$ . Thus, it is expected that  $\tilde{\boldsymbol{\theta}}$  provides better asymptotic properties than  $\hat{\boldsymbol{\theta}}$ .

#### B. ML estimator obtained by modified profile likelihood

As other way for obtaining improved ML estimator, some authors have chosen to use the modified profile likelihood. The following discussion presents an overview of this methodology.

Firstly, let  $\boldsymbol{\theta} = [\boldsymbol{\psi}^t, \boldsymbol{\lambda}^t]^t$  be a partition of vector  $\boldsymbol{\theta}$  such that  $\dim(\boldsymbol{\psi}) = q$  and  $\dim(\boldsymbol{\lambda}) = p - q$ . The proposal of statistical inference under a interest parameter vector,  $\boldsymbol{\psi}$ , in the presence of nuisance parameter vector,  $\boldsymbol{\lambda}$ , is a problem which has been addressed by several papers [19], [20]. In this case, the ideal situation would be working with the marginal or conditional likelihood, however the derivation of these functions is often a intractable task. Thus, it is common to approximate such functions by the profile log-likelihood which is defined by  $\tilde{\ell}(\boldsymbol{\psi}) = \ell(\boldsymbol{\psi}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{\psi}})$ , where  $\hat{\boldsymbol{\lambda}}_{\boldsymbol{\psi}}$  is the maximum likelihood estimate of  $\boldsymbol{\lambda}$  given the parameter  $\boldsymbol{\psi}$  fixed.

However, there is a problem in the utilization of  $\tilde{\ell}(\cdot)$ . The profile likelihood not has statistical properties similar to a genuine likelihood and, thus, the profile score and information biases are only guaranteed to be  $O(1)$ , i.e.,

$$E[\tilde{U}(\boldsymbol{\psi})] = O(1)$$

and

$$\mathbb{E}[\tilde{U}(\psi)\tilde{U}(\psi)^t] + \mathbb{E}\left[\frac{\partial\tilde{U}(\psi)}{\partial\psi^t}\right] = O(1),$$

where  $\tilde{U}(\psi) = \partial\tilde{\ell}(\psi)/\partial\psi$ . This fact implies that ML estimators based on profile log-likelihood can be more affected for small value of  $N$ . In order to solve this problem, several author have proposed modification for profile log-likelihood, such as Barndorff-Nielsen [11], Cox-Reid [19], McCullagh and Tibshirani [21], and Stern [20]. This paper considered only the profile likelihood modified by Barndorff-Nielsen, which is described in the following discussion.

Using the ' $p^*$ ' formula' to approximate the probability density function of ML estimator [22], Barndorff-Nielsen proposed a approximation for marginal log-likelihood with bias of order  $O(N^{-1})$  given by

$$\ell_{\text{BN}}(\psi) = \ell_p(\psi) - \log\left|\frac{\partial\hat{\lambda}_\psi}{\partial\hat{\lambda}}\right| - \frac{1}{2}\log|\mathcal{J}_{\lambda\lambda}(\psi, \hat{\lambda}_\psi)|,$$

where  $\mathcal{J}_{\lambda\lambda}(\psi, \hat{\lambda}) = -\partial^2\ell(\theta)/\partial\lambda\partial\lambda^t$ . For situations in which the parameters are orthogonal (i.e.,  $\partial\hat{\lambda}_\psi/\partial\hat{\lambda} = 1$ ), this formula is degenerated in

$$\ell_{\text{BN}}(\psi) = \ell_p(\psi) - \frac{1}{2}\log|\mathcal{J}_{\lambda\lambda}(\psi, \hat{\lambda}_\psi)|, \quad (4)$$

which is known as Cox-Reid modified profile log-likelihood [19]. The ML estimator under this log-likelihood for  $\psi$  (termed by  $\hat{\psi}_{\text{BN}}$ ) presents a bias of order  $O(N^{-3/2})$ , i.e.,  $\mathbb{E}(\hat{\psi}_{\text{BN}}) = \psi + O(N^{-3/2})$ , being more accurate than one based on  $\ell(\cdot)$ .

#### IV. RESULTS

Assuming the scaled complex Wishart distribution as an *a priori* model, we propose two methodologies for improving the ML estimators discussed by Anfinssen *et al.* [1]. Firstly, the correction based on Cox-Snell second-order bias was considered for obtaining more accurate estimates for the number of looks in such distribution. Subsequently, we employed the Barndorff-Nielsen adjustment to that end. Monte Carlo experiments were additionally assumed to quantify the performance of proposed estimators and these results are compared to ones of three others estimators derived by Anfinssen *et al.* [1]. In this case, the mean square error, the bias, and the coefficient of variation were utilized as comparison criterion. Finally, the considered methodology is applied to actual data.

##### A. Second-order bias for the ML estimator of number of looks

Let  $\{\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_N\}$  be a random sample from  $\mathbf{Z} \sim \mathcal{W}(\theta)$  such that  $\theta = (\sigma^t, L)^t$ , where  $\sigma = \text{vec}(\Sigma)$  and  $\text{vec}(\cdot)$  is the column stacking vectorisation operator. In this case, the total log-likelihood and the relative score function are given by

$$\begin{aligned} \tilde{\ell}(\theta) = & mNL \log L + (L - m) \sum_{k=1}^N \log |\mathbf{Z}_k| - LN \log |\Sigma| \\ & - N \frac{m(m-1)}{2} \log \pi - N \sum_{k=0}^{m-1} \log \Gamma(L - k) \\ & - NL \text{tr}(\Sigma^{-1} \bar{\mathbf{Z}}) \end{aligned} \quad (5)$$

and

$$\begin{aligned} \tilde{U}_\theta = & (U_\sigma^t \quad U_L)^t \\ = & \begin{cases} NL \text{vec}(\Sigma^{-1} \bar{\mathbf{Z}} \Sigma^{-1} - \Sigma^{-1}), \\ mN(\log L + 1) + \sum_{k=1}^N \log |\mathbf{Z}_k| - N \log |\Sigma| \\ -N \sum_{i=0}^{m-1} \psi^{(0)}(L - i) - N \text{tr}(\Sigma^{-1} \bar{\mathbf{Z}}), \end{cases} \end{aligned} \quad (6)$$

respectively.

From  $\tilde{U}_\theta = \mathbf{0}$ , the following results hold:

$$\begin{aligned} \hat{\Sigma} = & \frac{1}{N} \sum_{k=1}^N \mathbf{Z}_k = \bar{\mathbf{Z}}, \\ m \log \hat{L} + N^{-1} \sum_{k=1}^N \log |\mathbf{Z}_k| - \log |\bar{\mathbf{Z}}| \\ & - \sum_{i=0}^{m-1} \psi^{(0)}(\hat{L} - i) = 0, \end{aligned} \quad (7)$$

Thus, the ML estimators for  $\Sigma$  and  $L$  are given by the average of the elements of considered random sample and by the resolution of the non-linear equation given in (7), respectively. Since the estimator for the number of looks does not present closed formula, its ML estimates are obtained by the Newton-Raphson numerical optimization method [23].

From (6), the second and third-order derivatives are given by

$$\begin{aligned} U_{LL} = & \frac{Nm}{L} - N \sum_{i=1}^{m-1} \psi^{(1)}(L - i), \\ U_{L\sigma_i} = & N[\text{vec}(\Sigma^{-1} \bar{\mathbf{Z}} \Sigma^{-1} - \Sigma^{-1})]_i, \\ U_{\sigma_i \sigma_j} = & LN[(\Sigma^{-1} \otimes \Sigma^{-1}) - (\Sigma^{-1} \otimes \Sigma^{-1}) \bar{\mathbf{Z}} \Sigma^{-1} \\ & - (\Sigma^{-1} \otimes \Sigma^{-1}) \Sigma^{-1} \bar{\mathbf{Z}}]_{i,j}, \end{aligned}$$

$$U_{L\sigma_i\sigma_j} = L^{-1}U_{\sigma_i\sigma_j},$$

$$\begin{aligned} U_{\sigma_i\sigma_j\sigma_k} = & LN[3(\Sigma^{-1} \otimes \Sigma^{-1}) \otimes (\Sigma^{-1}\bar{\mathbf{Z}}\Sigma^{-1}) \\ & + 3(\Sigma^{-1} \otimes \Sigma^{-1}) \otimes (\Sigma^{-1}\Sigma^{-1}\bar{\mathbf{Z}}) \\ & - 2(\Sigma^{-1} \otimes \Sigma^{-1}) \otimes \Sigma^{-1} \\ & ]_{k+(i-1)m^2, k+(j-1)m^2}. \end{aligned}$$

Based on the above results, we derive the following expression for the cumulants,

$$\begin{aligned} \kappa^{L,L} &= \frac{1}{N} \left[ \sum_{i=0}^{m-1} \psi^{(0)}(L-i) - \frac{m}{L} \right]^{-1}, \\ \kappa^{L,\sigma_i} &= 0, \quad \kappa^{\sigma_i,\sigma_j} = \frac{1}{NL} (\Sigma \otimes \Sigma)_{i,j}, \\ \kappa_{LLL} &= \kappa_{LL}^{(L)} = -N \left[ \frac{m}{L^2} + \sum_{i=0} \psi^{(2)}(L-i) \right], \\ \kappa_{LL\sigma_i} &= \kappa_{LL}^{(\sigma_i)} = \kappa_{L\sigma_i}^{(\sigma_j)} = 0 \\ \kappa_{\sigma_i\sigma_j L} &= -N(\Sigma^{-1} \otimes \Sigma^{-1})_{i,j}, \end{aligned} \quad (8)$$

$$\begin{aligned} \kappa_{\sigma_i\sigma_j}^{(\sigma_k)} &= 2LN[(\Sigma^{-1} \otimes \Sigma^{-1}) \otimes \Sigma^{-1} \\ & ]_{k+(i-1)m^2, k+(j-1)m^2}, \end{aligned} \quad (9)$$

$$\begin{aligned} \kappa_{\sigma_i\sigma_j\sigma_k} &= 4LN[(\Sigma^{-1} \otimes \Sigma^{-1}) \otimes \Sigma^{-1} \\ & ]_{k+(i-1)m^2, k+(j-1)m^2}. \end{aligned} \quad (10)$$

Due to the orthogonality of the parameters  $L$  and  $\sigma_i$  for  $i = 1, 2, \dots, m^2$  (justified by result  $\kappa^{L\sigma_i} = 0$ ), the Eq. (2) is degenerated for the following formula in this case:

$$\begin{aligned} B(\hat{L}) &= \kappa^{L,L} \kappa^{L,L} \left( \kappa_{LL}^{(L)} - \frac{1}{2} \kappa_{LLL} \right) \\ &+ \sum_{i=1}^{m^2} \sum_{j=1}^{m^2} \kappa^{L,L} \kappa^{\sigma_i,\sigma_j} \left( \kappa_{L\sigma_i}^{(\sigma_i)} - \frac{1}{2} \kappa_{L\sigma_i\sigma_j} \right). \end{aligned}$$

In order to achieve a closed expression for this formula, the following result is pivotal:

$$\begin{aligned} \kappa^{\sigma_i,\sigma_j} \left( \kappa_{L\sigma_i}^{(\sigma_i)} - \frac{1}{2} \kappa_{L\sigma_i\sigma_j} \right) \\ = \frac{1}{NL} (\Sigma \otimes \Sigma)_{i,j} (0 + N \{ \Sigma^{-1} \otimes \Sigma^{-1} \}_{i,j}) \\ = \frac{1}{2L} (\Sigma \otimes \Sigma)_{i,j} \{ (\Sigma \otimes \Sigma)^{-1} \}_{i,j}. \end{aligned}$$

By means of minor algebraic manipulations, the following results holds

$$\sum_{i=1}^{m^2} \sum_{j=1}^{m^2} \{ \Sigma \otimes \Sigma \}_{i,j} \{ (\Sigma \otimes \Sigma)^{-1} \}_{i,j} = m^2.$$

Thus, the derivation for second-order bias is given by the expression

$$\begin{aligned} B(\hat{L}) &= -\frac{1}{2N} \frac{\frac{m}{2L} + \sum_{i=0}^{m-1} \psi^{(2)}(L-i)}{\left[ \sum_{i=0}^{m-1} \psi^{(0)}(L-i) - \frac{m}{L} \right]^2} \\ &+ \frac{m^2}{2NL \left[ \sum_{i=0}^{m-1} \psi^{(0)}(L-i) - \frac{m}{L} \right]}. \end{aligned} \quad (11)$$

With respect to the calculate of bias for  $\sigma_i$  for  $i = 1, 2, \dots, m^2$ , the orthogonality and the identities given in (8) reduce the Expression (2) in

$$B(\hat{\sigma}_i) = \sum_{r=1}^{m^2} \sum_{s=1}^{m^2} \sum_{t=1}^{m^2} \kappa^{\sigma_i,\sigma_r} \kappa^{\sigma_s,\sigma_t} \left( \kappa_{\sigma_r\sigma_s}^{(\sigma_t)} - \frac{1}{2} \kappa_{\sigma_r\sigma_s\sigma_t} \right).$$

Using the results given in Expression (9) and (10), one can easily show that  $\kappa_{\sigma_r\sigma_s}^{(\sigma_t)} - \frac{1}{2} \kappa_{\sigma_r\sigma_s\sigma_t} = 0$  and thus conclude that  $B(\hat{\sigma}_i) = 0$  for all  $i$ . This fact is expected since the ML estimator for covariance matrix is non-biased, i.e.,  $E(\hat{\Sigma}) = \Sigma$ .

### B. Modified profile log-likelihood for number of looks

Replacing the covariance matrix by its ML estimator, the log-likelihood function given in (5) can be rewritten as

$$\begin{aligned} \ell_p(L) &= mNL(\log L - 1) + (L - m) \sum_{i=1}^N \log |\mathbf{Z}_k| \\ &- NL \log |\bar{\mathbf{Z}}| - N \frac{m(m-1)}{2} \log \pi \\ &- N \sum_{k=0}^{m-1} \log \Gamma(L - k). \end{aligned}$$

Such function represents the profile likelihood. Thus the profile score is expressed by

$$\begin{aligned} U_p(L) &= mL \log L - N \sum_{k=0}^{m-1} \psi^{(0)}[\Gamma(L - k)] + N \log |\bar{\mathbf{Z}}| \\ &- N \log |\bar{\mathbf{Z}}|. \end{aligned}$$

Following the definition given in (4), the profile log-likelihood modified by Brandorff-Nielsen and its relative score function is expressed by

$$\begin{aligned} \ell_{\text{BN}}(L) &= \ell_p(L) - \frac{m^2}{2} (\log N + \log L) - \frac{1}{2} \log |\bar{\mathbf{Z}}^{-1} \otimes \bar{\mathbf{Z}}^{-1}| \\ &\text{and} \end{aligned}$$

$$U_{\text{BN}}(L) = U_p(L) - \frac{m^2}{2L}, \quad (12)$$

respectively.

TABLE I  
ML ESTIMATES AND THEIR PERFORMANCE STATISTICAL MEASURES FOR SYNTHETIC DATA

| $N$ | $\hat{L}$        | $L = 4$     |                  |                   | $L = 6$     |                  |                   | $L = 8$     |                  |                   | $L = 12$    |                  |                   |
|-----|------------------|-------------|------------------|-------------------|-------------|------------------|-------------------|-------------|------------------|-------------------|-------------|------------------|-------------------|
|     |                  | $\tilde{L}$ | $CV_{\tilde{L}}$ | $MSE_{\tilde{L}}$ | $\tilde{L}$ | $CV_{\tilde{L}}$ | $MSE_{\tilde{L}}$ | $\tilde{L}$ | $CV_{\tilde{L}}$ | $MSE_{\tilde{L}}$ | $\tilde{L}$ | $CV_{\tilde{L}}$ | $MSE_{\tilde{L}}$ |
| 9   | $\hat{L}_{MM1}$  | 6.278       | 0.676            | 23.174            | 9.344       | 0.689            | 52.605            | 12.478      | 0.698            | 95.978            | 18.119      | 0.683            | 190.432           |
|     | $\hat{L}_{MM2}$  | 4.957       | 0.291            | 2.993             | 7.401       | 0.285            | 6.399             | 9.849       | 0.294            | 11.800            | 14.606      | 0.292            | 24.977            |
|     | $\hat{L}_{ML}$   | 4.339       | 0.126            | 0.414             | 6.663       | 0.145            | 1.373             | 8.967       | 0.153            | 2.810             | 13.538      | 0.158            | 6.963             |
|     | $\hat{L}_{BN}$   | 4.090       | 0.118            | 0.243             | 6.150       | 0.140            | 0.760             | 8.197       | 0.148            | 1.518             | 12.259      | 0.155            | 3.700             |
|     | $\tilde{L}_{ML}$ | 3.998       | 0.118            | 0.221             | 6.000       | 0.139            | 0.695             | 7.989       | 0.148            | 1.398             | 11.937      | 0.155            | 3.435             |
| 49  |                  | 4.333       | 0.223            | 1.045             | 6.452       | 0.219            | 2.201             | 8.601       | 0.216            | 3.809             | 12.847      | 0.215            | 8.363             |
|     |                  | 4.165       | 0.124            | 0.294             | 6.235       | 0.122            | 0.633             | 8.313       | 0.120            | 1.093             | 12.435      | 0.119            | 2.388             |
|     |                  | 4.055       | 0.049            | 0.042             | 6.110       | 0.057            | 0.133             | 8.157       | 0.059            | 0.258             | 12.269      | 0.063            | 0.661             |
|     |                  | 4.014       | 0.048            | 0.037             | 6.026       | 0.057            | 0.117             | 8.031       | 0.059            | 0.225             | 12.059      | 0.062            | 0.568             |
|     |                  | 4.000       | 0.048            | 0.037             | 6.002       | 0.056            | 0.115             | 7.998       | 0.059            | 0.222             | 12.007      | 0.062            | 0.559             |
| 121 |                  | 4.131       | 0.140            | 0.350             | 6.182       | 0.136            | 0.738             | 8.212       | 0.134            | 1.261             | 12.310      | 0.134            | 2.820             |
|     |                  | 4.063       | 0.080            | 0.108             | 6.097       | 0.077            | 0.233             | 8.113       | 0.077            | 0.403             | 12.164      | 0.076            | 0.871             |
|     |                  | 4.023       | 0.031            | 0.016             | 6.041       | 0.036            | 0.048             | 8.064       | 0.038            | 0.096             | 12.100      | 0.039            | 0.237             |
|     |                  | 4.006       | 0.031            | 0.015             | 6.008       | 0.035            | 0.045             | 8.014       | 0.038            | 0.091             | 12.016      | 0.039            | 0.223             |
|     |                  | 4.001       | 0.031            | 0.015             | 5.998       | 0.035            | 0.045             | 8.001       | 0.038            | 0.090             | 11.995      | 0.039            | 0.222             |

TABLE II  
ML ESTIMATES AND THEIR PERFORMANCE STATISTICAL MEASURES FOR ACTUAL DATA

| Regions | $N$ | $\hat{L}_{ML}$ |                  |                   | $\tilde{L}_{ML}$ |                  |                   | $\hat{L}_{BN}$ |                  |                   |
|---------|-----|----------------|------------------|-------------------|------------------|------------------|-------------------|----------------|------------------|-------------------|
|         |     | $\tilde{L}$    | $CV_{\tilde{L}}$ | $MSE_{\tilde{L}}$ | $\tilde{L}$      | $CV_{\tilde{L}}$ | $MSE_{\tilde{L}}$ | $\tilde{L}$    | $CV_{\tilde{L}}$ | $MSE_{\tilde{L}}$ |
| Ocean   | 9   | 4.520          | 0.153            | 0.747             | 3.776            | 0.145            | 0.350             | 3.995          | 0.294            | 1.376             |
|         | 36  | 4.074          | 0.193            | 0.623             | 3.932            | 0.193            | 0.582             | 3.979          | 0.216            | 0.739             |
|         | 121 | 4.140          | 0.029            | 0.034             | 4.099            | 0.029            | 0.024             | 4.123          | 0.029            | 0.029             |
|         | 144 | 4.133          | 0.026            | 0.029             | 4.099            | 0.025            | 0.021             | 4.118          | 0.026            | 0.025             |
| Forest  |     | 3.293          | 0.095            | 0.599             | 3.069            | 0.087            | 0.938             | 3.159          | 0.092            | 0.792             |
|         |     | 3.119          | 0.043            | 0.794             | 3.075            | 0.042            | 0.872             | 3.089          | 0.049            | 0.852             |
|         |     | 2.809          | 0.281            | 2.041             | 2.798            | 0.281            | 2.065             | 2.818          | 0.273            | 1.987             |
|         |     | 3.082          | 0.017            | 0.845             | 3.072            | 0.017            | 0.864             | 3.075          | 0.017            | 0.858             |
| Urban   |     | 2.738          | 0.336            | 2.439             | 2.492            | 0.394            | 3.239             | 2.678          | 0.313            | 2.451             |
|         |     | 2.567          | 0.256            | 2.486             | 2.529            | 0.260            | 2.594             | 2.454          | 0.316            | 2.989             |
|         |     | 2.218          | 0.399            | 3.959             | 2.209            | 0.401            | 3.990             | 2.177          | 0.417            | 4.146             |
|         |     | 2.447          | 0.265            | 2.833             | 2.439            | 0.265            | 2.857             | 2.426          | 0.275            | 2.923             |

### C. Application

This section aims to compare the proposed estimators to others three derived and explored by Anfinssen *et al.* [1]. That paper proposed two estimators based on trace moment whose expressions are given by

$$\bar{L}_{MM1} = \frac{\text{tr}(\bar{\mathbf{Z}} \bar{\mathbf{Z}})}{\text{tr}(\mathbf{Z})^2 - \text{tr}(\bar{\mathbf{Z}} \bar{\mathbf{Z}})}$$

and

$$\bar{L}_{MM2} = \frac{\text{tr}(\bar{\mathbf{Z}})^2}{\text{tr}(\mathbf{Z}\mathbf{Z}) - \text{tr}(\bar{\mathbf{Z}} \bar{\mathbf{Z}})},$$

where  $\text{tr}(\bar{\mathbf{Z}})^2 = N^{-1} \sum_{i=1}^N \text{tr}(\mathbf{Z}_i)^2$  and  $\text{tr}(\mathbf{Z}\mathbf{Z}) = N^{-1} \sum_{i=1}^N \text{tr}(\mathbf{Z}_i \mathbf{Z}_i)$ . Moreover, a rich approach was

performed concerning the ML estimator obtained numerically by the solution of the Eq. (7). In the following discussion, these estimators are compared to two new proposals in both synthetic and actual study. In summary, the considered estimators are termed by

- 1) first and second trace moment estimators ( $\hat{L}_{MM1}$  and  $\hat{L}_{MM2}$ , respectively),
- 2) standard ML estimator ( $\hat{L}_{ML}$ ) [1],
- 3) improved ML estimator according to the Barndorff-Nielsen adjustment ( $\hat{L}_{BN}$ ), and
- 4) improved ML estimator based on the Cox-Snell methodology ( $\tilde{L}_{ML}$ ).

1) *Synthetic data:* In order to quantify the performance of five estimators listed above, a simulation

study was considered. To that end, Monte Carlo scenarios were employed under 5500 replications as given in [24]. Moreover, the following parametric configuration was used in this simulation: the sample sizes  $N \in \{9, 25, 49, 81, 121\}$ , the number of looks  $L \in \{4, 6, 8, 12\}$  (where the higher values of  $L$  represent scenarios with the smaller signal-noise ratios), and covariance matrix given by

$$\Sigma = \begin{bmatrix} 962892 & 19171 - 3579\mathbf{i} & -154638 + 191388\mathbf{i} \\ & 56707 & -5798 + 16812\mathbf{i} \\ & & 472251 \end{bmatrix},$$

where the omitted elements can be obtained taking the conjugation of their respective symmetric elements. Frery *et al.* [25] utilized such matrix for representing urban areas. This estimated covariance was obtained by the E-SAR sensor over Weßling, Germany. Additionally, such simulation employed three statistical measures for quantifying the performance of estimators: (i) the mean squared error ( $\text{MSE}_{\hat{L}}$ ), (ii) the bias ( $B_{\hat{L}}$ ), and (iii) the coefficient of variation ( $\text{CV}_{\hat{L}}$ ).

Table I presents the mean of estimates for  $L$  and the values for the comparison measures. In general terms, the increasing of sample size diminishes the values assumed by MSE and CV. This behavior is expected since the estimators are asymptotically unbiased. Results show evidence that the estimator  $\tilde{L}_{\text{ML}}$  presented the best performance for all the different pairs  $(N, L)$ . Additionally to these facts, Fig. 1 presents the values of bias for several sample sizes. One has that the following inequality was observed:

$$B_{\hat{L}_{\text{MM1}}} \geq B_{\hat{L}_{\text{MM2}}} \geq B_{\hat{L}_{\text{ML}}} \geq B_{\hat{L}_{\text{BN}}} \geq B_{\tilde{L}_{\text{ML}}}.$$

Moreover, it is possible to observe that the estimator  $\tilde{L}_{\text{ML}}$  is the less affected by the increasing of number of looks. Therefore, this estimator showed to be more robust to both the sample sizes and the considered numbers of looks.

2) *Analysis with actual data:* The discussed synthetic study is applied to a actual PolSAR image, considering only the three best estimators. To that end, we work on three areas of the San Francisco image. Fig. 2 presents the HH channel intensity data of the area with size  $150 \times 150$  represented by three main targets (ocean, forest, and urban area) and one subregion was selected for each one of them by means of visual inspection. This classical multi-look PolSAR data was obtained by the AIRSAR sensor and band  $L$ , with four nominal looks [26]. For each selected subregion, we considered 5500 subsamples without replacement and with sample size  $N \in \{9, 16, 25, 36, 49, 64, 81, 100, 121, 144\}$  in order to apply the above methodology to actual data.

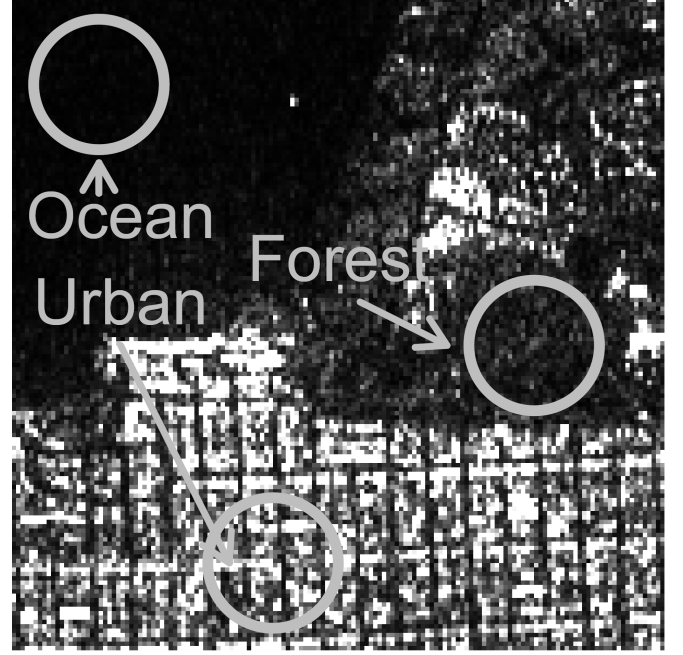


Fig. 2. San Francisco image with selected regions.

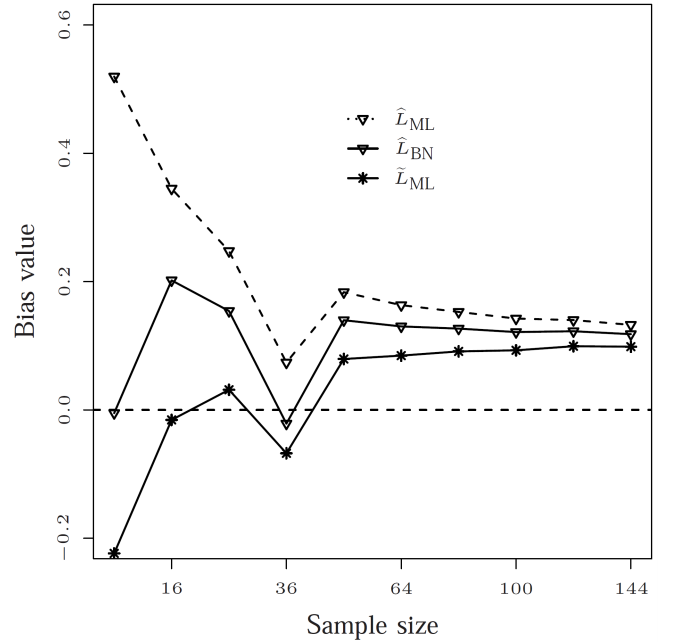


Fig. 3. Bias curves in terms of sample sizes for actual data.

Table II presents the values for the mean of estimated number of looks and for the measures MSE and CV under some adopted samples sizes. For the selected ocean subregion, the corrected estimator  $\tilde{L}_{\text{ML}}$  presented the smallest MSE and CV and the estimates more closed to the true value of parameter. As discussed by Anfinssen *et al.* [1], the textured regions underestimate the estimation and the corrections became less efficient

than the estimator under genuine likelihood,  $\hat{L}_{ML}$ .

Fig. 3 shows the behavior of biases for the ocean region. These curves show that the use of  $\tilde{L}_{ML}$  is more indicated than usual ML estimator in the modelling of homogeneous target.

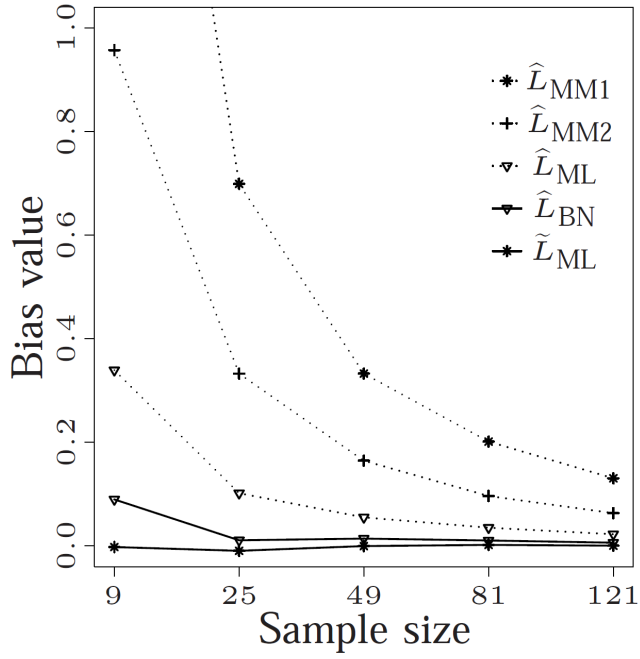
## V. CONCLUSION

In this work, we proposed improvements for the ML parameter estimation for the scaled complex Wishart distribution. We present closed expressions for second-order bias and modified profile likelihood relative to the number of looks according to Cox-Snell [3] and Barndorff-Nielsen [11] general proposals, respectively. Such proposed tools degenerate in two new estimators for the parameter  $L$ .

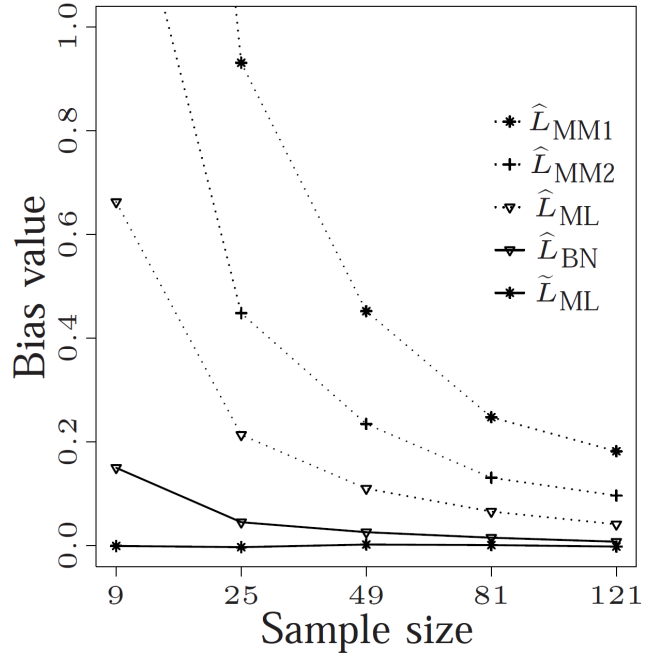
These methodologies have been compared to three recent estimators presented by Anfinsen *et al.* [1] by means of Monte Carlo experiments. Taking as comparison criterion the mean square error, the coefficient of variation, and the bias on the obtained estimates, the results show evidence that the corrected estimators outperform the others suggested by literature until now. Finally, an application to regions of San Francisco PolSAR image indicated that the derived estimators presented the best performance in homogeneous scenarios.

## REFERENCES

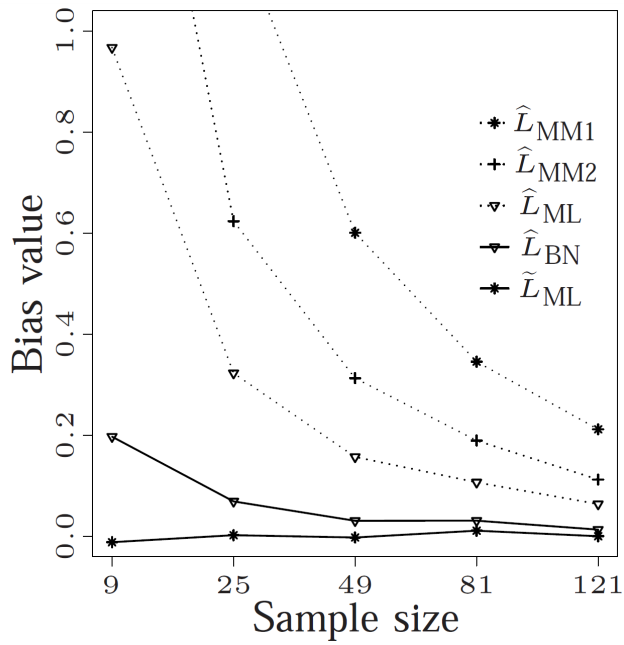
- [1] S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, "Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3795–3809, 2009.
- [2] C. H. Gierull and I. C. Sikaneta, "Estimating the effective number of looks in interferometric SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 8, pp. 1733–1742, 2002.
- [3] D. R. Cox and E. J. Snell, "A general definition of residuals," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 30, no. 2, pp. 248–275, 1968.
- [4] G. M. Cordeiro, "Bias correction," in *International Encyclopedia of Statistical Science*, M. Lovric, Ed. Springer Berlin Heidelberg, 2011, pp. 148–152.
- [5] K. L. P. Vasconcellos, A. C. Frery, and L. B. Silva, "Improving estimation in speckled imagery," *Computational Statistics*, vol. 20, no. 3, pp. 503–519, 2005.
- [6] D. M. Pianto and F. Cribari-Neto, "Dealing with monotone likelihood in a model for speckled data," *Computational Statistics & Data Analysis*, vol. 55, no. 3, pp. 1394–1409, 2011.
- [7] D. Firth, "Bias reduction of maximum likelihood estimates," *Biometrika*, vol. 80, no. 1, pp. 27–38, 1993.
- [8] B. Efron, "More efficient bootstrap computations," *Journal of the American Statistical Association*, vol. 85, no. 409, pp. 79–89, 1990.
- [9] T. A. Severini, *Likelihood Methods in Statistics*. Oxford: Oxford university Press, 2000.
- [10] D. A. S. Fraser and N. Reid, "Adjustments to profile likelihood," *Biometrika*, vol. 76, no. 3, pp. 477–488, 1989.
- [11] O. E. Barndorff-Nielsen, "Adjusted versions of profile likelihood and directed likelihood, and extended likelihood," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 56, no. 1, pp. 125–140, 1994.
- [12] A. H. M. A. Cysneiros, F. Cribari-Neto, and C. A. G. Araújo Jr., "On Birnbaum-Saunders inference," *Computational Statistics & Data Analysis*, vol. 52, no. 11, pp. 4939–4950, 2008.
- [13] T. Melo and S. Ferrari, "A modified signed likelihood ratio test in elliptical structural models," *AStA Advances in Statistical Analysis*, vol. 94, pp. 75–87, 2010, 10.1007/s10182-010-0123-4.
- [14] M. Silva, F. Cribari-Neto, and A. C. Frery, "Improved likelihood inference for the roughness parameter of the GA0 distribution," *Environmetrics*, vol. 19, no. 4, pp. 347–368, 2008.
- [15] F. T. Ulaby and C. Elachi, *Radar Polarimetry for Geoscience Applications*. Norwood: Artech House, 1990.
- [16] A. Freeman and S. L. Durden, "A three-component scattering model for polarimetric SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 3, pp. 963–973, May 1998.
- [17] N. R. Goodman, "Statistical analysis based on a certain complex Gaussian distribution (an introduction)," *The Annals of Mathematical Statistics*, vol. 34, pp. 152–177, 1963.
- [18] O. E. Barndorff-Nielsen and D. R. Cox, *Asymptotic Techniques for use in Statistics*. London: Chapman & Hall, 1989.
- [19] D. R. Cox and N. Reid, "A note on the difference between profile and modified profile likelihood," *Biometrika*, vol. 79, no. 2, pp. 408–411, 1992.
- [20] S. E. Stern, "A second-order adjustment to the profile likelihood in the case of a multidimensional parameter of interest," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, vol. 59, no. 3, pp. 653–665, 1997.
- [21] P. McCullagh and R. Tibshirani, "A simple method for the adjustment of profile likelihoods," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 52, no. 2, pp. 325–344, 1990.
- [22] O. E. Barndorff-Nielsen, "On a formula for the distribution of the maximum likelihood estimator," *Biometrika*, vol. 70, no. 2, pp. 343–365, 1983.
- [23] J. E. Gentle, *Elements of Computational Statistics*, ser. Statistics and computing. Springer, 2002.
- [24] A. D. C. Nascimento, R. J. Cintra, and A. C. Frery, "Hypothesis testing in speckled data with stochastic distances," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 1, pp. 373–385, 2010.
- [25] A. C. Frery, J. Jacobo-Berlles, J. Gambini, and M. Mejail, "Polarimetric SAR image segmentation with B-splines and a new statistical model," *Multidimensional Systems and Signal Processing*, vol. 21, pp. 319–342, 2010.
- [26] European Space Agency, *Polarimetric SAR Data Processing and Educational Tool (POLARPRO)*, 4th ed. European Space Agency, 2009.



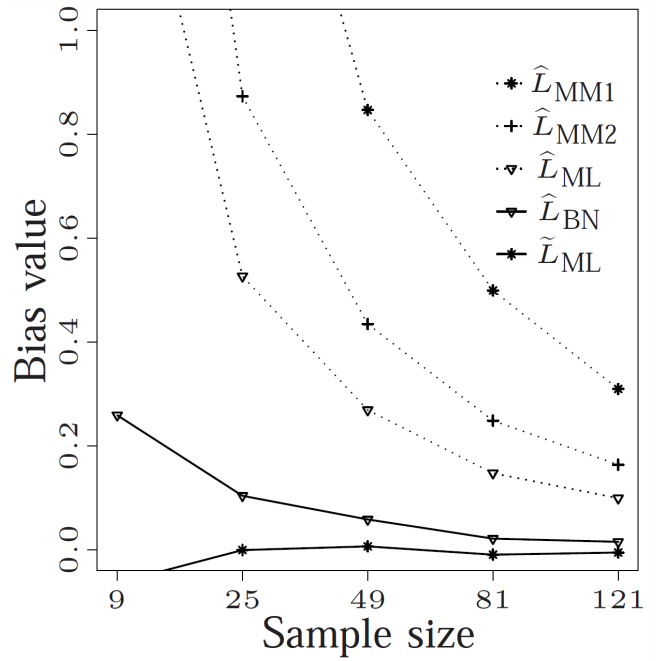
(a)  $L = 4$



(b)  $L = 6$



(c)  $L = 8$



(d)  $L = 12$

Fig. 1. Bias curves in terms of sample sizes for synthetic data.

### 5.3 Expressões para distâncias em dados PolSAR

Imagens resultantes de sistemas coerentes são sujeitas a um padrão de interferência particular, *ruído speckle* (Oliver & Quegan 1998). Este fenômeno pode afetar seriamente a interpretação de imagens PolSAR (López-Martínez et al. 2005). Assim, o processamento com tais imagens requer a proposta de técnicas especializadas para análise do sinal retornado.

Técnicas de segmentação (Beaulieu & Touzi 2004), classificação (Kersten et al. 2005), detecção de bordas (Frery, Jacobo-Berlles, Gambini & Mejail 2010, Gambini et al. 2008) e mudanças (Inglada & Mercier 2007) frequentemente empregam medidas de dissimilaridade para discriminar dados. Tais medidas têm sido utilizadas para quantificar diferenças entre as regiões de uma imagem, sendo por vezes chamada de “medidas de contraste”. Desta forma, a derivação analítica das medidas de contraste e de suas propriedades podem ser uma importante fonte de entendimento para análise de imagens.

Goudail & Réfrégier (2004) aplicaram estas medidas para caracterizar a performance em algoritmos de detecção de mudanças e segmentação em processamento de imagens PolSAR. Naquele estudo, as distâncias de Kullback-Leibler e de Bhattacharyya foram consideradas como ferramentas para discriminar elementos da família de distribuições gaussiana complexa circular (vide Seção 3.2). Como conclusão, a medida de Bhattacharyya superou a medida de Kullback-Leibler. Entretanto, as propriedades estatísticas destas medidas não foram explicitamente consideradas por Goudail & Réfrégier.

A tecnologia PolSAR considera que o sinal retornado é adequadamente representado por sua matriz de covariância. Baseando-se na proposta de López-Martínez et al. (2005), a distribuição Wishart complexa escalonada é um modelo bastante útil para descrever dados com natureza similar a da matriz de covariância, isto é, dados positivos definidos Hermitianos.

Conradsen et al. (2003) propuseram uma metodologia baseada sobre o teste da razão de verossimilhanças objetivando discriminar duas variáveis Wishart distribuídas. Esta proposta consistiu em um teste de hipótese envolvendo duas matrizes de covariâncias. Em caminho similar, testes de hipóteses para dados SAR monopolarizados foram propostos por Nascimento et al. (2010). Esta seção estende tal estudo no sentido que trabalha com  $m$  polarizações no mesmo contexto teórico, a saber a *classe*  $(h, \phi)$  de *medidas de divergências* proposta por Salicrú et al. (1994).

Neste estudo, nós:

- Derivamos expressões analíticas para quatro distâncias estocásticas [a saber, Kullback-Leibler ( $d_{KL}$ ), Rényi de ordem  $\beta$  ( $d_R^\beta$ ), Bhattacharyya ( $d_B$ ) e Hellinger ( $d_H$ )] entre os



elementos da família de distribuições do tipo Wishart complexa escalonada. Além do tratamento algébrico num sentido mais geral, importantes casos particulares foram considerados,

- ✎ Usando estas distâncias, propomos desigualdades, duas das quais resultaram em um novo e simples caminho de derivar duas distâncias usuais no tratamento de imagem PolSAR.
- ✎ Quantificamos a expressividade das quatro distâncias analíticas com relação a variação dos parâmetros dessa distribuição: matriz de covariâncias e número de looks.
- ✎ Com uma conveniente parametrização da distribuição Wishart complexa escalonada (vide Seção 3.2), outras distâncias estocásticas são formuladas em termos de seus coeficientes de correlação entre canais de polarização,
- ✎ Aplicamos estas medidas ao método  $k$ -médias proposto por Cao et al. (2007).
- ✎ Consideramos uma aplicação a dados reais a fim de ilustrar a capacidade de discriminação destas distâncias.

Estes requisitos foram alcançados e discutidos no trabalho

**“Analytic Expressions for Stochastic Distances Between Relaxed Complex Wishart Distributions: Testing Homogeneous in PolSAR Data”,**

o qual está anexo a esta seção. Muito embora os resumos teóricos deste trabalho estejam fortemente correlacionados com as Seções 3.1 e 4.2, este artigo tem uma notação própria e pode ser analisado de forma independente da tese.

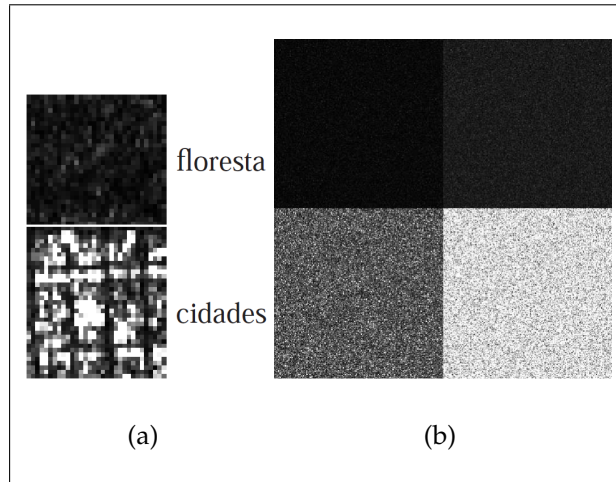
Na seguinte subseção, as expressões analíticas para as distâncias estocásticas  $d_{KL}$ ,  $d_R^\beta$ ,  $d_B$  e  $d_H$  entre duas variáveis com distribuição Wishart complexa escalonada de dimensão  $m \times m$  são derivadas. Em todos os casos, considerou-se distâncias entre variáveis com distribuições Wishart escalonada definidas em  $\theta_1 = (\Sigma_1, L_1)$  e  $\theta_2 = (\Sigma_2, L_2)$ . Nós examinamos três casos especiais:

- (i) a situação mais geral  $\Sigma_1 \neq \Sigma_2$  e  $L_1 \neq L_2$ ,
- (ii) situação em mesmo número de looks  $L_1 = L_2 = L$  e diferentes matrizes de covariância  $\Sigma_1 \neq \Sigma_2$ ,
- (iii) situação numa mesma matriz de covariância  $\Sigma_1 = \Sigma_2$  e diferentes números de looks  $L_1 \neq L_2$ .

O caso (ii) é provavelmente o mais utilizado na prática, uma vez que permite a comparação de duas áreas diferentes, possivelmente a partir da mesma imagem (para aplicações neste sentido, consultar Conradsen et al. 2003). No caso (iii), permite-se a

avaliação de mudanças na distribuição, considerando apenas processamentos multi-look sobre uma mesma área. A seguir, nós apresentaremos um estudo como imagens sintéticas a fim de exibir as distinções de cenários ao considerarmos diferentes pares  $(\Sigma, L)$ .

As Figuras 5.10(a) e 5.10(b) apresentam, respectivamente, (a) uma imagem PolSAR real bi-particionada por regiões de cidades e de florestas para a polarização HH e (b) uma imagem sintética a partir da real obedecendo a lei Wishart complexa escalonada. A imagem sintética possui quatro regiões de tamanhos  $200 \times 200$  cada uma, em que as duas superiores foram obtidas da parte de florestas para os números de looks de 3 e 8, considerando o sentido da esquerda para direita do leitor; enquanto que as duas da base foram geradas a partir da região de cidades, respeitando a mesma lógica de looks das superiores. Para que a imagem obtida via simulação de Monte Carlo obedecesse a região de cidades, nós consideramos, como matriz de covariância, a média amostral dos pixels desta região (a exemplo de Frery et al. 2007). O mesmo argumento foi usado para a obtenção de dados sintéticos a partir da região de florestas. É notável que as regiões sintéticas com maior número looks e/ou maiores determinantes da matriz de covariância amostral apresentam um brilho mais intenso. Esta ilustração reflete a importância da busca por medidas de contraste em imagem PolSAR.



**Figura 5.10** Imagens real com as matrizes de covariâncias amostrais  $S_U$  (para a parte urbana) e  $S_F$  (para floresta) e simulada tipo PolSAR. A imagem simulada foi gerada obedecendo a lei Wishart complexa escalonada sob a seguinte configuração paramétrica:  $\{\Sigma : |\Sigma| \in \{|S_F|, |S_U|\} = \{4.89 \times 10^{-5}, 4.21 \times 10^{-3}\}\}$  (de cima para baixo) e  $L \in (3, 8)$  (da esquerda para direita).

### 5.3.1 Contribuições

**Linha teórica:**

### 1) Distâncias para modelos polarimétricos:

Aplicando as distâncias discutidas na Seção 4.2 ao contexto em que  $\mathbf{X} \sim \mathcal{W}(\boldsymbol{\theta}_1)$  :  $\boldsymbol{\theta}_1 = (\boldsymbol{\Sigma}_1, L_1)$  e  $\mathbf{Y} \sim \mathcal{W}(\boldsymbol{\theta}_2)$  :  $\boldsymbol{\theta}_2 = (\boldsymbol{\Sigma}_2, L_2)$ , nós obtivemos as seguintes expressões analíticas:

$$d_{\text{KL}}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \frac{L_1 - L_2}{2} \left\{ \log \frac{|\boldsymbol{\Sigma}_1|}{|\boldsymbol{\Sigma}_2|} - m \log \frac{L_1}{L_2} + m [\psi^{(0)}(L_1 - m + 1) - \psi^{(0)}(L_2 - m + 1)] \right. \\ \left. + (L_2 - L_1) \sum_{i=1}^{m-1} \frac{i}{(L_1 - i)(L_2 - i)} \right\} \\ + \frac{\text{tr}(L_2 \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 + L_1 \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_2)}{2} - \frac{m(L_1 + L_2)}{2}, \quad (5.3)$$

em que  $\psi^{(0)}(\cdot)$  representa a função digama.

$$d_{\text{B}}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \sum_{k=0}^{m-1} \log \frac{\sqrt{\Gamma(L_1 - k)\Gamma(L_2 - k)}}{\Gamma(\frac{L_1 + L_2}{2} - k)} + \frac{L_1 \log |\boldsymbol{\Sigma}_1|}{2} + \frac{L_2 \log |\boldsymbol{\Sigma}_2|}{2} \\ - \frac{m}{2}(L_1 \log L_1 + L_2 \log L_2) - \frac{L_1 + L_2}{2} \log \left| \left( \frac{L_1 \boldsymbol{\Sigma}_1^{-1} + L_2 \boldsymbol{\Sigma}_2^{-1}}{2} \right)^{-1} \right|. \quad (5.4)$$

$$d_{\text{H}}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = 1 - \frac{\left| \left( \frac{L_1 \boldsymbol{\Sigma}_1^{-1} + L_2 \boldsymbol{\Sigma}_2^{-1}}{2} \right)^{-1} \right|^{\frac{L_1 + L_2}{2}}}{|\boldsymbol{\Sigma}_1|^{\frac{L_1}{2}} |\boldsymbol{\Sigma}_2|^{\frac{L_2}{2}}} \sqrt{L_1^{mL_1} L_2^{mL_2}} \prod_{k=0}^{m-1} \frac{\Gamma(\frac{L_1 + L_2}{2} - k)}{\sqrt{\Gamma(L_1 - k)\Gamma(L_2 - k)}}. \quad (5.5)$$

$$d_{\text{R}}^{\beta}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \frac{1}{\beta - 1} \log \frac{I(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)}{2}, \quad (5.6)$$

tal que

$$I(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \left[ \Gamma(L_1 - m + 1)^m |\boldsymbol{\Sigma}_1|^{L_1} L_1^{-mL_1} \prod_{i=1}^{m-1} (L_1 - i)^i \right]^{-\beta} \Gamma(E_1 - m + 1)^m |\boldsymbol{\Sigma}_{12}|^{E_1} \prod_{i=1}^{m-1} (E_1 - i)^i \\ \times \left[ \Gamma(L_2 - m + 1)^m |\boldsymbol{\Sigma}_2|^{L_2} L_2^{-mL_2} \prod_{i=1}^{m-1} (L_2 - i)^i \right]^{\beta-1} + \Gamma(E_2 - m + 1)^m |\boldsymbol{\Sigma}_{21}|^{E_2} \prod_{i=1}^{m-1} (E_2 - i)^i \\ \left[ \Gamma(L_2 - m + 1)^m |\boldsymbol{\Sigma}_2|^{L_2} L_2^{-mL_2} \prod_{i=1}^{m-1} (L_2 - i)^i \right]^{-\beta} \\ \times \left[ (\Gamma(L_1 - m + 1)^m |\boldsymbol{\Sigma}_1|^{L_1} L_1^{-mL_1} \prod_{i=1}^{m-1} (L_1 - i)^i)^{\beta-1} \right],$$

em que  $E_1 = \beta L_1 + (1 - \beta)L_2$ ,  $E_2 = \beta L_2 + (1 - \beta)L_1$ ,  $0 < \beta < 1$ , e  $\boldsymbol{\Sigma}_{ij} = |(L_i \beta \boldsymbol{\Sigma}_i^{-1} + L_j(1 - \beta) \boldsymbol{\Sigma}_j^{-1})^{-1}|$  para  $(i, j) = (1, 2), (2, 1)$ . O tratamento matemático desta seção

é apresentado no Apêndice E. Adicionalmente, a distância  $\chi^2$  também foi estudada por Frery, Nascimento & Cintra (2011), entretanto ela demonstrou ser numericamente instável. Portanto, ela não será considerada nesta tese.

A não negatividade destas distâncias implica nas seguintes desigualdades:

$$\text{tr}(\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1 + \mathbf{\Sigma}_1^{-1}\mathbf{\Sigma}_2) \geq 2m, \quad (5.7)$$

$$\begin{aligned} & \left( \frac{|\mathbf{\Sigma}_2|}{|\mathbf{\Sigma}_1|} \right)^{L\beta} |\mathbf{\Sigma}_2|^{-L} |(\beta\mathbf{\Sigma}_1^{-1} + (1-\beta)\mathbf{\Sigma}_2^{-1})^{-1}|^L \\ & + \left( \frac{|\mathbf{\Sigma}_1|}{|\mathbf{\Sigma}_2|} \right)^{L\beta} |\mathbf{\Sigma}_1|^{-L} |(\beta\mathbf{\Sigma}_2^{-1} + (1-\beta)\mathbf{\Sigma}_1^{-1})^{-1}|^L \leq 2, \end{aligned} \quad (5.8)$$

$$\frac{\log |\mathbf{\Sigma}_1| + \log |\mathbf{\Sigma}_2|}{2} \geq \log \left| \left( \frac{\mathbf{\Sigma}_1^{-1} + \mathbf{\Sigma}_2^{-1}}{2} \right)^{-1} \right|, \quad (5.9)$$

e

$$\sqrt{|\mathbf{\Sigma}_1||\mathbf{\Sigma}_2|} \geq \left| \left( \frac{\mathbf{\Sigma}_1^{-1} + \mathbf{\Sigma}_2^{-1}}{2} \right)^{-1} \right|, \quad (5.10)$$

respectivamente. Note que fixando  $\beta = 1/2$  em (5.8), nós obtemos (5.10) diretamente; tomando o logaritmo de ambos os lados de (5.10) resulta em (5.9). Esse resultado é justificado pelas duas seguintes propriedades:

1.  $d_B(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = d_R^{1/2}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$  e
2.  $d_H(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = 1 - \exp \left[ -\frac{1}{2}d_R^{1/2}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \right]$ .

As distâncias *Wishart revisada* proposta por Ersahin, Cumming & Ward (2010) e de *Bartlett* derivada por Kersten et al. (2005) podem ser obtidas em um novo e simples caminho. De fato, a distância Wishart revisada ( $d_{RW}$ ) pode ser derivada depois de simples manipulações da inequação (5.7), resultando:

$$d_{RW}(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) = \frac{\text{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1} + \mathbf{\Sigma}_2\mathbf{\Sigma}_1^{-1})}{2} - m \geq 0.$$

Por outro lado, a distância de Bartlett é obtida pela aplicação do logaritmo a ambos os lados da desigualdade (5.10). Álgebra simples leva a

$$\log \frac{|\mathbf{\Sigma}_1 + \mathbf{\Sigma}_2|^2}{|\mathbf{\Sigma}_1||\mathbf{\Sigma}_2|} - 2m \log 2 \geq 0.$$

O logaritmando acima é referenciado como a distância de Bartlett.

## 2) Distâncias para modelos bivariados:

Considerando a restrição (3.14) para as matrizes de covariância  $\Sigma_i$ ,  $i = 1, 2$ , temos que as distâncias correspondentes entre os parâmetros  $\theta_1 = (\rho_1, L)$  e  $\theta_2 = (\rho_2, L)$  são dadas por:

$$d_{\text{KL}}(\theta_1, \theta_2) = L \left[ \frac{(1 - |\rho_1||\rho_2|)(2 - |\rho_1|^2 - |\rho_2|^2)}{(1 - |\rho_1|^2)(1 - |\rho_2|^2)} - 2 \right]. \quad (5.11)$$

$$d_{\text{R}}^\beta(\theta_1, \theta_2) = \frac{\log 2}{1 - \beta} + \frac{1}{\beta - 1} \log \left\{ \left[ \frac{(1 - |\rho_1|^2)^{1-\beta}(1 - |\rho_2|^2)^\beta}{1 - \{|\rho_1| - \beta(|\rho_1| - |\rho_2|)\}^2} \right]^L + \left[ \frac{(1 - |\rho_2|^2)^{1-\beta}(1 - |\rho_1|^2)^\beta}{1 - \{|\rho_2| - \beta(|\rho_2| - |\rho_1|)\}^2} \right]^L \right\}. \quad (5.12)$$

$$d_{\text{B}}(\theta_1, \theta_2) = L \left\{ \frac{\log(1 - |\rho_1|^2) + \log(1 - |\rho_2|^2)}{2} - 2 \log 2 + \log \left[ \frac{4 - (|\rho_1| + |\rho_2|)^2}{(1 - |\rho_1|^2)(1 - |\rho_2|^2)} \right] \right\}. \quad (5.13)$$

$$d_{\text{H}}(\theta_1, \theta_2) = 1 - \left[ 4 \frac{\sqrt{(1 - |\rho_1|^2)(1 - |\rho_2|^2)}}{4 - (|\rho_1| + |\rho_2|)^2} \right]^L. \quad (5.14)$$

A seguir, propomos cenários, em que estas medidas são aplicadas com efeito. Estes cenários utilizam estas medidas como um parâmetro de discriminação bem como as inserem em contextos de inferência estatística: tais como na realização de testes de hipótese ou na construção de regiões de confiança para parâmetros importantes em tecnologia PolSAR.

#### **Linha aplicada:**

### **1) Testes de homogeneidade em dados polarimétricos:**

Baseando-se nas expressões para as distâncias combinadas à metodologia discutida na Seção 4.2, nós propomos quatro novos testes de hipótese para avaliar homogeneidade em imagens PolSAR. Afim de avaliar a influência dos parâmetros sobre os tamanhos estimados dos testes propostos, um estudo de simulação de Monte Carlo foi considerado. Neste caso, os experimentos de Monte Carlo foram conduzidos empregando os seguintes parâmetros de simulação:

**Estimáveis:** Números de looks  $L = L_1 = L_2 \in \{4, 8, 16\}$  e a matriz de covariância amostral de regiões de florestas [área “forest-1” dada na Figura 3.5 e discutida na Tabela 3.1].

**Não estimáveis:** Tamanhos de amostra relacionado às janelas quadradas  $7 \times 7$ ,  $11 \times 11$  e  $20 \times 20$  pixels e aos níveis nominais  $\alpha \in \{1\%, 5\%\}$ .

Seja  $T$  o número de replicações de Monte Carlo e  $C$  o número de cenários em que a hipótese nula é rejeitada no nível nominal  $\alpha$ . O tamanho empírico do teste é dado por  $\hat{\alpha}_{1-\alpha} = C/T$ . Similarmente à metodologia empregada por Nascimento et al. (2010), nós utilizamos  $T = 5500$  replicações.

A Tabela 5.9 apresenta os tamanhos empíricos nos níveis 1% e 5%, a média das estatísticas de teste ( $\bar{S}$ ) e seus coeficientes de variação (CV). Para cada caso, os melhores tamanhos dos testes estimados (isto é, os tamanhos estimados mais próximos aos níveis nominais adotados) estão destacados. Além disso, os resultados para  $N_1 = 49, N_2 = \{121, 400\}$  e  $N_1 = 121, N_2 = 400$  são consistentes com aqueles mostrados na Tabela 5.9 e foram omitidos por simplicidade.

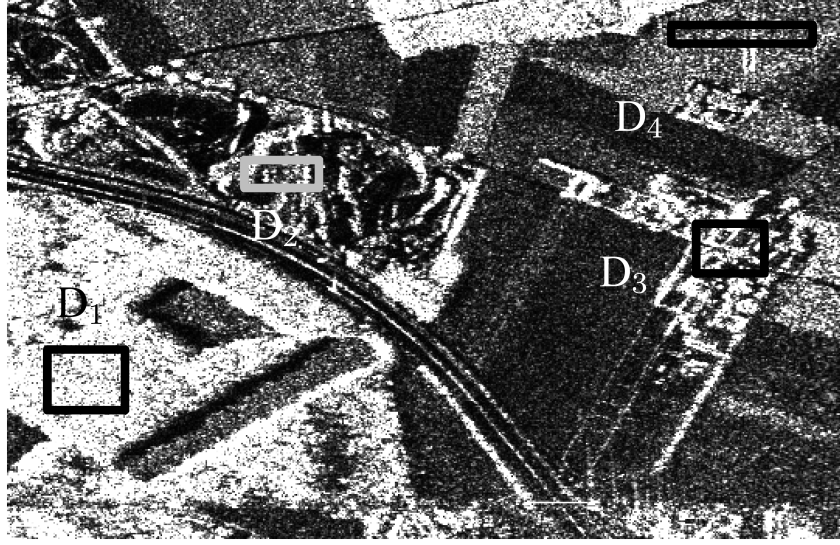
**Tabela 5.9** Tamanho empírico do teste para a região “forest-1”

| Fatores       |       |       | $L = 4$      |              |           |        | $L = 8$      |              |           |        | $L = 16$     |              |           |         |
|---------------|-------|-------|--------------|--------------|-----------|--------|--------------|--------------|-----------|--------|--------------|--------------|-----------|---------|
| $S_{\phi}^h$  | $N_1$ | $N_2$ | 1%           | 5%           | $\bar{S}$ | CV     | 1%           | 5%           | $\bar{S}$ | CV     | 1%           | 5%           | $\bar{S}$ | CV      |
| $S_{KL}$      | 49    | 49    | 1.109        | 5.291        | 10.132    | 44.661 | 1.036        | 5.855        | 10.243    | 44.797 | 1.036        | 5.236        | 10.005    | 45.048  |
|               | 121   | 121   | 1.036        | <b>4.927</b> | 10.041    | 44.849 | <b>0.891</b> | 5.200        | 10.070    | 44.148 | 2.655        | 6.727        | 14.618    | 250.833 |
|               | 400   | 400   | <b>0.836</b> | 5.164        | 10.026    | 44.483 | 1.036        | 5.273        | 10.022    | 45.312 | 1.036        | 5.036        | 10.091    | 44.069  |
| $S_R^{\beta}$ | 49    | 49    | 1.091        | 5.182        | 10.105    | 44.539 | <b>1.000</b> | 5.782        | 10.222    | 44.710 | <b>1.018</b> | 5.145        | 9.991     | 44.981  |
|               | 121   | 121   | 1.036        | 4.891        | 10.030    | 44.802 | <b>0.891</b> | 5.164        | 10.061    | 44.111 | 2.655        | 6.727        | 14.385    | 242.443 |
|               | 400   | 400   | <b>0.836</b> | 5.127        | 10.022    | 44.468 | 1.036        | 5.273        | 10.019    | 45.301 | <b>1.018</b> | <b>5.036</b> | 10.089    | 44.061  |
| $S_B$         | 49    | 49    | <b>1.018</b> | <b>5.036</b> | 10.058    | 44.329 | 0.927        | <b>5.691</b> | 10.186    | 44.561 | 0.945        | <b>4.963</b> | 9.965     | 44.865  |
|               | 121   | 121   | <b>0.982</b> | 4.763        | 10.011    | 44.719 | 0.855        | <b>5.091</b> | 10.047    | 44.047 | 2.655        | 6.654        | 14.035    | 229.490 |
|               | 400   | 400   | <b>0.836</b> | <b>5.109</b> | 10.016    | 44.441 | 1.036        | 5.273        | 10.015    | 45.281 | <b>1.018</b> | <b>5.036</b> | 10.086    | 44.048  |
| $S_H$         | 49    | 49    | 0.582        | 3.909        | 9.756     | 43.000 | 0.600        | 4.236        | 9.876     | 43.238 | 0.527        | 3.909        | 9.668     | 43.547  |
|               | 121   | 121   | 0.709        | 4.200        | 9.888     | 44.172 | 0.655        | 4.600        | 9.924     | 43.505 | <b>2.364</b> | <b>6.218</b> | 12.946    | 190.636 |
|               | 400   | 400   | 0.818        | 4.945        | 9.979     | 44.277 | <b>1.018</b> | <b>5.091</b> | 9.977     | 45.116 | 0.964        | 4.927        | 10.048    | 43.884  |

Pode-se verificar que os valores de  $\bar{S}$  estão convergindo para 10; isto é, para o grau de liberdade de uma distribuição  $\chi_{10}^2$ . Note que, para o caso de uma matriz de covariância  $3 \times 3$ , tem-se três parâmetros da diagonal principal a serem estimados, três referentes à parte real do triângulo superior e outros três à parte imaginária. Além disso, deve-se estimar o número de looks, somatizando 10 parâmetros a serem testados. Desta forma, as convergências obtidas são esperadas, pois, como (i) as estatísticas tem uma distribuição  $\chi_{10}^2$  *a priori* e (ii) a esperança desta distribuição é equivalente ao seu grau de liberdade, a média das estatísticas devem convergir para seu grau de liberdade pela lei fraca dos grandes números de Kintchin. A similaridade estocástica das estatísticas à distribuição  $\chi_{10}^2$  é bastante razoável até mesmo para tamanhos pequenos de amostras; isto é, os testes são recomendáveis para tamanhos pequenos de amostras. Em termos de competitividade, o teste baseado na distância de Bhattacharyya apresentou melhores resultados na maior parte dos casos.

## 2) Regiões de confiança para coeficientes de correlação entre canais de polarizações de diferentes regiões PolSAR :

Usando a parametrização proposta por Lee, Hoppel, Mango & Miller (1994) dis-



**Figura 5.11** Imagem PolSAR como quatro regiões selecionadas.

cutida na Seção 3.2, quatro medidas baseadas no coeficiente de correlação complexo entre os canais HH e VV foram obtidas, (5.11)-(5.14). Assim, quatro novos testes de hipóteses são derivados para checar se duas amostras, de tamanhos  $N_1$  e  $N_2$ , vem de regiões com coeficientes de correlação similares entre os canais de polarização HH e HV. Estas estatísticas podem ser reescritas como regiões de confiança. Note que as estatísticas baseadas nas distâncias de Kullback-Leibler e Hellinger apresentam as seguintes regiões de confiança:

$$R_{KL}(\rho_1, \rho_2) = \left\{ (\rho_1, \rho_2) \in [0, 1] \times [0, 1] : \right.$$

$$\underbrace{\frac{(1 - |\rho_1| |\rho_2|)(2 - |\rho_1|^2 - |\rho_2|^2)}{(1 - |\rho_1|^2)(1 - |\rho_2|^2)}}_{\xi_1} \leq \overbrace{\frac{(N_1 + N_2)}{2nN_1N_2} \chi_1^2(\alpha) + 2}^{t_{KL}} \left. \right\}$$

(5.15)

e

$$R_H(\rho_1, \rho_2) = \left\{ (\rho_1, \rho_2) \in [0, 1] \times [0, 1] : \right.$$

$$\underbrace{\frac{\sqrt{(1 - |\rho_1|^2)(1 - |\rho_2|^2)}}{4 - (|\rho_1| + |\rho_2|)^2}}_{\xi_2} \geq \overbrace{\frac{1}{4} \left[ 1 - \frac{(N_1 + N_2)}{2nN_1N_2} \chi_1^2(\alpha) \right]^{1/n}}^{t_H} \left. \right\},$$

(5.16)

em que  $\alpha$  é o nível nominal especificado.

Em termos práticos, a utilização destas regiões é bastante simples. Dadas duas amostras polarimétricas seus coeficientes de correlação podem ser obtidos pela equação (3.13). Adicionalmente, as quantidades  $\hat{\xi}_1$ ,  $\hat{\xi}_2$ ,  $t_{KL}$  e  $t_H$  são obtidas por (5.15)-(5.16). Assim, as seguintes regras de decisão podem ser definidas:

- Critério de Kullback-Leibler: Se  $\hat{\xi}_1 \leq t_{KL}$ , então nós temos evidência estatística que as amostras vêm de populações com coeficientes de correlação similares
- Critério de Hellinger: Se  $\hat{\xi}_2 \geq t_H$ , as amostras apresentam correlações entre os canais HH e VV equivalentes.

Deste modo, as regiões  $R_{KL}(\rho_1, \rho_2)$  e  $R_H(\rho_1, \rho_2)$  combinadas à equação (3.13) resultam em métodos para comparar duas regiões PolSAR bivariadas, baseando-se apenas em seus coeficientes de correlação complexos.

A fim de ilustrar esta metodologia, considere quatro regiões retiradas da uma amostra da imagem de Munique, Alemanha, na Figura 5.11. A Tabela 5.10 lista as quantidades estatísticas para avaliar homogeneidade em termos do coeficiente de correlação, juntamente com as decisões. Note que ambas as regras discriminam bem amostras, de fato, distintas.

**Tabela 5.10** Resultados baseados sobre coeficiente de correlação

| Regions     | $\hat{\xi}_1$ | $t_{KL}$ | Decisão          | $\hat{\xi}_2$ | $t_H$  | Decision         |
|-------------|---------------|----------|------------------|---------------|--------|------------------|
| $D_1 - D_2$ | 2.0220        | 2.0016   | <i>Diferente</i> | 0.2486        | 0.2499 | <i>Diferente</i> |
| $D_1 - D_3$ | 2.0068        | 2.0012   | <i>Diferente</i> | 0.2495        | 0.2499 | <i>Diferente</i> |
| $D_1 - D_4$ | 2.0734        | 2.0014   | <i>Diferente</i> | 0.2455        | 0.2499 | <i>Diferente</i> |
| $D_2 - D_3$ | 2.0534        | 2.0018   | <i>Diferente</i> | 0.2467        | 0.2499 | <i>Diferente</i> |
| $D_2 - D_4$ | 2.0149        | 2.0020   | <i>Diferente</i> | 0.2490        | 0.2498 | <i>Diferente</i> |
| $D_3 - D_4$ | 2.1254        | 2.0016   | <i>Diferente</i> | 0.2424        | 0.2499 | <i>Diferente</i> |

Adicionalmente, é importante ressaltar que a distância de Kullback-Leibler entre elementos da importante família exponencial foi considerada no Capítulo 4.2. A partir da expressão apresentada, é possível obter expressões para importantes distribuições SAR univariada, tais como entre distribuições gamma.

### 3) Análise de agrupamento em imagens polarimétricas:

Finalmente, nós aplicamos as medidas desenvolvidas em análise de agrupamento via método  $k$ -médias. Neste caso, uma área com tamanho  $182 \times 210$  foi selecionada da imagem AIRSAR de São Francisco, como ilustrada na Figura 5.12(a). Esta área é composta por regiões urbanas e florestais. A Figura 5.12(b) mostra o estágio inicial da análise de agrupamento, que foi aleatoriamente gerado (a exemplo de Horta 2009), enquanto que os resultados dos novos métodos são apresentados nas figuras 5.12(c)-5.12(f). Note que a distância  $d_R^{0.1}$  tende a capturar mais altos retornos (pixels regiões



urbanas) do que as outras distâncias. Adicionalmente, a distância de Kullback-Leibler apresentou a pior performance.

### 5.3.2 Considerações da Seção 5.3

Expressões analíticas para quatro medidas de contraste entre elementos da distribuição Wishart complexa escalonada foram derivadas. Da não negatividade destas medidas derivadas, quatro desigualdades dependentes de matrizes de covariâncias Hermitianas positivas definidas foram obtidas. Duas delas, resultam em derivações mais simples às distâncias clássicas *Wishart revisada* e *Bartlett*, sugerindo uma forma alternativa mais simples de obtê-las. Dados do sensor AIRSAR confirmaram o comportamento esperado de todas as distâncias: elas são menores quando aplicadas às amostras de rugosidade similar, e maiores da outra forma. As expressões derivadas se mostraram eficientes em diversas aplicações de processamento de imagens PolSAR, tais como classificação e segmentação. Adicionalmente, as distâncias estocásticas foram utilizadas como medidas de dissimilaridade no algoritmo *k*-médias, resultando em bons classificadores em dados polarimétricos.

Utilizando a parametrização proposta por Lee, Hoppel, Mango & Miller (1994), medidas de contraste dependentes do coeficiente de correlação complexo entre canais de polarizações também foram obtidas com base nas distâncias derivadas. Combinando estas novas medidas aos resultados assintóticos propostos Salicrú et al. (1994), regiões de confiança foram derivadas. Estas regiões se mostraram aptas a avaliar contraste considerando coeficientes de correlação complexo.

Tais medidas também podem ser consideradas como núcleos de estatísticas de teste com distribuição assintótica qui-quadrado, segundo a abordagem de Salicrú et al. (1994). Utilizando experimentos de Monte Carlo, todas as estatísticas propostas com base nas distâncias estocásticas apresentaram bons desempenhos para amostras de tamanho finito. Em particular, os resultados forneceram evidências de que o teste baseado na distância de Bhattacharyya tem tamanho empírico mais próximo ao nível nominal para a maioria dos cenários considerados.

### 5.3.3 Trabalhos futuros

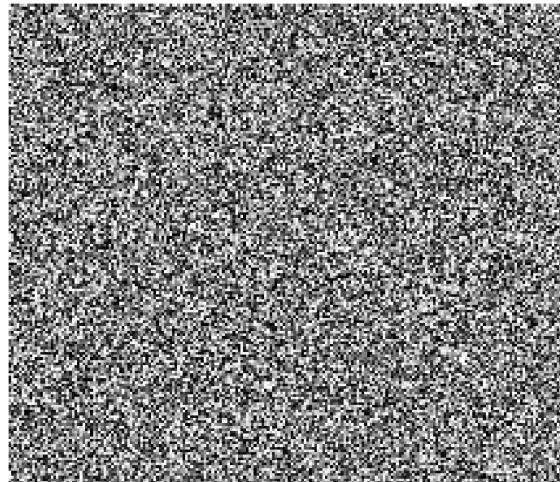
As seguintes frentes de trabalhos são consideradas a partir desta seção:

- ✎ Utilização das medidas de distâncias para detecção de outlier.
- ✎ Derivação de outras medidas de distâncias para as outras leis polarimétricas apresentadas na Tabela 4.1.

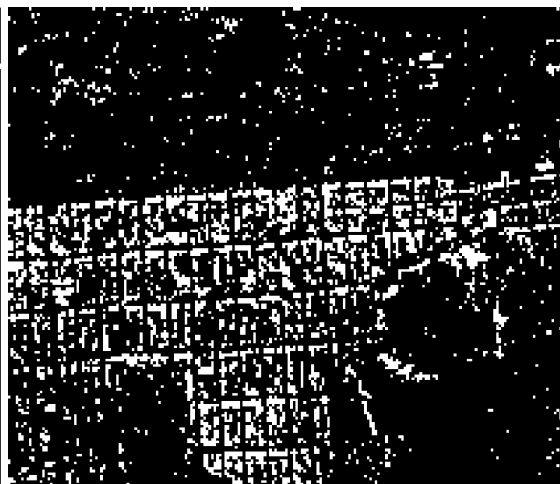
- ✎ Estudo mais apurado de análise de agrupamento: (i) comparação com outras técnicas de discriminação e (ii) análise de custo computacional.



(a) Imagem AIRSAR



(b) Inicialização aleatória

(c) Agrupamento com  $d_{KL}$ (d) Agrupamento com  $d_B$ (e) Agrupamento com  $d_H$ (f) Agrupamento com  $d_R^{0.1}$ **Figura 5.12** Análise de agrupamento envolvendo imagem PolSAR real.

---

## Anexo da Seção 5.3

### ARTIGO 3:

“Analytic Expressions for Stochastic Distances Between  
Relaxed Complex Wishart Distributions”

Submitted to *IEEE Transaction Geoscience and Remote Sensing*

---

# Analytic Expressions for Stochastic Distances Between Relaxed Complex Wishart Distributions

Alejandro C. Frery, *Member*, Abraão D. C. Nascimento, *Student Member*, and Renato J. Cintra, *Senior Member*

**Abstract**—The scaled complex Wishart distribution is a widely used model for multilook polarimetric SAR data whose adequacy has been attested in the literature. Classification, segmentation, and image analysis techniques which depend on this model have been devised, and many of them employ some type of dissimilarity measure. In this paper we derive analytic expressions for four stochastic distances between relaxed scaled complex Wishart distributions in their most general form and in important particular cases. Using these distances, inequalities are obtained which lead to new ways of deriving the Bartlett and revised Wishart distances. The expressiveness of the four analytic distances is assessed with respect to the variation of parameters. With a convenient parametrization of the Wishart distribution, the discussed distances are formulated in terms of the complex correlation coefficient between channels. Such distances are then used for deriving new tests statistics, which are proved to have asymptotic chi-square distribution. Adopting the test size as a comparison criterion, a sensitivity study is performed by means of Monte Carlo experiments suggesting that the Bhattacharyya statistic outperforms all the others. Applications to actual data illustrate the discrimination and homogeneity identification capabilities of these distances.

**Index Terms**—statistics, image analysis, information theory, polarimetric radar, contrast measures.

## I. INTRODUCTION

**P**OLARIMETRIC Synthetic Aperture Radar (PolSAR) devices transmit orthogonally polarized pulses towards a target, and the returned echo is recorded with respect to each polarization. Such remote sensing apparatus provides the means for a better capture of scene information when compared to its univariate counterpart, namely the conventional SAR technology, and

complementary information with respect to other remote sensing modalities [1], [2].

PolSAR can achieve high spatial resolution due to its coherent processing of the returned echoes [3]. Being multichanneled by design, PolSAR also allows individual characterization of the targets in various channels. Moreover, it enables the identification of correlation structures among channels.

Resulting images from coherent systems are prone to a particular interference pattern called speckle [3]. This phenomenon can seriously affect the interpretation of PolSAR imagery [2]. Thus, specialized signal analysis techniques are usually required.

Segmentation [4], classification [5], boundary detection [6], [7], and change detection [8] techniques often employ dissimilarity measures for data discrimination. Such measures have been used to quantify the difference between image regions, and are often called ‘contrast measures’. The analytical derivation of contrast measures and their properties can be an important venue for image understanding.

Goudail and Réfrégier [9] applied stochastic measures to characterize the performance of target detection and segmentation algorithms in PolSAR image processing. In that study, both Kullback-Leibler and Bhattacharyya distances were considered as tools for quantifying the dissimilarity between circular complex Gaussian distributions. Additionally, the Bhattacharyya measure was reported to possess better contrast capabilities than the Kullback-Leibler measure. However, the statistical properties of the measures were not explicitly considered [9].

PolSAR theory prescribes that the returned (backscattered) signal of distributed targets is adequately represented by its complex covariance matrix. Goodman [10] presents a comprehensive analysis of complex Gaussian models, along with the connection between the class of complex covariance matrices and the Wishart distribution. Indeed, the complex scaled Wishart distribution is widely adopted as a statistical model for multilook polarimetric data [2].

Conradsen *et al.* [11] proposed a methodology based on the likelihood-ratio test aiming at the discrimination

This work was supported by CNPq, Fapeal and FACEPE, Brazil.

A. C. Frery is with the Instituto de Computação, Universidade Federal de Alagoas, BR 104 Norte km 97, 57072-970, Maceió, AL, Brazil, email: acfrery@gmail.com

A. D. C. Nascimento is with the Departamento de Estatística, Universidade Federal de Pernambuco, Cidade Universitária, 50740-540, Recife, PE, Brazil, e-mail: abraao.susej@gmail.com

R. J. Cintra is with the Departamento de Estatística, Universidade Federal de Pernambuco, Cidade Universitária, 50740-540, Recife, PE, Brazil, e-mail: rjdesc@de.ufpe.br

of two Wishart distributed targets, leading to a test statistic that takes into account the complex covariance matrices of PolSAR images. In a similar fashion, hypothesis tests for monopolarized SAR data were proposed in [12].

In this paper, we present analytic expressions for the Kullback-Leibler, Rényi (of order  $\beta$ ), Bhattacharyya, and Hellinger distances between scaled complex Wishart distributions in their most general form and in important particular cases. In [13], Frery *et al* provided a brief analytical treatment for these distance as well as for the  $\chi^2$  distance, where it was demonstrated that the  $\chi^2$  distance is numerically unstable. Therefore, in the present work tests based on the  $\chi^2$  distance were not considered.

We also verify that those distances present scale invariance with respect to their covariance matrices. Using such distances, we derive inequalities which depend on covariance matrices; two among them, obtained from Kullback-Leibler and Hellinger distances, provide alternative forms for deriving the revised Wishart and Bartlett distances, respectively. Based in a reparametrization proposed by Lee *et al.* [14], we also propose contrast measures which depend on the complex correlation coefficient.

Besides advancing the comparison of samples by means of their covariance matrices, the proposed distances are a venue for contrasting images rendered by different numbers of looks.

Considering the hypothesis test methodology proposed by Salicrú *et al.* [15], the derived distances are multiplied by a coefficient which involves the sizes of two samples of PolSAR images. The asymptotic behavior of the resulting quantities is studied.

In order to quantify the sensitivity of the distances, we perform Monte Carlo experiments in several possible scenarios. We illustrate the behavior of these distances and their associated hypothesis tests with actual data.

This paper unfolds as follows. Section II presents the scaled and the relaxed complex Wishart distributions and estimators for their parameters. Section III recalls the background of stochastic dissimilarities. Section IV presents the analytic expressions of distances between Wishart models. Additionally, a new way to derive the Bartlett [5] and the modified Wishart [16] distances is introduced. Section V illustrates the application of these distances in PolSAR image discrimination. Section VI concludes the paper.

## II. THE COMPLEX WISHART DISTRIBUTION

PolSAR sensors record intensity and relative phase data which can be presented as complex scattering matrices.

In principle, these matrices consist of four complex elements  $S_{HH}, S_{HV}, S_{VH}$ , and  $S_{VV}$ , where H and V refer to the horizontal and vertical wave polarization states, respectively but, under the conditions of the reciprocity theorem [17], [18], we have that  $S_{HV} = S_{VH}$ . This scenario is realistic when natural targets are considered [11].

In general, we may consider systems with  $p$  polarization elements, which constitute a complex random vector denoted by:

$$\mathbf{y} = (S_1 \ S_2 \ \cdots \ S_p)^t, \quad (1)$$

where the superscript ‘ $t$ ’ indicates vector transposition. In PolSAR image processing,  $\mathbf{y}$  is often admitted to obey the multivariate complex circular Gaussian distribution with zero mean [10], whose probability density function is:

$$f_{\mathbf{y}}(\mathbf{y}; \Sigma) = \frac{1}{\pi^p |\Sigma|} \exp(-\mathbf{y}^* \Sigma^{-1} \mathbf{y}),$$

where  $|\cdot|$  is the determinant of a matrix or the absolute value of a scalar, the superscript ‘ $*$ ’ denotes the complex conjugate transpose of a vector,  $\Sigma$  is the covariance matrix of  $\mathbf{y}$  given by

$$\Sigma = E(\mathbf{y}\mathbf{y}^*) = \begin{bmatrix} E(S_1 S_1^*) & E(S_1 S_2^*) & \cdots & E(S_1 S_p^*) \\ E(S_2 S_1^*) & E(S_2 S_2^*) & \cdots & E(S_2 S_p^*) \\ \vdots & \vdots & \ddots & \vdots \\ E(S_p S_1^*) & E(S_p S_2^*) & \cdots & E(S_p S_p^*) \end{bmatrix},$$

and  $E\{\cdot\}$  is the statistical expectation operator. Besides being Hermitian and positive definite, the covariance matrix  $\Sigma$  contains all the necessary information to characterize the backscattered data under analysis [2].

In order to enhance the signal-to-noise ratio,  $L$  independent and identically distributed (iid) samples are usually averaged in order to form the  $L$ -looks covariance matrix [19]:

$$\mathbf{Z} = \frac{1}{L} \sum_{i=1}^L \mathbf{y}_i \mathbf{y}_i^*,$$

where  $\mathbf{y}_i$ ,  $i = 1, 2, \dots, L$ , are realizations of (1). Under the aforementioned hypotheses,  $\mathbf{Z}$  follows a scaled complex Wishart distribution. Having  $\Sigma$  and  $L$  as parameters, such law is characterized by the following probability density function:

$$f_{\mathbf{Z}}(\mathbf{Z}; \Sigma, L) = \frac{L^p |\Sigma| |\mathbf{Z}|^{L-p}}{|\Sigma|^L \Gamma_p(L)} \exp(-L \operatorname{tr}(\Sigma^{-1} \mathbf{Z})), \quad (2)$$

where  $\Gamma_p(L) = \pi^{p(p-1)/2} \prod_{i=0}^{p-1} \Gamma(L-i)$ ,  $L \geq p$ ,  $\Gamma(\cdot)$  is the gamma function, and  $\operatorname{tr}(\cdot)$  is the trace operator. This situation is denoted  $\mathbf{Z} \sim \mathcal{W}(\Sigma, L)$ , and this distribution satisfies  $E\{\mathbf{Z}\} = \Sigma$ , which is a Hermitian positive definite matrix [19]. In practice,  $L$  is treated as a parameter and must be estimated. In [20], Anfinson *et al.*

removed the restriction  $L \geq p$ . The resulting distribution has the same form as in (2) and is termed the relaxed Wishart distribution denoted as  $\mathcal{W}_{\mathcal{R}}(\Sigma, n)$ . This model accepts variations of  $n$  along the image. Henceforth, the scaled complex Wishart law  $\mathcal{W}_{\mathcal{R}}(\Sigma, n)$  will represent its relaxed version.

Due to its optimal asymptotic properties, the maximum likelihood (ML) estimation is employed to estimate  $\Sigma$  [21] and  $n$  [19]. Let  $\{\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_N\}$  be a random sample of size  $N$  obeying the  $\mathcal{W}_{\mathcal{R}}(\Sigma, n)$  distribution. If (i) it is assumed that the parameter  $n$  is a known quantity and (ii) the profile likelihood of  $f_{\mathbf{Z}}$  is considered in terms of  $\Sigma$ , we establish the following estimator for  $\Sigma$  [21]:

$$\hat{\Sigma} = \frac{1}{N} \sum_{i=1}^N \mathbf{Z}_i.$$

Deriving the profile likelihood from (2) with respect to  $n$  we obtain:

$$\begin{aligned} \frac{\partial}{\partial n} \ln[f_{\mathbf{Z}}(\mathbf{Z}; \hat{\Sigma}, n)] &= p[\log(n) + 1] + \log \frac{|\mathbf{Z}|}{|\hat{\Sigma}|} \\ &- \text{tr}(\hat{\Sigma}^{-1} \mathbf{Z}) - \sum_{i=0}^{p-1} \psi^{(0)}(n - i), \quad (3) \end{aligned}$$

where  $\psi^{(0)}(\cdot)$  is the digamma function [22, p. 258]. Thus, the solution of above nonlinear equation provides the ML estimator for  $n$ . Several estimation methods for  $n$  are discussed in [19].

Fig. 1 presents a polarimetric SAR image obtained by an E-SAR sensor over surroundings of Weßling, Germany. The informed (nominal) number of looks is 3. The area exhibits two distinct types of target roughness: (i) homogeneous (pasture) and (ii) heterogeneous (forest). Table I lists the resulting ML parameter estimates, as well as the sample sizes. The estimates of  $n$  are smaller than number of channels,  $p = 3$ , due to the pronounced presence of texture. This fact suggests the  $\mathcal{W}(\Sigma, 3)$  distribution as an inappropriate statistical model.

TABLE I  
PARAMETER ESTIMATES

| Regions   | $\hat{n}$ | $ \hat{\Sigma} $ | # pixels |
|-----------|-----------|------------------|----------|
| pasture-1 | 1.005     | 784050.800       | 2106     |
| pasture-2 | 1.121     | 143.437          | 2352     |
| pasture-3 | 1.000     | 190.130          | 3340     |
| forest-1  | 1.003     | 3854056.000      | 2870     |
| forest-2  | 2.480     | 10752870.000     | 1020     |
| forest-3  | 1.050     | 14568.150        | 2496     |

Fig. 2 depicts the empirical densities of data samples from the selected forest and pasture regions. Additionally, the associated fitted marginal densities  $\mathcal{W}_{\mathcal{R}}(\hat{\Sigma}, \hat{n})$



Fig. 1. E-SAR image (HH channel) with selected regions.

TABLE II  
AIC VALUES FOR THE HH CHANNEL FROM RELAXED AND ORIGINAL WISHART DISTRIBUTIONS ON REAL DATA

| Regions   | AIC                         |               | $\frac{\text{AIC}(\mathcal{W}_{\mathcal{R}})}{\text{AIC}(\mathcal{W})}$ |
|-----------|-----------------------------|---------------|-------------------------------------------------------------------------|
|           | $\mathcal{W}_{\mathcal{R}}$ | $\mathcal{W}$ |                                                                         |
| pasture-1 | 23881.52                    | 27454.97      | 0.87                                                                    |
| pasture-2 | 14885.62                    | 17345.38      | 0.86                                                                    |
| pasture-3 | 21998.75                    | 22112.37      | 0.99                                                                    |
| forest-1  | 36823.14                    | 40464.83      | 0.91                                                                    |
| forest-2  | 14108.18                    | 14481.67      | 0.97                                                                    |
| forest-3  | 22761.28                    | 25630.94      | 0.89                                                                    |

and  $\mathcal{W}(\hat{\Sigma}, 3)$  are displayed for comparison. In this case, the scaled Wishart density collapses to a gamma density as demonstrated in [23]:

$$f_{Z_i}(Z'_i; n/\sigma_i^2, n) = \frac{n^n Z_i'^{n-1}}{\sigma_i^{2n} \Gamma(n)} \exp(-n Z'_i / \sigma_i^2),$$

for  $i \in \{\text{HH}, \text{HV}, \text{VV}\}$ , where  $\sigma_i^2$  is the  $(i, i)$  entry of  $\Sigma$ , and  $Z'_i$  is the  $(i, i)$  entry of the random matrix  $\mathbf{Z}$ . In order to quantitatively compare the data fittings, we obtained the Akaike information criterion (AIC) values. Table II presents the AICs for two elements of the family of scaled Wishart distributions:  $\mathcal{W}_{\mathcal{R}}(\hat{\Sigma}, \hat{n})$  and  $\mathcal{W}(\hat{\Sigma}, 3)$ . In all cases, the  $\mathcal{W}_{\mathcal{R}}$  distribution presented the best fit. These samples will be used to validate our proposed methods in Section IV-E.

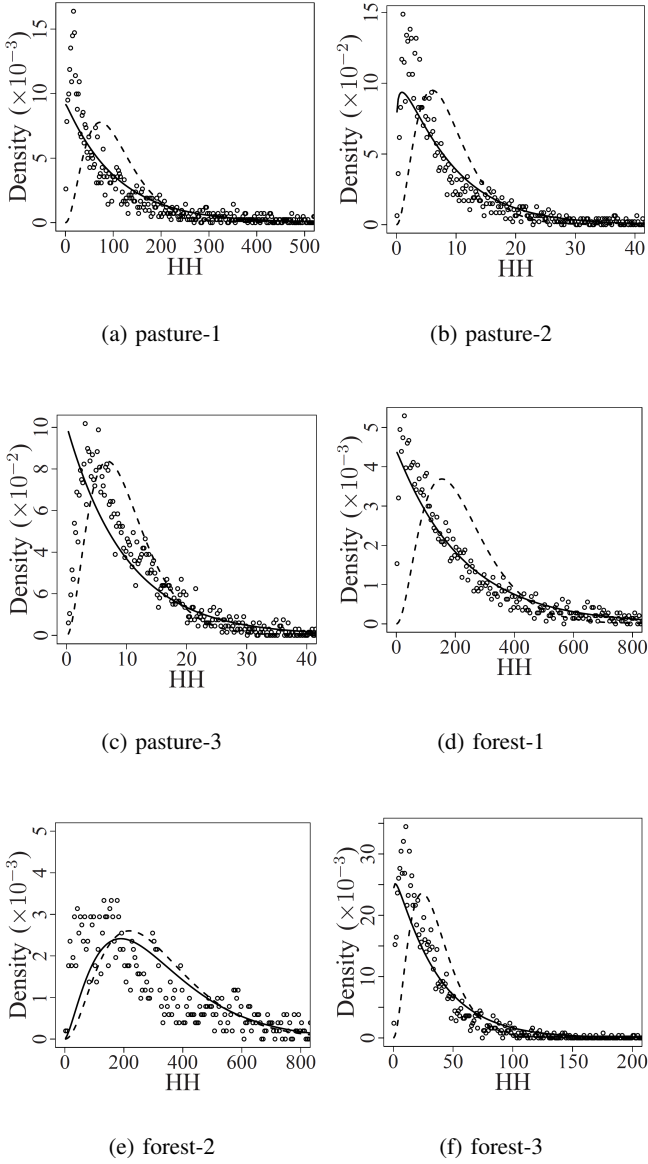


Fig. 2. Histograms and empirical relaxed (solid curve) and original (dashed curve) densities of samples.

### III. STOCHASTIC DISSIMILARITIES

In the following we adhere to the convention that a ‘divergence’ is any non-negative function between two probability measures which obeys the identity of definiteness property [24, ch. 11, p. 328]. If the function is also symmetric, it is called a ‘distance’. Finally, we understand ‘metric’ as a distance which also satisfies the triangular inequality [25, ch. 1 and 14].

An image can be understood as a set of regions, in which the enclosed pixels are observations of random variables following a certain distribution. Therefore, stochastic dissimilarity measures can be used as image processing tools, since they may be able to assess the dif-

ference between the distributions that describe different image areas [12]. Dissimilarity measures were submitted to a systematic and comprehensive treatment in [26], [27] and, as a result, the class of  $(h, \phi)$ -divergences was proposed [15].

Assume that  $\mathbf{X}$  and  $\mathbf{Y}$  are random matrices associated with densities  $f_{\mathbf{X}}(Z; \theta_1)$  and  $f_{\mathbf{Y}}(Z; \theta_2)$ , respectively, where  $\theta_1$  and  $\theta_2$  are parameter vectors. The densities are assumed to share a common support given by the cone of Hermitian positive definite matrices  $\mathcal{A}$ . The  $(h, \phi)$ -divergence between  $f_{\mathbf{X}}$  and  $f_{\mathbf{Y}}$  is defined by

$$D_{\phi}^h(\mathbf{X}, \mathbf{Y}) = h \left( \int_{\mathcal{A}} \phi \left( \frac{f_{\mathbf{X}}(Z; \theta_1)}{f_{\mathbf{Y}}(Z; \theta_2)} \right) f_{\mathbf{Y}}(Z; \theta_2) dZ \right), \quad (4)$$

where  $h: (0, \infty) \rightarrow [0, \infty)$  is a strictly increasing function with  $h(0) = 0$  and  $\phi: (0, \infty) \rightarrow [0, \infty)$  is a convex function such that  $0 \phi(0/0) = 0$  and  $0 \phi(x/0) = \lim_{x \rightarrow \infty} \phi(x)/x$ . The differential element  $dZ$  is given by

$$dZ = dZ_{11} dZ_{22} \cdots dZ_{pp} \prod_{\substack{i, j = 1 \\ i < j}}^p d\Re\{Z_{ij}\} d\Im\{Z_{ij}\},$$

where  $Z_{ij}$  is the  $(i, j)$  entry of matrix  $Z$ , and operators  $\Re\{\cdot\}$  and  $\Im\{\cdot\}$  return real and imaginary parts of their arguments, respectively [10].

Well-known divergences arise with adequate choices of  $h$  and  $\phi$ . Among them, the following were examined: (i) the Kullback-Leibler divergence [28], (ii) the Rényi divergence, (iii) the Bhattacharyya distance [29], and (iv) the Hellinger distance [12]. As the triangular inequality is not necessarily satisfied, not every divergence measure is a metric [30]. Additionally, the symmetry property is not followed by some of these divergence measures. Nevertheless, such tools are mathematically appropriate for comparing the distribution of random variables [31]. Thus, the following expression has been suggested as a possible solution for this problem [28]:

$$d_{\phi}^h(\mathbf{X}, \mathbf{Y}) = \frac{D_{\phi}^h(\mathbf{X}, \mathbf{Y}) + D_{\phi}^h(\mathbf{Y}, \mathbf{X})}{2}.$$

Functions  $d_{\phi}^h: \mathcal{A} \times \mathcal{A} \rightarrow \mathbb{R}$  are distances over  $\mathcal{A}$  since, for all  $\mathbf{X}, \mathbf{Y} \in \mathcal{A}$ , the following properties hold:

- 1)  $d_{\phi}^h(\mathbf{X}, \mathbf{Y}) \geq 0$  (Non-negativity).
- 2)  $d_{\phi}^h(\mathbf{X}, \mathbf{Y}) = d_{\phi}^h(\mathbf{Y}, \mathbf{X})$  (Symmetry).
- 3)  $d_{\phi}^h(\mathbf{X}, \mathbf{Y}) = 0 \Leftrightarrow \mathbf{X} = \mathbf{Y}$  (Definiteness).

Table III shows the functions  $h$  and  $\phi$  which lead to the distances considered in this work.

In the following we discuss integral expressions of these  $(h, \phi)$ -distances. For simplicity, we suppress the



TABLE III  
( $h, \phi$ )-DISTANCES AND THEIR FUNCTIONS

| ( $h, \phi$ )-distance | $h(y)$                                                               | $\phi(x)$                                                                  |
|------------------------|----------------------------------------------------------------------|----------------------------------------------------------------------------|
| Kullback-Leibler       | $y/2$                                                                | $(x-1) \log x$                                                             |
| Rényi (order $\beta$ ) | $\frac{1}{\beta-1} \log((\beta-1)y+1), 0 \leq y < \frac{1}{1-\beta}$ | $\frac{x^{1-\beta} + x^\beta - \beta(x-1) - 2}{2(\beta-1)}, 0 < \beta < 1$ |
| Bhattacharyya          | $-\log(-y+1), 0 \leq y < 1$                                          | $-\sqrt{x} + \frac{x+1}{2}$                                                |
| Hellinger              | $y/2, 0 \leq y < 2$                                                  | $(\sqrt{x}-1)^2$                                                           |

explicit dependence on  $Z$  and  $(\theta_1, \theta_2)$ , reminding that the integration is with respect to  $Z$  on  $\mathcal{A}$ .

(i) The Kullback-Leibler distance:

$$\begin{aligned} d_{\text{KL}}(\mathbf{X}, \mathbf{Y}) &= \frac{1}{2} [D_{\text{KL}}(\mathbf{X}, \mathbf{Y}) + D_{\text{KL}}(\mathbf{Y}, \mathbf{X})] \\ &= \frac{1}{2} \left[ \int f_{\mathbf{X}} \log \frac{f_{\mathbf{X}}}{f_{\mathbf{Y}}} + \int f_{\mathbf{Y}} \log \frac{f_{\mathbf{Y}}}{f_{\mathbf{X}}} \right] \\ &= \frac{1}{2} \int (f_{\mathbf{X}} - f_{\mathbf{Y}}) \log \frac{f_{\mathbf{X}}}{f_{\mathbf{Y}}}. \end{aligned}$$

The divergence  $D_{\text{KL}}$  has a close relationship with the Neyman-Pearson lemma [32] and its symmetrization has been suggested as a correction form of the Akaike information criterion [28].

(ii) The Rényi distance of order  $\beta$ :

$$\begin{aligned} \widetilde{d}_{\text{R}}^{\beta}(\mathbf{X}, \mathbf{Y}) &= \frac{1}{2} [D_{\text{R}}^{\beta}(\mathbf{X}, \mathbf{Y}) + D_{\text{R}}^{\beta}(\mathbf{Y}, \mathbf{X})] \\ &= \frac{\log \int f_{\mathbf{X}}^{\beta} f_{\mathbf{Y}}^{1-\beta} + \log \int f_{\mathbf{X}}^{1-\beta} f_{\mathbf{Y}}^{\beta}}{2(\beta-1)}, \end{aligned}$$

where  $0 < \beta < 1$ . The divergence  $D_{\text{R}}^{\beta}$  has been used for analysing geometric characteristics with respect to probability laws [33]. By the Fejér inequality [34], we have that

$$\begin{aligned} d_{\text{R}}^{\beta}(\mathbf{X}, \mathbf{Y}) &\triangleq \frac{1}{\beta-1} \log \frac{\int f_{\mathbf{X}}^{\beta} f_{\mathbf{Y}}^{1-\beta} + \int f_{\mathbf{X}}^{1-\beta} f_{\mathbf{Y}}^{\beta}}{2} \\ &\leq \widetilde{d}_{\text{R}}^{\beta}(\mathbf{X}, \mathbf{Y}). \end{aligned}$$

Considering algebraic manipulations with the complex Wishart density, distance  $\widetilde{d}_{\text{R}}^{\beta}$  proves to be more tractable than  $d_{\text{R}}^{\beta}$ . Thus, we select the former for our subsequent analysis.

(iii) The Bhattacharyya distance:

$$d_{\text{B}}(\mathbf{X}, \mathbf{Y}) = -\log \int \sqrt{f_{\mathbf{X}} f_{\mathbf{Y}}}.$$

Goudail *et al.* showed that this distance is an efficient tool for contrast definition in algorithms for image processing [35].

(iv) The Hellinger distance:

$$d_{\text{H}}(\mathbf{X}, \mathbf{Y}) = 1 - \int \sqrt{f_{\mathbf{X}} f_{\mathbf{Y}}}.$$

Estimation methods based on the minimization of  $d_{\text{H}}$  have been successfully employed in the context of stochastic differential equations [36]. This is the only bounded distance among the ones considered in this paper.

When considering the distance between particular cases of the same distribution, only parameters are relevant. In this case, the parameters  $\theta_1$  and  $\theta_2$  replace the random variables  $\mathbf{X}$  and  $\mathbf{Y}$  as arguments of the discussed distances. This notation is in agreement with that of [15].

In the following, the hypothesis test based on stochastic distances proposed by Salicrú *et al.* [15] is introduced. Let  $M$ -point vectors  $\widehat{\theta}_1 = (\widehat{\theta}_{11}, \dots, \widehat{\theta}_{1M})$  and  $\widehat{\theta}_2 = (\widehat{\theta}_{21}, \dots, \widehat{\theta}_{2M})$  be the ML estimators of parameters  $\theta_1$  and  $\theta_2$  based on independent samples of sizes  $N_1$  and  $N_2$ , respectively. Under the regularity conditions discussed in [15, p. 380] the following lemma holds:

*Lemma 1:* If  $\frac{N_1}{N_1+N_2} \xrightarrow[N_1, N_2 \rightarrow \infty]{} \lambda \in (0, 1)$  and  $\theta_1 = \theta_2$ , then

$$S_{\phi}^h(\widehat{\theta}_1, \widehat{\theta}_2) = \frac{2N_1N_2}{N_1+N_2} \frac{d_{\phi}^h(\widehat{\theta}_1, \widehat{\theta}_2)}{h'(0)\phi''(1)} \xrightarrow[N_1, N_2 \rightarrow \infty]{\mathcal{D}} \chi_M^2, \quad (5)$$

where “ $\xrightarrow{\mathcal{D}}$ ” denotes convergence in distribution and  $\chi_M^2$  represents the chi-square distribution with  $M$  degrees of freedom.

Based on Lemma 1, statistical hypothesis tests for the null hypothesis  $\theta_1 = \theta_2$  can be derived in the form of following proposition.

*Proposition 1:* Let  $N_1$  and  $N_2$  be large and  $S_{\phi}^h(\widehat{\theta}_1, \widehat{\theta}_2) = s$ , then the null hypothesis  $\theta_1 = \theta_2$  can be rejected at level  $\alpha$  if  $\Pr(\chi_M^2 > s) \leq \alpha$ .

We denote the statistics based on the Kullback-Leibler, Rényi, Bhattacharyya, and Hellinger distances as  $S_{\text{KL}}$ ,  $S_{\text{R}}^{\beta}$ ,  $S_{\text{B}}$ , and  $S_{\text{H}}$ , respectively.

#### IV. ANALYTIC EXPRESSIONS, SENSITIVITY, INEQUALITIES, DEPENDENCE ON THE COMPLEX CORRELATION COEFFICIENT, AND FINITE SAMPLE SIZE BEHAVIOR

In the following, analytic expressions for the stochastic distances  $d_{\text{KL}}$ ,  $d_{\text{R}}^{\beta}$ ,  $d_{\text{B}}$ , and  $d_{\text{H}}$  between two re-

laxed complex Wishart distributions are derived (Section IV-A). We examine three special cases: (i) the most general situation, i.e.,  $\Sigma_1 \neq \Sigma_2$  and  $n_1 \neq n_2$ , (ii) same equivalent number of looks  $n_1 = n_2 = n$  and different covariance matrices  $\Sigma_1 \neq \Sigma_2$ , and (iii) same covariance matrix  $\Sigma_1 = \Sigma_2$  and different equivalent number of looks  $n_1 \neq n_2$ . Case (ii) is likely to be the most frequently used in practice since it allows the comparison of two possibly different areas from the same image. Case (iii) allows the assessment of a change in distribution due only to multilook processing on the same area.

The sensitivity of the tests based on these distances to variations of parameters is qualitatively assessed and discussed in Section IV-B.

In Section IV-C we derive inequalities which  $\Sigma_1$  and  $\Sigma_2$  must obey. These inequalities lead to the Bartlett and revised Wishart distances in a different and simple way when compared to a well-known method available in literature [37]. Distances are also shown to satisfy scale invariance with respect to  $\Sigma$ .

Section IV-D presents and discusses the distances under the parameterization proposed by Lee *et al.* [14] which employs the complex correlation coefficient among different polarization channels, instead of the covariance matrix.

The performance of the tests for finite size samples is quantified by means of (i) Monte Carlo simulation and (ii) true data analysis in Section IV-E.

#### A. Analytic expressions

##### 1) Kullback-Leibler distance:

Case (i):

$$\begin{aligned} d_{\text{KL}}(\theta_1, \theta_2) = & \frac{n_1 - n_2}{2} \left\{ \log \frac{|\Sigma_1|}{|\Sigma_2|} - p \log \frac{n_1}{n_2} \right. \\ & + p[\psi^{(0)}(n_1 - p + 1) - \psi^{(0)}(n_2 - p + 1)] \\ & + (n_2 - n_1) \sum_{i=1}^{p-1} \frac{i}{(n_1 - i)(n_2 - i)} \left. \right\} \\ & + \frac{\text{tr}(n_2 \Sigma_2^{-1} \Sigma_1 + n_1 \Sigma_1^{-1} \Sigma_2)}{2} - \frac{p(n_1 + n_2)}{2}. \end{aligned} \quad (6)$$

Details of this derivation are given in Appendix A.

Case (ii):

$$d_{\text{KL}}(\theta_1, \theta_2) = n \left[ \frac{\text{tr}(\Sigma_1^{-1} \Sigma_2 + \Sigma_2^{-1} \Sigma_1)}{2} - p \right].$$

This result was also derived by Lee and Bretschneider [38] and applied to real PolSAR data for assessing separability of target classes.

Case (iii):

$$\begin{aligned} d_{\text{KL}}(\theta_1, \theta_2) = & \frac{n_1 - n_2}{2} \left\{ -p \log \frac{n_1}{n_2} \right. \\ & + p[\psi^{(0)}(n_1 - p + 1) - \psi^{(0)}(n_2 - p + 1)] \\ & + (n_2 - n_1) \sum_{i=1}^{p-1} \frac{i}{(n_1 - i)(n_2 - i)} \left. \right\}. \end{aligned}$$

##### 2) Rényi distance of order $0 < \beta < 1$ :

Case (i):

$$d_{\text{R}}^{\beta}(\theta_1, \theta_2) = \frac{1}{\beta - 1} \log \frac{I(\theta_1, \theta_2)}{2}, \quad (7)$$

where

$$\begin{aligned} I(\theta_1, \theta_2) = & \left[ \Gamma(n_1 - p + 1)^p |\Sigma_1|^{n_1} n_1^{-pn_1} \prod_{i=1}^{p-1} (n_1 - i)^i \right]^{-\beta} \\ & \times \left[ \Gamma(n_2 - p + 1)^p |\Sigma_2|^{n_2} n_2^{-pn_2} \prod_{i=1}^{p-1} (n_2 - i)^i \right]^{\beta-1} \\ & \times \Gamma(E_1 - p + 1)^p |\Sigma_{12}|^{E_1} \prod_{i=1}^{p-1} (E_1 - i)^i \\ & + \left[ (\Gamma(n_1 - p + 1)^p |\Sigma_1|^{n_1} n_1^{-pn_1} \prod_{i=1}^{p-1} (n_1 - i)^i)^{\beta-1} \right. \\ & \times \left. \left[ \Gamma(n_2 - p + 1)^p |\Sigma_2|^{n_2} n_2^{-pn_2} \prod_{i=1}^{p-1} (n_2 - i)^i \right]^{-\beta} \right. \\ & \times \left. \Gamma(E_2 - p + 1)^p |\Sigma_{21}|^{E_2} \prod_{i=1}^{p-1} (E_2 - i)^i \right], \end{aligned}$$

where  $E_1 = \beta n_1 + (1 - \beta) n_2$ ,  $E_2 = \beta n_2 + (1 - \beta) n_1$ , and  $\Sigma_{ij} = |(n_i \beta \Sigma_i^{-1} + n_j (1 - \beta) \Sigma_j^{-1})^{-1}|$  for  $(i, j) = (1, 2), (2, 1)$ .

Case (ii):

$$\begin{aligned} d_{\text{R}}^{\beta}(\theta_1, \theta_2) = & \frac{\log 2}{1 - \beta} + \frac{1}{\beta - 1} \log \left\{ \right. \\ & \left[ |\Sigma_1|^{-\beta} |\Sigma_2|^{(\beta-1)} |(\beta \Sigma_1^{-1} + (1 - \beta) \Sigma_2^{-1})^{-1}|^n \right. \\ & \left. + \left[ |\Sigma_1|^{(\beta-1)} |\Sigma_2|^{-\beta} |(\beta \Sigma_2^{-1} + (1 - \beta) \Sigma_1^{-1})^{-1}|^n \right] \right\}. \end{aligned}$$

Case (iii):

$$\begin{aligned}
d_R^\beta(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = & \frac{\log 2}{1-\beta} + \frac{1}{\beta-1} \log \left\{ \right. \\
& \left[ \Gamma(n_1 - p + 1)^p n_1^{-pn_1} \prod_{i=1}^{p-1} (n_1 - i)^i \right]^{-\beta} \\
& \times \left[ \Gamma(n_2 - p + 1)^p n_2^{-pn_2} \prod_{i=1}^{p-1} (n_2 - i)^i \right]^{\beta-1} \\
& \times \Gamma(E_1 - p + 1)^p E_1^{-pE_1} \prod_{i=1}^{p-1} (E_1 - i)^i \\
& + \left[ \Gamma(n_1 - p + 1)^p n_1^{-pn_1} \prod_{i=1}^{p-1} (n_1 - i)^i \right]^{\beta-1} \\
& \times \left[ \Gamma(n_2 - p + 1)^p n_2^{-pn_2} \prod_{i=1}^{p-1} (n_2 - i)^i \right]^{-\beta} \\
& \times \Gamma(E_2 - p + 1)^p E_2^{-pE_2} \prod_{i=1}^{p-1} (E_2 - i)^i \left. \right\}
\end{aligned}$$

3) *Bhattacharyya distance*:

Case (i):

$$\begin{aligned}
d_B(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = & \sum_{k=0}^{p-1} \log \frac{\sqrt{\Gamma(n_1 - k)\Gamma(n_2 - k)}}{\Gamma(\frac{n_1+n_2}{2} - k)} \\
& + \frac{n_1 \log |\boldsymbol{\Sigma}_1|}{2} + \frac{n_2 \log |\boldsymbol{\Sigma}_2|}{2} \\
& - \frac{p}{2} (n_1 \log n_1 + n_2 \log n_2) \\
& - \frac{n_1 + n_2}{2} \log \left| \left( \frac{n_1 \boldsymbol{\Sigma}_1^{-1} + n_2 \boldsymbol{\Sigma}_2^{-1}}{2} \right)^{-1} \right|.
\end{aligned} \tag{8}$$

Case (ii):

$$\begin{aligned}
d_B(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = & n \left[ \frac{\log |\boldsymbol{\Sigma}_1| + \log |\boldsymbol{\Sigma}_2|}{2} \right. \\
& \left. - \log \left| \left( \frac{\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1}}{2} \right)^{-1} \right| \right].
\end{aligned}$$

Case (iii):

$$\begin{aligned}
d_B(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = & \sum_{k=0}^{p-1} \log \frac{\sqrt{\Gamma(n_1 - k)\Gamma(n_2 - k)}}{\Gamma(\frac{n_1+n_2}{2} - k)} \\
& - \frac{p}{2} (n_1 \log n_1 + n_2 \log n_2) \\
& + p \frac{n_1 + n_2}{2} \log \frac{n_1 + n_2}{2}.
\end{aligned}$$

4) *Hellinger distance*:

Case (i):

$$\begin{aligned}
d_H(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = & 1 - \prod_{k=0}^{p-1} \frac{\Gamma(\frac{n_1+n_2}{2} - k)}{\sqrt{\Gamma(n_1 - k)\Gamma(n_2 - k)}} \\
& \times \frac{|2^{-1}(n_1 \boldsymbol{\Sigma}_1^{-1} + n_2 \boldsymbol{\Sigma}_2^{-1})^{-1}|^{\frac{n_1+n_2}{2}}}{|\boldsymbol{\Sigma}_1|^{\frac{n_1}{2}} |\boldsymbol{\Sigma}_2|^{\frac{n_2}{2}}} \sqrt{n_1^{pn_1} n_2^{pn_2}}.
\end{aligned} \tag{9}$$

Case (ii):

$$d_H(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = 1 - \left[ \frac{|2^{-1}(\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1})^{-1}|}{\sqrt{|\boldsymbol{\Sigma}_1| |\boldsymbol{\Sigma}_2|}} \right]^n.$$

Case (iii):

$$\begin{aligned}
d_H(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = & 1 - \prod_{k=0}^{p-1} \frac{\Gamma(\frac{n_1+n_2}{2} - k)}{\sqrt{\Gamma(n_1 - k)\Gamma(n_2 - k)}} \\
& \times \left( \frac{n_1 + n_2}{2} \right)^{-p \frac{n_1+n_2}{2}} \sqrt{n_1^{pn_1} n_2^{pn_2}}.
\end{aligned}$$

## B. Sensitivity analysis

Now we examine the behavior of the statistics presented in Lemma 1 with respect to parameter variations, i.e., under the alternative hypotheses. These statistics are directly comparable since they all have the same asymptotic distribution; we used  $N_1 = N_2 = 100$ . Two simple alternative hypotheses are illustrated: changes in an entry in the diagonal of the covariance matrix, and changes in the number of looks.

Firstly, we assumed  $n = 8$  looks,  $\boldsymbol{\theta}_1 = (\boldsymbol{\Sigma}(360932), 8)$  and  $\boldsymbol{\theta}_2 = (\boldsymbol{\Sigma}(x), 8)$ , where

$$\boldsymbol{\Sigma}(x) = \begin{bmatrix} x & 11050 + 3759\mathbf{i} & 63896 + 1581\mathbf{i} \\ & 98960 & 6593 + 6868\mathbf{i} \\ & & 208843 \end{bmatrix}. \tag{10}$$

Since the covariance matrix is Hermitian, only the upper triangle and the diagonal are displayed. The fixed covariance matrix  $\boldsymbol{\Sigma}(360932)$  was analyzed in [6] in PolSAR data of forested areas.

Fig. 3(a) shows the statistics for  $x \in [160000, 560000]$ . They present roughly the same behavior.

Secondly, we considered fixed covariance matrices with varying equivalent number of looks:  $\boldsymbol{\theta}_1 = (\boldsymbol{\Sigma}(360932), 8)$  and  $\boldsymbol{\theta}_2 = (\boldsymbol{\Sigma}(360932), m)$ , for  $3 \leq m \leq 13$ . Fig. 3(b) shows the statistics. It is noticeable that the test statistics are steeper to the left of the minimum. This is in accordance with the notion that small values of the number of looks significantly affect the analysis [19]. In other words, the difference between  $\mathcal{W}_R(\boldsymbol{\Sigma}, n)$  and  $\mathcal{W}_R(\boldsymbol{\Sigma}, kn)$ , for any fixed  $k > 1$  and any  $\boldsymbol{\Sigma}$ , becomes smaller when  $n$  increases.

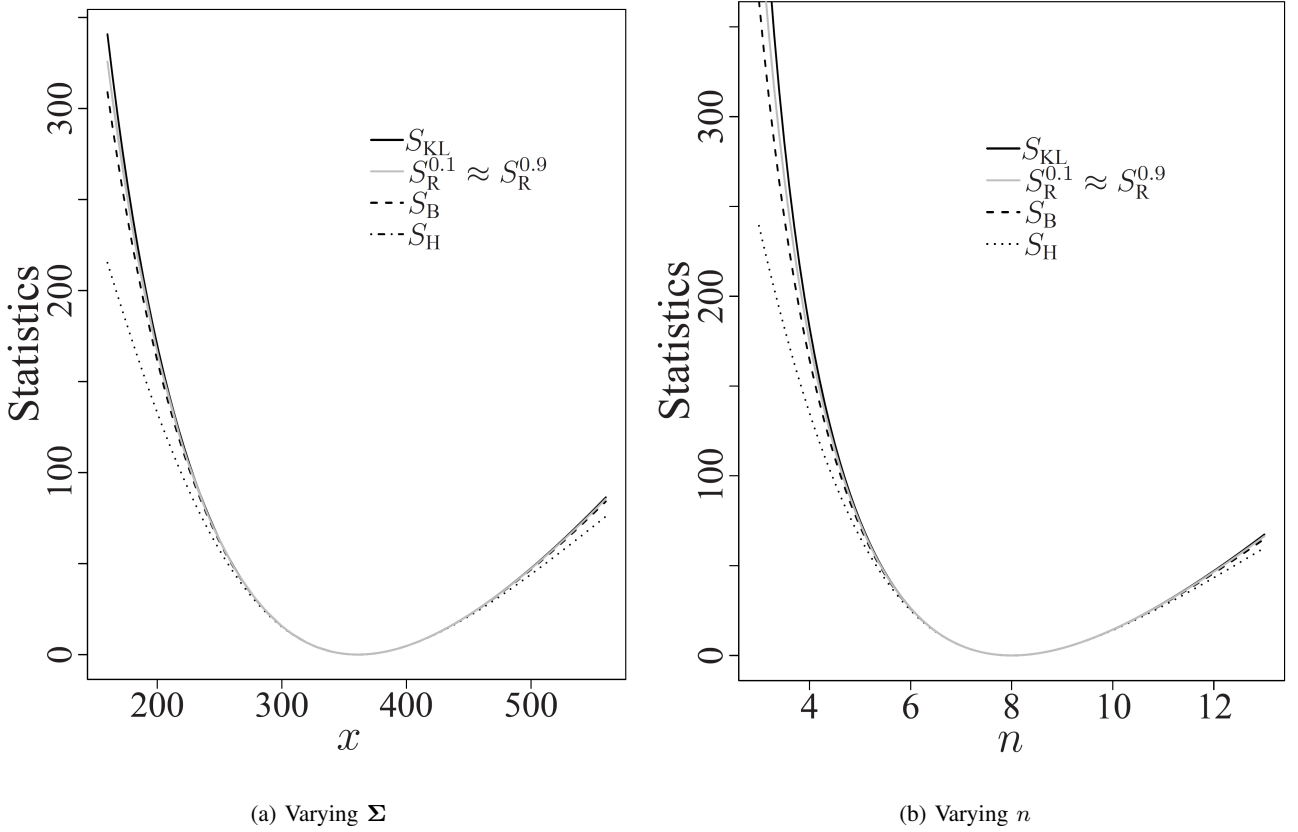


Fig. 3. Sensitivity of statistics.

### C. Invariance and inequalities

The derived distances are invariant under scalings of the covariance matrix  $\Sigma$ . In fact, it can be shown that

$$d_{\mathcal{M}}[(a\Sigma_1, n_1), (a\Sigma_2, n_2)] = d_{\mathcal{M}}[(\Sigma_1, n_1), (\Sigma_2, n_2)],$$

where  $a$  is a positive real value and  $\mathcal{M} \in \{\text{KL}, \text{R}, \text{B}, \text{H}\}$ . This fact stems directly from the mathematical definition of these distances.

De Maio and Alfano [39] derived a new estimator for the covariance matrix under the complex Wishart model using inequalities relating the sought parameters. In the following we derive new inequalities for this model. Due to the major role of the covariance matrix in polarimetry [11], we limit our analysis to inequalities that depend on  $\Sigma$ .

Case (ii) described in previous subsections paved the way for the new inequalities. The following results stem from the nonnegativity of the four distances:

$$\text{tr}(\Sigma_2^{-1}\Sigma_1 + \Sigma_1^{-1}\Sigma_2) \geq 2p, \quad (11)$$

$$\begin{aligned} & \left(\frac{|\Sigma_2|}{|\Sigma_1|}\right)^{n\beta} |\Sigma_2|^{-n} |(\beta\Sigma_1^{-1} + (1-\beta)\Sigma_2^{-1})^{-1}|^n \\ & + \left(\frac{|\Sigma_1|}{|\Sigma_2|}\right)^{n\beta} |\Sigma_1|^{-n} |(\beta\Sigma_2^{-1} + (1-\beta)\Sigma_1^{-1})^{-1}|^n \leq 2, \end{aligned} \quad (12)$$

$$\frac{\log |\Sigma_1| + \log |\Sigma_2|}{2} \geq \log \left| \left( \frac{\Sigma_1^{-1} + \Sigma_2^{-1}}{2} \right)^{-1} \right|, \quad (13)$$

and

$$\sqrt{|\Sigma_1||\Sigma_2|} \geq \left| \left( \frac{\Sigma_1^{-1} + \Sigma_2^{-1}}{2} \right)^{-1} \right|, \quad (14)$$

respectively. Fixing  $\beta = 1/2$  in (12), we obtain (14) directly; taking the logarithm of both sides of (14) yields (13). This result is justified by the following two relations:

- 1)  $d_{\text{B}}(\theta_1, \theta_2) = d_{\text{R}}^{1/2}(\theta_1, \theta_2)$  and
- 2)  $d_{\text{H}}(\theta_1, \theta_2) = 1 - \exp(-\frac{1}{2}d_{\text{R}}^{1/2}(\theta_1, \theta_2))$ .

The revised Wishart [16] and Bartlett distances [5] can be obtained in a new and simple manner. Indeed, the revised Wishart distance ( $d_{\text{RW}}$ ) can be derived after simple manipulations of inequality (11), yielding:

$$d_{\text{RW}}(\Sigma_1, \Sigma_2) = \frac{\text{tr}(\Sigma_1\Sigma_2^{-1} + \Sigma_2\Sigma_1^{-1})}{2} - p \geq 0.$$

The Bartlett distance arises after taking the logarithm of both sides of inequality (14). Straightforward algebra leads to:

$$\ln \frac{|\mathbf{\Sigma}_1 + \mathbf{\Sigma}_2|^2}{|\mathbf{\Sigma}_1||\mathbf{\Sigma}_2|} - 2p \ln 2 \geq 0.$$

The rightmost term in the inequality above is referred to as the Bartlett distance [5], [16].

#### D. Dependence on the complex correlation coefficient

In PolSAR image processing, it is important to quantify the influence of the complex correlation coefficient between channels upon classification algorithms [40]. For such assessment, Lee *et al.* [14] presented a reparametrization of a particular case of the complex Wishart distribution based on a two-element scattering vector  $\mathbf{y}^{(k)}$  at the  $k$ th look, for which the covariance matrix is written as

$$\mathbf{Z} = \begin{bmatrix} z_{11} & \alpha e^{i\delta} \\ \alpha e^{-i\delta} & z_{22} \end{bmatrix},$$

where  $\alpha$  and  $\delta$  are the multilook magnitude and phase, respectively, and  $z_{ii} = n^{-1} \sum_{k=1}^n y_i^{(k)} y_i^{(k)*}$ , for  $i = 1, 2$ . The resulting covariance matrix is

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_{11} & \sqrt{\sigma_{11}\sigma_{22}}|\rho_c|e^{i\delta} \\ \sqrt{\sigma_{11}\sigma_{22}}|\rho_c|e^{-i\delta} & \sigma_{22} \end{bmatrix}, \quad (15)$$

where  $\sigma_{ii} = E(z_{ii})$  and  $\rho_c$  is the complex correlation coefficient between  $z_{11}$  and  $z_{22}$ . The normalized quantities

$$B_1 = \frac{z_{11}}{\sigma_{11}}, \quad B_2 = \frac{z_{22}}{\sigma_{22}}, \quad \eta = \frac{\alpha}{\sqrt{\sigma_{11}\sigma_{22}}},$$

and  $\delta$  obey the following joint probability density function:

$$f(B_1, B_2, \eta, \delta; \rho_c, n) = \frac{\eta(B_1 B_2 - \eta^2)^{n-2}}{\pi(1 - |\rho_c|^2)^n \Gamma(n) \Gamma(n-1)} \times \exp\left(-\frac{B_1 + B_2 - 2\eta|\rho_c|\cos(\delta - \theta)}{1 - |\rho_c|^2}\right), \quad (16)$$

where  $\theta$  is the pulse incidence angle, which is assumed to be constant [14].

The presence of the complex correlation coefficient  $\rho_c$  as a model parameter [11], [14] offers a venue for the proposition of divergence measures based on such parameter.

Consider two random vectors distributed according to Eq. (16) with parameters  $\boldsymbol{\theta}_1 = (\rho_1, n_1)$  and  $\boldsymbol{\theta}_2 = (\rho_2, n_2)$ . Invoking the most general form of the distances as in Section IV, we derived closed form expressions for the considered distances. It is known that normalization procedures often seek transformed polarization channels with variance equal to one. Thus, when  $\sigma_{11} = \sigma_{22} = 1$

in (15) and  $n_1 = n_2 = n$ , the distances are given by Eqs. (20)-(23). Mathematical details are given in Appendix B.

Note that the change of correlation coefficient between polarization channels can be captured when one considers pixels of different regions. Lee *et al.* [41] studied the preservation of polarimetric properties in filtering processes, such as the correlation coefficient between polarization channels which can be obtained by

$$\rho_c = \frac{E(S_{HH}S_{VV}^*)}{\sqrt{E(|S_{HH}|^2)E(|S_{VV}|^2)}}. \quad (17)$$

Using this parameter in Lemma 1, four new hypothesis tests are derived for checking whether two samples, of sizes  $N_1$  and  $N_2$ , come from regions with statistically similar correlation coefficient between channels. The statistics can be also used as confidence regions.

The statistics based on Kullback-Leibler and Hellinger distances present the following confidence regions:

$$R_{KL}(\rho_1, \rho_2) = \left\{ (\rho_1, \rho_2) \in [0, 1] \times [0, 1] : \underbrace{\frac{(1 - |\rho_1||\rho_2|)(2 - |\rho_1|^2 - |\rho_2|^2)}{(1 - |\rho_1|^2)(1 - |\rho_2|^2)}}_{\xi_1} \leq \underbrace{\frac{(N_1 + N_2)}{2nN_1N_2} \chi_1^2(\alpha) + 2}_{t_{KL}} \right\} \quad (18)$$

and

$$R_H(\rho_1, \rho_2) = \left\{ (\rho_1, \rho_2) \in [0, 1] \times [0, 1] : \underbrace{\frac{\sqrt{(1 - |\rho_1|^2)(1 - |\rho_2|^2)}}{4 - (|\rho_1| + |\rho_2|)^2}}_{\xi_2} \geq \underbrace{\frac{1}{4} \left[ 1 - \frac{(N_1 + N_2)}{2nN_1N_2} \chi_1^2(\alpha) \right]^{1/n}}_{t_H} \right\}, \quad (19)$$

where  $\alpha$  is the specified nominal level.

Given two polarimetric data samples, their correlation coefficients can be estimated according to Eq. (17). Quantities  $\hat{\xi}_1$ ,  $\hat{\xi}_2$ ,  $t_{KL}$ , and  $t_H$  are evaluated by means of (18)-(19). We then propose the following decision rules:

- Kullback-Leibler criterion: If  $\hat{\xi}_1 \leq t_{KL}$ , then we have statistical evidence that the samples come from populations with similar correlation between HH and VV channels.
- Hellinger criterion: If  $\hat{\xi}_2 \geq t_H$ , the samples present equivalent correlation between such channels.

Thus, regions  $R_{KL}(\rho_1, \rho_2)$  and  $R_H(\rho_1, \rho_2)$  along with (17) are methods for comparing two regions based on their correlation coefficients.

In order to illustrate this methodology, consider the samples highlighted in Fig. 4 (which is extracted from the ESAR image shown in Fig. 1). Table IV lists the

$$d_{\text{KL}}(\theta_1, \theta_2) = n \left[ \frac{(1 - |\rho_1| |\rho_2|)(2 - |\rho_1|^2 - |\rho_2|^2)}{(1 - |\rho_1|^2)(1 - |\rho_2|^2)} - 2 \right] \quad (20)$$

$$d_{\text{R}}^\beta(\theta_1, \theta_2) = \frac{\log 2}{1 - \beta} + \frac{1}{\beta - 1} \log \left\{ \left[ \frac{(1 - |\rho_1|^2)^{1-\beta}(1 - |\rho_2|^2)^\beta}{1 - \{|\rho_1| - \beta(|\rho_1| - |\rho_2|)\}^2} \right]^n + \left[ \frac{(1 - |\rho_2|^2)^{1-\beta}(1 - |\rho_1|^2)^\beta}{1 - \{|\rho_2| - \beta(|\rho_2| - |\rho_1|)\}^2} \right]^n \right\} \quad (21)$$

$$d_{\text{B}}(\theta_1, \theta_2) = n \left\{ \frac{\log(1 - |\rho_1|^2) + \log(1 - |\rho_2|^2)}{2} - 2 \log 2 + \log \left[ \frac{4 - (|\rho_1| + |\rho_2|)^2}{(1 - |\rho_1|^2)(1 - |\rho_2|^2)} \right] \right\} \quad (22)$$

$$d_{\text{H}}(\theta_1, \theta_2) = 1 - \left[ 4 \frac{\sqrt{(1 - |\rho_1|^2)(1 - |\rho_2|^2)}}{4 - (|\rho_1| + |\rho_2|)^2} \right]^n \quad (23)$$

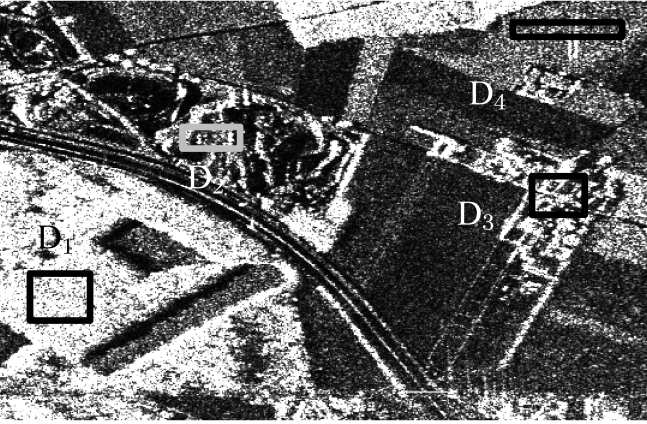


Fig. 4. PolSAR image (HH channel) with four samples.

TABLE IV  
RESULTS BASED ON CONFIDENCE REGIONS

| Regions                         | $\hat{\xi}_1$ | $t_{\text{KL}}$ | Decision        | $\hat{\xi}_2$ | $t_{\text{H}}$ | Decision        |
|---------------------------------|---------------|-----------------|-----------------|---------------|----------------|-----------------|
| D <sub>1</sub> – D <sub>2</sub> | 2.0220        | 2.0016          | <i>Distinct</i> | 0.2486        | 0.2499         | <i>Distinct</i> |
| D <sub>1</sub> – D <sub>3</sub> | 2.0068        | 2.0012          | <i>Distinct</i> | 0.2495        | 0.2499         | <i>Distinct</i> |
| D <sub>1</sub> – D <sub>4</sub> | 2.0734        | 2.0014          | <i>Distinct</i> | 0.2455        | 0.2499         | <i>Distinct</i> |
| D <sub>2</sub> – D <sub>3</sub> | 2.0534        | 2.0018          | <i>Distinct</i> | 0.2467        | 0.2499         | <i>Distinct</i> |
| D <sub>2</sub> – D <sub>4</sub> | 2.0149        | 2.0020          | <i>Distinct</i> | 0.2490        | 0.2498         | <i>Distinct</i> |
| D <sub>3</sub> – D <sub>4</sub> | 2.1254        | 2.0016          | <i>Distinct</i> | 0.2424        | 0.2499         | <i>Distinct</i> |

quantities observed, along with the decisions they led to. Notice that both rules discriminate well.

Additionally, statistics derived from these distances can be used, for instance, to measure the influence of the number of looks on the complex correlation between  $z_{11}$  and  $z_{22}$  in  $\mathbf{Z}$  [11].

Fig. 5 illustrates the resulting statistics for  $|\rho_1|^2 = 0.5$  and  $n_1 = n_2 = n \in \{2, 5\}$  (Figures 5(a) and 5(b), respectively), while  $|\rho_2|$  varies on  $[0, 1]$ .

As expected, the curves have their minimum value, zero, at  $\rho_1 = \rho_2$ . Notice, however, that the curves are steeper on the interval  $(|\rho_1|, 1)$  than on  $(|\rho_1|, 0)$ . Thus for discrimination purposes, the statistics provide better

sensitivity capabilities when  $|\rho_2| > |\rho_1|$ .

The larger the number of looks is, the larger the test statistics are in the alternative hypothesis.

In accordance with the results presented in Fig. 3, all statistics exhibit similar behavior in the alternative hypothesis.

#### E. Finite sample size behavior

We assessed the influence of estimation on the size of the new hypothesis tests using simulated data. To that end, the study was conducted considering the following simulation parameters: number of looks  $n = n_1 = n_2 \in \{4, 8, 16\}$  and the forest covariance matrix shown in (10) with  $x = 360932$ . The sample sizes relate to square windows of size  $7 \times 7$ ,  $11 \times 11$ , and  $20 \times 20$  pixels, i.e.,  $N_1, N_2 \in \{49, 121, 400\}$ . Nominal significance levels  $\alpha \in \{1\%, 5\%\}$  were verified.

Let  $T$  be the number of Monte Carlo replicas and  $C$  the number of cases for which the null hypothesis is rejected at nominal level  $\alpha$ . The empirical test size is given by  $\hat{\alpha}_{1-\alpha} = C/T$ . Following the methodology described in [12], we employed  $T = 5500$  replicas.

Table V presents the empirical test sizes at 1% and 5% nominal levels, the execution time in milliseconds<sup>1</sup>, the test statistic mean ( $\bar{S}$ ) and coefficient of variation (CV). For each case, the best obtained empirical sizes and distance means are in boldface. Results for  $N_1 = 49, N_2 = \{121, 400\}$  and  $N_1 = 121, N_2 = 400$  are consistent with the ones shown, and are omitted for brevity.

We tested ten parameters: nine related to the covariance matrix of order  $p = 3$  and the number of looks, leading to 10 degrees of freedom. Thus, the statistics expected value should converge in probability to 10, as a consequence of the weak law of large numbers since it is asymptotically  $\chi_{10}^2$  distributed. In Table V, notice that

<sup>1</sup>Running on a PC with an Intel Core 2 Duo processor 2.10 GHz, 4GB of RAM, Windows XP, and R 2.8.1

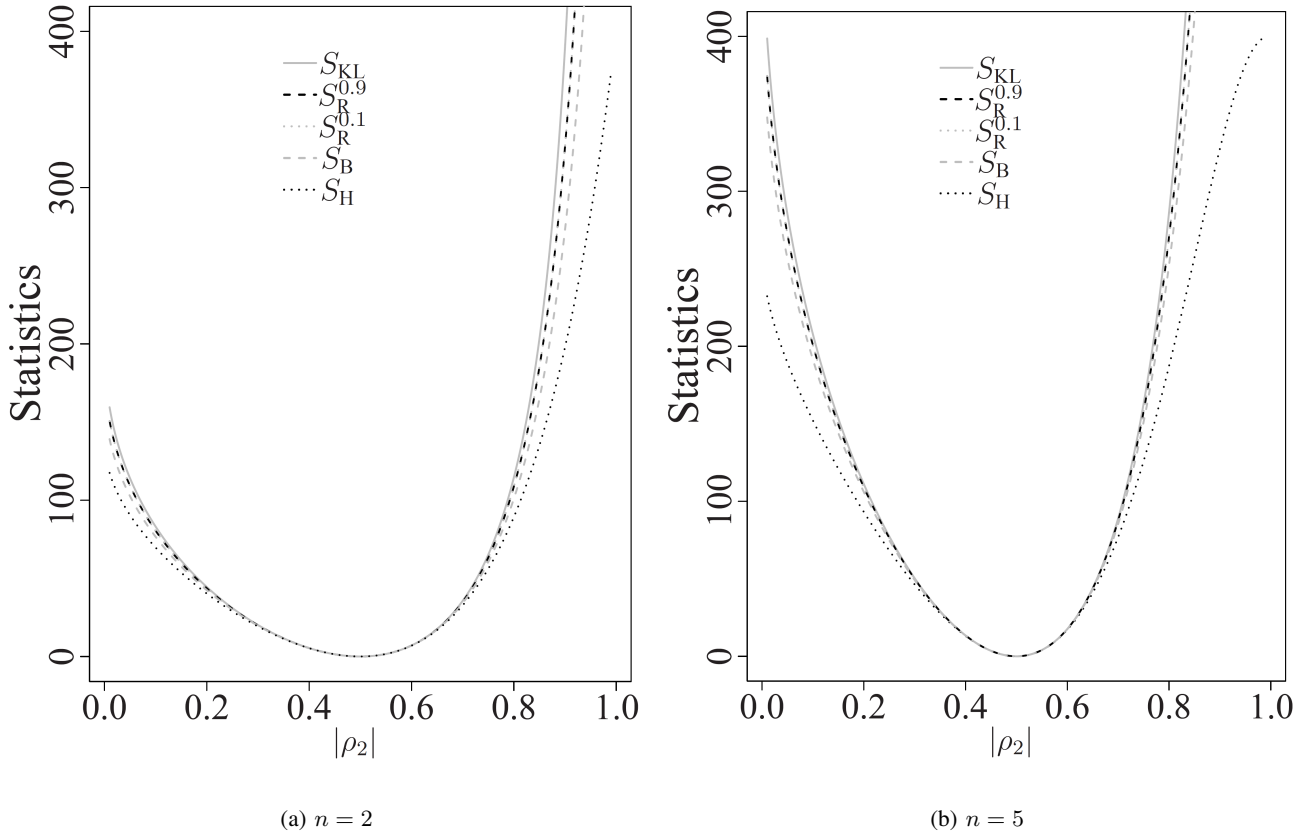


Fig. 5. Distances expressed in (20)-(23) for  $n_1 = n_2 = n = \{2, 5\}$ , and  $|\rho_1|^2 = 0.5$ .

$\bar{S}$  tends to 10 as the sample size increases. Except for the case  $(N_1, N_2, n) = (121, 121, 16)$ , the CVs varied in the interval  $(43.00, 44.98)$  regardless the sample sizes. By fixing the sample size while varying the number of looks  $n$ , test sizes obey the inequalities  $S_H \leq S_B \leq S_R^\beta \leq S_{KL}$ , as illustrated in Fig. 3.

In general terms, proposed hypothesis tests presented good results even for small sample sizes. The Kullback-Leibler-based test presented the best performance in terms of execution time, while the test based on the Bhattacharyya distance showed the best empirical test 55.55% of cases.

The presented methodology for assessing test sizes was also applied to the three forest samples from the E-SAR image shown in Fig. 1. Each sample was submitted to the following procedure [12]:

- (i) split the sample in disjoint blocks of size  $N_1$ ;
- (ii) for each block from (i), split the remaining sample in disjoint blocks of size  $N_2$ ;
- (iii) perform the hypothesis test as described in Proposition 1 for each pair of samples with sizes  $N_1$  and  $N_2$ .

Table VI presents the results, omitting some entries

as in Table V. All test sizes were smaller than the nominal level, i.e., the proposed tests do not reject the null hypothesis when similar samples are considered.

## V. APPLICATIONS

This section presents two applications of the tests based on stochastic distances. Firstly, a discrimination analysis was performed in order to assess the influence of image texture on the tests. It is known that the complex Wishart distribution is more appropriated for describing homogeneous regions. However, other polarimetric distributions, potentially more apt to describing textured areas, yield intractable expressions which depend on special functions, such as the hypergeometric and modified Bessel functions. In order to quantify textures, we considered distances between relaxed scaled complex Wishart laws as proposed by Anfinson *et al.* [20]. Secondly, stochastic distances are embedded into the  $k$ -means method in order to identify groups in PolSAR data. The performance of four distances was assessed by means of a synthetic image generated from the relaxed scaled complex Wishart distribution.

TABLE V  
EMPIRICAL SIZES FOR FOREST-1

| Factors                |       |       | $n = 4$      |              |           |           |        |              | $n = 8$      |           |           |        |              |              | $n = 16$  |           |         |  |  |  |
|------------------------|-------|-------|--------------|--------------|-----------|-----------|--------|--------------|--------------|-----------|-----------|--------|--------------|--------------|-----------|-----------|---------|--|--|--|
| $S_{\phi}^h$           | $N_1$ | $N_2$ | 1%           | 5%           | time (ms) | $\bar{S}$ | CV     | 1%           | 5%           | time (ms) | $\bar{S}$ | CV     | 1%           | 5%           | time (ms) | $\bar{S}$ | CV      |  |  |  |
| $S_{\text{KL}}$        | 49    | 49    | 1.109        | 5.291        | 0.44      | 10.132    | 44.661 | 1.036        | 5.855        | 0.49      | 10.243    | 44.797 | 1.036        | 5.236        | 0.39      | 10.005    | 45.048  |  |  |  |
|                        | 121   | 121   | 1.036        | <b>4.927</b> | 0.43      | 10.041    | 44.849 | <b>0.891</b> | 5.200        | 0.49      | 10.070    | 44.148 | 2.655        | 6.727        | 0.48      | 14.618    | 250.833 |  |  |  |
|                        | 400   | 400   | <b>0.836</b> | 5.164        | 0.42      | 10.026    | 44.483 | 1.036        | 5.273        | 0.50      | 10.022    | 45.312 | 1.036        | 5.036        | 0.46      | 10.091    | 44.069  |  |  |  |
| $S_{\text{R}}^{\beta}$ | 49    | 49    | 1.091        | 5.182        | 0.54      | 10.105    | 44.539 | <b>1.000</b> | 5.782        | 0.51      | 10.222    | 44.710 | <b>1.018</b> | 5.145        | 0.58      | 9.991     | 44.981  |  |  |  |
|                        | 121   | 121   | 1.036        | 4.891        | 0.57      | 10.030    | 44.802 | <b>0.891</b> | 5.164        | 0.61      | 10.061    | 44.111 | 2.655        | 6.727        | 0.56      | 14.385    | 242.443 |  |  |  |
|                        | 400   | 400   | <b>0.836</b> | 5.127        | 0.57      | 10.022    | 44.468 | 1.036        | 5.273        | 0.57      | 10.019    | 45.301 | <b>1.018</b> | <b>5.036</b> | 0.65      | 10.089    | 44.061  |  |  |  |
| $S_{\text{B}}$         | 49    | 49    | <b>1.018</b> | <b>5.036</b> | 1.10      | 10.058    | 44.329 | 0.927        | <b>5.691</b> | 1.12      | 10.186    | 44.561 | 0.945        | <b>4.963</b> | 1.07      | 9.965     | 44.865  |  |  |  |
|                        | 121   | 121   | <b>0.982</b> | 4.763        | 1.11      | 10.011    | 44.719 | 0.855        | <b>5.091</b> | 0.99      | 10.047    | 44.047 | 2.655        | 6.654        | 1.12      | 14.035    | 229.490 |  |  |  |
|                        | 400   | 400   | <b>0.836</b> | <b>5.109</b> | 0.99      | 10.016    | 44.441 | 1.036        | 5.273        | 1.13      | 10.015    | 45.281 | <b>1.018</b> | <b>5.036</b> | 1.20      | 10.086    | 44.048  |  |  |  |
| $S_{\text{H}}$         | 49    | 49    | 0.582        | 3.909        | 1.10      | 9.756     | 43.000 | 0.600        | 4.236        | 1.12      | 9.876     | 43.238 | 0.527        | 3.909        | 1.07      | 9.668     | 43.547  |  |  |  |
|                        | 121   | 121   | 0.709        | 4.200        | 1.11      | 9.888     | 44.172 | 0.655        | 4.600        | 0.99      | 9.924     | 43.505 | <b>2.364</b> | <b>6.218</b> | 1.12      | 12.946    | 190.636 |  |  |  |
|                        | 400   | 400   | 0.818        | 4.945        | 0.99      | 9.979     | 44.277 | <b>1.018</b> | <b>5.091</b> | 1.13      | 9.977     | 45.116 | 0.964        | 4.927        | 1.20      | 10.048    | 43.884  |  |  |  |

TABLE VI  
EMPIRICAL SIZES FOR FORESTS

| Factors       |       |       | forest-1 |      |                                |       | forest-2 |      |                                |       | forest-3 |      |                                |        |
|---------------|-------|-------|----------|------|--------------------------------|-------|----------|------|--------------------------------|-------|----------|------|--------------------------------|--------|
| $S_{\phi}^h$  | $N_1$ | $N_2$ | 1%       | 5%   | $\bar{S}$ ( $\times 10^{-1}$ ) | CV    | 1%       | 5%   | $\bar{S}$ ( $\times 10^{-1}$ ) | CV    | 1%       | 5%   | $\bar{S}$ ( $\times 10^{-1}$ ) | CV     |
| $S_{KL}$      | 49    | 49    | 0.00     | 0.00 | 27.44                          | 65.53 | 0.00     | 0.00 | 25.34                          | 71.71 | 0.82     | 1.88 | 35.55                          | 123.33 |
|               | 121   | 121   | 0.00     | 0.00 | 10.20                          | 52.36 | 0.00     | 0.00 | 13.32                          | 41.37 | 0.00     | 0.00 | 25.53                          | 112.68 |
|               | 400   | 400   | 0.00     | 0.00 | 2.56                           | 38.61 | 0.00     | 0.00 | 4.95                           | —     | 0.00     | 0.00 | 19.50                          | 108.19 |
| $S_R^{\beta}$ | 49    | 49    | 0.00     | 0.00 | 18.39                          | 62.90 | 0.00     | 0.00 | 17.03                          | 68.78 | 0.00     | 0.16 | 22.74                          | 112.75 |
|               | 121   | 121   | 0.00     | 0.00 | 7.03                           | 51.46 | 0.00     | 0.00 | 9.14                           | 40.26 | 0.00     | 0.00 | 16.83                          | 108.80 |
|               | 400   | 400   | 0.00     | 0.00 | 1.78                           | 38.48 | 0.00     | 0.00 | 3.45                           | —     | 0.00     | 0.00 | 13.10                          | 106.62 |
| $S_B$         | 49    | 49    | 0.00     | 0.00 | 18.43                          | 59.37 | 0.00     | 0.00 | 17.13                          | 64.85 | 0.00     | 0.00 | 21.75                          | 102.36 |
|               | 121   | 121   | 0.00     | 0.00 | 7.32                           | 50.04 | 0.00     | 0.00 | 9.47                           | 38.62 | 0.00     | 0.00 | 16.54                          | 103.62 |
|               | 400   | 400   | 0.00     | 0.00 | 1.90                           | 38.26 | 0.00     | 0.00 | 3.65                           | —     | 0.00     | 0.00 | 13.16                          | 104.24 |
| $S_H$         | 49    | 49    | 0.00     | 0.00 | 14.30                          | 46.14 | 0.00     | 0.00 | 13.42                          | 48.14 | 0.00     | 0.00 | 14.62                          | 69.088 |
|               | 121   | 121   | 0.00     | 0.00 | 6.63                           | 45.78 | 0.00     | 0.00 | 8.42                           | 34.53 | 0.00     | 0.00 | 11.91                          | 81.619 |
|               | 400   | 400   | 0.00     | 0.00 | 1.86                           | 37.22 | 0.00     | 0.00 | 3.51                           | —     | 0.00     | 0.00 | 10.06                          | 91.374 |

#### A. Discrimination analysis

AIRSAR was an airborne mission with PolSAR capabilities, designed and built by the Jet Propulsion Laboratory, which operated at P-, L-, and C-bands [42]. Fig. 6 shows a  $550 \times 645$  pixels image (HH channel) of San Francisco recorded by this sensor, acquired with four nominal looks. Nine areas were chosen to represent three different degrees of roughness: homogeneous, heterogeneous, and extremely heterogeneous, labeled as  $A_i$ ,  $B_i$ , and  $C_i$ , respectively, for  $i = 1, 2, 3$ .

The parameters of the complex Wishart distribution were estimated by maximum likelihood, cf. Eq. (3). Table VII presents the estimated number of looks and determinants of the complex covariance matrices for each area, along with the number of observations.

Goodman defined the determinant of the complex covariance matrix as a generalized variance [21] which

is associated with the degree of heterogeneity. As the heterogeneity increases, the elements of the covariance matrices become larger along with the determinant. The most homogeneous region,  $A_1$ , has the covariance matrix with smallest determinant. Sample  $B_1$  has the largest determinant among heterogeneous regions. This suggests that this sample is the most heterogeneous among them, even with the presence of a double bounce [43] in sample  $B_3$ . As expected, urban areas produce the largest determinants since they are extremely heterogeneous targets. The estimated number of looks decreases with the heterogeneity, as expected.

Stochastic distances were computed between pairs of these estimated distributions. Table VIII shows the distances between regions of the same class. In all but one case, the values were found to be ordered as follows:  $d_{KL} > d_R^{0.9} \geq d_B \geq d_H > d_R^{0.1}$ . The only discrepancy



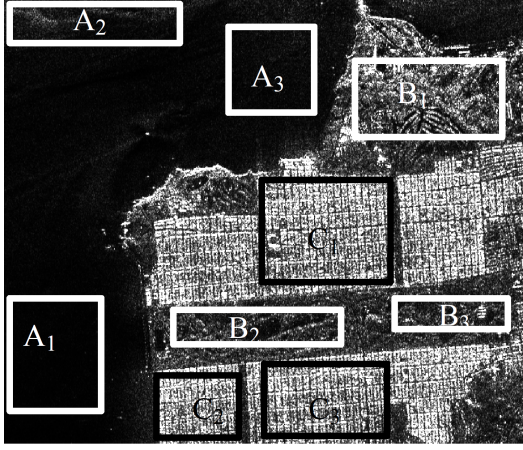


Fig. 6. AIRSAR HH data and samples

TABLE VII  
ESTIMATED NUMBER OF LOOKS AND GENERALIZED VARIANCE

| Estimates Region |       | Subscript of regions   |                        |                        |
|------------------|-------|------------------------|------------------------|------------------------|
|                  |       | $i = 1$                | $i = 2$                | $i = 3$                |
| $\hat{n}$        | $A_i$ | 4.04                   | 3.75                   | 3.98                   |
| $ \hat{\Sigma} $ |       | $3.24 \times 10^{-9}$  | $59.71 \times 10^{-9}$ | $18.83 \times 10^{-9}$ |
| # pixels         |       | 15960                  | 10339                  | 11449                  |
|                  | $B_i$ | 3.05                   | 3.21                   | 3.15                   |
|                  |       | $35.86 \times 10^{-5}$ | $7.38 \times 10^{-5}$  | $10.87 \times 10^{-5}$ |
|                  |       | 17385                  | 9152                   | 5499                   |
|                  | $C_i$ | 3.03                   | 3.12                   | 3.09                   |
|                  |       | $1.97 \times 10^{-3}$  | $1.46 \times 10^{-3}$  | $1.17 \times 10^{-3}$  |
|                  |       | 20320                  | 8034                   | 13770                  |

occurs when comparing homogeneous regions  $A_1$  and  $A_2$ , where the last inequality is not preserved.

TABLE VIII  
DISTANCES BETWEEN REGIONS OF SIMILAR ROUGHNESS

| Regions   | $d_{KL}$ | $d_R^{0.9}$ | $d_B$ | $d_H$ | $d_R^{0.1}$ |
|-----------|----------|-------------|-------|-------|-------------|
| $A_1-A_2$ | 19.83    | 13.13       | 2.66  | 0.93  | 1.46        |
| $A_1-A_3$ | 7.34     | 5.91        | 1.41  | 0.76  | 0.66        |
| $A_2-A_3$ | 2.01     | 1.73        | 0.45  | 0.36  | 0.19        |
| $B_1-B_2$ | 1.83     | 1.58        | 0.41  | 0.34  | 0.18        |
| $B_1-B_3$ | 1.11     | 0.97        | 0.26  | 0.23  | 0.11        |
| $B_2-B_3$ | 0.26     | 0.23        | 0.06  | 0.06  | 0.03        |
| $C_1-C_2$ | 0.35     | 0.31        | 0.09  | 0.08  | 0.03        |
| $C_1-C_3$ | 0.21     | 0.18        | 0.05  | 0.05  | 0.02        |
| $C_2-C_3$ | 0.19     | 0.17        | 0.05  | 0.05  | 0.02        |

Table IX presents the distances between regions of different roughness. Similarly to the univariate case [12], in these cases regions become more distinguishable. In all cases, the distances satisfy

$$d_{\mathcal{D}}[(\hat{\Sigma}_k, \hat{n}), (\hat{\Sigma}_\ell, \hat{n})] < d_{\mathcal{D}}[(\hat{\Sigma}_k, \hat{n}), (\hat{\Sigma}_m, \hat{n})],$$

$$\text{if } ||\hat{\Sigma}_k| - |\hat{\Sigma}_\ell|| < ||\hat{\Sigma}_k| - |\hat{\Sigma}_m||.$$

TABLE IX  
DISTANCES BETWEEN REGIONS OF DIFFERENT ROUGHNESS

| Regions   | $d_{KL}$ | $d_R^{0.9}$ | $d_B$ | $d_H$ | $d_R^{0.1}$ |
|-----------|----------|-------------|-------|-------|-------------|
| $A_1-B_1$ | 621.16   | 91.09       | 12.38 | 1.00  | 10.12       |
| $A_1-B_2$ | 318.09   | 74.13       | 10.61 | 1.00  | 8.24        |
| $A_1-B_3$ | 397.15   | 79.29       | 11.08 | 1.00  | 8.81        |
| $A_2-B_1$ | 222.93   | 60.04       | 8.72  | 1.00  | 6.67        |
| $A_2-B_2$ | 119.27   | 46.15       | 7.23  | 1.00  | 5.13        |
| $A_2-B_3$ | 152.90   | 51.65       | 7.81  | 1.00  | 5.74        |
| $A_3-B_1$ | 361.81   | 72.51       | 10.14 | 1.00  | 8.06        |
| $A_3-B_2$ | 185.07   | 57.19       | 8.50  | 1.00  | 6.35        |
| $A_3-B_3$ | 239.22   | 62.80       | 9.05  | 1.00  | 6.98        |
| $A_1-C_1$ | 1559.32  | 110.64      | 14.65 | 1.00  | 12.29       |
| $A_1-C_2$ | 1469.62  | 109.18      | 14.57 | 1.00  | 12.13       |
| $A_1-C_3$ | 1393.29  | 106.23      | 14.16 | 1.00  | 11.80       |
| $A_2-C_1$ | 621.61   | 77.63       | 10.70 | 1.00  | 8.62        |
| $A_2-C_2$ | 575.70   | 75.85       | 10.56 | 1.00  | 8.43        |
| $A_2-C_3$ | 554.52   | 74.31       | 10.29 | 1.00  | 8.26        |
| $A_3-C_1$ | 1010.30  | 90.89       | 12.26 | 1.00  | 10.10       |
| $A_3-C_2$ | 954.88   | 89.29       | 12.14 | 1.00  | 9.92        |
| $A_3-C_3$ | 907.62   | 87.21       | 11.81 | 1.00  | 9.69        |
| $B_1-C_1$ | 4.76     | 3.91        | 0.95  | 0.61  | 0.43        |
| $B_1-C_2$ | 4.23     | 3.54        | 0.88  | 0.59  | 0.39        |
| $B_1-C_3$ | 3.88     | 3.25        | 0.81  | 0.55  | 0.36        |
| $B_2-C_1$ | 13.41    | 9.50        | 2.02  | 0.87  | 1.06        |
| $B_2-C_2$ | 12.44    | 8.97        | 1.93  | 0.86  | 1.00        |
| $B_2-C_3$ | 11.18    | 8.10        | 1.75  | 0.83  | 0.90        |
| $B_3-C_1$ | 9.42     | 7.24        | 1.65  | 0.81  | 0.80        |
| $B_3-C_2$ | 8.71     | 6.79        | 1.57  | 0.79  | 0.75        |
| $B_3-C_3$ | 7.59     | 5.96        | 1.38  | 0.75  | 0.66        |

As expected, distances between samples of different classes are much larger than those between samples with similar roughness.

### B. Clustering analysis in synthetic and actual images

A common characteristic of segmentation and classification algorithms is their sensitivity to the dissimilarity measure they employ [16], [44]. As already presented, stochastic distances present good discriminatory properties and, therefore, can be used for identifying clusters in PolSAR data. To that end, we use of the  $k$ -means method with these measures applied to synthetic data.

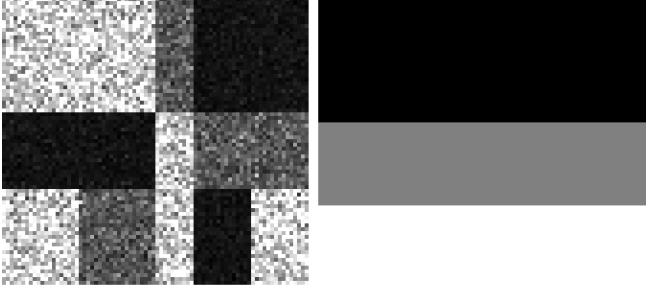
Fig. 7(a) presents a simulated PolSAR image of  $75 \times 80$  pixels with ten regions generated from the scaled complex Wishart law and 12 looks. The regions have low, intermediate, and high brightness. The middle brightness has covariance matrix given in Eq. (10), while the largest and smallest brightnesses are simulated from the following matrices, respectively:

$$\begin{bmatrix} 962892 & 19171 - 3579\mathbf{i} & -154638 + 191388\mathbf{i} \\ & 56707 & -5798 + 16812\mathbf{i} \\ & & 472251 \end{bmatrix}$$

and

$$\begin{bmatrix} 32556 & 556 + 787i & 24046 - 27287i \\ & 1647 & -146 - 482i \\ & & 61028 \end{bmatrix}.$$

These two matrices were observed in [6], in urban and pasture regions, respectively.



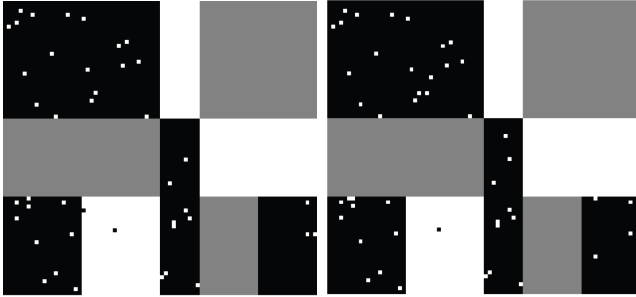
(a) Synthetic image

(b) Initial solution



(c)  $d_{KL}$  clusters

(d)  $d_B$  clusters



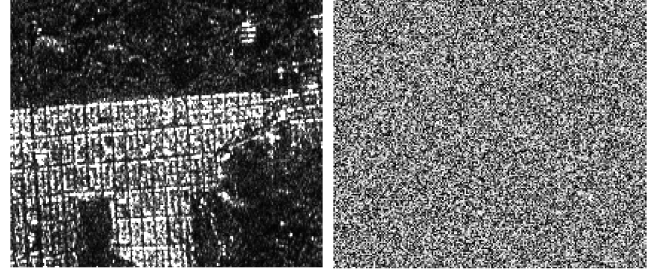
(e)  $d_H$  clusters

(f)  $d_R^{0.1}$  clusters

Fig. 7. Analysis of clustering in synthetic PolSAR images.

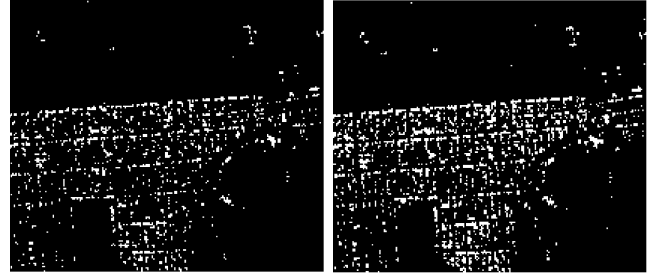
Fig. 7(b) shows the initial stage in the clustering process, which is quite far from the ideal solution. Figs 7(c) to 7(f) present the results of using the  $k$ -means algorithm proposed by Cao *et al.* [45] using the stochastic distances already presented. Notice that all the distances were able to identify clusters accurately with a few spurious spots.

We applied this methodology to a  $182 \times 210$  pixels area from the San Francisco AIRSAR image (Fig. 8(a)). This area is composed of urban and forest regions. Fig. 8(b) shows the initial stage of the clustering analysis, which was randomly generated. Notice that  $d_R^{0.1}$  gathers more pixels of urban regions than the other distances. The Kullback-Leibler distance presented the worst performance.



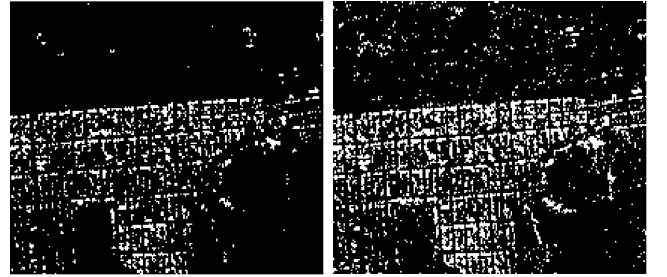
(a) AIRSAR image

(b) Random initialization



(c)  $d_{KL}$  clusters

(d)  $d_B$  clusters



(e)  $d_H$  clusters

(f)  $d_R^{0.1}$  clusters

Fig. 8. Analysis of clustering in actual PolSAR images.

## VI. CONCLUSIONS

Analytic expressions of four contrast measures between relaxed complex Wishart distribution were derived for the most general case (different number of looks and different covariance matrices), along with the particular cases of same looks and same covariance matrix. These

measures are shown to be scale invariant, and they lead to test statistics with asymptotic  $\chi^2$  distribution under the null hypothesis. Novel inequalities which relate covariance matrices and distances were derived, leading to a new and simple derivation of the revised Wishart and Bartlett distances. Using the parametrization proposed by Lee *et al.* [14], contrast measures based on the derived distances, which depend on the complex correlation coefficients between channels, were also obtained. These new expressions can be used in a variety of applications as, for instance, segmentation, and classification.

Those stochastic distances were successfully used as dissimilarities in a  $k$ -means algorithm. Data from AIR-SAR sensors confirmed the expected behavior of all the distances: distances are smaller when applied to samples of similar roughness, and larger otherwise.

Considering Monte Carlo experiments, all the proposed statistics based stochastic distances present good performance with finite size samples. In particular, the results provided evidence that the test based on the  $d_B$  has the smallest empirical test size in a variety of situations. This behavior was confirmed with samples from a PolSAR sensor.

We presented numerical evidence that the statistics based on Bhattacharyya distance overcome all other distances based on statistics. It is noteworthy that the tests here considered tend to reject more than their nominal levels when dealing with small samples and small number of looks. Correcting these tests statistics is a venue for new research.

#### ACKNOWLEDGMENT

The authors are grateful to CNPq, Facepe, Fapeal, and Capes for their support.

#### APPENDIX A THE KULLBACK-LEIBLER DISTANCE IN GENERAL FORM

The Kullback-Leibler distance is given by

$$\begin{aligned} d_{KL}(\theta_1, \theta_2) = & \frac{n_1 - n_2}{2} [E(\log |Z_1|) - E(\log |Z_2|)] \\ & - \frac{1}{2} E[n_1 \text{tr}(\Sigma_1^{-1} Z_1) - n_2 \text{tr}(\Sigma_2^{-1} Z_1)] \\ & + \frac{1}{2} E[n_1 \text{tr}(\Sigma_1^{-1} Z_2) - n_2 \text{tr}(\Sigma_2^{-1} Z_2)], \end{aligned} \quad (24)$$

where  $Z_i \sim \mathcal{WR}(\Sigma_i, n_i)$ ,  $i = 1, 2$ . According to Anfinson *et al.* [19], we have:

$$\begin{aligned} E(\log |Z_i|) &= \log |\Sigma_i| + \sum_{k=0}^{p-1} \psi^{(0)}(n_i - k) - p \log n_i \\ &= \log |\Sigma_i| + p\psi^{(0)}(n_i - p + 1) + \sum_{k=1}^{p-1} \frac{k}{n_i - k} - p \log n_i, \end{aligned} \quad (25)$$

since  $\psi^{(0)}(x + 1) = \psi^{(0)}(x) + x^{-1}$ , for any  $x$  real. Additionally,

$$\begin{aligned} E[\text{tr}(\Sigma_j^{-1} Z_i)] &= E\left(\sum_{k=1}^p \sum_{\ell=1}^p \delta_{k\ell j} z_{k\ell i}\right) \\ &= \sum_{k=1}^p \sum_{\ell=1}^p \delta_{k\ell j} E(z_{k\ell i}) = \text{tr}[\Sigma_j^{-1} E(Z_i)] \\ &= \begin{cases} \text{tr}(\Sigma_j^{-1} \Sigma_i), & \text{if } i \neq j, \\ p, & \text{if } i = j, \end{cases} \end{aligned} \quad (26)$$

where  $\delta_{k\ell j}$  and  $z_{k\ell i}$  are the  $(k, \ell)$  entry of the matrices  $\Sigma_j^{-1}$  and  $Z_i$ , respectively. Hence, applying (25) and (26) into (24) yields (7).

#### APPENDIX B DISTANCES AND THE COMPLEX CORRELATION COEFFICIENT

Consider  $\theta_1 = (\Sigma_1, n_1)$  and  $\theta_2 = (\Sigma_2, n_2)$ , where

$$\Sigma_k = \begin{bmatrix} \sigma_{11}^{(k)} & \sqrt{\sigma_{11}^{(k)} \sigma_{22}^{(k)}} |\rho_k| e^{i\delta} \\ \sqrt{\sigma_{11}^{(k)} \sigma_{22}^{(k)}} |\rho_k| e^{-i\delta} & \sigma_{22}^{(k)} \end{bmatrix},$$

for  $k = 1, 2$ . The following results hold:

- 1)  $|\Sigma_i| = \sigma_{11}^{(i)} \sigma_{22}^{(i)} (1 - |\rho_i|^2)$  and
- 2)

$$\begin{aligned} \text{tr}(\Sigma_i^{-1} \Sigma_j) &= \\ \frac{1}{1 - |\rho_i|^2} \left( \frac{\sigma_{11}^{(j)}}{\sigma_{11}^{(i)}} + \frac{\sigma_{22}^{(j)}}{\sigma_{22}^{(i)}} - 2|\rho_i| |\rho_j| \sqrt{\frac{\sigma_{11}^{(j)} \sigma_{22}^{(j)}}{\sigma_{11}^{(i)} \sigma_{22}^{(i)}}} \right), \end{aligned}$$

for  $i, j = 1, 2$ .

The following result is pivotal for the derivation of the remainder distances, except for the Kullback-Leibler

distance:

$$\begin{aligned}\zeta(k_i, k_j, i, j) &= |(k_i \Sigma_i^{-1} + k_j \Sigma_j^{-1})^{-1}| \\ &= \left[ \frac{1}{\sigma_{11}^{(i)} \sigma_{22}^{(i)}} \left( \frac{k_i}{1 - |\rho_i|^2} \right)^2 \left( 1 - \frac{|\rho_i|^2}{\sigma_{11}^{(i)} \sigma_{22}^{(i)}} \right) \right. \\ &\quad + \frac{1}{\sigma_{11}^{(j)} \sigma_{22}^{(j)}} \left( \frac{k_j}{1 - |\rho_j|^2} \right)^2 \left( 1 - \frac{|\rho_j|^2}{\sigma_{11}^{(j)} \sigma_{22}^{(j)}} \right) \\ &\quad + \frac{k_i k_j}{(1 - |\rho_i|^2)(1 - |\rho_j|^2)} \\ &\quad \left. \times \frac{\sigma_{11}^{(j)} \sigma_{22}^{(i)} + \sigma_{11}^{(i)} \sigma_{22}^{(j)} - 2|\rho_i| |\rho_j|}{\sigma_{11}^{(i)} \sigma_{22}^{(i)} \sigma_{11}^{(j)} \sigma_{22}^{(j)}} \right]^{-1}.\end{aligned}$$

Applying this expression in (6) and (8)-(10), one obtains the particular forms of  $d_R^\beta$ ,  $d_B$ , and  $d_H$ .

In the following, the distances are derived and their formulas are specified for the situation  $(\sigma_{11}^{(1)}, \sigma_{22}^{(1)}, n_1) = (\sigma_{11}^{(2)}, \sigma_{22}^{(2)}, n_2) = (1, 1, n)$ .

- Expression for  $d_{KL}$ : Applying results 1) and 2) in (6), the Kullback-Leibler distance becomes

$$\begin{aligned}d_{KL}(\theta_1, \theta_2) &= \frac{n_1 - n_2}{2} \left\{ \log \frac{\sigma_{11}^{(1)} \sigma_{22}^{(1)}}{\sigma_{11}^{(2)} \sigma_{22}^{(2)}} - 2 \log \frac{n_1}{n_2} \right. \\ &\quad + \log \frac{(1 - |\rho_1|^2)}{(1 - |\rho_2|^2)} + \frac{n_2 - n_1}{(n_1 - 1)(n_2 - 1)} \\ &\quad \left. + 2[\psi^{(0)}(n_1 - 1) - \psi^{(0)}(n_2 - 1)] \right\} \\ &\quad + (1 - |\rho_1| |\rho_2|) \left( \frac{n_1}{1 - |\rho_1|^2} + \frac{n_2}{1 - |\rho_2|^2} \right) \\ &\quad - (n_1 + n_2).\end{aligned}$$

and

$$d_{KL}(\theta_1, \theta_2) = n \left[ \frac{(1 - |\rho_1| |\rho_2|)(2 - |\rho_1|^2 - |\rho_2|^2)}{(1 - |\rho_1|^2)(1 - |\rho_2|^2)} - 2 \right].$$

- Expression for  $d_R^\beta$ :

$$\begin{aligned}I(\theta_1, \theta_2) &= \left[ \frac{(n_1 - 1)\Gamma(n_1 - 1)^2}{n_1^{2n_1}} (\sigma_{11}^{(1)} \sigma_{22}^{(1)})^{n_1} (1 - |\rho_1|^2)^{n_1} \right]^{-\beta} \\ &\quad \times \left[ \frac{(n_2 - 1)\Gamma(n_2 - 1)^2}{n_2^{2n_2}} (\sigma_{11}^{(2)} \sigma_{22}^{(2)})^{n_2} (1 - |\rho_2|^2)^{n_2} \right]^{\beta-1} \\ &\quad \times (E_1 - 1)\Gamma(E_1 - 1)^2 \zeta(n_1 \beta, n_2(1 - \beta), 1, 2)^{E_1} \\ &\quad + \left[ \frac{(n_1 - 1)\Gamma(n_1 - 1)^2}{n_1^{2n_1}} (\sigma_{11}^{(1)} \sigma_{22}^{(1)})^{n_1} (1 - |\rho_1|^2)^{n_1} \right]^{\beta-1} \\ &\quad \times \left[ \frac{(n_2 - 1)\Gamma(n_2 - 1)^2}{n_2^{2n_2}} (\sigma_{11}^{(2)} \sigma_{22}^{(2)})^{n_2} (1 - |\rho_2|^2)^{n_2} \right]^{-\beta} \\ &\quad \times (E_2 - 1)\Gamma(E_2 - 1)^2 \zeta(n_2 \beta, n_1(1 - \beta), 2, 1)^{E_2}\end{aligned}$$

and

$$\begin{aligned}d_R^\beta(\theta_1, \theta_2) &= \frac{\log 2}{1 - \beta} + \frac{1}{\beta - 1} \\ &\quad \times \log \left\{ \left[ \frac{(1 - |\rho_1|^2)^{1-\beta} (1 - |\rho_2|^2)^\beta}{1 - \{|\rho_1| - \beta(|\rho_1| - |\rho_2|)\}^2} \right]^n \right. \\ &\quad \left. + \left[ \frac{(1 - |\rho_2|^2)^{1-\beta} (1 - |\rho_1|^2)^\beta}{1 - \{|\rho_2| - \beta(|\rho_2| - |\rho_1|)\}^2} \right]^n \right\}.\end{aligned}$$

- Expression for  $d_B$ :

$$\begin{aligned}d_B(\theta_1, \theta_2) &= \sum_{k=0}^1 \log \frac{\sqrt{\Gamma(L_1 - k)\Gamma(L_2 - k)}}{\Gamma((L_1 + L_2)/2 - k)} \\ &\quad + \frac{n_1 [\log(1 - |\rho_1|^2) + \log(\sigma_{11}^{(1)} \sigma_{22}^{(1)})]}{2} \\ &\quad + \frac{n_2 [\log(1 - |\rho_2|^2) + \log(\sigma_{11}^{(2)} \sigma_{22}^{(2)})]}{2} \\ &\quad - [n_1 \log n_1 + n_2 \log n_2] \\ &\quad - \frac{n_1 + n_2}{2} \log \zeta(n_1/2, n_2/2, 1, 2)\end{aligned}$$

and

$$\begin{aligned}d_B(\theta_1, \theta_2) &= n \left\{ \frac{\log(1 - |\rho_1|^2) + \log(1 - |\rho_2|^2)}{2} \right. \\ &\quad \left. - 2 \log 2 + \log \frac{4 - (|\rho_1| + |\rho_2|)^2}{(1 - |\rho_1|^2)(1 - |\rho_2|^2)} \right\}.\end{aligned}$$

- Expression for  $d_H$ :

$$\begin{aligned}d_H(\theta_1, \theta_2) &= 1 - \left\{ \prod_{k=0}^1 \frac{\Gamma(\frac{n_1+n_2}{2} - k)}{\sqrt{\Gamma(n_1 - k)\Gamma(n_2 - k)}} \right. \\ &\quad \times \frac{\zeta(n_1/2, n_2/2, 1, 2)^{\frac{n_1+n_2}{2}}}{[\sigma_{11}^{(1)} \sigma_{22}^{(1)} (1 - |\rho_1|^2)]^{\frac{n_1}{2}} [\sigma_{11}^{(2)} \sigma_{22}^{(2)} (1 - |\rho_2|^2)]^{\frac{n_2}{2}}} \\ &\quad \left. \times \sqrt{n_1^{2n_1} n_2^{2n_2}} \right\}.\end{aligned}$$

and

$$d_H(\theta_1, \theta_2) = 1 - \left[ 4 \frac{\sqrt{(1 - |\rho_1|^2)(1 - |\rho_2|^2)}}{4 - (|\rho_1| + |\rho_2|)^2} \right]^n.$$

## REFERENCES

- [1] J. S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton: CRC, 2009.
- [2] C. López-Martínez, X. Fábregas, and E. Pottier, "Multidimensional speckle noise model," *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 20, pp. 3259–3271, 2005.
- [3] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*, ser. The SciTech radar and defense series. SciTech Publishing, 1998.

- [4] J. M. Beaulieu and R. Touzi, "Segmentation of textured polarimetric SAR scenes by likelihood approximation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 10, pp. 2063–2072, 2004.
- [5] P. R. Kersten, J. S. Lee, and T. L. Ainsworth, "Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and EM clustering," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 519–527, 2005.
- [6] A. C. Frery, J. Jacobo-Berlles, J. Gambini, and M. Mejail, "Polarimetric SAR image segmentation with B-splines and a new statistical model," *Multidimensional Systems and Signal Processing*, vol. 21, pp. 319–342, 2010.
- [7] J. Gambini, M. Mejail, J. Jacobo-Berlles, and A. C. Frery, "Accuracy of edge detection methods with local information in speckled imagery," *Statistics and Computing*, vol. 18, no. 1, pp. 15–26, 2008.
- [8] J. Inglada and G. Mercier, "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*, vol. 45, no. 5, pp. 1432–1445, 2007.
- [9] F. Goudail and P. Réfrégier, "Contrast definition for optical coherent polarimetric images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 7, pp. 947–951, jul 2004.
- [10] N. R. Goodman, "Statistical analysis based on a certain complex Gaussian distribution (an introduction)," *The Annals of Mathematical Statistics*, vol. 34, pp. 152–177, 1963.
- [11] 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 Transactions on Geoscience and Remote Sensing*, vol. 41, no. 1, pp. 4–19, 2003.
- [12] A. D. C. Nascimento, R. J. Cintra, and A. C. Frery, "Hypothesis testing in speckled data with stochastic distances," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 1, pp. 373–385, 2010.
- [13] A. C. Frery, A. D. C. Nascimento, and R. J. Cintra, "Information theory and image understanding: An application to polarimetric SAR imagery," *Chilean Journal of Statistics*, vol. 2, no. 2, pp. 81–100, 2011.
- [14] J. S. Lee, K. W. Hoppel, S. A. Mango, and A. R. Miller, "Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 5, pp. 1017–1028, sep 1994.
- [15] M. Salicrú, M. L. Menéndez, L. Pardo, and D. Morales, "On the applications of divergence type measures in testing statistical hypothesis," *Journal of Multivariate Analysis*, vol. 51, pp. 372–391, 1994.
- [16] K. Ersahin, I. Cumming, and R. K. Ward, "Segmentation and classification of polarimetric SAR data using spectral graph partitioning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 1, pp. 164–174, jan 2010.
- [17] F. T. Ulaby and C. Elachi, *Radar Polarimetry for Geoscience Applications*. Norwood: Artech House, 1990.
- [18] C. Lopez-Martinez and X. Fabregas, "Polarimetric SAR speckle noise model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 10, pp. 2232–2242, 2003.
- [19] S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, "Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3795–3809, 2009.
- [20] S. N. Anfinsen, T. Eltoft, and A. P. Doulgeris, "A relaxed Wishart model for polarimetric SAR data," in *Proceedings of the 4th International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLIN-SAR'2009)*, vol. ESA SP-668, no. 8, Frascati, Italy, 26–30, jan 2009.
- [21] N. R. Goodman, "The distribution of determinant of a complex Wishart distributed matrix," *The Annals of Mathematical Statistics*, vol. 34, pp. 178–180, 1963.
- [22] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, t. G. p. ninth Dover printing, Ed. Dover, 1964.
- [23] M. Hagedorn, P. J. Smith, P. J. Bones, R. P. Millane, and D. Pairman, "A trivariate chi-squared distribution derived from the complex Wishart distribution," *Journal of Multivariate Analysis*, vol. 97, no. 3, pp. 655–674, 2006.
- [24] R. G. Bartle and D. R. Sherbert, *Introduction to Real Analysis*, ser. Matemáticas (Limusa). Wiley, 1982.
- [25] M. M. Deza and E. Deza, *Encyclopedia of Distances*, 1st ed. Springer, Aug. 2009.
- [26] S. M. Ali and S. D. Silvey, "A general class of coefficients of divergence of one distribution from another," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 26, pp. 131–142, 1996.
- [27] I. Csiszár, "Information type measures of difference of probability distributions and indirect observations," *Studia Scientiarum Mathematicarum Hungarica*, vol. 2, pp. 299–318, 1967.
- [28] A. K. Seghouane and S. I. Amari, "The AIC criterion and symmetrizing the Kullback-Leibler divergence," *IEEE Transactions on Neural Networks*, vol. 18, no. 1, pp. 97–106, 2007.
- [29] T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," *IEEE Transactions on Communication Technology*, vol. 15, pp. 52–60, feb 1967.
- [30] J. Burbea and C. Rao, "Entropy differential metric, distance and divergence measures in probability spaces: a unified approach," *Journal of Multivariate Analysis*, vol. 12, pp. 575–596, 1982.
- [31] L. Jager and J. A. Wellner, "Goodness-of-fit tests via phi-divergences," *The Annals of Statistics*, vol. 35, no. 5, pp. 2018–2053, 2007.
- [32] S. Eguchi and J. Copas, "Interpreting Kullback-Leibler divergence with the Neyman-Pearson lemma," *Journal of Multivariate Analysis*, vol. 97, pp. 2034–2040, 2006.
- [33] A. Andai, "On the geometry of generalized Gaussian distributions," *Journal of Multivariate Analysis*, vol. 100, no. 4, pp. 777–793, 2009.
- [34] E. Neuman, "Inequalities involving multivariate convex functions II," *Proceedings of the American Mathematical Society*, vol. 109, no. 4, pp. 965–974, 1990.
- [35] F. Goudail, P. Réfrégier, and G. Delyon, "Bhattacharyya distance as a contrast parameter for statistical processing of noisy optical images," *Journal of the Optical Society of America A*, vol. 21, no. 7, pp. 1231–1240, 2004.
- [36] L. Giet and M. Lubrano, "A minimum Hellinger distance estimator for stochastic differential equations: An application to statistical inference for continuous time interest rate models," *Computational Statistics & Data Analysis*, vol. 52, no. 6, pp. 2945–2965, 2008.
- [37] J. S. Lee, M. R. Grunes, and R. Kwok, "Classification of multilook polarimetric SAR imagery based on complex Wishart distribution," *International Journal of Remote Sensing*, vol. 15, no. 11, pp. 2299–2311, sep 1994.
- [38] K. Y. Lee and T. R. Bretschneider, "Derivation of separability measures based on central complex gaussian and wishart distributions," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, jul 2011, pp. 3740–3743.
- [39] A. De Maio and G. Alfano, "Polarimetric adaptive detection in non-Gaussian noise," *Signal Processing*, vol. 83, no. 2, pp. 297–306, 2003.

- [40] T. L. Ainsworth, D. L. Schuler, and J. S. Lee, "Polarimetric SAR characterization of man-made structures in urban areas using normalized circular-pol correlation coefficients," *Remote Sensing of Environment*, vol. 112, no. 6, pp. 2876–2885, 2008.
- [41] J.-S. Lee, M. Grunes, and G. de Grandi, "Polarimetric SAR speckle filtering and its implication for classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 5, pp. 2363–2373, sep 1999.
- [42] European Space Agency, *Polarimetric SAR Data Processing and Educational Tool (POLARPRO)*, 4th ed. European Space Agency, 2009.
- [43] R. J. Cintra, A. C. Frery, and A. D. C. Nascimento, "Parametric and nonparametric tests for speckled imagery," *Pattern Analysis and Applications*, pp. 1–21, 2011.
- [44] A. P. Doulgeris, S. N. Anfinsen, and T. Eltoft, "Automated non-gaussian clustering of polarimetric synthetic aperture radar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3665–3676, oct 2011.
- [45] F. Cao, W. Hong, Y. Wu, and E. Pottier, "An unsupervised segmentation with an adaptive number of clusters using the SPAN/H/ $\alpha$ /A space and the complex Wishart clustering for fully polarimetric SAR data analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 11, pp. 3454–3467, 2007.



**Abraão D. C. Nascimento** holds B.Sc. and M.Sc. degrees in Statistics from the Universidade Federal de Pernambuco (UFPE), Brazil, and he is currently a doctorate student in Statistics at the same University. His research interests are stochastic models and distances.



**Renato J. Cintra (SM'2010)** earned his B.Sc., M.Sc., and D.Sc. degrees in Electrical Engineering from Universidade Federal de Pernambuco (UFPE), Brazil, in 1999, 2001, and 2005, respectively. In 2005, he joined the Department of Statistics at UFPE. His long term topics of research include theory and methods for digital signal processing, communications systems, and applied mathematics. He performs routinely reviewing work for major peer-review journals. During 2008-2009, he worked at the University of Calgary, Canada, as a visiting research fellow. In 2010, he was a visiting professor at the University of Akron, OH. He is also member of AMS and SIAM.



**Alejandro C. Frery** graduated in Electronic and Electrical Engineering from the Universidad de Mendoza, Argentina. His M.Sc. degree was in Applied Mathematics (Statistics) from the Instituto de Matemática Pura e Aplicada (Rio de Janeiro) and his Ph.D. degree was in Applied Computing from the Instituto Nacional de Pesquisas Espaciais (São José dos Campos, Brazil). He is currently with the Instituto de Computação, Universidade Federal de Alagoas, Maceió, Brazil. His research interests are statistical computing and stochastic modelling.

## 5.4 Novas propostas de detecção de bordas em imagens PolSAR

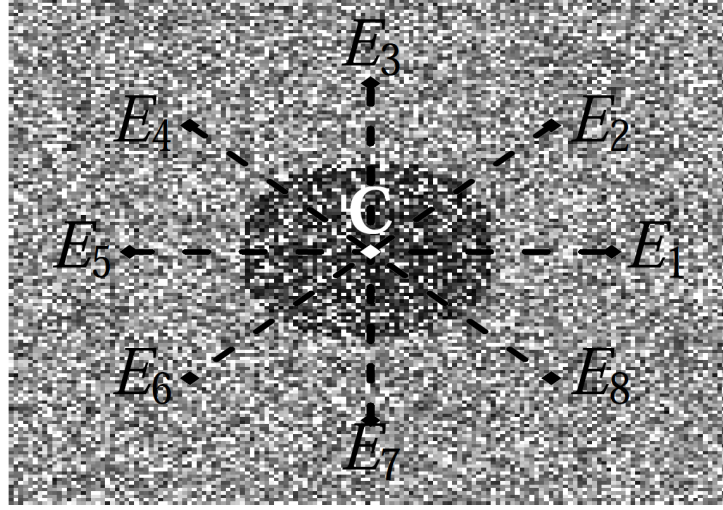
Imagens PolSAR são corrompidas pelo ruído speckle, dificultando a realização de métodos de processamento de imagem, tal como detecção de bordas. Este campo do processamento objetiva procurar bordas entre as diferentes regiões de uma imagem, que apresentam pixels com propriedades similares. Metodologias bem sucedidas têm sido propostas por Gambini et al. (2006), para dados monopolarizado, e por Frery, Jacobo-Berlles, Gambini & Mejail (2010) para dados polarimétricos. Os processos de detecção em tais artigos estão fundamentados basicamente em três passos: (i) a proposta de regiões iniciais, (ii) a detecção de alguns pontos para o controle do processo (sendo, alguns deles, os centróides das regiões iniciais) e, finalmente, (iii) a definição da região de contorno pelo método *B-spline* com respeito aos pontos detectados. Esta seção compara quatro abordagens alternativas baseadas em distâncias estocásticas com uma outra proposta por Gambini et al. (2008), que leva em conta a verossimilhança conjunta de uma distribuição a priori associada aos dados. Em todos os casos, assumimos que as imagens polarimétricas são bem descritas pela lei Wishart complexa escalonada (esta suposição também foi adotada em Lee & Pottier 2009). Nos procedimentos utilizados, enquanto uma proposta usual procura a divisão que maximiza a verossimilhança, as novas abordagens buscam os pontos que maximizam as distâncias. As distâncias empregadas neste trabalho são aquelas propostas por Kullback-Leibler, Bhattacharyya, Hellinger e Rényi (de ordem  $\beta$ ). Experimento com imagem sintética apresenta evidência que as metodologias baseadas em distâncias estocásticas são aptas a encontrar pontos de bordas em imagens PolSAR. Finalmente, uma aplicação a dados reais da cidade de São Francisco é apresentada a fim de comparar os desempenhos das metodologias. Neste caso, as regiões iniciais são obtidas pelo algoritmo de k-médias adaptado segundo Cao et al. (2007).

### 5.4.1 Detectores de bordas

Uma vez definidas as regiões iniciais, a proposta original consiste em formar segmentos de pixels cujas extremidades sejam os centróides destas regiões e alguns pontos externos a elas (chamados de *pontos de controle*). A Figura 5.13 ilustra a sentença anterior propondo como cenário uma imagem PolSAR sintética com duas regiões iniciais distintas separadas por uma curva convexa.

Assim, objetivando identificar a borda de separação, traça-se segmentos  $s^{(i)} = \overline{C E_i}$  do centróide  $C$  até pontos  $E_i$  externos à região convexa para  $i = 1, 2, \dots, 8$ . Os elementos externos são definidos pela seguinte expressão  $E_i = (r_i \cos(\theta_i), r_i \sin(\theta_i))$  em que (i)  $\theta_{i+1} = \angle(s^{(i)}, s^{(i+1)})$  é o ângulo formado entre dois segmentos consecutivos e (ii)  $r_i$





**Figura 5.13** Imagem PolSAR sintética com as subregiões externa e interna distribuídas segundo  $\mathcal{W}(S_U, 5)$  e  $\mathcal{W}(S_U, 3)$  sendo  $S_U$  dado na Figura 5.10, respectivamente.

é o raio de uma circunferência centrada em  $C$  e que passa pelos pontos externos  $E_i$ .

O próximo passo é definir um ponto móvel sobre cada segmento de modo que, ao caminhar sobre suas posições, ele convergirá para o ponto de borda nominal (denominado por *ponto de transição* (PT), conforme Gambini et al. 2006) segundo alguns critérios discutidos nos seguintes itens.

**(a) Metodologia baseada na máxima verossimilhança conjunta:**

Este método consiste na procura do PT numa faixa de pixels chamada de “segmento”. De acordo com a proposta de Gambini et al. (2006), o PT é o ponto que divide o segmento em duas regiões com diferentes parâmetros. Por exemplo, em todos os segmentos do tipo  $\overline{CE_i}$  da Figura 5.13, os PTs são aqueles que os divide em dois sub-segmentos que são realizações misturadas das distribuições  $\mathcal{W}(S_U, 5)$  e  $\mathcal{W}(S_U, 3)$ . Assim, considere um PT na posição  $j$  de um segmento  $s^{(i)}$  distribuído conforme uma lei Wishart complexa escalonada para  $1 \leq i \leq N_s$ , sendo  $N_s$  a quantidade de segmentos especificados em um determinado estudo. Então este segmento pode ser representado por

$$\begin{aligned} \mathbf{Z}_k^{(i)} &\sim \mathcal{W}(\Sigma_X, L_X), \text{ para } k = 1, \dots, j, \\ \mathbf{Z}_k^{(i)} &\sim \mathcal{W}(\Sigma_Y, L_Y), \text{ para } k = j + 1, \dots, N, \end{aligned} \quad (5.17)$$

em que  $\mathbf{Z}_k'^{(i)}$  é uma realização da matriz aleatória  $\mathbf{Z}_k^{(i)}$  para  $1 \leq k \leq N$  e  $N$  é a quantidade de matrizes estocásticas em um segmento  $s^{(i)}$ . Desta forma, um estimador para o PT em  $s^{(i)}$  é dado pelo índice que maximiza a verossimilhança



conjunta

$$L(j) = \prod_{k=1}^j f_Z(\mathbf{Z}_k^{(i)}; \hat{\Sigma}_X, \hat{L}_X) \prod_{k=j+1}^N f_Z(\mathbf{Z}_k^{(i)}; \hat{\Sigma}_Y, \hat{L}_Y),$$

ou, alternativamente,

$$\ell(j) = \log(L(j)) = \sum_{k=1}^j \log(f_Z(\mathbf{Z}_k^{(i)}; \hat{\Sigma}_X, \hat{L}_X)) + \sum_{k=j+1}^N \log(f_Z(\mathbf{Z}_k^{(i)}; \hat{\Sigma}_Y, \hat{L}_Y)),$$

em que  $\hat{\Sigma}_h$  e  $\hat{L}_h$  são os estimadores de MV (segundo abordagem discutida na Seção 3.1.3) para  $\Sigma_h$  e  $L_h$  baseados nas amostras aleatórias  $\{\mathbf{Z}_1^{(i)}, \mathbf{Z}_2^{(i)}, \dots, \mathbf{Z}_j^{(i)}\}$  e  $\{\mathbf{Z}_{j+1}^{(i)}, \mathbf{Z}_{j+2}^{(i)}, \dots, \mathbf{Z}_N^{(i)}\}$  para  $h = X, Y$ , respectivamente. De acordo com a densidade da Wishart dada na equação (3.7), temos que

$$\begin{aligned} \ell(j) = & \sum_{k=1}^j \log \left( \frac{L_X^{mL_X} |\mathbf{Z}_k^{(i)}|^{L_X-m}}{|\Sigma_X|^{L_X} \Gamma_m(L_X)} \text{etr}(-L_X \Sigma_X^{-1} \mathbf{Z}_k^{(i)}) \right) \\ & + \sum_{k=j+1}^N \log \left( \frac{L_Y^{mL_Y} |\mathbf{Z}_k^{(i)}|^{L_Y-m}}{|\Sigma_Y|^{L_Y} \Gamma_m(L_Y)} \text{etr}(-L_Y \Sigma_Y^{-1} \mathbf{Z}_k^{(i)}) \right). \end{aligned}$$

Em outras palavras, a expressão para o estimador  $\hat{j}$  do PT em  $\mathbf{s}^{(i)}$  é dada por

$$\hat{j}_V = \arg \max_{1 \leq j \leq N} \ell(j).$$

Essa abordagem foi empregada com sucesso por Gambini et al. (2006) para dados speckle univariado, e por Frery, Jacobo-Berlles, Gambini & Mejail (2010) para um modelo que generaliza a distribuição Wishart complexa.

#### (b) Metodologia baseada em distâncias estocásticas:

Considere duas matrizes aleatórias  $\mathbf{X}$  e  $\mathbf{Y}$  distribuídas conforme uma lei Wishart complexa escalonada cujos vetores de parâmetros são  $\theta_1 = (\Sigma_X, L_X)$  e  $\theta_2 = (\Sigma_Y, L_Y)$ . Assim, considerando a definição na Equação (6), tome  $\hat{\theta}_1^{(j)}$  e  $\hat{\theta}_2^{(j)}$  como estimadores de MV sob as amostras aleatórias  $\{\mathbf{Z}_1^{(i)}, \mathbf{Z}_2^{(i)}, \dots, \mathbf{Z}_j^{(i)}\}$  e  $\{\mathbf{Z}_{j+1}^{(i)}, \mathbf{Z}_{j+2}^{(i)}, \dots, \mathbf{Z}_N^{(i)}\}$ , respectivamente. Então, quatro detectores são propostos através da expressão

$$\hat{j}_{\mathcal{I}} = \arg \max_j d_{\mathcal{I}}(\hat{\theta}_1^{(j)}, \hat{\theta}_2^{(j)}) = \arg \max_j d_{\mathcal{I}}(j),$$

em que  $\mathcal{I} = \{\text{KL}, \text{B}, \text{H}, \text{R}\}$  e as distâncias estocásticas  $d_{\mathcal{I}}$  são definidas na Seção 5.3.1. Aqui, no caso da Rényi, considerou-se uma ordem  $\beta = 0.1$ .

A discussão da seguinte subseção pretende examinar o comportamento dos detectores discutidos com relação à variação dos parâmetros da distribuição Wishart complexa.

### 5.4.2 Performance dos detectores

De acordo com o estudo realizado por Frery, Jacobo-Berlles, Gambini & Mejail (2010), é razoável representar os brilhos referentes às áreas urbanas ( $U$ ) e de flores-tas ( $F$ ) por meio das matrizes complexas definidas na Figura 5.14.

---


$$U = \begin{bmatrix} 962892 & 19171 - 3579i & -154638 + 191388i \\ 19171 + 3579i & 56707 & -5798 + 16812i \\ -154638 - 191388i & -5798 - 16812i & 472251 \end{bmatrix},$$

e

$$F = \begin{bmatrix} 360932 & 11050 + 3759i & 63896 + 1581i \\ 11050 - 3759i & 98960 & 6593 + 6868i \\ 63896 - 1581i & 6593 - 6868i & 208843 \end{bmatrix}.$$


---

**Figura 5.14** Matrizes utilizadas em experimento.

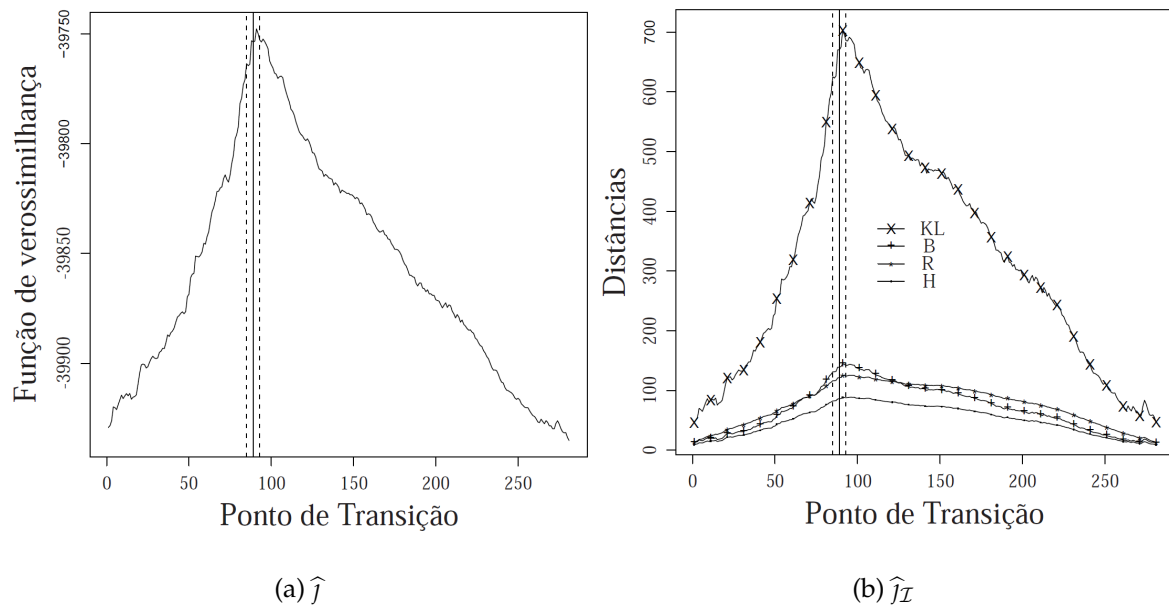
Baseando-se sobre tais matrizes, a Figura 5.15(c) mostra uma faixa retangular sintética com blocos de tamanhos  $90 \times 10$  e  $190 \times 10$ , obedecendo a dados gerados em conformidade com as leis  $\mathcal{W}_{\mathcal{R}}(U, 3)$  e  $\mathcal{W}_{\mathcal{R}}(F, 3)$ , respectivamente.

A performance dos detectores é avaliada sobre esta faixa e a Figura 5.15 mostra as curvas de cada detector. Neste caso, espera-se que os estimadores sejam não viciados à borda nominal, isto é, ao índice 90 para o nosso exemplo. A técnica baseada na verossimilhança conjunta apresenta boa performance, similarmente aos métodos baseados em distâncias estocásticas, para os quais é válida a seguinte desigualdade:  $d_{\text{KL}}(\hat{j}_{\text{KL}}) > d_{\text{B}}(\hat{j}_{\text{B}}) > d_{\text{R}}(\hat{j}_{\text{R}}) > d_{\text{H}}(\hat{j}_{\text{H}})$ .

### 5.4.3 Método B-spline

Uma vez obtidos os pontos de transição em cada segmento conforme a discussão da última subseção é necessário ajustar uma curva que se acomode bem a esses pontos. O ajuste B-spline para descrever contornos em cenas de uma imagem PolSAR tem se mostrado uma técnica eficiente. A seguir, é apresentada uma exposição deste assunto (para maiores detalhes, consultar Blake & Isard 1998).

Seja um conjunto de pontos de controle  $\{q_0, q_1, \dots, q_{N_Y-1}\}$  em que  $q_k = (x_k, y_k) \in \mathbb{R}^2$  para  $k = 0, 1, \dots, N_Y - 1$ . Considere também um conjunto  $\{s_0, s_1, \dots, s_{V-1}\}$  de ta-



(c) Imagem obtida obedecendo as leis  $\mathcal{W}_{\mathcal{R}}(U, 3)$  (esquerda) e  $\mathcal{W}_{\mathcal{R}}(F, 3)$  (direita)

**Figura 5.15** Estudo particular para avaliar a capacidade dos detectores sob regiões com diferentes matrizes de covariância.

manho  $V$  tal que  $s_0 < s_1 < \dots < s_{V-1} \in \mathbb{R}$ . Nesta conjectura, estes elementos são comumente chamados de ‘nós’<sup>2</sup>. Uma curva B-spline de ordem  $d$  é definida como uma soma ponderada de  $N_B$  funções polinomiais,  $B_{k,d}(s)$ , de grau  $d - 1$  em que os pesos são os pontos de controle, com  $s_i \leq s \leq s_{i+1}$  para  $i = 0, 1, \dots, V - 1$ .

Desta forma, a curva B-spline  $r(s) = (x(s), y(s))^T$  pode ser escrita pela seguinte expressão

$$r(s) = \left\{ \sum_{k=0}^{N_Y-1} B_{n,d}(s)x_n, \sum_{k=0}^{N_Y-1} B_{n,d}(s)y_n \right\}^T = \begin{pmatrix} \mathbf{b}(s) & \mathbf{0} \\ \mathbf{0} & \mathbf{b}(s) \end{pmatrix} \begin{pmatrix} \mathbf{q}_x \\ \mathbf{q}_y \end{pmatrix}, \quad (5.18)$$

em que  $\mathbf{b}(s) = \{B_{0,d}(s), B_{1,d}(s), \dots, B_{N_Y-1,d}(s)\}$ ,  $\mathbf{q}_x = \{x_0, x_1, \dots, x_{N_Y-1}\}^T$  e  $\mathbf{q}_y = \{y_0, y_1, \dots, y_{N_Y-1}\}^T$ .

A discussão acima partiu dos parâmetros para explicar a Expressão (5.18). Consideremos agora o problema de ajustar a curva B-spline a partir de um conjunto de pontos  $\mathbf{D} = \{\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_{k-1}\}^T$  em uma imagem plana, em que  $\mathbf{d}_i = (x_i, y_i)^T$  para  $i = 0, 1, \dots, k - 1$  e  $N_B \leq k$ . Da Equação (5.18), temos que o problema do ajustamento é similar a encontrar a matriz  $\mathbf{Q} = (\mathbf{q}_x, \mathbf{q}_y)$  solução do seguinte sistema de equação:

1. Para o caso em que  $k = N_B$ ,  $\mathbf{K} = \{K_{ij} = B_{j,d}(t_i)\}$  em que  $i, j = 0, 1, \dots, N_B - 1$ . Assim, assumindo que a matriz quadrada  $\mathbf{K}$  é não singular,

$$\mathbf{Q} = \mathbf{K}^{-1} \mathbf{D}.$$

2. Para o caso onde  $N_B < k = N'_B$ ,  $\mathbf{K} = \{K_{ij} = B_{j,d}(t_i)\}$  para  $i = 0, 1, \dots, N'_B - 1$  e  $0 \leq j \leq N_B - 1$  em que  $\mathbf{K}^+$  é uma matriz-pseudo inversa. Daí, segue-se o resultado dado por

$$\mathbf{K}^+ \mathbf{D} = (\mathbf{K}^T \mathbf{K})^{-1} \mathbf{K}^T \mathbf{D} = \mathbf{Q}.$$

Em ambos os casos, 1 e 2, têm-se que

$$t_0 = 0, \quad t_\ell = \frac{\sum_{i=1}^{\ell} \|\mathbf{d}_i - \mathbf{d}_{i-1}\|}{\sum_{i=1}^{k-1} \|\mathbf{d}_i - \mathbf{d}_{i-1}\|},$$

em que  $\|\cdot\|$  representa a norma de um vetor. E, quanto a expressão do polinômio adotado, utiliza-se uma base funcional chamada de B-spline de ordem  $d$  definida como

$$B_{n,1}(s) = \begin{cases} 1, & t_n \leq s < t_{n+1} \\ 0, & \text{caso contrário,} \end{cases}$$

e, como passo iterativo em  $d$ ,

$$B_{n,d}(s) = \frac{(s - t_n)B_{n,d-1}(s)}{t_{n+d-1} - t_n} + \frac{(t_{n+d} - s)B_{n+1,d-1}(s)}{t_{n+d} - t_{n+1}},$$

para  $d > 1$ , em que  $0 \leq t_n \leq V - 1$  e  $0 \leq n \leq N_B - 1$ .

<sup>2</sup>Na prática, esse conjunto de nós é escolhido arbitrariamente, conforme metodologia proposta por Gambini et al. (2008).

#### 5.4.4 Aplicação a dados reais

A seguinte discussão está baseada em dois passos: (i) proposta de regiões iniciais e (ii) detecção de contornos.

##### 1. Regiões iniciais

A definição de regiões iniciais é um passo fundamental na estimação de contornos. Estas sementes podem ser definidas de diferentes modos: automático, semiautomático ou manual. Nesta tese, este passo é realizado através de um algoritmo semiautomático, o qual combina técnicas de agrupamento em distribuições Wishart complexa escalonada (como proposto por Cao et al. 2007) com as idéias de detecção apresentadas por Frery, Jacobo-Berlles, Gambini & Mejail (2010) e por Gambini et al. (2006).

Assumindo que uma imagem PolSAR é bem descrita pela distribuição Wishart complexa escalonada, o método de segmentação emprega o algoritmo de  $k$ -médias proposto por Cao et al. (2007). Subsequentemente, divide-se a imagem em blocos e procura pelas regiões de interesse convexas por meio do agrupamento destes blocos usando “limitadores” (seguindo a proposta de Gambini et al. 2006).

Para este fim, quatro passos são necessários: seleção das regiões de interesse, a busca por grupos de blocos concatenados que pertençam a cada uma destas regiões, eliminação dos grupos de ruído dos blocos concatenados e definição das linhas radiais (faixas de dados) para detecção. Estes passos estão descritos em detalhes na Figura 5.16.

A Figura 5.17 perfila o processo de segmentação dando como exemplo a detecção da região de florestas da imagem PolSAR de São Francisco (Figura 5.19(a))

O resultado da segmentação por  $k$ -média para 3 grupos é plotado na Figura 5.17(a). Então, a Figura 5.17(b) mostra a imagem binária, sendo uma região segmentada aquela selecionada pelo usuário (florestas, no presente caso) e outra contendo as demais regiões. A Figura 5.17(c) apresenta os blocos selecionados, definindo  $n = 11$  e  $T_s = 70$  pixels. A região inicial adotada pelo limitador  $T_n = 20$  blocos é plotada na Figura 5.17(d). Finalmente, nós devemos obter raios de controle a partir dos centróides das regiões encontradas. Este fato é ilustrado através da Figura 5.18.

##### 2. Detecção em imagem PolSAR

A detecção é feita sobre uma pequena área da cidade de São Francisco. Tal imagem foi obtida pelo sensor AIRSAR operando na banda L, com número nominal de looks 4. A Figura 5.19 apresenta uma área ( $150 \times 150$ ) com três regiões: oceano (área escura), florestas (cinza) e cidades (clara).

A Figura 5.20 mostra as curvas B-spline para os diferentes detectores com respeito as três regiões. Embora Eguchi & Copas (2006) tenham apresentado evi-

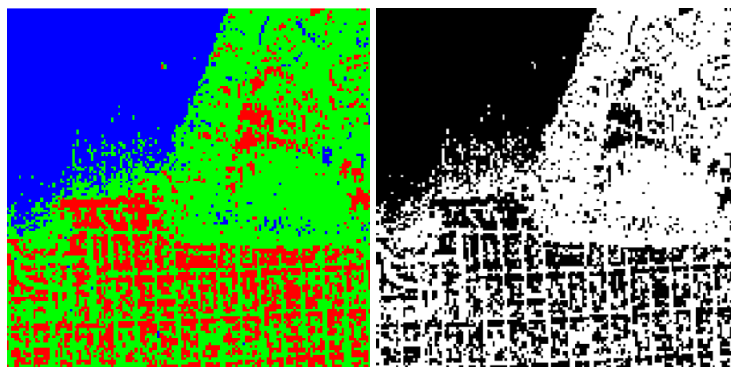
- 
- (a) Executar o método  $k$ -médias (ver Figura 5.17(a)) tendo a medida de dissimilaridade do centróide do  $i$ -ésimo cluster  $C_i$  ao  $j$ -ésimo pixel  $Z_j$  dada por

$$d_{ij}(Z_j, C_i) = \ln |C_i| + \text{Tr}(C_i^{-1} Z_j),$$

em que  $C_i = \frac{1}{N_i} \sum_{k=1}^{N_i} Z_k$ ,  $N_i$  é o número de elementos dentro do  $i$ -ésimo cluster. E selecione a região a ser analisada (de acordo com a Figura 5.17(b));

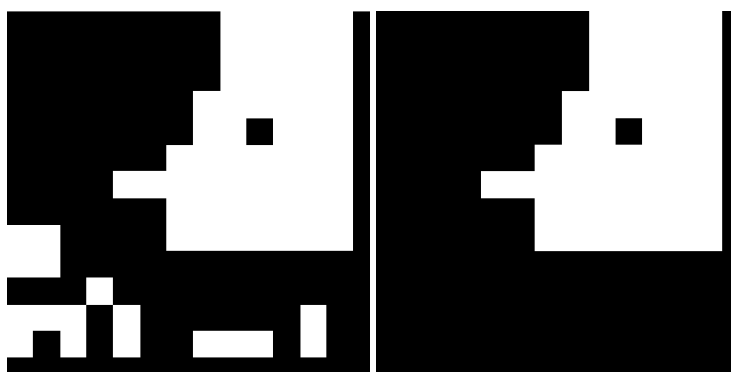
- (b) Dividir a imagem segmentada em blocos de tamanho  $n \times n$  (para esta aplicação, especificamos  $n = 11$  conforme Frery, Jacobo-Berlles, Gambini & Mejail 2010). Em seguida, considerar quais são os blocos que fazem parte da região a ser analisada. A seleção dos blocos é realizada com um limitador especificado  $T_s$ : Se  $\# \{S_r \cap S_b\} > T_s$  então  $S_b \subset S_r$ , ou seja, o bloco  $S_b$  faz parte da região  $S_r$ , onde  $S_r$  é o conjunto de pixels da região selecionada,  $S_b$  é o conjunto de pixels de um bloco e  $\#$  representa o operador de cardinalidade de um conjunto (ilustrado na Figura 5.17(c));
- (c) Agrupar os blocos adjacentes:  $S_{b1} \in N_8(S_{b2}) \rightarrow S_{b1} \cap S_{b2} = S_g$ , onde  $S_{bi}$  é o  $i$ -ésimo bloco. Em seguida, definir quais grupos serão analisados. No exemplo, especificamos os grupos de acordo com o limiar  $T_a$ : Se  $\#S_{gi} > T_a$  então  $S_{gi} \subset S_r$ , onde  $S_{gi}$  é o  $i$ -ésimo grupo (como na Figura 5.17(d)).
- (d) Especificar os centros de cada grupo  $S_{gi}$ . Em seguida, especificar os pontos de controle de cada grupo de acordo com o centro, ângulo e raio (conforme Figura 5.18). Os pontos de controle localizados no limiar da imagem e dentro das regiões foram desconsiderados.
- 

**Figura 5.16** Passo a passo da obtenção dos pontos de controle.



(a) Resultado k-médias com 3 grupos

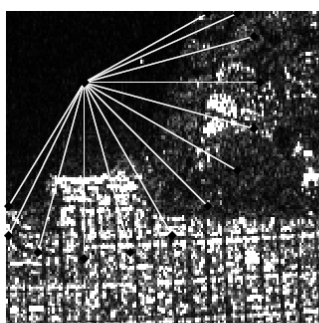
(b) Região de florestas selecionada (branco)



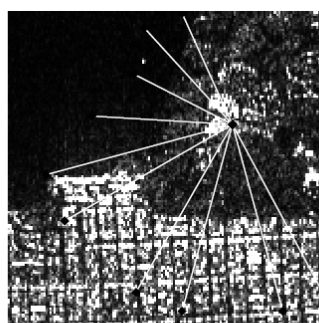
(c) Seleção dos blocos da região de florestas

(d) Seleção dos grupos da região de florestas

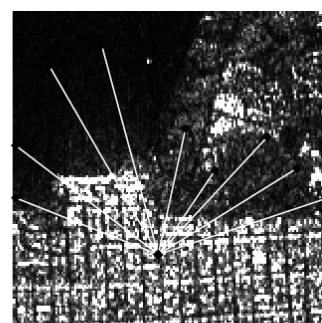
**Figura 5.17** Passos do método de segmentação para a região de florestas.



(a) Oceano

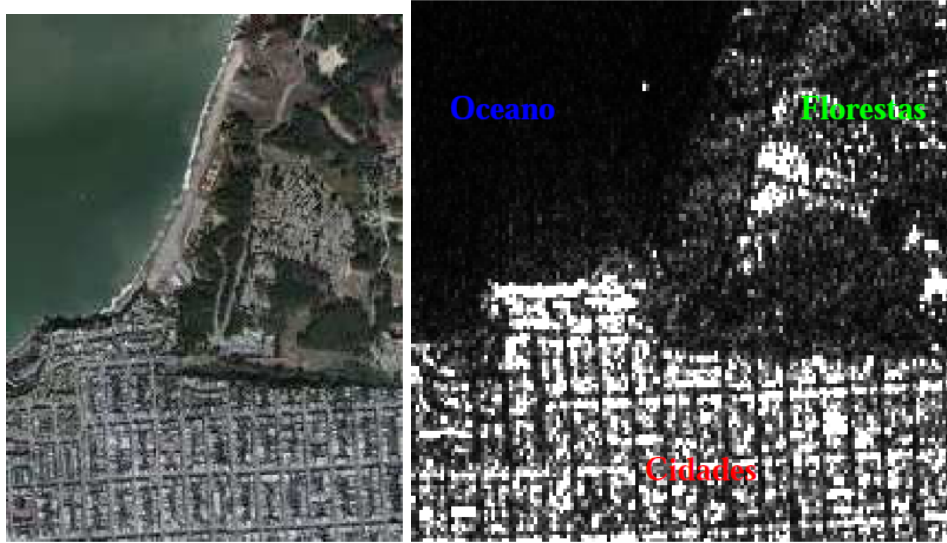


(b) Florestas



(c) Cidades

**Figura 5.18** Definição dos segmentos de controle.



(a) Foto Aérea

(b) Imagem  $I_{HH}$  com regiões**Figura 5.19** Região  $150 \times 150$  de São Francisco.

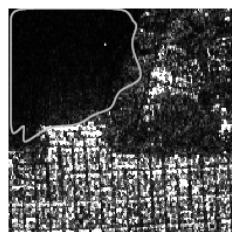
dência de estreitas relações entre a verossimilhança conjunta sob uma amostra aleatória e a divergência de Kullback-Leibler, os resultados obtidos pelos detectores  $\hat{j}_{KL}$  e  $\hat{j}_V$  foram bastante diferentes. Os melhores resultados foram obtidos pelos detectores  $\hat{j}_B$  e  $\hat{j}_R$ , superando a proposta usual  $\hat{j}_V$ . Além disso, eles se mostraram robustos a diferentes graus de heterogeneidade.

#### 5.4.5 Considerações da Seção 5.4

Cinco metodologias para detecção de bordas em imagens PolSAR foram comparadas. Tais algoritmos utilizam faixas de dados para encontrar bordas em dados speckle. As novas abordagens determinam os pontos que maximizam as distâncias estocásticas entre dois conjuntos de dados.

Dados obtidos por simulação mostram evidências que a detecção de bordas baseada sobre distâncias estocásticas são aptas a estimar pontos de transição em segmentos de uma imagem PolSAR. Aplicações a dados reais mostram que os métodos baseados sobre as distâncias de Rényi e Bhattacharyya apresentaram melhores desempenhos do que aqueles baseados sobre a distância clássica de Kullback-Leibler e do que a metodologia apoiada na máxima verossimilhança.

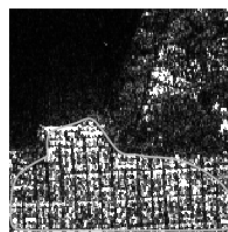




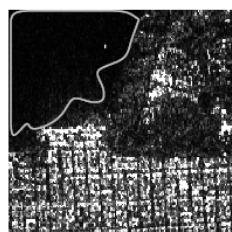
(a) Oceano:V



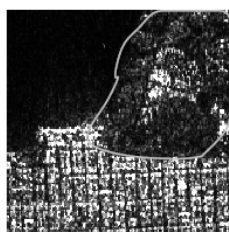
(b) Floresta:V



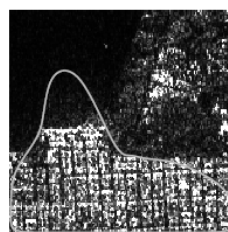
(c) Cidades:V



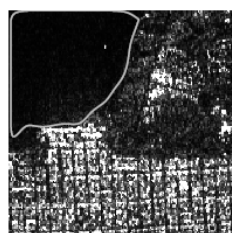
(d) Oceano:KL



(e) Floresta:KL



(f) Cidades:KL



(g) Oceano:B



(h) Floresta:B



(i) Cidades:B



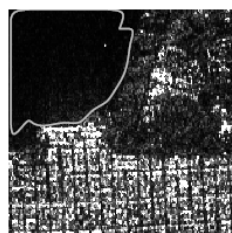
(j) Oceano:H



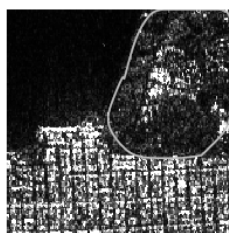
(k) Floresta:H



(l) Cidades:H



(m) Oceano:R



(n) Floresta:R



(o) Cidades:R

**Figura 5.20** Resultados com as regiões da imagem São Francisco.

### 5.4.6 Trabalhos futuros

Como complemento desta seção, objetiva-se

- ✎ Particularizar este estudo para cenários com distribuições dadas na Tabela 3.6 dada na página 53;
- ✎ Utilizar entropias em detecção de borda para imagens PolSAR e compará-las aos detectores baseados em distâncias estocásticas;
- ✎ Estender este estudo para cenários polarimétricos descritos pelas leis da Tabela 3.5 dada na página 50.

## 5.5 Análise estatística para dados PolSAR baseada em medidas de entropias

Analisar dados de imagens PolSAR é equivalente (vide Seção 3) a lidar com um ruído multidimensional resultante de seus processamentos, conforme proposta de López-Martínez et al. (2005). Assim, esta etapa exige que uma modelagem adequada seja definida *a priori* à utilização de quaisquer métodos estatísticos. A distribuição Wishart complexa escalonada tem sido proposta como um importante modelo para representar dados PolSAR (Freitas et al. 2005). Várias técnicas de segmentação (Beaulieu & Touzi 2004), classificação (Kersten et al. 2005) e detecção de bordas (Schou et al. 2003) têm considerado tal distribuição como uma hipótese em seus processamentos. Esta seção propõe ferramentas analíticas baseadas em Teoria da Informação para este modelo.

Depois que Shannon (1948) introduziu os conceitos de *informação* e *entropia*, a derivação de tais medidas tem se tornado tópico de pesquisa em vários artigos. Por exemplo, Zografos & Nadarajah propuseram fórmulas fechadas para as entropias de Rényi e de Shannon para várias distribuições univariadas (Nadarajah & Zografos 2003), bivariadas (Nadarajah & Zografos 2005) e multivariadas (Zografos & Nadarajah 2005). Medidas de entropia de Rényi foram derivadas por Abbasnejad & Arghami (2011) e por Baratpour, Ahmadi & Arghami (2008) em contextos de estatística de ordem. Em anos recentes, estas ferramentas têm sido aplicadas ao processamento de imagens PolSAR. Morio et al. (2008) analisaram a entropia de Shannon para caracterização de imagens polarimétricas. Naquele artigo, as medidas derivadas são comparadas à entropia de Von Neumann. Esta última medida foi estudada por Cloude & Pottier (1997) e por Yan, Yang, Liu & Sun (2010) no contexto de classificação em imagens SAR polarimétricas.

Esta seção objetiva

- ✎ derivar as expressões analíticas das entropias de Shannon, Rényi e Tsallis para a distribuição Wishart complexa escalonada;
- ✎ propor expressões para as variâncias assintóticas das entropias estocásticas segundo Shannon e Rényi, em que os parâmetros dentro da classe de entropias ( $h, \phi$ ) são substituídos pelos seus estimadores MV;
- ✎ sugerir metodologias para testar hipótese e construir intervalos de confiança (Salicrú & Vives 2005), a fim de quantificar contraste em uma imagem PolSAR;
- ✎ aplicar a metodologia desenvolvida a dados PolSAR reais.

Para este fim o trabalho intitulado por

### “Entropy-based Statistical Analysis of PolSAR Data”

será anexado a esta seção. Muito embora os resumos teóricos deste trabalho estejam fortemente correlacionados com as Seções 4.1 e 3.1, este artigo tem uma notação própria e pode ser analisado de forma independente da tese. A seguinte discussão resume os seus principais resultados.

#### 5.5.1 Contribuições

##### Linha teórica:

##### 1) Expressões para medidas de entropia:

Seja  $\mathbf{Z} \sim \mathcal{W}(\boldsymbol{\theta})$  tal que  $\boldsymbol{\theta} = [\text{vec}(\boldsymbol{\Sigma})^t, L]^t$ . Nós derivamos as seguintes expressões para as entropias de Shannon ( $H_S$ ), Rényi ( $H_R^\beta$  de ordem  $\beta$ ) e Tsallis ( $H_T^\beta$  de ordem  $\beta$ ):

$$H_S(\boldsymbol{\theta}) = \frac{m(m-1)}{2} \log \pi - m^2 \log L + m \log |\boldsymbol{\Sigma}| + mL + (m-L) \psi_m^{(0)}(L) + \sum_{k=0}^{m-1} \log \Gamma(L-k), \quad (5.19)$$

$$H_R^\beta(\boldsymbol{\theta}) = \frac{m(m-1)}{2} \log \pi - m^2 \log L + m \log |\boldsymbol{\Sigma}| - \frac{mq \log \beta}{1-\beta} + \frac{\sum_{i=0}^{m-1} [\log \Gamma(q-i) - \beta \log \Gamma(L-i)]}{1-\beta}, \quad (5.20)$$

em que  $q = L + (1-\beta)(m-L)$  e

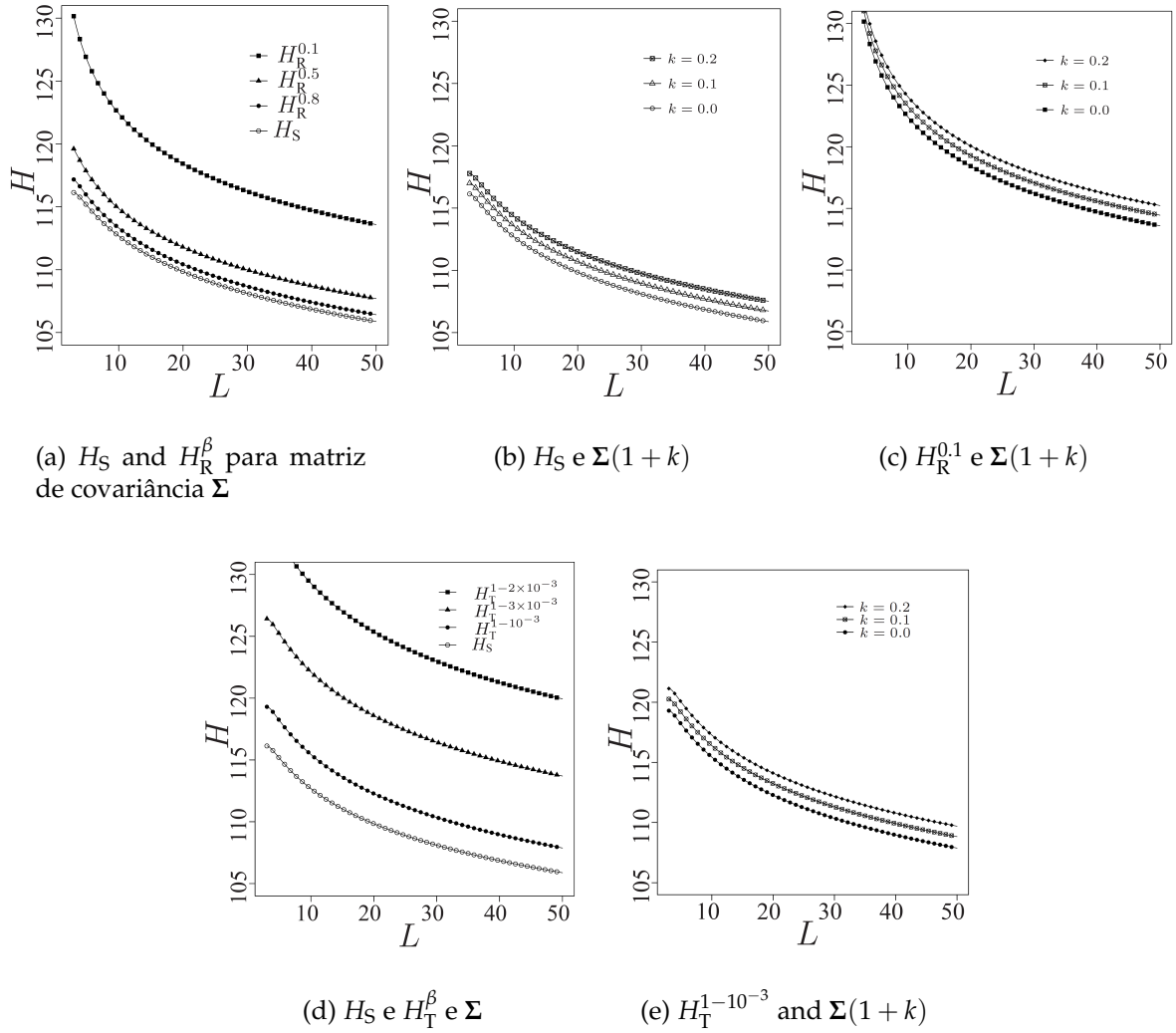
$$H_T^\beta(\boldsymbol{\theta}) = (1-\beta)^{-1} \left[ \frac{\Gamma_m(L + (1-\beta)(m-L))}{\Gamma_m^\beta(L) \beta^{m[L+(1-\beta)(m-L)]}} \frac{|\boldsymbol{\Sigma}|^{(1-\beta)m}}{L^{m^2(1-\beta)}} - 1 \right].$$

Estas quantidades são aptas a fornecer a influência do número de looks e/ou da matriz de covariância (parâmetros comuns a qualquer distribuição polarimétrica conforme Tabela 4.1) sobre o grau de desordem ou variabilidade de um evento probabilístico no domínio Hermitiano positivo definido, por exemplo, em uma região de uma imagem PolSAR. A seguir, uma discussão objetiva construir cenários para ilustrar essas influências

Cosideremos, inicialmente, a matriz

$$\boldsymbol{\Sigma}_U = \begin{bmatrix} 962892 & 19171 - 3579i & -154638 + 191388i \\ & 56707 & -5798 + 16812i \\ & & 472251 \end{bmatrix},$$

que foi estimada por Frery, Jacobo-Berlles, Gambini & Mejail (2010), baseando-se em uma área urbana de uma imagem E-SAR de Weßling (Munique, Alemanha). A Figura 5.21(a) apresenta gráficos das entropias discutidas para  $3 \leq L \leq 50$ ,  $\beta \in \{0.1, 0.5, 0.8\}$  e  $\Sigma_U$ . Considerando o mesmo intervalo para o número de looks, as Figuras 5.21(b), 5.21(e) e 5.21(c) apresentam o comportamento das entropias de Shannon, Tsallis (de ordem  $\beta = 1 - 10^{-3}$ ) e Rényi (de ordem  $\beta = 0.1$ ) para a matriz de covariância  $(1+k)\Sigma_U$ ,  $k \in \{0, 0.1, 0.2\}$ , respectivamente.



**Figura 5.21** Entropias de Shannon, Tsallis e Rényi para várias matrizes de covariâncias e número de looks.

As Figuras 5.21(a) e 5.21(d) ilustram a seguinte propriedade clássica de convergência (maiores detalhes são encontrados em Cover & Thomas 1991):

$$\lim_{\beta \rightarrow 1} H_R^\beta(\theta) = H_S(\theta) = \lim_{\beta \rightarrow 1} H_T^\beta(\theta).$$

Nos casos de convergência considerados, a entropia de Shannon funciona como um limite inferior às entropias  $H_R^\beta$  e  $H_T^\beta$ .

As Figuras 5.21(b), 5.21(c) e 5.21(e) sugerem que a multiplicação da matriz de covariância por uma constante (aumentando seu determinante), resulta em maiores valores de entropias. Adicionalmente, como esperado, o aumento do número de looks implica em menores valores para as entropias devido ao melhoramento da razão sinal-ruído.

Embora a entropia de Tsallis possa ser usada em vários campos de processamento de imagem, a derivação de sua variância assintótica não se mostrou tratável. Assim, a partir de então, o foco desta seção será sobre as entropias de Shannon e Rényi. Diferentemente de  $H_T^\beta$ , estas entropias apresentam formas analíticas tratáveis com respeito à metodologia discutida na Seção 4.1. A seguir, nós apresentamos as variâncias assintóticas das entropias estocásticas de Shannon e Rényi.

## 2) Expressões para as variâncias assintóticas de medidas estocásticas baseadas nas entropias de Shannon e Rényi:

Substituindo  $\theta = [\text{vec}(\Sigma)^t, L]$  por  $\hat{\theta} = [\text{vec}(\hat{\Sigma})^t, \hat{L}]$  nas expressões (5.19) e (5.20) acima, nós mostramos que suas variâncias assintóticas relativas são dadas por:

(a) Para a entropia de Shannon:

$$\sigma_S^2 = \frac{[(m-L)\psi_m^{(1)}(L) + m - \frac{m^2}{L}]^2}{\psi_m^{(1)}(L) - \frac{m}{L}} + \frac{m^2}{L} \text{vec}(\Sigma^{-1})^* (\Sigma \otimes \Sigma) \text{vec}(\Sigma^{-1}). \quad (5.21)$$

(b) Para a entropia de Rényi:

$$\begin{aligned} \sigma_{R,\beta}^2 = & \frac{\left\{ \frac{\beta}{1-\beta} [\psi_m^{(0)}(q) - \psi_m^{(0)}(L)] - \frac{m\beta \log(\beta)}{1-\beta} - \frac{m^2}{L} \right\}^2}{\psi_m^{(1)}(L) - \frac{m}{L}} \\ & + \frac{m^2}{L} \text{vec}(\Sigma^{-1})^* (\Sigma \otimes \Sigma) \text{vec}(\Sigma^{-1}). \end{aligned} \quad (5.22)$$

As expressões (5.21) e (5.22) combinadas às entropias (5.19) e (5.20), respectivamente, resultam em duas estatísticas segundo a Proposição 1 da Seção 4.1:  $S_S$  e  $S_R^\beta$ . Note que  $S_R^\beta$  é uma classe de estatísticas que inclui  $S_S$  (quando  $\beta \rightarrow 1$ ). Estas estatísticas são bastante amplas uma vez que podem ser utilizadas nas seguintes frentes: (i) medidas de qualidade de ajuste como uma alternativa às medidas propostas por Anfinsen, Doulgieris & Eltoft (2011), (ii) medidas de homogeneidade como uma extensão à proposta de Nascimento et al. (2010) e, num contexto ainda pouco explorado, (iii) homogeneidade em múltiplas populações. Esta seção tem o interesse em avançar na frente (ii). A fim contextualizar as contribuições teóricas, nós checamos a eficiência das medidas derivadas em vários cenários como dados PolSAR.

**Linha aplicada:**

Novos testes de hipótese e intervalos de confiança baseados sobre a família de entropias  $(h, \phi)$  (vide Seção 4.1) foram propostos para a distribuição Wishart complexa escalonada. As ferramentas propostas visam contrastar amostras e verificar se elas vêm de uma mesma distribuição populacional. A fim de avaliar os testes propostos, nós consideramos um estudo de simulação de Monte Carlo e aplicações a dados reais.

A Figura 5.22 mostra o canal HH de uma imagem de Munique (Alemanha) obtida pelo sensor E-SAR sob um número de equivalente de looks 3.2. Três regiões foram selecionadas:  $A_1$ ,  $A_2$  e  $A_3$ , correspondendo a áreas de retornos com textura intensa, moderada e fraca (ou mais homogêneas). A Tabela 5.11 mostra os tamanhos de amostra utilizados  $N$  e as estimativas MV dos parâmetros da distribuição Wishart para cada região. As Figuras 5.22(b)-5.22(d) apresentam os histogramas e as densidades ajustadas de duas distribuições Wishart escalonada:  $\mathcal{W}(\hat{\Sigma}, 3.2)$  e  $\mathcal{W}_R(\hat{\Sigma}, \hat{L})$ . Adicionalmente, a Tabela 5.11 mostra os valores dos AICs destas distribuições para os dados considerados. Em todos os casos, o modelo  $\mathcal{W}_R$  se ajustou melhor aos dados. Como mostrado na Tabela 5.11, a amostra mais homogênea apresentou maior estimativa para o número de equivalente de looks e menor determinante.

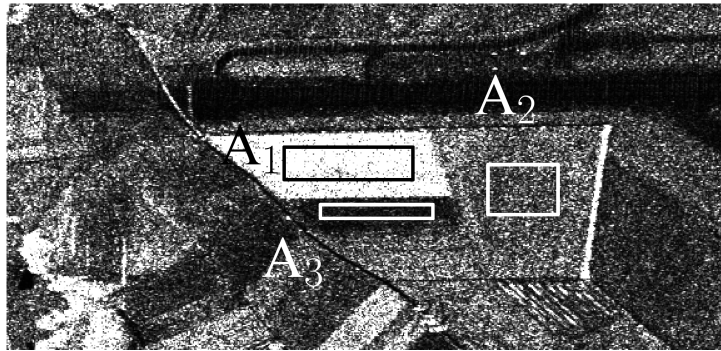
**Tabela 5.11** Estimativas MV

| Regiões | $N$  | $ \hat{\Sigma} $ | $\hat{L}$ | AIC                              |                                        |
|---------|------|------------------|-----------|----------------------------------|----------------------------------------|
|         |      |                  |           | $\mathcal{W}(\hat{\Sigma}, 3.2)$ | $\mathcal{W}_R(\hat{\Sigma}, \hat{L})$ |
| $A_1$   | 3708 | 355494.500       | 1.361     | 50769.93                         | 49856.90                               |
| $A_2$   | 2088 | 3321.241         | 1.657     | 18353.35                         | 17931.51                               |
| $A_3$   | 1079 | 274.189          | 2.557     | 6749.56                          | 6629.15                                |

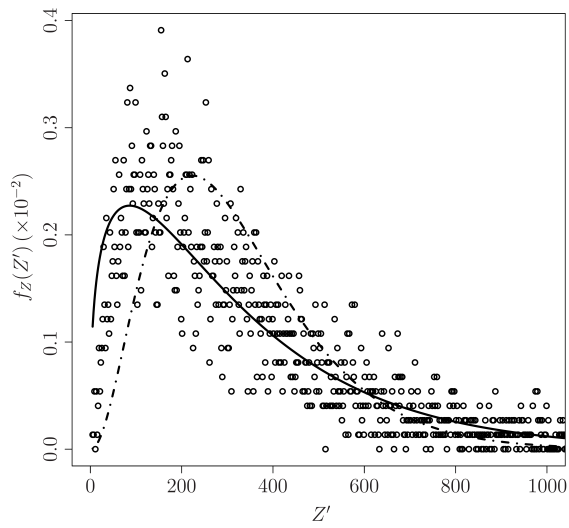
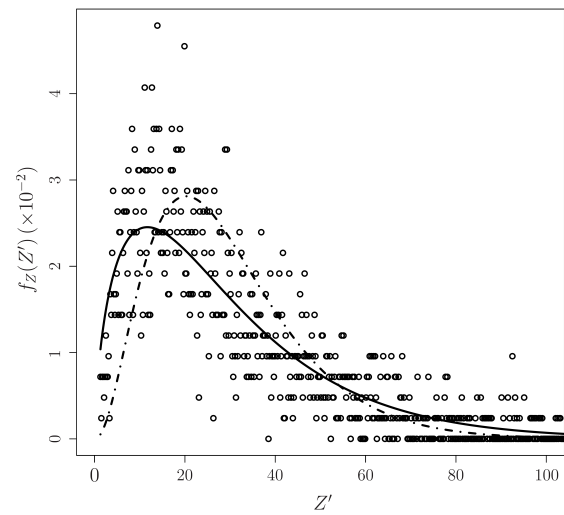
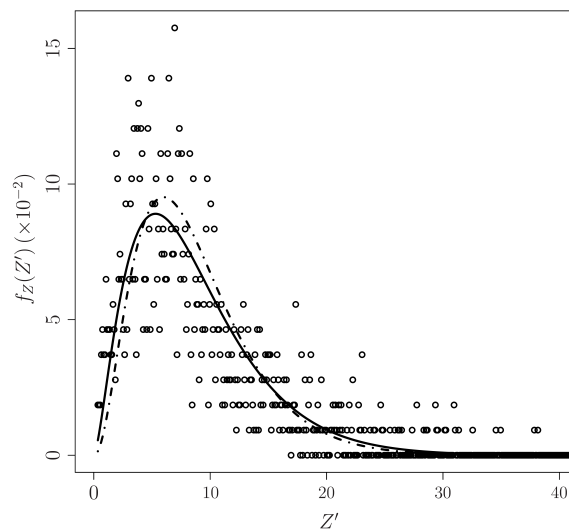
A Tabela 5.12 apresenta os limites de confiança assintóticos para as entropias de Shannon e Rényi em dados reais a um nível de confiança de 95%, como discutido na Seção 4.1.2. Em termos gerais, os intervalos das entropias são disjuntos entre as regiões, o que qualifica o uso de intervalos de confiança baseados em entropias como uma metodologia apropriada para diferenciar regiões PolSAR.

**Tabela 5.12** Intervalos assintóticos estimados para as entropias de Shannon e Rényi no nível 95%

| Região | $\widehat{H}_S$ |          | $\widehat{H}_R^{0.1}$ |          | $\widehat{H}_R^{0.8}$ |          |
|--------|-----------------|----------|-----------------------|----------|-----------------------|----------|
|        | Inferior        | Superior | Inferior              | Superior | Inferior              | Superior |
| $A_1$  | 37.979          | 38.432   | 61.083                | 61.332   | 44.045                | 44.364   |
| $A_2$  | 30.079          | 30.541   | 45.563                | 45.867   | 36.124                | 37.049   |
| $A_3$  | 19.611          | 19.949   | 35.000                | 35.346   | 20.901                | 21.230   |



(a) Imagem E-SAR

(b) Ajuste para  $A_1$ (c) Ajuste para  $A_2$ (d) Ajuste para  $A_3$ 

**Figura 5.22** Imagem PolSAR real com regiões selecionadas e curvas de densidades ajustadas para distribuições Wishart relaxada (ponto tracejada) e original (sólida).



Além disso, os resultados fornecem evidências de que as menores entropias estimadas estão associadas às regiões cujos determinantes das matrizes de covariância estimadas são menores. Este fato pode ser justificado por meio das evidências do estudo gráfico de entropias na discussão sobre as contribuições teóricas: os valores maiores para o determinante das matrizes de covariância estão associados com os maiores valores de entropias, enquanto os maiores números de looks às menores entropias. Uma vez que (i) a entropia é uma medida de variabilidade e (ii) áreas PolSAR com altos valores para o retorno são mais heterogêneas, é intuitiva a seguinte conclusão dos dados: as grandezas “determinante da matriz de covariância” e “entropia” são diretamente proporcionais. Goodman (1963a) definiu o determinante da matriz de covariância como *variância generalizada*, a qual pode ser associada com o grau de heterogeneidade. Assim, embora não explorada na literatura de processamento destas imagens, esta relação não deve ser descartada.

### 1) Estudos com dados sintéticos:

Usando experimentos Monte Carlo e parâmetros obtidos de dados reais, os testes estatísticos baseados nas entropias de Shannon e Rényi (de ordem  $\beta = 0.1, 0.8$ ) foram avaliados. Com esse objetivo, dados sintéticos seguindo uma distribuição Wishart complexa escalonada foram considerados conforme algoritmo proposto por Lee, Grunes & Kwok (1994). Esta simulação empregou os seguintes parâmetros: (i)  $L = 3.2$ ; (ii) matriz de covariância estimada das regiões  $A_1$ ,  $A_2$  e  $A_3$  da Figura 5.22(a) (discutida adiante), referenciadas como  $\tilde{\Sigma}_1$ ,  $\tilde{\Sigma}_2$  e  $\tilde{\Sigma}_3$ , respectivamente; (iii) tamanhos de amostras  $N = \{9, 49, 81, 121, 400\}$  (isto é, janelas quadradas de tamanhos  $\{3, 7, 9, 11, 20\}$ ); e (iv) níveis de significância  $\alpha \in \{1\%, 5\%, 10\%\}$ . Uma vez geradas duas amostras independentes de tamanhos  $N$  das matrizes populacionais  $\mathbf{X} \sim \mathcal{W}(\tilde{\Sigma}_i, L)$  e  $\mathbf{Y} \sim \mathcal{W}(\tilde{\Sigma}_j, L)$ , nós testamos a hipótese nula  $(\tilde{\Sigma}_i, L) = (\tilde{\Sigma}_j, L)$  para  $i, j = 1, 2, 3$ . Seguindo Nascimento et al. (2010), nós utilizamos 5500 replicações de Monte Carlo para todos os casos. Como critérios de avaliação, valores para os tamanhos e poderes dos testes propostos foram considerados.

A Tabela 5.13 apresenta os resultados para os tamanhos empíricos dos testes. Do ponto de vista estatístico, a situação ideal para os dados sintéticos é ter um teste com tamanhos empíricos similares aos níveis nominais especificados. Assim, nós concluímos que o teste sobre a entropia de Rényi para  $\beta = 0.1$  é geralmente superado pelos outros testes, exceto para  $N = 9$ . Os testes baseados sobre as entropias de Shannon e Rényi com ordem 0.8 apresentaram valores próximos aos nominais, até mesmo para tamanhos de amostras pequenos.

Para as situações consideradas, as amostras de diferentes matrizes de covariância sempre rejeitaram a hipótese nula com os testes estatísticos propostos, resultando em poderes do teste estimados com valores 1. A Tabela 5.14 mostra os valores médios das

**Tabela 5.13** Tamanhos empíricos do teste

| N                                   | A <sub>1</sub> |      |       | A <sub>2</sub> |      |       | A <sub>3</sub> |      |       |
|-------------------------------------|----------------|------|-------|----------------|------|-------|----------------|------|-------|
|                                     | 1%             | 5%   | 10%   | 1%             | 5%   | 10%   | 1%             | 5%   | 10%   |
| Entropia de Shannon                 |                |      |       |                |      |       |                |      |       |
| 9                                   | 1.35           | 6.15 | 11.76 | 1.36           | 6.16 | 12.02 | 1.44           | 6.20 | 12.05 |
| 49                                  | 1.38           | 5.53 | 10.44 | 1.38           | 5.65 | 10.35 | 1.35           | 5.47 | 10.44 |
| 81                                  | 1.33           | 5.51 | 9.75  | 1.35           | 5.47 | 9.84  | 1.35           | 5.36 | 10.07 |
| 121                                 | 1.18           | 5.56 | 10.58 | 1.15           | 5.51 | 10.51 | 1.15           | 5.49 | 10.44 |
| 400                                 | 1.35           | 5.38 | 10.33 | 1.38           | 5.47 | 10.40 | 1.40           | 5.51 | 10.22 |
| Entropia de Rényi com $\beta = 0.8$ |                |      |       |                |      |       |                |      |       |
|                                     | 1.38           | 6.22 | 11.65 | 1.44           | 6.18 | 11.93 | 1.45           | 6.09 | 11.78 |
|                                     | 1.27           | 5.62 | 10.36 | 1.31           | 5.65 | 10.31 | 1.33           | 5.38 | 10.55 |
|                                     | 1.31           | 5.47 | 9.76  | 1.20           | 5.47 | 9.91  | 1.29           | 5.36 | 9.95  |
|                                     | 1.05           | 5.35 | 10.45 | 1.15           | 5.44 | 10.44 | 1.13           | 5.53 | 10.55 |
|                                     | 1.35           | 5.27 | 10.31 | 1.36           | 5.29 | 10.29 | 1.27           | 5.33 | 10.42 |
| Entropia de Rényi com $\beta = 0.1$ |                |      |       |                |      |       |                |      |       |
|                                     | 0.69           | 4.04 | 8.80  | 0.76           | 4.15 | 8.65  | 0.82           | 4.04 | 8.55  |
|                                     | 0.71           | 3.62 | 7.82  | 0.75           | 3.87 | 7.89  | 0.69           | 3.45 | 7.73  |
|                                     | 0.53           | 3.69 | 7.51  | 0.56           | 3.84 | 7.49  | 0.55           | 3.62 | 7.51  |
|                                     | 0.60           | 3.73 | 7.58  | 0.60           | 3.40 | 7.47  | 0.53           | 3.58 | 7.64  |
|                                     | 0.56           | 3.84 | 7.65  | 0.67           | 3.82 | 7.64  | 0.67           | 3.62 | 7.65  |

estatísticas

$$\bar{S}(A_i, A_j) = 5500^{-1} \sum_{k=1}^{5500} S_{\phi}^h([\hat{\Sigma}_i, \hat{L}_i]_k, [\hat{\Sigma}_j, \hat{L}_j]_k),$$

em que  $S_{\phi}^h$  é  $S_S$  ou  $S_R$  como dado na Tabela 4.1 e  $[\hat{\Sigma}_i, \hat{L}_i]_k$  são as estimativas MV baseadas sobre os dados gerados da população  $A_i$  na  $k$ -ésima replicação de Monte Carlo, e o coeficiente de variação (CV) das estatísticas na hipótese alternativa:

$$CV(A_i, A_j) = \bar{S}(A_i, A_j)^{-1} \sqrt{5500^{-1} \sum_{k=1}^{5500} [S_{\phi}^h([\hat{\Sigma}_i, \hat{L}_i]_k, [\hat{\Sigma}_j, \hat{L}_j]_k) - \bar{S}(A_i, A_j)]^2}.$$

**Tabela 5.14** Média e coeficiente de variação sob a hipótese alternativa

| N                                   | A <sub>1</sub> -A <sub>2</sub> |           | A <sub>1</sub> -A <sub>3</sub> |           | A <sub>2</sub> -A <sub>3</sub> |           |
|-------------------------------------|--------------------------------|-----------|--------------------------------|-----------|--------------------------------|-----------|
|                                     | CV                             | $\bar{S}$ | CV                             | $\bar{S}$ | CV                             | $\bar{S}$ |
| Entropia de Shannon                 |                                |           |                                |           |                                |           |
| 9                                   | 20.59                          | 108.16    | 13.83                          | 248.53    | 38.42                          | 29.89     |
| 49                                  | 8.65                           | 576.23    | 5.82                           | 1329.19   | 16.38                          | 156.21    |
| 81                                  | 6.70                           | 949.51    | 4.52                           | 2191.84   | 12.71                          | 256.54    |
| 121                                 | 5.51                           | 1418.22   | 3.71                           | 3274.03   | 10.43                          | 383.17    |
| 400                                 | 3.01                           | 4687.29   | 2.02                           | 10817.80  | 5.72                           | 1265.69   |
| Entropia de Rényi com $\beta = 0.8$ |                                |           |                                |           |                                |           |
|                                     | 20.56                          | 106.72    | 13.72                          | 245.22    | 38.45                          | 29.50     |
|                                     | 8.63                           | 569.69    | 5.76                           | 1314.11   | 16.40                          | 154.43    |
|                                     | 6.67                           | 938.93    | 4.47                           | 2167.39   | 12.72                          | 253.68    |
|                                     | 5.48                           | 1402.43   | 3.66                           | 3237.61   | 10.42                          | 378.93    |
|                                     | 3.00                           | 4635.65   | 2.00                           | 10698.61  | 5.71                           | 1251.74   |
| Entropia de Rényi com $\beta = 0.1$ |                                |           |                                |           |                                |           |
|                                     | 21.70                          | 78.42     | 14.27                          | 180.12    | 40.60                          | 21.79     |
|                                     | 9.11                           | 423.29    | 6.00                           | 976.32    | 17.44                          | 114.77    |
|                                     | 7.02                           | 698.47    | 4.63                           | 1612.06   | 13.50                          | 188.81    |
|                                     | 5.73                           | 1043.26   | 3.77                           | 2408.63   | 11.00                          | 282.06    |
|                                     | 3.15                           | 3450.84   | 2.07                           | 7963.98   | 6.03                           | 931.79    |

Neste caso a maior média  $\bar{S}(A_i, A_j)$  corresponde ao maior valor  $\|\tilde{\Sigma}_i\| - \|\tilde{\Sigma}_j\|$ , resultando na seguinte relação:

$$\bar{S}(A_1, A_3) > \bar{S}(A_1, A_2) > \bar{S}(A_2, A_3). \quad (5.23)$$

A estatística de Shannon produz maiores valores quando contrastamos amostras de áreas diferentes, indicando que esta estatística tem um maior potencial de discriminação.

## 2) Aplicação a dados reais:

Aplicando a metodologia descrita a dados reais, nós obtivemos os resultados apresentados nas Figuras 5.23 e 5.25. Para este fim, as seguintes etapas foram seguidas:

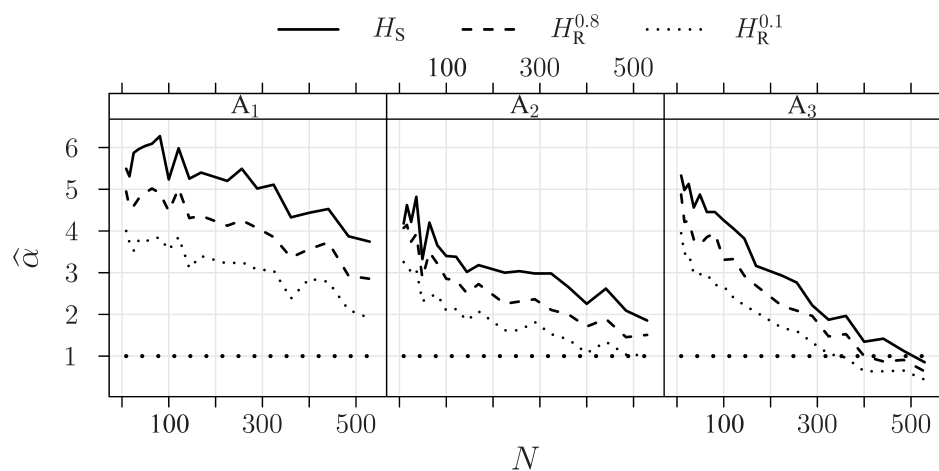
1. Para  $j = 1, 2, \dots, 5500$  realize os passos 2) a 5).
2. Amostre duas regiões  $U_j$  e  $V_j$  das áreas  $A_1, A_2$ , e  $A_3$ .
3. Utilizando uma amostragem sem reposição, gere dois vetores com  $N$  pontos,  $u^{(j)}$  e  $v^{(j)}$  de  $U_j$  e  $V_j$ , respectivamente, garantindo que  $u^{(i)} \neq u^{(j)}$  e  $v^{(i)} \neq v^{(j)}$  para  $i \neq j$ .
4. Estime os vetores de parâmetros  $\hat{\theta}_1^{(j)}$  e  $\hat{\theta}_2^{(j)}$  baseado sobre  $u^{(j)}$  e  $v^{(j)}$ , respectivamente.
5. Tome decisão baseado na Proposição 1 para  $\alpha = \{1\%, 5\%, 10\%\}$ .
6. Seja  $T$  o número de vezes em que a hipótese nula é rejeitada. Calcule o tamanho empírico do teste ( $\hat{\alpha}_{1-\alpha}$ ) e o poder ( $1 - \hat{\beta}_{1-\alpha}$ ) no nível  $\alpha$  quando

$$\begin{cases} \hat{\alpha}_{1-\alpha} = T/5500, & \text{se } V_j = U_j, \\ 1 - \hat{\beta}_{1-\alpha} = T/5500, & \text{se } V_j \neq U_j. \end{cases}$$

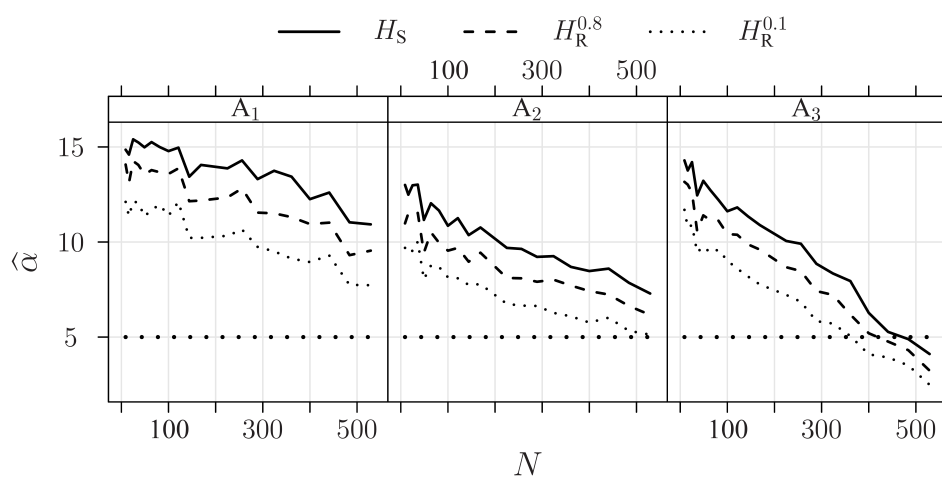
Nós consideramos  $N$  o número de observações nas janelas quadradas  $\{3 \times 3, 4 \times 4, \dots, 23 \times 23\}$ .

A Figura 5.23 apresenta os tamanhos empíricos dos testes, com o tamanho nominal  $\alpha$  representado pela linha pontilhada. O comportamento de  $\hat{\alpha}_{1-\alpha}$  é consistente através dos três valores de  $\alpha$ : o tamanho empírico do teste é maior para a entropia de Shannon do que para a de Rényi com  $\beta = 0.8$ , a qual é maior do que a de Rényi com ordem  $\beta = 0.1$ . A mais rápida convêrgencia a  $\alpha$  com respeito aos tamanhos de amostras é consistentemente observada na área  $A_3$ . Menhum teste alcança o nível nominal especificado na área  $A_1$ . Este fato é devido ao brilho pronunciado desta região, como notado na Tabela 5.11.

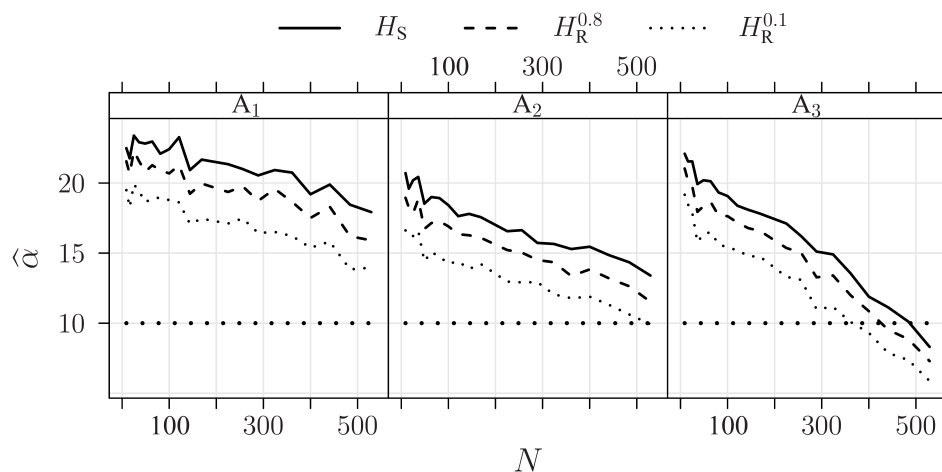
No estudo de poder do teste, nós objetivamos selecionar imagens visualmente mais similares do que as coletadas na Figura 5.22, mas com propriedades estatísticas mais distintas. A este fim, a Figura 5.24 (a) mostra uma imagem EMISAR de Foulum (Dinamarca) (para maiores detalhes, consulte o Apêndice A), a qual foi obtida no número de looks 8. Essa imagem representa áreas de agricultura e, para nosso estudo, três regiões foram selecionadas. A Tabela 5.15 apresenta uma breve análise descritiva das estimativas MV das regiões amostradas. A Figuras 5.24(b)-(d) fornecem as densidades empíricas para diferentes áreas e canais de polarização. Como esperado dos valores assumidos pelas estimativas de  $L$ , nota-se que as densidades empíricas de  $B_2$  and  $B_3$



(a) Tamanho a 1%



(b) Tamanho a 5%



(c) Tamanho a 10%

**Figura 5.23** Tamanho empírico do teste para diferentes regiões em vários tamanhos de amostras.

são mais próximas para todos os canais de polarização. A seguir, estuda-se as estimativas para o poder dos testes considerados.

Baseando-se no algoritmo apresentado no início desta aplicação, nós apresentamos o estudo do poder do teste entre as regiões selecionadas para as estatísticas  $S_S$  e  $S_R^{0.1}$ . A Figuras 5.25(a)–5.25(c) fornecem estimativas para o poder do teste considerando pares de amostras nas regiões selecionadas. Em termos gerais, o poder estimado aumenta rapidamente quando o tamanho de amostra cresce. O teste baseado sobre a entropia de Shannon supera aqueles obtidos para o teste na entropia de Rényi. Esses resultados sugerem o teste baseado na entropia de Shannon como uma ferramenta mais apropriada para detectar bordas e mudanças em imagens PolSAR do que aquele em função da entropia de Rényi.

**Tabela 5.15** estimativas MV

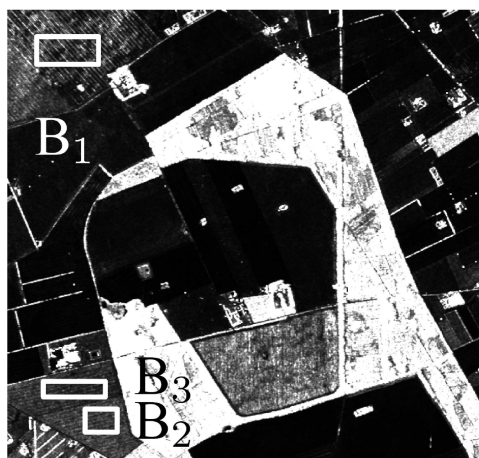
| Regiões | $N$  | $ \hat{\Sigma} $       | $\hat{L}$ |
|---------|------|------------------------|-----------|
| $B_1$   | 2280 | $2.333 \times 10^{-5}$ | 6.060     |
| $B_2$   | 1408 | $1.112 \times 10^{-6}$ | 11.937    |
| $B_3$   | 1848 | $5.814 \times 10^{-7}$ | 10.752    |

### 5.5.2 Considerações da Seção 5.5

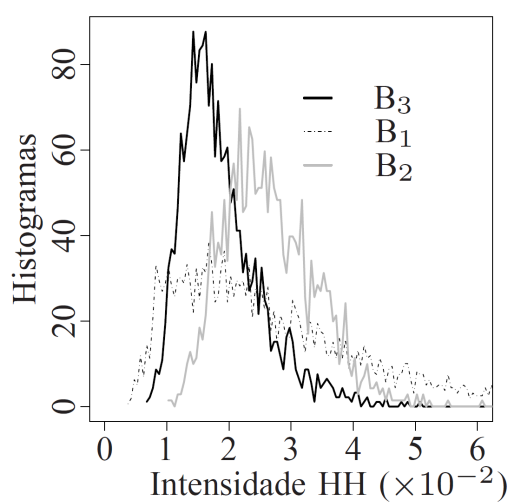
Nesta seção, nós derivamos expressões para as entropias de Shannon, Tsallis e Rényi para distribuição Wishart complexa escalonada. Estes parâmetros, funções do número de looks e da matriz de covariância, carregam o efeito dos parâmetros da lei Wishart complexa escalonada sobre o grau de desordem. Além disso, avaliando as entropias nos estimadores de MV para  $L$  e  $\Sigma$ , as variâncias assintóticas de seus estimadores de MV foram derivadas. Baseando-se nos resultados acima, novas metodologias para testar hipótese e construir intervalos de confiança baseadas na classe de entropias  $(h, \phi)$  foram propostas. Tais métodos consistem de testes simples para checar se duas amostras matriciais Hermitianas positivas definidas vem de uma mesmas distribuição.

Considerando dados reais de um imagem E-SAR, os intervalos de confiança baseados em entropias se mostraram válidos para identificar regiões distintas com 95% de confiança. Evidências com dados reais indicam que os valores das entropias são maiores para dados com maiores determinantes da covariância amostral, isto é, com maiores variâncias generalizadas. Baseando-se em um estudo de Monte Carlo, a performance dos testes foi quantificada. Adotando como critério de comparação seus tamanhos, o teste baseado na entropia de Shannon apresentou melhores resultados do que o dependente da entropia de Rényi. Em termos de poder, é possível verificar que a estatística de Shannon tem um poder de discriminação maior.

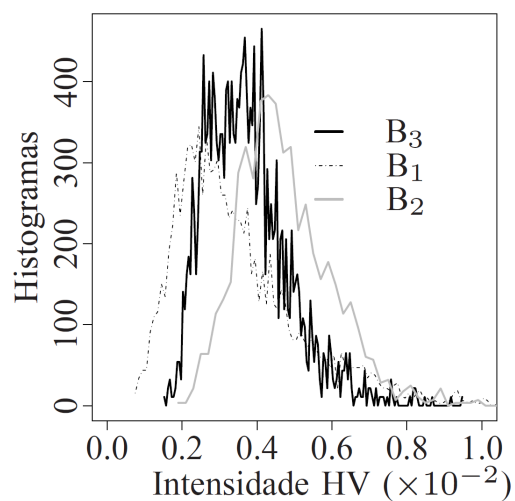
Duas aplicações a dados reais foram realizadas. Os resultados se mostraram alinhados com os experimentos em dados sintéticos. Desta forma, sugerimos a entropia



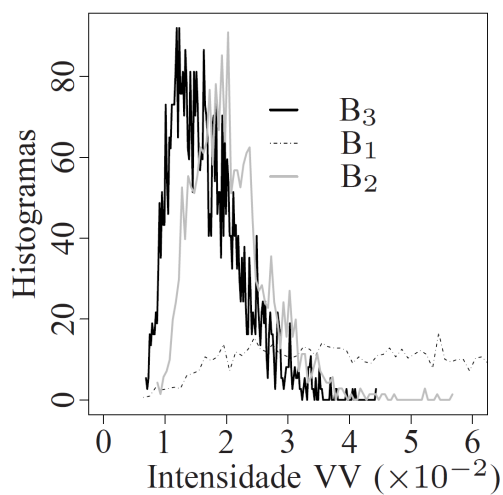
(a) Imagem EMISAR de Follum



(b) Canal HH

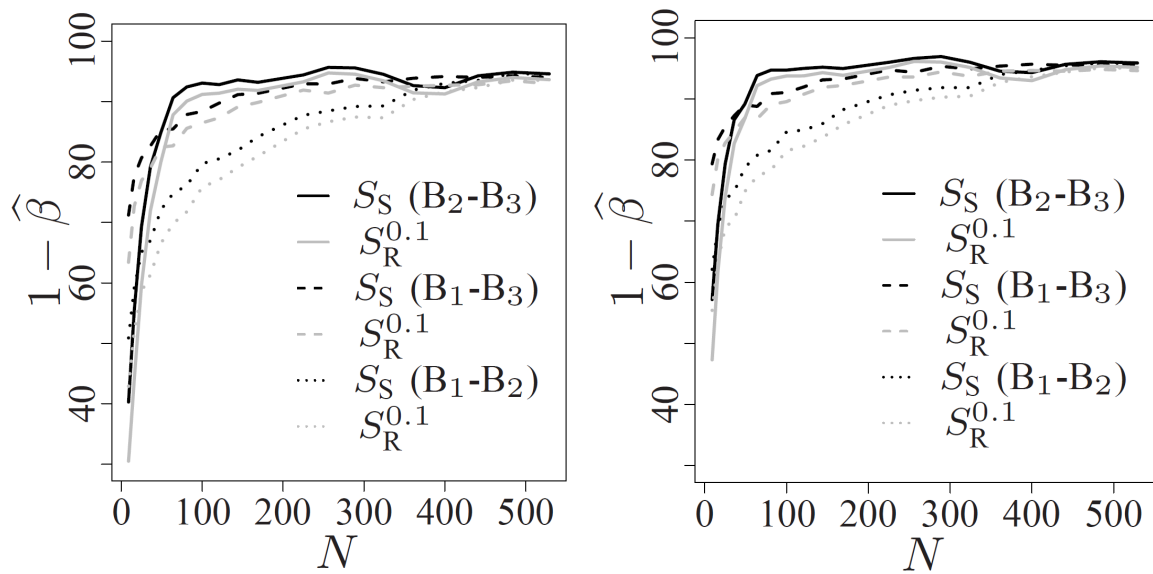


(c) Canal HV



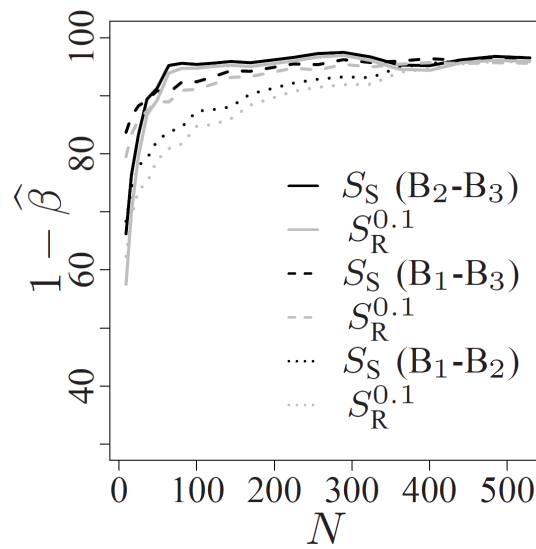
(d) Canal VV

**Figura 5.24** Imagem PolSAR com regiões selecionadas e densidades empíricas para várias áreas e canais de polarização.



(a) 1%

(b) 5%



(c) 10%

**Figura 5.25** Poder empírico do teste para diferentes regiões nos tamanhos de amostras considerados.



de Shannon como a melhor fonte de discriminação e a entropia de Rényi com ordem 0.1 como mais apta a reconhecer homogeneidade.

### 5.5.3 Trabalhos futuros

De uma crítica científica desta seção, as seguintes frentes de trabalhos podem ser consideradas:

- ✎ Estudo destas medidas no contexto de qualidade de ajuste para dados PolSAR e sua comparação com outras propostas, tais com as sugeridas por Anfinsen et al. (2011);
- ✎ Testes para classificar um pixels PolSAR como uma reflexão de canto (*outlier*) em imagens PolSAR;
- ✎ Proposta dos testes derivados a múltiplas populações ( $> 2$ ).

---

## Anexo da Seção 5.5

### ARTIGO 4:

“Entropy-based Statistical Analysis of PolSAR Data”

Submitted to *IEEE Transaction Geoscience and Remote Sensing*

---

# Entropy-based Statistical Analysis of PolSAR Data

Alejandro C. Frery, *Member*, Renato J. Cintra, *Senior Member*, and Abraão D. C. Nascimento, *Student Member*

**Abstract**—Identifying the image contrast between different regions is a pivotal task in image processing. Images obtained from coherent illumination processes are inexorably contaminated with speckle noise. Multilook polarimetric synthetic aperture radar (PolSAR) imagery is a prominent example. With an adequacy widely attested in the literature, the scaled complex Wishart distribution has been suggested as a acceptable model for PolSAR data. In this perspective, we derive analytic expressions for the Shannon, Rényi, and restricted Tsallis entropies under such statistical model, including the asymptotic variance of the first two when replacing distribution parameters by maximum likelihood estimators. Relationships of derived measures with the parameters of the scaled Wishart law (equivalent number of looks and covariance matrix) are discussed. Subsequently, confidence intervals based on entropies are also proposed and indicated as a new methodology of capturing contrast. New hypothesis tests are additionally proposed using these results, and their performance is assessed using simulated and real data. Finally, in general terms, the test based on the Shannon entropy outperforms one proposed by Rényi in considered cases.

**Index Terms**—information theory, SAR polarimetry, contrast measures.

## I. INTRODUCTION

**P**OLARIMETRIC synthetic aperture radar (PolSAR) has been used to describe earth surface phenomena [1]. This technology uses coherent illumination which causes the interference pattern called ‘speckle’ [2], which is non-Gaussian and multiplicative in nature. This fact precludes the use of conventional tools to PolSAR image analysis, requiring specialized techniques.

The scaled complex Wishart has been successfully employed as a statistical model for homogeneous regions in PolSAR images [3]. This distribution is at the core

of segmentation [4], classification [5], and boundary detection [6] techniques.

The concepts of ‘information’ and ‘entropy’ were given formal mathematical definitions in the context of data communications by Shannon in 1948 [7]. Thenceforth, the proposition and application of information and entropy measures have become an active research field in several areas. Zografos and Nadarajah proposed closed expressions for Shannon and Rényi entropies for several univariate [8], bivariate [9], and multivariate [10] distributions.

Exploring relationships associated to the log-likelihood function, Zong [11] applied the Rényi entropy to several univariate distributions. In fact, this entropy measure has been applied to image processing problems such as data mining, detection, segmentation, classification [12]–[14]. Another prominent entropy measure is the restricted Tsallis entropy. This tool was introduced by Tsallis in [15], [16] and is related to the Rényi entropy. The restricted Tsallis entropy has found applications in statistical physics [17].

Among these information theoretical tools, the Shannon entropy has been applied to PolSAR imagery. Morio *et al.* [18] analyzed such entropy for the characterization of polarimetric targets under the circular multidimensional Gaussian distribution. Subsequently, this measure was compared with the von Neumann entropy by Cloude and Pottier [19], [20] for classification purposes.

Additionally, in order to apply information theory tools to extract conclusions from speckle data, the utilization of stochastic distances has been sought. A comprehensive examination of these measures is presented and applied to intensity SAR data in [21], [22].

In this paper, we derive analytic expressions for the Shannon, Rényi, and restricted Tsallis entropies under the scaled complex Wishart law. These measures are analyzed as particular cases of the  $(h, \phi)$ -entropy proposed by Salicrú *et al.* [23]. Considering the Shannon and Rényi entropies and replacing their parameters by maximum likelihood (ML) estimators, expressions for the asymptotic variances of the entropy estimates are derived. Novel methodologies for testing hypotheses and constructing confidence intervals are proposed for quan-

This work was supported by CNPq, Fapeal and FACEPE, Brazil.

A. C. Frery is with the Instituto de Computação, Universidade Federal de Alagoas, BR 104 Norte km 97, 57072-970, Maceió, AL, Brazil, email: acfrery@gmail.com

R. J. Cintra and A. D. C. Nascimento are with the Departamento de Estatística, Universidade Federal de Pernambuco, Cidade Universitária, 50740-540, Recife, PE, Brazil, e-mail: rjdc@stat.ufpe.org, abraao.susej@gmail.com

tifying ‘contrast’ in PolSAR imagery. Such measures can also be adapted to proposing the new SAR image processing techniques, such as segmentation [4], classification [5], boundary detection [24], [25], and change detection [26] methods. Monte Carlo experiments are performed for assessing the performance of the discussed measures in synthetic data, and an application to real PolSAR data is performed.

The remainder of this paper is organized as follows. Section II recalls the scaled Wishart law. Selected information theoretic tools are summarized in Section III. Section IV presents the proposed entropies and the asymptotic variance of their estimates, along with an application. Section VI concludes the paper.

## II. THE COMPLEX WISHART DISTRIBUTION

Full-polarimetric SAR sensors record the four combinations of the received and transmitted complex linear polarizations:  $S_{HH}$  (horizontal-horizontal),  $S_{HV}$  (horizontal-vertical),  $S_{VH}$  (vertical-horizontal), and  $S_{VV}$  (vertical-vertical). When natural targets are considered [27], the conditions of the reciprocity theorem [2], [28] are satisfied and it can be assumed that  $S_{HV} = S_{VH}$ .

In general, we may consider systems with  $m$  polarization elements, which constitute a complex random vector denoted by:

$$\mathbf{y} = [S_1 \ S_2 \ \cdots \ S_m]^t, \quad (1)$$

where  $(\cdot)^t$  is the transposition operator. It is commonly assumed that the scattering vector  $\mathbf{y}$  is distributed according to the circular complex Gaussian law [29].

Multilook PolSAR data are usually formed in order to enhance the signal-to-noise ratio (SNR):

$$\mathbf{Z} = \frac{1}{L} \sum_{i=1}^L \mathbf{y}_i \mathbf{y}_i^H,$$

where  $(\cdot)^H$  is the Hermitian operator,  $\mathbf{y}_i$  represents the scattering vector in the  $i$ th look, and  $L$  is the number of looks which is a parameter that controls the noise effect on PolSAR images. Matrix  $\mathbf{Z}$  is defined over the set of positive-definite Hermitian matrices  $\mathcal{A}$ . Moreover, Goodman showed that  $L\mathbf{Z}$  follows the ordinary complex Wishart law [30], whose density of the scaled random matrix  $\mathbf{Z}$  is

$$f_{\mathbf{Z}}(\mathbf{Z}'; \boldsymbol{\Sigma}, L) = \frac{L^m |\mathbf{Z}'|^{L-m}}{|\boldsymbol{\Sigma}|^L \Gamma_m(L)} \exp[-L \operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}')], \quad (2)$$

where  $m$  is the order of  $\boldsymbol{\Sigma}$ ,  $m \leq L$ ,  $\Gamma_m(L) = \pi^{m(m-1)/2} \prod_{k=0}^{m-1} \Gamma(L-k)$ ,  $\Gamma(\cdot)$  is the gamma function,  $\boldsymbol{\Sigma} = \mathbb{E}(\mathbf{y}\mathbf{y}^H)$ , and  $\mathbb{E}\{\cdot\}$  is the expectation operator. This is denoted as  $\mathbf{Z} \sim \mathcal{W}_m(\boldsymbol{\Sigma}, L)$ .

## III. INFORMATION THEORY

Information theory has provided important tools for statistical inference [31], data compression [32], and image processing [18], to name a few applications. In particular, the entropy is a fundamental concept related to the notion of disorder in mechanical statistics [33]. Salicrú *et al.* [23] proposed the  $(h, \phi)$ -entropy class, which generalizes the concept. In the following, we recall this definition and we derive entropies for positive-definite Hermitian random matrices.

Let  $f_{\mathbf{Z}}(\mathbf{Z}'; \boldsymbol{\theta})$  be a probability density function with parameter vector  $\boldsymbol{\theta}$  which characterizes the distribution of the random matrix  $\mathbf{Z}$ . The  $(h, \phi)$ -entropy relative to  $\mathbf{Z}$  is defined by

$$H_{\phi}^h(\boldsymbol{\theta}) = h\left(\int_{\mathcal{A}} \phi(f_{\mathbf{Z}}(\mathbf{Z}'; \boldsymbol{\theta})) d\mathbf{Z}'\right),$$

where either  $\phi : [0, \infty) \rightarrow \mathbb{R}$  is concave and  $h : \mathbb{R} \rightarrow \mathbb{R}$  is increasing, or  $\phi$  is convex and  $h$  is decreasing. The differential element  $d\mathbf{Z}'$  is given by

$$d\mathbf{Z}' = \prod_{i=1}^m dZ_{ii} \prod_{\substack{i,j=1 \\ i < j}}^m d\Re\{Z_{ij}\} d\Im\{Z_{ij}\},$$

where  $Z_{ij}$  is the  $(i, j)$ -th entry of matrix  $\mathbf{Z}'$ , and  $\Re$  and  $\Im$  denote the real and imaginary parts, respectively [30]. Table I shows the specification of the three entropies we use in this article: Shannon, Rényi, and restricted Tsallis.

TABLE I  
( $h, \phi$ )-ENTROPIES AND RELATED FUNCTIONS

| ( $h, \phi$ )-entropy                                                    | $h(y)$                  | $\phi(x)$                     |
|--------------------------------------------------------------------------|-------------------------|-------------------------------|
| Shannon [23]                                                             | $y$                     | $-x \ln x$                    |
| Restricted Tsallis (order $\beta \in \mathbb{R}_+ : \beta \neq 1$ ) [34] | $y$                     | $\frac{x^\beta - x}{1-\beta}$ |
| Rényi (order $\beta \in \mathbb{R}_+ : \beta \neq 1$ ) [35]              | $\frac{\ln y}{1-\beta}$ | $x^\beta$                     |

The following result, derived by Pardo *et al.* [36], paves the way for the proposal of asymptotic statistical inference methods based on entropy.

*Lemma 1:* Let  $\hat{\boldsymbol{\theta}} = [\hat{\theta}_1 \ \hat{\theta}_2 \ \cdots \ \hat{\theta}_p]^t$  be the ML estimate of the parameter vector  $\boldsymbol{\theta} = [\theta_1 \ \theta_2 \ \cdots \ \theta_p]^t$  based on a random sample of size  $N$  under the model  $f(\mathbf{Z}'; \boldsymbol{\theta})$ . Then

$$\sqrt{N} [H_{\hat{\boldsymbol{\theta}}}^{\phi}(\hat{\boldsymbol{\theta}}) - H_{\hat{\boldsymbol{\theta}}}^{\phi}(\boldsymbol{\theta})] \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \sigma_H^2(\boldsymbol{\theta})),$$

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

$$\sigma_H^2(\boldsymbol{\theta}) = \boldsymbol{\delta}^t \mathcal{K}(\boldsymbol{\theta})^{-1} \boldsymbol{\delta}, \quad (3)$$

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

In the following we introduce a methodology for hypothesis tests and confidence intervals based on entropy.

#### A. Hypothesis test

Let  $\mathbf{Z}_i$  be a positive-definite random matrix with probability density function defined over  $\mathcal{A}$  with parameter vector  $\boldsymbol{\theta}_i$  for  $i = 1, 2, \dots, r$ , where  $r$  is the number of populations to be assessed. We are interested in testing the following hypotheses:

$$\begin{cases} \mathcal{H}_0: H_h^\phi(\boldsymbol{\theta}_1) = H_h^\phi(\boldsymbol{\theta}_2) = \dots = H_h^\phi(\boldsymbol{\theta}_r) = v, \\ \mathcal{H}_1: H_h^\phi(\boldsymbol{\theta}_i) \neq H_h^\phi(\boldsymbol{\theta}_j) \text{ for some } i \text{ and } j. \end{cases}$$

In other words, statistical evidence is sought for assessing whether at least one of the  $r$  regions of a PolSAR image has different entropy when compared to the remaining regions.

Let  $\hat{\boldsymbol{\theta}}_i$  be the ML estimate for  $\boldsymbol{\theta}_i$  based on a random sample of size  $N_i$  under  $\mathbf{Z}_i$ , for  $i = 1, 2, \dots, r$ . From Lemma 1 we have that

$$\frac{\sqrt{N_i}(H_h^\phi(\hat{\boldsymbol{\theta}}_i) - v)}{\sigma_H(\hat{\boldsymbol{\theta}}_i)} \xrightarrow[N_i \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, 1)$$

for  $i = 1, 2, \dots, r$ . Therefore,

$$\sum_{i=1}^r \frac{N_i(H_h^\phi(\hat{\boldsymbol{\theta}}_i) - v)^2}{\sigma_H^2(\hat{\boldsymbol{\theta}}_i)} \xrightarrow[N_i \rightarrow \infty]{\mathcal{D}} \chi_r^2. \quad (4)$$

Since  $v$  is, in practice, unknown, in the following we modify this test statistic in order to take this into account. Considering an application of Cochran's theorem [37], we obtain:

$$\begin{aligned} \sum_{i=1}^r \frac{N_i(H_h^\phi(\hat{\boldsymbol{\theta}}_i) - v)^2}{\sigma_H^2(\hat{\boldsymbol{\theta}}_i)} &= \\ \sum_{i=1}^r \frac{N_i(H_h^\phi(\hat{\boldsymbol{\theta}}_i) - \bar{v})^2}{\sigma_H^2(\hat{\boldsymbol{\theta}}_i)} + \sum_{i=1}^r \frac{N_i(\bar{v} - v)^2}{\sigma_H^2(\hat{\boldsymbol{\theta}}_i)}, \end{aligned} \quad (5)$$

where

$$\bar{v} = \left[ \sum_{i=1}^r \frac{N_i}{\sigma_H^2(\hat{\boldsymbol{\theta}}_i)} \right]^{-1} \sum_{i=1}^r \frac{N_i H_h^\phi(\hat{\boldsymbol{\theta}}_i)}{\sigma_H^2(\hat{\boldsymbol{\theta}}_i)}.$$

Salicrú *et al.* [23] showed that the second summation in the right-hand side of Equation (5) is chi-square distributed with one degree of freedom. Since the left-hand side of (5) is chi-square distributed with  $r$  degrees of freedom (cf. Equation (4)), we conclude that:

$$\sum_{i=1}^r \frac{N_i(H_h^\phi(\hat{\boldsymbol{\theta}}_i) - \bar{v})^2}{\sigma_H^2(\hat{\boldsymbol{\theta}}_i)} \xrightarrow[N_i \rightarrow \infty]{\mathcal{D}} \chi_{r-1}^2.$$

In particular, consider the following test statistic:

$$S_\phi^h(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2, \dots, \hat{\boldsymbol{\theta}}_r) = \sum_{i=1}^r \frac{N_i(H_h^\phi(\hat{\boldsymbol{\theta}}_i) - \bar{v})^2}{\sigma_H^2(\hat{\boldsymbol{\theta}}_i)}. \quad (6)$$

We are now in the position to state the following result.

*Proposition 1:* Let  $N_i$ ,  $i = 1, 2, \dots, r$ , be sufficiently large. If  $S_\phi^h(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2, \dots, \hat{\boldsymbol{\theta}}_r) = s$ , then the null hypothesis  $\mathcal{H}_0$  can be rejected at a level  $\alpha$  if  $\Pr(\chi_{r-1}^2 > s) \leq \alpha$ .

#### B. Confidence intervals

Let  $\hat{\boldsymbol{\theta}}$  be the ML estimate of  $\boldsymbol{\theta}$  for a sufficiently large sample  $N$ . An approximate confidence interval for  $H_h^\phi(\boldsymbol{\theta})$  at nominal level  $\alpha$  is

$$H_h^\phi(\hat{\boldsymbol{\theta}}) \pm z_{\alpha/2} \sqrt{\frac{\sigma_H^2(\hat{\boldsymbol{\theta}})}{N}}, \quad (7)$$

where  $z_{\alpha/2}$  is the  $\alpha/2$  quantile of the standard Gaussian distribution.

Consider now  $\hat{\boldsymbol{\theta}}_1$  and  $\hat{\boldsymbol{\theta}}_2$ , ML estimates based on large samples  $N_1$  and  $N_2$ , respectively. An approximate confidence interval for  $H_h^\phi(\boldsymbol{\theta}_1) - H_h^\phi(\boldsymbol{\theta}_2)$  is given by [36]

$$[H_h^\phi(\hat{\boldsymbol{\theta}}_1) - H_h^\phi(\hat{\boldsymbol{\theta}}_2)] \pm z_{\alpha/2} \sqrt{\frac{\sigma_H^2(\hat{\boldsymbol{\theta}}_1)}{N_1} + \frac{\sigma_H^2(\hat{\boldsymbol{\theta}}_2)}{N_2}}.$$

## IV. RESULTS

In the following, we derive results for the Shannon, Rényi, and restricted Tsallis entropies, denoted as  $H_S$ ,  $H_R^\beta$ , and  $H_T^\beta$ , respectively, under the scaled complex Wishart law. In particular, these measures are algebraically expressed, numerically evaluated, and assessed. We adopt  $\boldsymbol{\theta} = [L, \text{vec}(\boldsymbol{\Sigma})^t]^t$ , where  $\text{vec}(\cdot)$  is the vectorisation operator, as the working parameter vector. Additionally, asymptotic results for these measures are computed. To that end, we derive analytic expressions for the variance of the considered entropies. The entropies were evaluated at the ML estimate values. Subsequently, hypothesis tests and confidence intervals are proposed based on Shannon and Rényi entropies.

#### A. Expressions for the complex Wishart distribution

To the following, we present three schemes concerning the derivations of Shannon, restricted Tsallis, and Rényi entropies.

1) *Shannon entropy*: Applying the expression of the Shannon entropy obtained by  $h$  and  $\phi$  functions in Table I to the density given in Equation (2), we have that

$$\begin{aligned} H_S(\boldsymbol{\theta}) &= - \int_{\mathcal{A}} f_{\mathbf{Z}}(\mathbf{Z}'; \boldsymbol{\Sigma}, L) \ln f_{\mathbf{Z}}(\mathbf{Z}'; \boldsymbol{\Sigma}, L) d\mathbf{Z}' \\ &= E\{-\ln f_{\mathbf{Z}}(\mathbf{Z})\}. \end{aligned}$$

Thus, minor manipulations yield the following result:

$$\begin{aligned} H_S(\boldsymbol{\theta}) &= -mL \ln L + (m-L) E\{\ln |\mathbf{Z}|\} \\ &\quad + L \ln |\boldsymbol{\Sigma}| + \frac{m(m-1)}{2} \ln \pi + \sum_{k=0}^{m-1} \ln \Gamma(L-k) \\ &\quad + L E\{\text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z})\}. \end{aligned}$$

In [38], Anfinsen *et al.* proposes the following identity:

$$E\{\ln |\mathbf{Z}|\} = \ln |\boldsymbol{\Sigma}| + \psi_m^{(0)}(L) - m \ln L,$$

where  $\psi_m^{(0)}(\cdot)$  represents the term of zero order of the  $v$ th-order multivariate polygamma function given by

$$\psi_m^{(v)}(L) = \sum_{i=0}^{m-1} \psi^{(v)}(L-i),$$

$\psi^{(v)}(\cdot)$  is the ordinary polygamma function expressed by

$$\psi^{(v)}(L) = \frac{\partial^{v+1} \ln \Gamma(L)}{\partial L^{v+1}},$$

for  $v \geq 0$  (in this case,  $\psi^{(0)}(\cdot)$  is known as the digamma function). By the linearity of the expectation operator, the following holds true:

$$E\{\text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z})\} = \text{tr}(\boldsymbol{\Sigma}^{-1} E\{\mathbf{Z}\}) = \text{tr}(\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}) = m.$$

Thus, the Shannon entropy relative to the random variable  $\mathbf{Z} \sim \mathcal{W}_m(\boldsymbol{\Sigma}, L)$  is expressed by

$$\begin{aligned} H_S(\boldsymbol{\theta}) &= \frac{m(m-1)}{2} \ln \pi - m^2 \ln L + m \ln |\boldsymbol{\Sigma}| + mL \\ &\quad + (m-L) \psi_m^{(0)}(L) + \sum_{k=0}^{m-1} \ln \Gamma(L-k). \end{aligned} \quad (8)$$

2) *Restricted Tsallis entropy*: Based on Table I, the restricted Tsallis entropy is defined by

$$H_T^\beta(\boldsymbol{\theta}) = (1-\beta)^{-1}(\tilde{\mu}_\beta - 1),$$

where  $\tilde{\mu}_\beta \triangleq E\{f_{\mathbf{Z}}^{\beta-1}(\mathbf{Z})\}$ . The quantity  $\tilde{\mu}_\beta$  can be explicitly calculated according to the following manipu-

lations:

$$\begin{aligned} \tilde{\mu}_\beta &= \int_{\mathcal{A}} \left[ \frac{L^m L}{|\boldsymbol{\Sigma}|^L \Gamma_m(L)} \right]^\beta |\mathbf{Z}'|^{\beta(L-m)} \\ &\quad \times \exp[-\beta L \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}')] d\mathbf{Z}' \\ &= \left[ \frac{L^m L}{|\boldsymbol{\Sigma}|^L \Gamma_m(L)} \right]^\beta \int_{\mathcal{A}} |\mathbf{Z}'|^{(\beta-1)(L-m)} |\mathbf{Z}'|^{L-m} \\ &\quad \times \exp[-L \text{tr}(\beta \boldsymbol{\Sigma}^{-1} \mathbf{Z}')] d\mathbf{Z}' \\ &= \left[ \frac{L^m L}{|\boldsymbol{\Sigma}|^L \Gamma_m(L)} \right]^\beta \left[ \frac{|\boldsymbol{\Sigma}/\beta|^L \Gamma_m(L)}{L^m L} \right] E\{|\mathbf{X}|^{(1-\beta)(m-L)}\} \\ &= \frac{L^{mL(\beta-1)} \beta^{-mL}}{|\boldsymbol{\Sigma}|^{L(\beta-1)} \Gamma_m^{\beta-1}(L)} E\{|\mathbf{X}|^{(m-L)(1-\beta)}\}, \end{aligned} \quad (9)$$

where  $\mathbf{X} \sim \mathcal{W}_m(\boldsymbol{\Sigma}/\beta, L)$ . Moreover, Anfinsen *et al.* [39] showed that

$$E\{|\mathbf{Z}|^{s-m}\} = \frac{\Gamma_m(L+s-m)}{\Gamma_m(L)} (L^{-m} |\boldsymbol{\Sigma}|)^{s-m}, \quad (11)$$

where  $\mathbf{Z} \sim \mathcal{W}_m(\boldsymbol{\Sigma}, L)$ . Thus, applying (11) in (10), we have that

$$\tilde{\mu}_\beta = \frac{\Gamma_m(L + (1-\beta)(m-L))}{\Gamma_m^\beta(L) \beta^{m[L+(1-\beta)(m-L)]}} \frac{|\boldsymbol{\Sigma}|^{(1-\beta)m}}{L^{m^2(1-\beta)}}. \quad (12)$$

3) *Rényi entropy*: From Table I, the Rényi entropy is given by

$$H_R^\beta(\boldsymbol{\theta}) = (1-\beta)^{-1} \ln \tilde{\mu}_\beta.$$

Notice that this measure also depends on  $\tilde{\mu}_\beta$ , which was already computed in Equation (12).

Therefore, denoting  $q = L + (1-\beta)(m-L)$ , the Rényi entropy is expressed by

$$\begin{aligned} H_R^\beta(\boldsymbol{\theta}) &= \frac{m(m-1)}{2} \ln \pi - m^2 \ln L + m \ln |\boldsymbol{\Sigma}| \\ &\quad - \frac{mq \ln \beta}{1-\beta} + \frac{\sum_{i=0}^{m-1} [\ln \Gamma(q-i) - \beta \ln \Gamma(L-i)]}{1-\beta}. \end{aligned} \quad (13)$$

It is known that, as  $\beta \rightarrow 1$ , both the Rényi [40, p. 676] and Tsallis [41] entropies converge to the Shannon entropy. Thus:

$$\lim_{\beta \rightarrow 1} H_R^\beta(\boldsymbol{\theta}) \stackrel{(a)}{=} H_S(\boldsymbol{\theta}) \stackrel{(b)}{=} \lim_{\beta \rightarrow 1} H_T^\beta(\boldsymbol{\theta}). \quad (14)$$

These convergences hold true regardless the number of looks. In the following, we examine the behavior of  $H_S$ ,  $H_T^\beta$ , and  $H_R^\beta$  in terms of  $\boldsymbol{\Sigma}$  and  $L$ .

#### 4) Case study:

Frery *et al.* [24] observed the following covariance matrix on an urban area from the E-SAR image of Weßling (Bayern, Germany):

$$\Sigma_U = \begin{bmatrix} 962892 & 19171 - 3579\mathbf{i} & -154638 + 191388\mathbf{i} \\ & 56707 & -5798 + 16812\mathbf{i} \\ & & 472251 \end{bmatrix},$$

only the diagonal and the upper triangle values are shown. Fig. 1(a) depicts plots of the discussed entropies for  $3 \leq L \leq 50$ ,  $\beta \in \{0.1, 0.5, 0.8\}$ , and  $\Sigma_U$ . Considering the same interval for the number of looks, Figs. 1(b), 1(e), and 1(c) show the Shannon, Tsallis (of order  $\beta = 1 - 10^{-3}$ ), and Rényi (of order  $\beta = 0.1$ ) entropies for the covariance matrix  $(1+k)\Sigma_U$ ,  $k \in \{0, 0.1, 0.2\}$ , respectively.

Figs. 1(a) and 1(d) illustrate the property stated in (14). In the case shown here, the convergences are from above, i.e.,  $H_S$  is always smaller than  $H_R^\beta$  and  $H_T^\beta$ .

Figs. 1(b) and 1(c) suggest that multiplying the covariance matrix by a constant — hence increasing its determinant — also increases both Shannon and Rényi entropies. As expected, increasing the number of looks leads to smaller entropy values due to the increase of the image SNR.

Although the restricted Tsallis entropy can be used in several fields of image processing [18], the derivation of its variance does not lead to a mathematically tractable expression. Thus, henceforth we focus our attention on the Shannon and Rényi entropies, which allow the necessary algebraic manipulations for the method described in Section III. In the next section, we derive asymptotic variances for the Shannon and Rényi entropy estimates

#### B. Asymptotic variances

Let  $\mathbf{Z}$  be a random matrix which follows a scaled complex Wishart distribution with parameter  $\boldsymbol{\theta}$  as defined in the previous subsection. The associated log-likelihood function is expressed by

$$\ell(\boldsymbol{\theta}) = mL \ln L + (L - m) \ln |\mathbf{Z}| - L \ln |\Sigma| - \frac{m(m-1)}{2} \ln \pi - \sum_{k=0}^{m-1} \ln \Gamma(L - k) - L \operatorname{tr}(\Sigma^{-1} \mathbf{Z}).$$

Since  $\partial \ell(\boldsymbol{\theta}) / \partial \operatorname{vec}(\Sigma) = \operatorname{vec}(\partial \ell(\boldsymbol{\theta}) / \partial \Sigma)$  holds true [42], the score functions are given by

$$\ell_L = m(\ln L + 1) + \ln |\mathbf{Z}| - \ln |\Sigma| - \psi_m^{(0)}(L) - \operatorname{tr}(\Sigma^{-1} \mathbf{Z})$$

and

$$\ell_{\operatorname{vec}(\Sigma)} = L \operatorname{vec}(\Sigma^{-1} \mathbf{Z} \Sigma^{-1} - \Sigma^{-1}).$$

The Hessian matrix  $\mathcal{J}(\boldsymbol{\theta})$ , the Fisher information matrix  $\mathcal{K}(\boldsymbol{\theta})$ , and the biased version (according to Anfinson *et al.* [38]) for Cramér-Rao lower bound  $\mathcal{C}(\boldsymbol{\theta})$  are necessary to obtain closed form expressions for the asymptotic entropy variance used in Equation (3). In particular, the following quantity plays a central role:

$$\mathcal{J}_{\Sigma\Sigma} = \frac{\partial}{\partial \operatorname{vec}(\Sigma)^*} \operatorname{vec} \left( \frac{\partial \ell(\boldsymbol{\theta})}{\partial \Sigma} \right)^t,$$

where  $(\cdot)^*$  represents complex conjugation. Anfinson *et al.* [38] showed that

$$\begin{aligned} T_1 &= -\frac{\partial(\Sigma^{-1} \mathbf{Z} \Sigma^{-1})}{\partial \Sigma} \\ &= -\Sigma^{-1} \otimes \Sigma^{-1} \mathbf{Z} \Sigma^{-1} - \Sigma^{-1} \otimes \Sigma^{-1} \Sigma^{-1} \mathbf{Z}. \end{aligned}$$

Moreover, it is known that [42]

$$T_2 = \partial \Sigma^{-1} / \partial \Sigma = -\Sigma^{-1} \otimes \Sigma^{-1}.$$

Thus, we have that

$$\begin{aligned} \mathcal{J}_{\Sigma\Sigma} &= L(T_2 - T_1) = L(\Sigma^{-1} \otimes \Sigma^{-1} - \Sigma^{-1} \otimes \Sigma^{-1} \mathbf{Z} \Sigma^{-1} \\ &\quad - \Sigma^{-1} \otimes \Sigma^{-1} \Sigma^{-1} \mathbf{Z}). \end{aligned} \quad (15)$$

From Equation (15) we obtain:

$$\mathcal{K}_{\Sigma\Sigma} = \mathbb{E}\{-\mathcal{J}_{\Sigma\Sigma}\} = L\Sigma^{-1} \otimes \Sigma^{-1}. \quad (16)$$

The analytical expressions for  $\mathcal{J}(\boldsymbol{\theta})$ ,  $\mathcal{K}(\boldsymbol{\theta})$ , and  $\mathcal{C}(\boldsymbol{\theta})$  are, thus,

$$\begin{aligned} \mathcal{J}(\boldsymbol{\theta}) &= \begin{bmatrix} \frac{m}{L} - \psi_m^{(1)}(L) & \operatorname{vec}(\Sigma^{-1} \mathbf{Z} \Sigma^{-1} - \Sigma^{-1})^t \\ \operatorname{vec}(\Sigma^{-1} \mathbf{Z} \Sigma^{-1} - \Sigma^{-1})^* & \mathcal{J}_{\Sigma\Sigma} \end{bmatrix}, \\ \mathcal{K}(\boldsymbol{\theta}) &= \mathbb{E}\{-\mathcal{J}(\boldsymbol{\theta})\} = \begin{bmatrix} \psi_m^{(1)}(L) - \frac{m}{L} & \operatorname{vec}(\mathbf{0}_m)^t \\ \operatorname{vec}(\mathbf{0}_m) & L\Sigma^{-1} \otimes \Sigma^{-1} \end{bmatrix}, \\ \mathcal{C}(\boldsymbol{\theta}) &= \mathcal{K}(\boldsymbol{\theta})^{-1} = \begin{bmatrix} [\psi_m^{(1)}(L) - \frac{m}{L}]^{-1} & \operatorname{vec}(\mathbf{0}_m)^t \\ \operatorname{vec}(\mathbf{0}_m) & L^{-1} \Sigma \otimes \Sigma \end{bmatrix}, \end{aligned}$$

where  $\mathbf{0}_m$  is the null square matrix of order  $m$ .

Anfinson *et al.* [38] derived the Fisher information matrix for the unscaled complex Wishart law. That approach found that the parameters of such distribution are not orthogonal. Based on  $\mathcal{J}(\boldsymbol{\theta})$ , we conclude that parameters  $L$  and  $\Sigma$  become orthogonal under the scaling of the complex Wishart law. Among other benefits, such scaling makes the likelihood equations separable, as shown in the following.

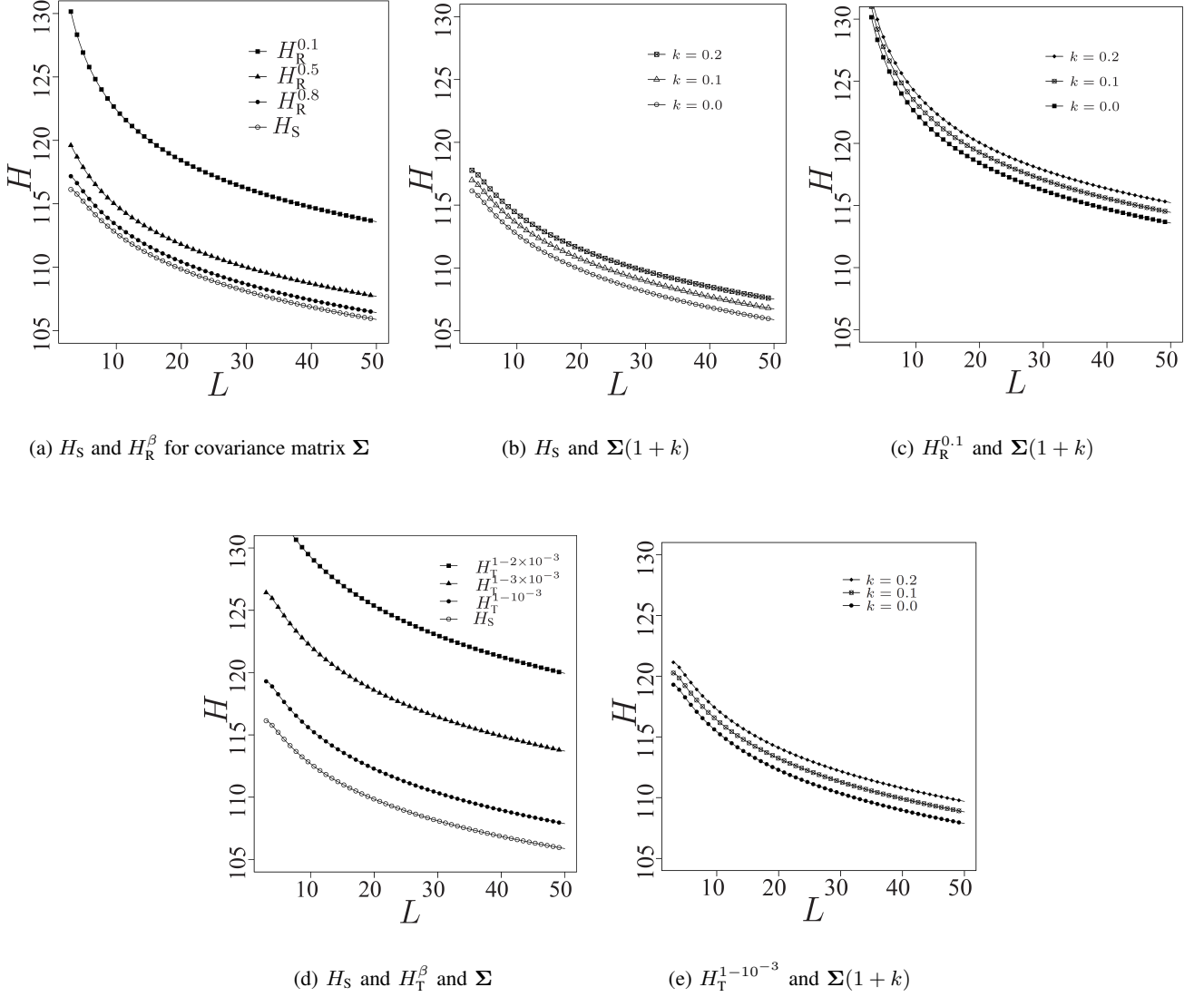


Fig. 1. Shannon, Tsallis, and Rényi entropies for several covariance matrices and number of looks.

Consider  $\{\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_N\}$  a random sample of size  $N$  obtained from  $\mathbf{Z} \sim \mathcal{W}_m(\Sigma, L)$ . Since  $N^{-1} \sum_{k=1}^N \nabla \ell_k(\hat{\theta}) = \mathbf{0}$ , we have that

$$\hat{\Sigma} = \frac{1}{N} \sum_{k=1}^N \mathbf{Z}_k = \bar{\mathbf{Z}},$$

and

$$m \ln \hat{L} + \frac{1}{N} \sum_{k=1}^N \ln |\mathbf{Z}_k| - \ln |\bar{\mathbf{Z}}| - \psi_m^{(0)}(\hat{L}) = 0. \quad (17)$$

Thus, the ML estimator of  $\Sigma$  is the sample mean, while  $\hat{L}$  is derived by solving the system shown in Equation (17). The Newton-Raphson iterative method [43] can be used to solve this nonlinear system.

The asymptotic variance given by Equation (3) is determined by  $\mathcal{C}(\theta)$  and by the term [36]

$$\delta = \left[ \frac{\partial H_{\phi}^h(\theta)}{\partial L} \quad \text{vec} \left( \frac{H_{\phi}^h(\theta)}{\partial \Sigma} \right)^t \right]^t.$$

We denote the resulting  $\delta$  as  $\delta_S$  and  $\delta_{R,\beta}$ , when  $H_S$  and  $H_R^{\beta}$  are considered, respectively. Analogously, entropy variances are denoted as  $\sigma_S^2$  and  $\sigma_{R,\beta}^2$ . These quantities are given by expressions (18)-(21).

Note that, as expected from Equation (14),  $\lim_{\beta \rightarrow 1} \sigma_{R,\beta}^2 = \sigma_S^2$ .

As a potential application, the two derived measures can be considered as an alternative goodness-of-fit proposed in [44] for the Wishart distribution, specifying  $r = 1$  in (6). In this paper, we aim to quantify contrast in PolSAR images and situations with  $r = 2$  are considered.



(a) Shannon entropy:

$$\delta_S = \left[ \frac{(m-L)\psi_m^{(1)}(L) + m - \frac{m^2}{L}}{m \text{vec}(\Sigma^{-1})} \right] \quad (18)$$

and

$$\sigma_S^2 = \frac{[(m-L)\psi_m^{(1)}(L) + m - \frac{m^2}{L}]^2}{\psi_m^{(1)}(L) - \frac{m}{L}} + \frac{m^2}{L} \text{vec}(\Sigma^{-1})^t (\Sigma \otimes \Sigma) \text{vec}(\Sigma^{-1}). \quad (19)$$

(b) Rényi entropy:

$$\delta_{R,\beta} = \left[ \frac{\frac{\beta}{1-\beta} [\psi_m^{(0)}(q) - \psi_m^{(0)}(L)] - \frac{m\beta \ln(\beta)}{1-\beta} - \frac{m^2}{L}}{m \text{vec}(\Sigma^{-1})} \right] \quad (20)$$

and

$$\sigma_{R,\beta}^2 = \frac{\left\{ \frac{\beta}{1-\beta} [\psi_m^{(0)}(q) - \psi_m^{(0)}(L)] - \frac{m\beta \ln(\beta)}{1-\beta} - \frac{m^2}{L} \right\}^2}{\psi_m^{(1)}(L) - \frac{m}{L}} + \frac{m^2}{L} \text{vec}(\Sigma^{-1})^t (\Sigma \otimes \Sigma) \text{vec}(\Sigma^{-1}). \quad (21)$$

A proposal of contrast measures is an important SAR image preprocessing step because they are often used for improving imagery techniques [21]. Cintra *et al.* [22] showed the such tools outperform the Kolmogorov-Smirnov nonparametric test in SAR imagery. To the following, the proposed tools are applied to the PolSAR imagery.

## V. APPLICATION

In this section, we combine the entropies and variances derived in previous section to form statistical tests. Monte Carlo experiments were used in order to quantify their performance with respect to test size and power. Finally, the discussed methodology is applied to real data.

Fig. 2 shows the HH band of a PolSAR image obtained by an E-SAR sensor with 3.2 nominal equivalent number of looks in the L-band image over Weßling, Germany [45]. Three regions were selected: A<sub>1</sub>, A<sub>2</sub>, and A<sub>3</sub>, corresponding to areas with strong, moderate, and weak return, respectively. Table II shows the sample size  $N$  and the ML estimates of the scaled Wishart distribution parameters for each region.

As shown in Table II, the smallest estimative for number of looks is presented by region A<sub>1</sub>, where the texture is more intense.

Table III presents the asymptotic lower and upper bounds for the Shannon and Rényi entropies at the 95% level of confidence, as computed according to Equation (7). The Shannon entropy has the smallest confidence intervals at the considered level and, therefore, its estimates are the most precise, at least for the samples here analyzed.

TABLE II  
ML ESTIMATES

| Regions        | $N$  | $ \hat{\Sigma} $ | $\hat{L}$ |
|----------------|------|------------------|-----------|
| A <sub>1</sub> | 3708 | 355494.500       | 1.361     |
| A <sub>2</sub> | 2088 | 3321.241         | 1.657     |
| A <sub>3</sub> | 1079 | 274.189          | 2.557     |

TABLE III  
ESTIMATED ASYMPTOTIC INTERVALS FOR SHANNON AND RÉNYI ENTROPIES AT 95%

| Region         | $\hat{H}_S$ |        | $\hat{H}_R^{0.1}$ |        | $\hat{H}_R^{0.8}$ |        |
|----------------|-------------|--------|-------------------|--------|-------------------|--------|
|                | Lower       | Upper  | Lower             | Upper  | Lower             | Upper  |
| A <sub>1</sub> | 37.979      | 38.432 | 61.083            | 61.332 | 44.045            | 44.364 |
| A <sub>2</sub> | 30.079      | 30.541 | 45.563            | 45.867 | 36.124            | 37.049 |
| A <sub>3</sub> | 19.611      | 19.949 | 35.000            | 35.346 | 20.901            | 21.230 |

These results present evidence that the lowest estimated entropies are associated with regions whose estimated covariance matrices have small determinant values. This fact can be justified by means of the case study presented in Section IV-A4 which illustrated that larger covariance matrices and/or equivalent numbers of looks are associated with areas of smaller entropies. Since (i) the entropy is a variability measure and (ii) PolSAR areas with a high reflectance are more heterogeneous, it is intuitive the conclusion obtained from data: determinant of covariance matrix and entropy are two measure directly proportional. Goodman [46] defined the stochastic covariance matrix determinant as a *generalized variance* which is associated with its heterogeneity degree. Thus, although poorly explored in the treatment of these images, this relationship may be highlighted.

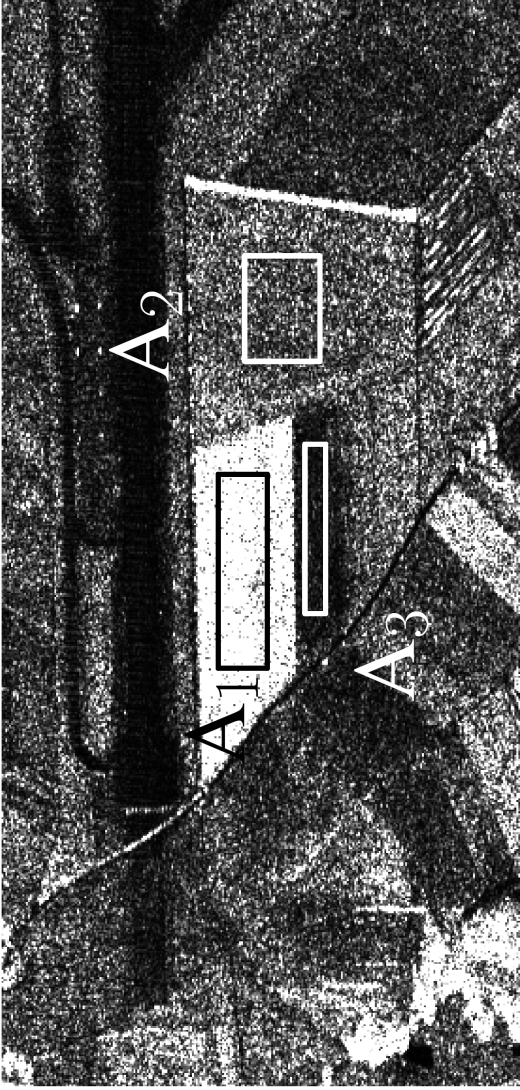


Fig. 2. PolSAR image and selected regions.

The following discussion applies the derived theory to synthetic and real data. Using Monte Carlo simulation and the estimated parameters from real data, statistical tests for identifying homogeneity in PolSAR images based on Shannon and Rényi (for  $\beta = 0.1, 0.8$ ) entropies were assessed. Subsequently, our methodology was applied to real data.

1) *Synthetic Data*: In this subsection, the performance of hypothesis tests based on Shannon and Rényi entropies was quantified. To that end, synthetic data following the scaled complex Wishart distribution were generated as suggested in [47]. The simulation employed the following parameters: (i)  $L = 3.2$ ; (ii) the estimated covariance matrices from regions  $A_1$ ,  $A_2$ , and  $A_3$ , denoted as  $\tilde{\Sigma}_1$ ,  $\tilde{\Sigma}_2$ , and  $\tilde{\Sigma}_3$ , respectively; (iii) sample sizes  $N = \{9, 49, 81, 121, 400\}$  (i.e., squared windows of size  $\{3, 7, 9, 11, 20\}$ ); and (iv) significance levels  $\alpha \in$

$\{1\%, 5\%, 10\%\}$ . After generating two samples of size  $N$  each from  $\mathbf{X} \sim \mathcal{W}_m(\tilde{\Sigma}_i, L)$  and  $\mathbf{Y} \sim \mathcal{W}_m(\tilde{\Sigma}_j, L)$ , we tested the null hypothesis  $(\tilde{\Sigma}_i, L) = (\tilde{\Sigma}_j, L)$  for  $i, j = 1, 2, 3$ . Following [21], we run 5500 replicas of the Monte Carlo simulation for every case. Empirical test size and power were employed as figures of merit.

Table IV presents the resulting empirical test sizes. From a purely statistical viewpoint, the ideal test is the one whose empirical size is exactly the nominal one for simulated data. Therefore, we conclude that the Rényi entropy for  $\beta = 0.1$  is generally outperformed by the Shannon and Rényi of order 0.8, except for  $N = 9$ . The two last hypothesis tests have empirical test sizes close to the nominal levels even for small sample sizes.

TABLE IV  
EMPIRICAL TEST SIZES

| $N$                              | $A_1$ |      |       | $A_2$ |      |       | $A_3$ |      |       |
|----------------------------------|-------|------|-------|-------|------|-------|-------|------|-------|
|                                  | 1%    | 5%   | 10%   | 1%    | 5%   | 10%   | 1%    | 5%   | 10%   |
| Shannon entropy                  |       |      |       |       |      |       |       |      |       |
| 9                                | 1.35  | 6.15 | 11.76 | 1.36  | 6.16 | 12.02 | 1.44  | 6.20 | 12.05 |
| 49                               | 1.38  | 5.53 | 10.44 | 1.38  | 5.65 | 10.35 | 1.35  | 5.47 | 10.44 |
| 81                               | 1.33  | 5.51 | 9.75  | 1.35  | 5.47 | 9.84  | 1.35  | 5.36 | 10.07 |
| 121                              | 1.18  | 5.56 | 10.58 | 1.15  | 5.51 | 10.51 | 1.15  | 5.49 | 10.44 |
| 400                              | 1.35  | 5.38 | 10.33 | 1.38  | 5.47 | 10.40 | 1.40  | 5.51 | 10.22 |
| Rényi entropy with $\beta = 0.8$ |       |      |       |       |      |       |       |      |       |
|                                  | 1.38  | 6.22 | 11.65 | 1.44  | 6.18 | 11.93 | 1.45  | 6.09 | 11.78 |
|                                  | 1.27  | 5.62 | 10.36 | 1.31  | 5.65 | 10.31 | 1.33  | 5.38 | 10.55 |
|                                  | 1.31  | 5.47 | 9.76  | 1.20  | 5.47 | 9.91  | 1.29  | 5.36 | 9.95  |
|                                  | 1.05  | 5.35 | 10.45 | 1.15  | 5.44 | 10.44 | 1.13  | 5.53 | 10.55 |
|                                  | 1.35  | 5.27 | 10.31 | 1.36  | 5.29 | 10.29 | 1.27  | 5.33 | 10.42 |
| Rényi entropy with $\beta = 0.1$ |       |      |       |       |      |       |       |      |       |
|                                  | 0.69  | 4.04 | 8.80  | 0.76  | 4.15 | 8.65  | 0.82  | 4.04 | 8.55  |
|                                  | 0.71  | 3.62 | 7.82  | 0.75  | 3.87 | 7.89  | 0.69  | 3.45 | 7.73  |
|                                  | 0.53  | 3.69 | 7.51  | 0.56  | 3.84 | 7.49  | 0.55  | 3.62 | 7.51  |
|                                  | 0.60  | 3.73 | 7.58  | 0.60  | 3.40 | 7.47  | 0.53  | 3.58 | 7.64  |
|                                  | 0.56  | 3.84 | 7.65  | 0.67  | 3.82 | 7.64  | 0.67  | 3.62 | 7.65  |

Samples from different covariance matrices always led to rejecting the null hypothesis with the proposed statistical tests, thus leading to tests with unitary empirical test power. Table V shows the mean value of the statistics

$$\bar{S}(A_i, A_j) = \frac{1}{5500} \sum_{k=1}^{5500} S_{\phi}^h([\hat{\Sigma}_i, \hat{L}_i]_k, [\hat{\Sigma}_j, \hat{L}_j]_k),$$

where  $S_{\phi}^h$  is  $S_S$  or  $S_R$  as given in (6) and  $[\hat{\Sigma}_i, \hat{L}_i]_k$  are the ML estimates based on generated data from population  $A_i$  at the  $k$ th Monte Carlo replica, and coefficient of variation (CV) of the test statistics under alternative

hypotheses:

$$CV(A_i, A_j) = \frac{\sqrt{\sum_{k=1}^{5500} [S_{\phi}^h([\hat{\Sigma}_i, \hat{L}_i]_k, [\hat{\Sigma}_j, \hat{L}_j]_k) - \bar{S}(A_i, A_j)]^2}}{\bar{S}(A_i, A_j)\sqrt{5500}}.$$

TABLE V  
MEAN AND COEFFICIENT OF VARIATION UNDER ALTERNATIVE HYPOTHESES

| $N$                              | $A_1-A_2$ |           | $A_1-A_3$ |           | $A_2-A_3$ |           |
|----------------------------------|-----------|-----------|-----------|-----------|-----------|-----------|
|                                  | CV        | $\bar{S}$ | CV        | $\bar{S}$ | CV        | $\bar{S}$ |
| Shannon entropy                  |           |           |           |           |           |           |
| 9                                | 20.59     | 108.16    | 13.83     | 248.53    | 38.42     | 29.89     |
| 49                               | 8.65      | 576.23    | 5.82      | 1329.19   | 16.38     | 156.21    |
| 81                               | 6.70      | 949.51    | 4.52      | 2191.84   | 12.71     | 256.54    |
| 121                              | 5.51      | 1418.22   | 3.71      | 3274.03   | 10.43     | 383.17    |
| 400                              | 3.01      | 4687.29   | 2.02      | 10817.80  | 5.72      | 1265.69   |
| Rényi entropy with $\beta = 0.8$ |           |           |           |           |           |           |
|                                  | 20.56     | 106.72    | 13.72     | 245.22    | 38.45     | 29.50     |
|                                  | 8.63      | 569.69    | 5.76      | 1314.11   | 16.40     | 154.43    |
|                                  | 6.67      | 938.93    | 4.47      | 2167.39   | 12.72     | 253.68    |
|                                  | 5.48      | 1402.43   | 3.66      | 3237.61   | 10.42     | 378.93    |
|                                  | 3.00      | 4635.65   | 2.00      | 10698.61  | 5.71      | 1251.74   |
| Rényi entropy with $\beta = 0.1$ |           |           |           |           |           |           |
|                                  | 21.70     | 78.42     | 14.27     | 180.12    | 40.60     | 21.79     |
|                                  | 9.11      | 423.29    | 6.00      | 976.32    | 17.44     | 114.77    |
|                                  | 7.02      | 698.47    | 4.63      | 1612.06   | 13.50     | 188.81    |
|                                  | 5.73      | 1043.26   | 3.77      | 2408.63   | 11.00     | 282.06    |
|                                  | 3.15      | 3450.84   | 2.07      | 7963.98   | 6.03      | 931.79    |

The largest means  $\bar{S}(A_i, A_j)$  correspond to the largest absolute values  $\text{abs}(|\hat{\Sigma}_i| - |\hat{\Sigma}_j|)$ , leading to the following relations:

$$\bar{S}(A_1, A_3) > \bar{S}(A_1, A_2) > \bar{S}(A_2, A_3). \quad (22)$$

Shannon test statistics produced the largest test statistics when contrasting samples from different areas.

2) *Real Data*: Applying the methodology described above to real data, we obtained the results shown in Fig. 3. To that end, the following steps were followed:

- 1) For  $j = 1, 2, \dots, 5500$  perform steps 2) to 5).
- 2) Extract two regions  $U_j$  and  $V_j$  from areas  $A_1, A_2$ , and  $A_3$ .
- 3) By sampling without replacement, generate two  $N$ -point vectors  $u^{(j)}$  and  $v^{(j)}$  from  $U_j$  and  $V_j$ , respectively, granting  $u^{(i)} \neq u^{(j)}$  and  $v^{(i)} \neq v^{(j)}$  for  $i \neq j$ .
- 4) Estimate the parameter vectors  $\hat{\theta}_1^{(j)}$  and  $\hat{\theta}_2^{(j)}$  based on  $u^{(j)}$  and  $v^{(j)}$ , respectively.
- 5) Compute the decision from Proposition 1 for  $\alpha = \{1\%, 5\%, 10\%\}$ .

TABLE VI  
ML ESTIMATES

| Regions        | $N$  | $ \hat{\Sigma} $       | $\hat{L}$ |
|----------------|------|------------------------|-----------|
| B <sub>1</sub> | 3192 | $1.609 \times 10^{-5}$ | 6.925     |
| B <sub>2</sub> | 1408 | $1.112 \times 10^{-6}$ | 11.937    |
| B <sub>3</sub> | 1848 | $5.814 \times 10^{-7}$ | 10.752    |

- 6) Let  $T$  be the number of times that the null hypothesis was rejected. Calculate the empirical test size  $(\hat{\alpha}_{1-\alpha})$  and power  $(\hat{\beta}_{1-\alpha})$  at level  $\alpha$  as

$$\begin{cases} \hat{\alpha}_{1-\alpha} = T/5500, & \text{if } V_j = U_j, \\ \hat{\beta}_{1-\alpha} = T/5500, & \text{if } V_j \neq U_j. \end{cases}$$

We considered  $N$  the number of observations in squared windows of size  $\{3 \times 3, 4 \times 4, \dots, 23 \times 23\}$ .

Fig. 3 shows the empirical test sizes, with the nominal test size  $\alpha$  in dotted line. The behavior of  $\hat{\alpha}_{1-\alpha}$  is consistent across the three values of  $\alpha$ : the empirical test size is larger in Shannon than in Rényi  $\beta = 0.8$  which, in turn, is larger than Rényi  $\beta = 0.1$ . The fastest convergence to  $\alpha$  with respect to the sample size is consistently observed in area  $A_3$ . No test attains the nominal size in area  $A_1$ . This is probably due to its pronounced brightness, as noted in Table II.

In order to extend the study concerning the test powers, we aim to select regions visually similar, but with distinct statistical properties. To that end, Fig. 4(a) shows EMISAR image of Foulum (Denmark) which was obtained with the nominal number of looks  $L = 8$ . This PolSAR image represents agricultural areas and, for our study, three regions were selected. Table VI presents a descriptive analysis of ML estimates of sampled areas. Figs. 4(b)-(d) presents empirical densities for different areas and polarization channels. As expected by estimates of  $L$ , notice that the empirical densities of B<sub>2</sub> and B<sub>3</sub> are closest for all polarization channels. The following discussion complements the test power study.

Based on algorithm presented in the begin of this section, we quantify the discrimination between selected regions by means of  $S_S$  and  $S_R^{0.1}$  statistics. Fig. 5 (a)-(c) give the empirical test power considering pairs of selected regions. In general terms, the estimated power increase quickly when sample sizes increase. The test based on the Shannon entropy outperforms one proposed by Rényi. These results suggest the test statistics in terms of the Shannon entropy as an more appropriate tool for change and boundary detection in PolSAR images than the measure on the Rényi entropy..

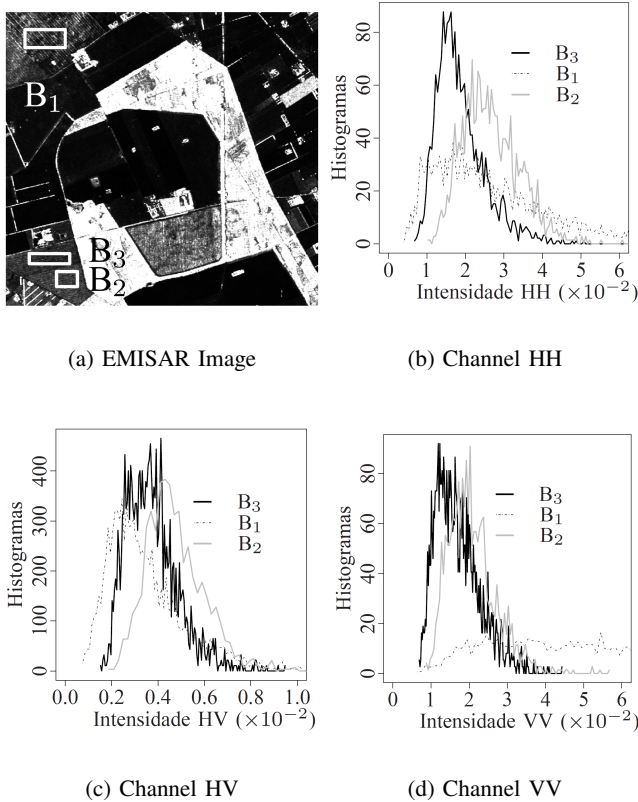


Fig. 4. PolSAR image with selected regions and empirical densities for several areas and polarization channel.

## VI. CONCLUSIONS

New hypothesis tests and confidence intervals based on the  $(h, \phi)$ -family of entropies were derived for the scaled complex Wishart distribution. These new statistical tools provide means for contrasting samples and verifying if they come from the same distribution. Such tools were derived by obtaining orthogonal maximum likelihood estimators for the scaled complex Wishart law, by characterizing their asymptotic behavior.

Monte Carlo experiments were employed for quantifying the performance of the proposed hypotheses tests. The results provided evidence that the statistic based on the Shannon entropy is the most efficient in terms of empirical test size and power.

An application to actual data was also considered in order to assess the performance of the proposed hypothesis tests. The empirical test sizes observed with real data are relatively high when dealing with small samples, but they decrease as the sample size increases. As expected, the tests presented the worst results in the most heterogeneous situations. Nevertheless, they perform correctly when the sample size is increased.

The hypothesis test based on the Shannon entropy

presented the smallest empirical test size. All test statistics detected differences among different regions at the specified levels. This is important for PolSAR image analysis, such as boundary detection [24] and change detection [26].

As an overall conclusion, the Shannon entropy can be safely used for discriminating areas in PolSAR imagery. However, care must be taken when small samples are analyzed. Indeed, in this case, the proposed tests are prone to classifying regions of similar natures as distinct, i.e., to incurring Type I error. In practice, however, this issue is not severe, since PolSAR image processing often handles data with a large number of pixels.

## REFERENCES

- [1] J. S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton: CRC, 2009.
- [2] F. T. Ulaby and C. Elachi, *Radar Polarimetry for Geoscience Applications*. Norwood: Artech House, 1990.
- [3] A. C. Frery, R. J. Cintra, and A. D. C. Nascimento, "Hypothesis test in complex Wishart distributions," in *Proceedings of the 5th International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLinSAR'2011)*, Frascati, Italy, 24-28 January 2011.
- [4] J. M. Beaulieu and R. Touzi, "Segmentation of textured polarimetric SAR scenes by likelihood approximation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 10, pp. 2063–2072, 2004.
- [5] P. R. Kersten, J. S. Lee, and T. L. Ainsworth, "Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and EM clustering," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 519–527, 2005.
- [6] J. Schou, H. Skriver, A. H. Nielsen, and K. Conradsen, "CFAR edge detector for polarimetric SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 1, pp. 20–32, 2003.
- [7] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, Jul 1948.
- [8] S. Nadarajah and K. Zografos, "Formulas for Rényi information and related measures for univariate distributions," *Information Science*, vol. 155, pp. 119–138, 2003.
- [9] —, "Expressions for Rényi and Shannon entropies for bivariate distributions," *Information Sciences*, vol. 170, no. 2-4, pp. 173–189, 2005.
- [10] K. Zografos and S. Nadarajah, "Expressions for Rényi and Shannon entropies for multivariate distributions," *Statistics & Probability Letters*, vol. 71, no. 1, pp. 71–84, 2005.
- [11] K. S. Song, "Rényi information, log-likelihood and an intrinsic distribution measure," *Journal of Statistical Planning and Inference*, vol. 93, pp. 51–69, 2001.
- [12] M. Basseville, "Distance measures for signal processing and pattern recognition," *Signal Processing*, vol. 18, pp. 349–369, 1989.
- [13] S. J. Maybank, "Application of the Fisher-Rao metric to ellipse detection," *International Journal of Computer Vision*, vol. 72, no. 3, pp. 287–307, 2007.
- [14] M. Liuni, A. Robel, M. Romito, and X. Rodet, "Rényi information measures for spectral change detection," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'2011)*, may 2011, pp. 3824–3827.

- [15] C. Tsallis, "Possible generalization of Boltzmann–Gibbs statistics," *Journal of Statistical Physics*, vol. 52, pp. 479–487, 1988.
- [16] —, "Entropic nonextensivity: a possible measure of complexity," *Chaos, Solitons, & Fractals*, vol. 13, pp. 371–391, 2002.
- [17] G. Wilk and Z. Włodarczyk, "Example of a possible interpretation of Tsallis entropy," *Physica A: Statistical Mechanics and its Applications*, vol. 387, no. 19–20, pp. 4809–4813, 2008.
- [18] J. Morio, P. Réfrégier, F. Goudail, P. C. Dubois Fernandez, and X. Dupuis, "A characterization of Shannon entropy and Bhattacharyya measure of contrast in polarimetric and interferometric SAR image," *Proceedings of the IEEE (PIEEE)*, vol. 97, no. 6, pp. 1097–1108, jun 2009.
- [19] S. R. Cloude and E. Pottier, "An entropy based classification scheme for land application of polarimetric SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, pp. 68–78, 1997.
- [20] W. Yan, W. Yang, Y. Liu, and H. Sun, "Unsupervised classification of PolinSAR image based on Shannon entropy characterization," in *Proceedings of the 2010 IEEE International Conference on Signal Processing (ICSP 2010)*, oct 2010, pp. 2192–2195.
- [21] A. D. C. Nascimento, R. J. Cintra, and A. C. Frery, "Hypothesis testing in speckled data with stochastic distances," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 1, pp. 373–385, 2010.
- [22] R. J. Cintra, A. C. Frery, and A. D. C. Nascimento, "Parametric and nonparametric tests for speckled imagery," *Pattern Analysis and Applications*, pp. 1–21, 2011.
- [23] M. Salicrú, M. L. Mendéndez, and L. Pardo, "Asymptotic distribution of  $(h, \phi)$ -entropy," *Communications in Statistics - Theory Methods*, vol. 22, no. 7, pp. 2015–2031, 1993.
- [24] A. C. Frery, J. Jacobo-Berlles, J. Gambini, and M. Mejail, "Polarimetric SAR image segmentation with B-splines and a new statistical model," *Multidimensional Systems and Signal Processing*, vol. 21, pp. 319–342, 2010.
- [25] J. Gambini, M. Mejail, J. Jacobo-Berlles, and A. C. Frery, "Accuracy of edge detection methods with local information in speckled imagery," *Statistics and Computing*, vol. 18, no. 1, pp. 15–26, 2008.
- [26] J. Inglada and G. Mercier, "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*, vol. 45, no. 5, pp. 1432–1445, 2007.
- [27] 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 Transactions on Geoscience and Remote Sensing*, vol. 41, no. 1, pp. 4–19, 2003.
- [28] C. López-Martínez and X. Fabregas, "Polarimetric SAR speckle noise model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 10, pp. 2232–2242, 2003.
- [29] F. Goudail and P. Réfrégier, "Contrast definition for optical coherent polarimetric images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 7, pp. 947–951, jul 2004.
- [30] N. R. Goodman, "Statistical analysis based on a certain complex Gaussian distribution (an introduction)," *The Annals of Mathematical Statistics*, vol. 34, pp. 152–177, 1963.
- [31] D. Blatt and A. O. Hero, "On tests for global maximum of the log-likelihood function," *IEEE Transactions on Information Theory*, vol. 53, pp. 2510–2525, July 2007.
- [32] D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies, "Data compression and harmonic analysis," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2435–2476, oct 1998.
- [33] S. Kullback, *Information Theory and Statistics*. Dover, 1978.
- [34] J. Havrda and F. Chalvat, "Concept of structural  $\alpha$ -entropy," *Kybernetika*, vol. 3, pp. 30–35, 1967.
- [35] A. Rényi, "On measures of entropy and information," in *4th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, 1961, pp. 547–561.
- [36] L. Pardo, D. Morales, M. Salicrú, and M. L. Menéndez, "Large sample behavior of entropy measures when parameters are estimated," *Communications in Statistics - Theory and Methods*, vol. 26, no. 2, pp. 483–501, 1997.
- [37] W. G. Cochran, "The distribution of quadratic forms in a normal system, with applications to the analysis of covariance," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 30, no. 2, pp. 178–191, 1934.
- [38] S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, "Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3795–3809, 2009.
- [39] S. N. Anfinsen and T. Eltoft, "Application of the matrix-variate Mellin transform to analysis of polarimetric radar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 2281–2295, jun 2011.
- [40] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley-Interscience, 1991.
- [41] D. Jiulin, "Property of Tsallis entropy and principle of entropy increase," *Bulletin of the Astronomical Society of India*, vol. 35, pp. 691–696, 2010.
- [42] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2740–2746, jun 2007.
- [43] J. E. Gentle, *Elements of Computational Statistics*, ser. Statistics and computing. Springer, 2002.
- [44] S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, "Goodness-of-fit tests for multilook polarimetric radar data based on the Mellin transform," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 7, pp. 2764–2781, jul 2011.
- [45] R. Horn, "The DLR airborne SAR project E-SAR," in *Geoscience and Remote Sensing Symposium*, vol. 3. IEEE Press, 1996, pp. 1624–1628.
- [46] N. R. Goodman, "The distribution of determinant of a complex Wishart distributed matrix," *The Annals of Mathematical Statistics*, vol. 34, pp. 178–180, 1963.
- [47] J. S. Lee, M. R. Grunes, and R. Kwok, "Classification of multilook polarimetric SAR imagery based on complex Wishart distribution," *International Journal of Remote Sensing*, vol. 15, no. 11, pp. 2299–2311, sep 1994.



**Alejandro C. Frery** graduated in Electronic and Electrical Engineering from the Universidad de Mendoza, Argentina. His M.Sc. degree was in Applied Mathematics (Statistics) from the Instituto de Matemática Pura e Aplicada (Rio de Janeiro) and his Ph.D. degree was in Applied Computing from the Instituto Nacional de Pesquisas Espaciais (São José dos Campos, Brazil). He is currently with the

Instituto de Computação, Universidade Federal de Alagoas, Maceió, Brazil. His research interests are statistical computing and stochastic modelling.

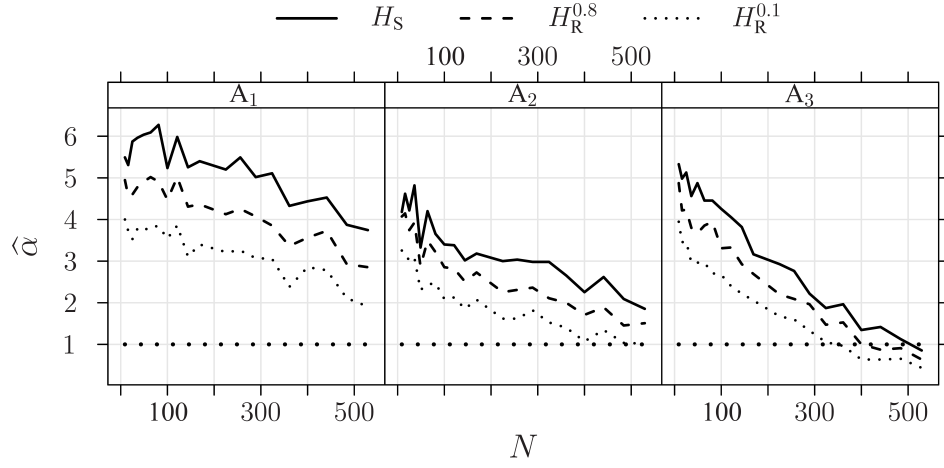


**Renato J. Cintra** earned his B.Sc., M.Sc., and D.Sc. degrees in Electrical Engineering from Universidade Federal de Pernambuco, Brazil, in 1999, 2001, and 2005, respectively. In 2005, he joined the Department of Statistics at UFPE. During 2008-2009, he worked at the University of Calgary, Canada, as a visiting research fellow. He is also a graduate faculty member of the Department of Electrical and

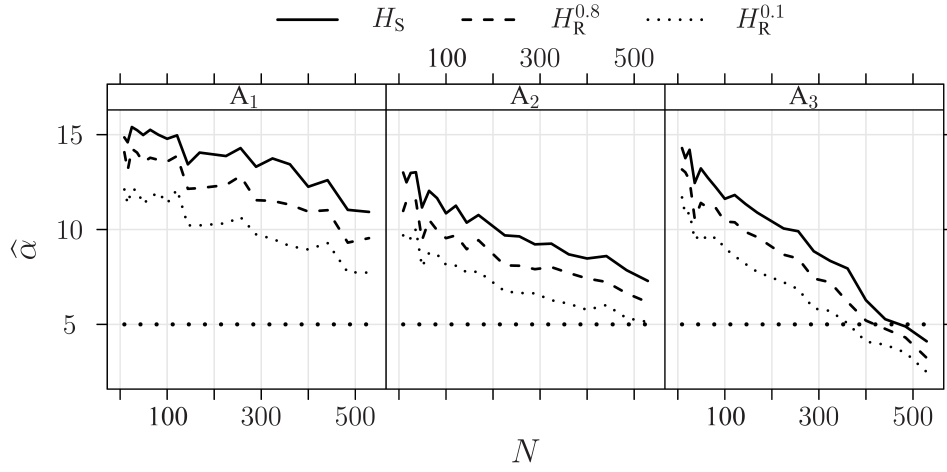
Computer Engineering, University of Akron, OH. His long term topics of research include theory and methods for digital signal processing, communications systems, and applied mathematics.



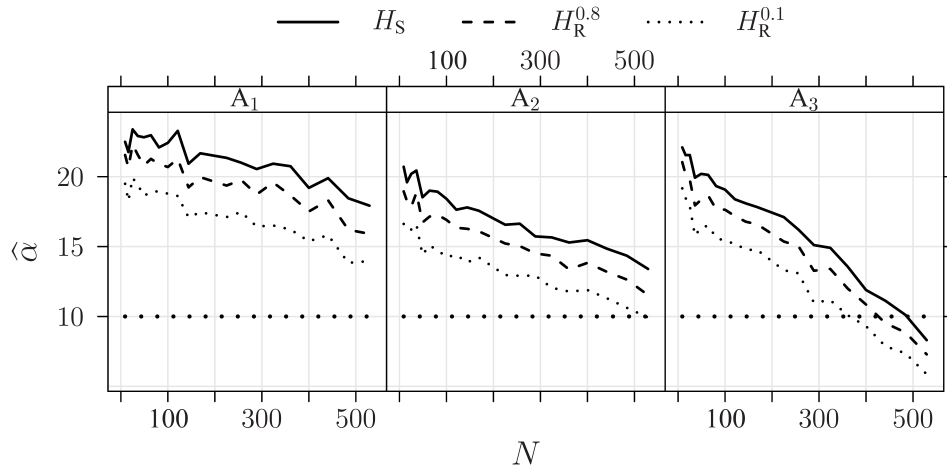
**Abraão D. C. Nascimento** holds B.Sc. and M.Sc. degrees in Statistics from the Universidade Federal de Pernambuco (UFPE), Brazil, and he is currently a doctorate student in Statistics at the same University. His research interests are stochastic models and distances.



(a) Size 1%

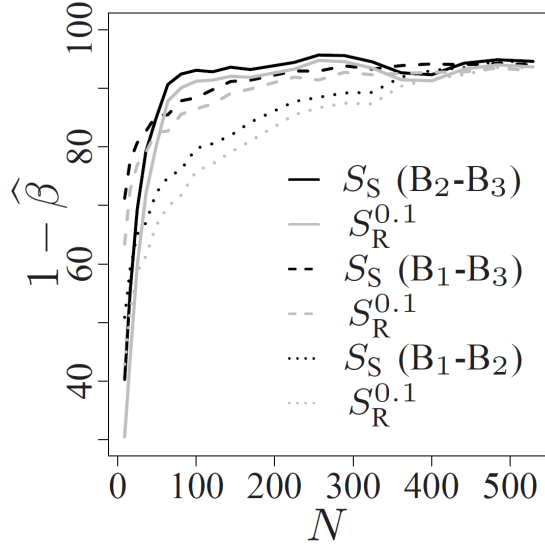


(b) Size 5%

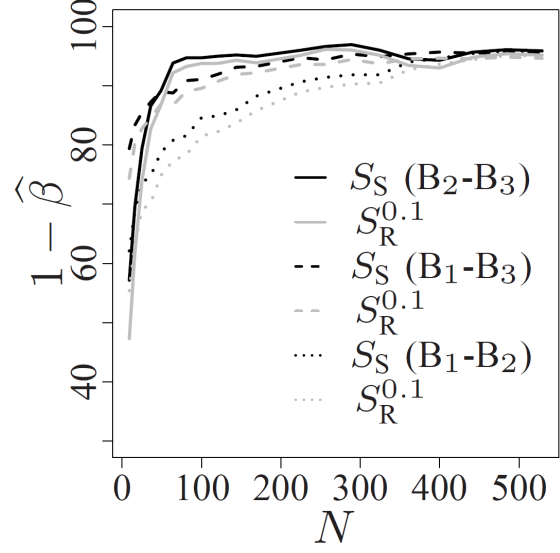


(c) Size 10%

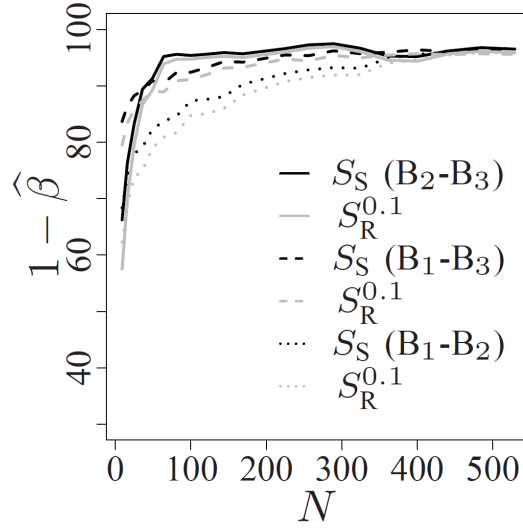
Fig. 3. Empirical test size for different regions at several sample sizes.



(a) 1%



(b) 5%



(c) 10%

Fig. 5. Empirical test power for different regions at several sample sizes.



## 5.6 Testes de hipóteses em imageamento PolSAR: Teoria da informação vs. Razão de verossimilhanças

Sistemas PolSAR operam registrando a amplitude e a fase do sinal retornado para todas as combinações de polarizações de transmissão e recepção linear: HH, HV, VH e VV (H para polarização horizontal e V para, vertical). Como resultado desta dinâmica, os pixels são representados por matrizes Hermitianas positivas definidas. O valor esperado destas matrizes estocásticas, denominado por matriz de covariância, guarda tanto as intensidades de cada canal de polarização quanto suas covariâncias. Dentro disto, a proposta de testes de hipótese baseados na matriz de covariância tem um papel central na análise de dados PolSAR: ela não só fornece meios para avaliar homogeneidade em dados PolSAR como também funciona como suporte para técnicas de processamento destas imagens, tal como detecção de bordas (Schou et al. 2003).

Algumas ferramentas estatísticas têm sido desenvolvidas neste sentido. Conradsen et al. (2003) propuseram uma metodologia baseada na razão de verossimilhanças objetivando discriminar duas regiões distribuídas conforme uma lei Wishart complexa original (ou não escalonada). Neste caso, a estatística de teste é avaliada em estimadores baseados nas hipóteses nula e alternativa, o que torna o seu uso computacionalmente custoso. Por outro lado, considerando ferramentas da Teoria da Informação, testes de hipóteses para duas amostras de dados PolSAR monopolarizados foram propostos por Nascimento et al. (2010). Subsequentemente, esses resultados foram estendidos por Frery, Nascimento & Cintra (2011) e aplicados a detecção de borda em imagens PolSAR por Nascimento et al. (2011) (como discutido na Seção 5.4). As estatísticas dos testes baseadas em medidas da Teoria da Informação são mais simples do que a proposta de Conradsen et al., dependendo apenas das estimativas por MV com relação a duas amostras consideradas, como discutido na Proposição 2 da página 74.

Esta seção objetiva

- derivar a expressão da estatística de razão de verossimilhança em termos da distribuição Wishart complexa escalonada. Neste caso, o resultado obtido é original visto que se baseia no teste da razão de verossimilhanças para a distribuição Wishart escalonada (diferentemente da proposta de Conradsen et al. (2003), baseada na Wishart original). É importante ressaltar, com base no Capítulo 3.1.2, que a lei Wishart escalonada, embora tenha uma densidade um pouco mais complexa, tem duas propriedades que a fazem estatisticamente mais razoável do que a não escalonada: (i) a média amostral é um estimador centrado para a matriz de covariância e (ii) os parâmetros são ortogonais e, portanto, as equações de verossimilhanças são separáveis. Adicionalmente, a estatística resultante é relacionada algebricamente às estatísticas marginais dos elementos da diagonal principal das matrizes amostrais, isto é, às estatísticas para as intensidades dos canais de polarização.

✎ , através de experimentos de Monte Carlo, quantificar a performance de quatro metodologias paramétricas para avaliar homogeneidade baseadas na matriz de covariância: duas delas consideram as entropias de Shannon e Rényi (tal como discutidas na Seção 5.5), outra dependente da distância de Kullback-Leibler ( Seção 5.3) e uma outra baseada na estatística da razão de verossimilhança.

A seguir, a Seção 5.6.1 apresenta uma discussão do teste da razão de verossimilhança para duas amostras distribuídas segundo uma lei Wishart complexa escalonada. Um estudo da performance de quatro testes de hipótese em dados sintéticos e reais é realizado na Seção 5.6.2. Finalmente, a Seção 5.6.3 elenca os principais resultados envolvidos na Seção 5.6.

### 5.6.1 Teste da razão de verossimilhanças para duas amostras distribuídas em uma lei Wishart complexa escalonada

Esta subseção fornece uma discussão sobre o teste da razão de verossimilhança para duas amostras no contexto de dados PolSAR. Para tal fim, nós assumiremos que estas amostras serão retiradas de matrizes populacionais com distribuição Wishart complexa escalonada.

A estatística da razão de verossimilhanças, denominada por  $S_{RV}$ , tem uma grande importância em inferência sobre modelos paramétricos. Seja  $S_{RV}$  a estatística RV para avaliar uma dada hipótese simples  $\mathcal{H}_0$ . Um resultado assintótico importante é que  $S_{RV}$  baseada em  $\mathcal{H}_0$  converge em distribuição para  $\chi^2_q$ , em que  $q$  representa a diferença entre as dimensões dos espaços paramétricos nas hipóteses alternativa e nula (para uma discussão mais detalhada, consultar Casella & Berger 2002). A fim de formalizar este teste de hipótese para distribuições Wishart complexa escalonada considere a seguinte discussão.

#### 5.6.1.1 Discussão para Wishart escalonada

Sejam  $\underline{X} = \{X_1, X_2, \dots, X_{N_1}\}$  e  $\underline{Y} = \{Y_1, Y_2, \dots, Y_{N_2}\}$  duas amostras iid obtidas das distribuições populacionais  $\mathcal{W}(\Sigma_1, L_1)$  e  $\mathcal{W}(\Sigma_2, L_2)$  cuja densidade é dada na equação (3.7) da página 28 com tamanhos  $N_1$  e  $N_2$ , respectivamente. Neste caso, a estatística LR para testar  $\mathcal{H}_0 : (\Sigma_1, L_1) = (\Sigma_2, L_2)$  vs.  $\mathcal{H}_1 : (\Sigma_1, L_1) \neq (\Sigma_2, L_2)$  é dada por

$$S_{RV} = -2\{\ell(\tilde{\theta}) - \ell(\hat{\theta})\} = -2 \log \lambda_{\mathcal{W}(\Sigma, L)},$$

em que  $\ell(\tilde{\theta})$  e  $\ell(\hat{\theta})$  representam as log-verossimilhanças conjuntas avaliadas nos estimadores MV para  $\theta = [L_1, L_2, \text{vec}(\Sigma_1)^t, \text{vec}(\Sigma_2)^t]^t$  nas hipóteses nula ( $\tilde{\theta}$ ) e alternativa ( $\hat{\theta}$ ), respectivamente, e  $\lambda_{\mathcal{W}(\Sigma, L)}$  representa a razão entre as funções de log-

verossimilhança conjuntas nas hipóteses nula e alternativa cujo logaritmo é dado por

$$\begin{aligned}
 \log \lambda_{\mathcal{W}(\Sigma, L)} &= \mathcal{A}(m) + \log \frac{|\bar{\mathbf{X}}|^{N_1 \hat{L}_1} |\bar{\mathbf{Y}}|^{N_2 \hat{L}_2}}{|\bar{\mathbf{Z}}|^{(N_1 + N_2) \tilde{L}_c}} + \underbrace{\sum_{i=1}^{N_1} \text{tr}[(\hat{L}_1 \bar{\mathbf{X}}^{-1} - \tilde{L}_c \bar{\mathbf{Z}}^{-1}) \mathbf{X}_i]}_{m \hat{L}_1 N_1 - \tilde{L}_c N_1 \text{tr}(\bar{\mathbf{Z}}^{-1} \bar{\mathbf{X}})} \\
 &\quad + (\tilde{L}_c - \hat{L}_1) \sum_{i=1}^{N_1} \log |\mathbf{X}_i| + \underbrace{\sum_{i=1}^{N_2} \text{tr}[(\hat{L}_2 \bar{\mathbf{Y}}^{-1} - \tilde{L}_c \bar{\mathbf{Z}}^{-1}) \mathbf{Y}_i]}_{m \hat{L}_2 N_2 - \tilde{L}_c N_2 \text{tr}(\bar{\mathbf{Z}}^{-1} \bar{\mathbf{Y}})} + (\tilde{L}_c - \hat{L}_2) \sum_{i=1}^{N_2} \log |\mathbf{Y}_i| \\
 \log \lambda_{\mathcal{W}(\Sigma, L)} &= \mathcal{A}(m) + \log \frac{|\bar{\mathbf{X}}|^{N_1 \hat{L}_1} |\bar{\mathbf{Y}}|^{N_2 \hat{L}_2}}{|\bar{\mathbf{Z}}|^{(N_1 + N_2) \tilde{L}_c}} + m(N_1 \hat{L}_1 + N_2 \hat{L}_2) - m(N_1 + N_2) \tilde{L}_c \\
 &\quad + (\tilde{L}_c - \hat{L}_1) \sum_{i=1}^{N_1} \log |\mathbf{X}_i| + (\tilde{L}_c - \hat{L}_2) \sum_{i=1}^{N_2} \log |\mathbf{Y}_i| \tag{5.24}
 \end{aligned}$$

Na equação (5.24),

- $\bar{\mathbf{X}}$  e  $\bar{\mathbf{Y}}$  (estimadores MV para  $\Sigma_1$  e  $\Sigma_2$  em  $\mathcal{H}_1$ ) são as médias amostrais baseadas nas amostras  $\underline{\mathbf{X}}$  e  $\underline{\mathbf{Y}}$ , respectivamente,
- $\bar{\mathbf{Z}} = (N_1 + N_2)^{-1} [N_1 \bar{\mathbf{X}} + N_2 \bar{\mathbf{Y}}]$  representa o estimador MV para  $\Sigma_1 = \Sigma_2 = \Sigma$  em  $\mathcal{H}_0$ ,
- e a quantidade  $\mathcal{A}(m)$  é dada por

$$\mathcal{A}(m) = m \log \frac{\tilde{L}_c^{\tilde{L}_c(N_1 + N_2)}}{\hat{L}_1^{\hat{L}_1 N_1} \hat{L}_2^{\hat{L}_2 N_2}} + \log \frac{\Gamma_m(\hat{L}_1)^{N_1} \Gamma_m(\hat{L}_2)^{N_2}}{\Gamma_m(\tilde{L}_c)^{N_1 + N_2}},$$

em que  $\hat{L}_1$  e  $\hat{L}_2$  são as estimativas MV na hipótese alternativa (como discutido na Seção 3.1.3) e  $\tilde{L}_c$  representa a estimativa MV em  $H_0$ , a qual é dada pela solução da seguinte equação não linear:

$$\begin{aligned}
 m \log \tilde{L}_c - \psi_m^{(0)}(\tilde{L}_c) &= \log |\bar{\mathbf{Z}}| - m \\
 &\quad + (N_1 + N_2)^{-1} \left\{ \sum_{i=1}^{N_1} [\text{tr}(\bar{\mathbf{Z}}^{-1} \mathbf{X}_i) - \log |\mathbf{X}_i|] + \sum_{i=1}^{N_2} [\text{tr}(\bar{\mathbf{Z}}^{-1} \mathbf{Y}_i) - \log |\mathbf{Y}_i|] \right\}.
 \end{aligned}$$

Como a distribuição Wishart complexa se degenera na distribuição gamma dada em (3.12), a estatística  $S_{\text{LR}}$  pode ser reduzida a  $S_{\text{RV}}(k) = -2 \log \lambda_{\Gamma(L, \theta_k)}$ , para  $k \in$

$\{1(\text{HH}), 2(\text{HV}), 3(\text{VV})\}$ , tal que

$$\begin{aligned} \log \lambda_{\Gamma(L, \theta_k)} = & \mathcal{A}(1) + \log \frac{\bar{X}_k^{N_1 \hat{L}_1} \bar{Y}_k^{N_2 \hat{L}_2}}{\bar{Z}_k^{(N_1+N_2) \tilde{L}_c}} + (\tilde{L}_c - \hat{L}_1) \sum_{i=1}^{N_1} \log X_{ik} + (\tilde{L}_c - \hat{L}_2) \sum_{i=1}^{N_2} \log Y_{ik} \\ & + \underbrace{\frac{\hat{L}_1 \bar{Z}_k - \tilde{L}_c \bar{X}_k}{\bar{Z}_k \bar{X}_k} \sum_{i=1}^{N_1} X_{ik} + \frac{\hat{L}_2 \bar{Z}_k - \tilde{L}_c \bar{Y}_k}{\bar{Z}_k \bar{Y}_k} \sum_{i=1}^{N_2} Y_{ik}}_{N_1 \hat{L}_1 + N_2 \hat{L}_2 - (N_1 + N_2) \tilde{L}_c}, \end{aligned}$$

em que  $X_{ik}$ ,  $Y_{ik}$ ,  $\bar{X}_k$ ,  $\bar{Y}_k$  e  $\bar{Z}_k$  são as entradas  $(k, k)$  das matrizes  $\mathbf{X}_i$ ,  $\mathbf{Y}_i$ ,  $\bar{\mathbf{X}}$ ,  $\bar{\mathbf{Y}}$  e  $\bar{\mathbf{Z}}$ , respectivamente. Embora esta abordagem não seja explorada nesta seção, ela é importante pois sugere um teste de hipótese para  $\mathcal{H}_0 : (\theta_1, L_1) = (\theta_2, L_2)$  (em que  $\theta_1$  e  $\theta_2$  representam as médias de um canal de polarização para duas regiões diferentes).

### 5.6.1.2 Comparação com a estatística proposta por Conradsen

Nesta seção, objetiva-se mostrar a relação entre a estatística proposta por Conradsen et al. (2003) e aquela derivada na seção anterior. Na verdade, é possível provar que a proposta de Conradsen et al. (2003) é um caso particular quando (i) os números de looks são conhecidos e os tamanhos de amostras são similares.

Consideremos, inicialmente, a expressão para razão de verossimilhanças baseada nas amostras aleatórias  $\underline{\mathbf{X}}$  e  $\underline{\mathbf{Y}}$  quando os números de looks são conhecidos e  $\mathcal{H}_0 : \Sigma_1 = \Sigma_2$ :

$$\lambda_{\mathcal{W}(\Sigma, L)} = \frac{|\bar{\mathbf{X}}|^{N_1 L_1} |\bar{\mathbf{Y}}|^{N_2 L_2}}{|\bar{\mathbf{Z}}|^{N_1 L_1 + N_2 L_2}}.$$

Assim, considerando  $N_1 = N_2 = 1$ , temos que

$$\lambda_{\mathcal{W}(\Sigma, L)} = \frac{|\frac{L_1}{L_1} \mathbf{X}|^{L_1} |\frac{L_2}{L_2} \mathbf{Y}|^{L_2}}{|\frac{L_1+L_2}{L_1+L_2} (\mathbf{X} + \mathbf{Y})|^{L_1+L_2}} = \frac{|\frac{\mathbf{X}_1}{L_1}|^{L_1} |\frac{\mathbf{Y}_2}{L_2}|^{L_2}}{|\frac{(\mathbf{X}_1 + \mathbf{Y}_2)}{L_1+L_2}|^{L_1+L_2}} = \frac{(L_1 + L_2)^{m(L_1+L_2)}}{L_1^{m L_1} L_2^{m L_2}} \frac{|\mathbf{X}_1|^{L_1} |\mathbf{Y}_2|^{L_2}}{|\mathbf{X}_1 + \mathbf{Y}_2|^{L_1+L_2}},$$

em que  $\mathbf{X}_1$  e  $\mathbf{Y}_2$  tem distribuições Wishart não escalonada definida nos parâmetros  $(L_1, \Sigma_1)$  e  $(L_2, \Sigma_2)$ . Esta última expressão é a estatística obtida por Conradsen et al. (2003) e, portanto, seu logaritmo é uma caso particular da estatística dada na equação (5.24).

## 5.6.2 Estudo de simulação: Razão de verossimilhança vs. Teoria da Informação

A fim de comparar as metodologias propostas nas Seções 5.3 e 5.5 com a subseção anterior, nós consideramos dados sintéticos Wishart distribuídos e informações PolSAR reais. Neste caso, nós adotamos níveis de significância 1%, 5% e 10%.

## 5.6.2.1 Estudo com dados sintéticos

Dados polarimétricos gerados da distribuição Wishart complexa foram utilizado para comparar testes com as seguintes estatísticas:

- Estatística da razão de verossimilhança:  $S_{RV}$ ;
- Estatística de Kullback-Leibler:  $S_{KL}$  (discutido na Seção 5.3);
- Estatísticas baseadas nas entropias de Shannon e Rényi, respectivamente:  $S_S$  e  $S_R^\beta$  (discutidas na Seção 5.5).

Diferentemente dos esquemas teóricos da subseção anterior e das abordagens das Seções 5.3 e 5.5, a seguinte discussão considera que o número de looks é conhecido (não estimáveis) e, portanto, procuraremos testar  $\mathcal{H}_0 : \Sigma_1 = \Sigma_2$ . Nós adotamos esta restrição pois a estatística da razão de verossimilhança se mostrou muito sensível a mudanças no número de looks,  $L$ , apresentando estimativas inaceitáveis para o tamanho do teste. Este fato já coloca os testes baseados em medidas da Teoria da Informação em vantagem com relação a LR. A seguir, nós discutimos o processo de simulação. Depois que duas amostras de dados foram geradas a partir de  $\mathbf{X} \sim \mathcal{W}(\Sigma_1, L)$  e  $\mathbf{Y} \sim \mathcal{W}(\Sigma_2, L)$ , nós procuramos evidência estatística para rejeitar a hipótese nula  $\mathcal{H}_0 : \Sigma_1 = \Sigma_2$ .

Para tal fim, seja  $T$  o número de réplicas de Monte Carlo e  $C$  o número de caso em que a hipótese nula é rejeitada no nível nominal  $\alpha \in \{1\%, 5\%, 10\%\}$ . Assim o tamanho empírico do teste é dado por

$$\hat{\alpha}_{1-\alpha} = \frac{C}{T}.$$

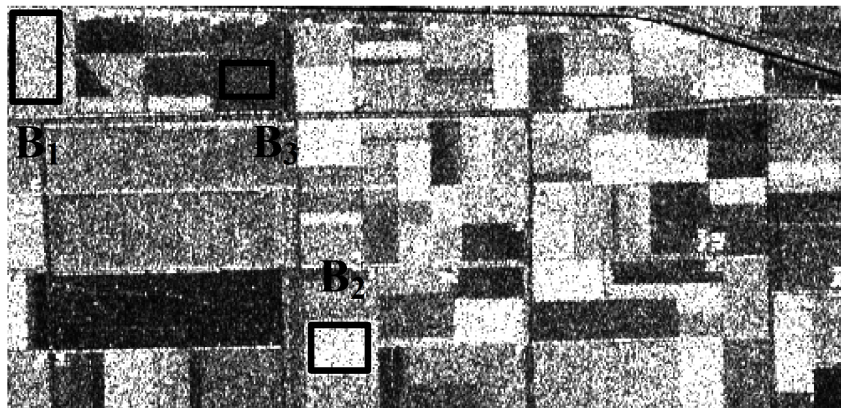
Seguindo a metodologia descrita por Nascimento et al. (2010), nós utilizamos  $T = 5500$  replicações de Monte Carlo. Em simulações com dados PolSAR, é comum que as matrizes de covariâncias utilizadas sejam obtidas dos dados reais. Para tal fim, considere a seguinte aplicação a dados reais. A Figura 5.26 mostra uma área de uma imagem AIRSAR de Flevoland, Holanda (obtida em agosto de 1989, consultar Doulgeris, Anfinsen & Eltoft 2011). Esta imagem foi obtida considerando número nominal de looks 4 e apresenta três regiões selecionadas (conforme discussão no Apêndice A.1.3).

A Tabela 5.17 apresenta um resumo das estimativas. Baseando-se nos valores das variâncias generalizadas,  $|\hat{\Sigma}|$ , a região  $B_2$  tem a textura mais intensa, seguindo-se as regiões  $B_2$  (com média intensidade) e  $B_3$  (a menos intensa). Como o tom de cinza mais claro indica cenários com textura mais intensa, então as estimativas estão em concordância com a Figura 5.26.

Neste estudo a matriz de covariância amostral obtida da região “ $B_1$ ”,  $\Sigma_B$ , e o número de looks  $L = 4$  foram considerados como parâmetros da simulação. Além disso assumimos os tamanhos de amostra  $N_1 = N_2 = N \in \{10, 11, \dots, 50\}$ .

**Tabela 5.16** Tamanhos empíricos dos testes

| $N$                                       | Valores médios |      |       |           |        |
|-------------------------------------------|----------------|------|-------|-----------|--------|
|                                           | 1%             | 5%   | 10%   | $\bar{S}$ | CV     |
| Razão de Verossimilhança ( $S_{RV}$ )     |                |      |       |           |        |
| 10 – 20                                   | 1.21           | 5.76 | 11.16 | 9.25      | 47,01  |
| 21 – 30                                   | 1.10           | 5.42 | 10.66 | 9.12      | 47,20  |
| 31 – 40                                   | 1.03           | 5.13 | 10.43 | 9.09      | 47,10  |
| 41 – 50                                   | 1.06           | 5.21 | 10.28 | 9.08      | 46,97  |
| Entropia de Shannon ( $S_S$ )             |                |      |       |           |        |
|                                           | 1.00           | 4.59 | 9.47  | 1.00      | 172,27 |
|                                           | 0.99           | 4.64 | 9.21  | 1.02      | 219,26 |
|                                           | 0.99           | 4.44 | 9.24  | 1.05      | 280,30 |
|                                           | 0.93           | 4.54 | 9.37  | 1.07      | 321,46 |
| Entropia de Rényi ( $S_R^{0,1}$ )         |                |      |       |           |        |
|                                           | 0.00           | 1.71 | 4.53  | 0.71      | 208,71 |
|                                           | 0.33           | 1.83 | 4.53  | 0.74      | 278,51 |
|                                           | 0.32           | 1.79 | 4.46  | 0.77      | 360,24 |
|                                           | 0.29           | 1.78 | 4.56  | 0.79      | 419,10 |
| Distância de Kulback-Leibler ( $S_{KL}$ ) |                |      |       |           |        |
|                                           | 1.83           | 7.06 | 12.89 | 9.53      | 48,43  |
|                                           | 1.43           | 6.16 | 11.66 | 9.27      | 47,99  |
|                                           | 1.24           | 5.67 | 11.17 | 9.20      | 47,65  |
|                                           | 1.24           | 5.55 | 10.85 | 9.16      | 47,39  |

**Figura 5.26** Imagem AIRSAR de Flevoland (canal HH).

**Tabela 5.17** Estimativas dos parâmetros

| Regiões        | $\hat{L}$ | $ \hat{\Sigma} $       | # pixels |
|----------------|-----------|------------------------|----------|
| B <sub>1</sub> | 3.470     | $7.78 \times 10^{-8}$  | 1566     |
| B <sub>2</sub> | 3.514     | $9.45 \times 10^{-7}$  | 980      |
| B <sub>3</sub> | 3.530     | $7.22 \times 10^{-10}$ | 651      |

A Tabela 5.16 apresenta os comportamentos das seguintes medidas: (i) tamanhos empírico nos níveis nominais 1%, 5% e 10%, (iii) médias dos valores das estatísticas de teste ( $\bar{S}$ ) e o coeficiente de variação (CV) das estatísticas consideradas.

Os tamanhos estimados mais próximos aos níveis nominais adotados foram apresentados pelos testes  $S_{RV}$  e  $S_S$ . Este fato pode ser melhor observado na Figura 5.27, gráficos das estimativas dos níveis nominais para diferentes tamanhos de amostra. Em termos gerais, os resultados apontam para a satisfação da seguinte desigualdade:

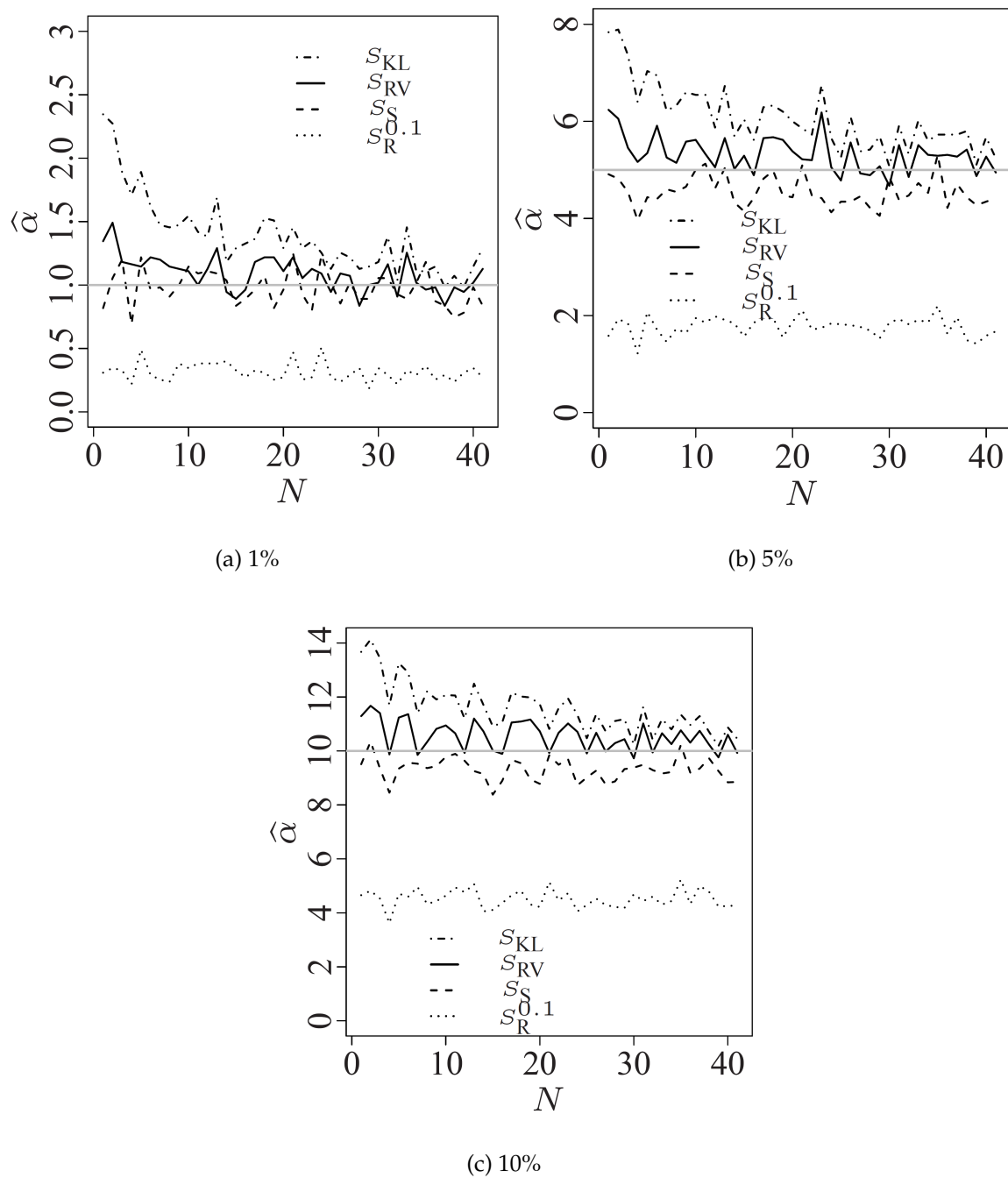
$$\hat{\alpha}(S_{KL}) \geq \hat{\alpha}(S_{RV}) \geq \hat{\alpha}(S_S) \geq \hat{\alpha}(S_R^{0.1}), \quad (5.25)$$

em que  $\hat{\alpha}(S_{\bullet})$  representa a estimativa do tamanho de um teste com estatística  $S_{\bullet}$ . Além disso, as estatísticas baseadas nas entropias apresentam maior variabilidade em torno do seu valor médio, isto é, possuem maiores CV. Esta situação se agrava quando aumentamos o tamanho de amostra, entretanto as estimativas para o tamanho destes testes tende a subestimá-lo. Este fato sugere que os testes dependentes de entropias tendem a errar menos do que nível nominal prefixado quando se aumenta o tamanho de amostra. As estatísticas  $S_{RV}$  e  $S_{KL}$  tem um comportamento esperado. Pela lei fraca dos grandes números, espera-se que a média amostral das estatísticas (que neste caso tem uma distribuição assintótica  $\chi_p^2$  com  $p = 9$ , números de parâmetros de uma matriz  $\Sigma$   $3 \times 3$ ) possa convergir para a média populacional, isto é, para o grau de liberdade de uma lei  $\chi_9^2$ , 9. Neste caso, a razão de verossimilhança supera a estatística de Kullback-Leibler. Mas, vale apenas frisar que  $S_{RV}$  dependente do termo dado na equação (5.24), que tem um custo computacional adicional á expressão  $S_{KL}$  dependente da distância de Kullback-Leibler dada na (5.3) da página 137. Assim, para situações em que os testes das estatísticas  $S_{KL}$  e  $S_{RV}$  apresentam estimativas próxima, como na Figura 5.27(a) para  $N > 30$ , pode ser mais recomendável o uso do teste de Kullback-Leibler.

Adicionalmente, um estudo do poder do teste foi realizado. Com este fim, nós objetivamos rejeitar a hipótese  $\mathcal{H}_0$  a partir de amostras das distribuições  $\mathbf{X} \sim \mathcal{W}(\Sigma_B, 4)$  e  $\mathbf{Y} \sim \mathcal{W}(\Sigma_B \cdot (1 + k), 4)$  para  $k = 0.2, 0.3, 0.4$ . Neste caso, a taxa

$$\hat{\beta} = \frac{T - C}{T}$$

é conhecida como erro do tipo II e nós estaremos interessados em avaliar o poder do teste,  $1 - \hat{\beta}$ .



**Figura 5.27** Tamanhos empíricos dos testes baseados em dados sintético para diferentes cenários nos níveis 1%, 5% e 10%.



A Figura 5.28 apresenta os resultados dos poderes estimados para vários tamanhos de amostra. Note que o teste baseado na entropia de Shannon apresenta as melhores estimativas para o poder. Outro evento interessante é que a ordem que se verifica na equação (5.25) se repete apenas dentro dos grupos  $\{S_{KL}, S_{RV}\}$  e  $\{S_S, S_R\}$ . Este fato pode ser explicado pela (i) relação existente entre o Lemma de Neyman Pearson e a distância de Kullback-Leibler (para detalhes, consultar Eguchi & Copas 2006) e (ii) pelo limite  $\lim_{\beta \rightarrow 1} S_R^\beta = S_S$ . Este fato combinado com a análise dos tamanhos estimados indicam  $S_S$  como o melhor discriminador e indentificador de homogeneidade em cenários modelado pela distribuição Wishart complexa escalonada.

### 5.6.2.2 Estudo com dados reais

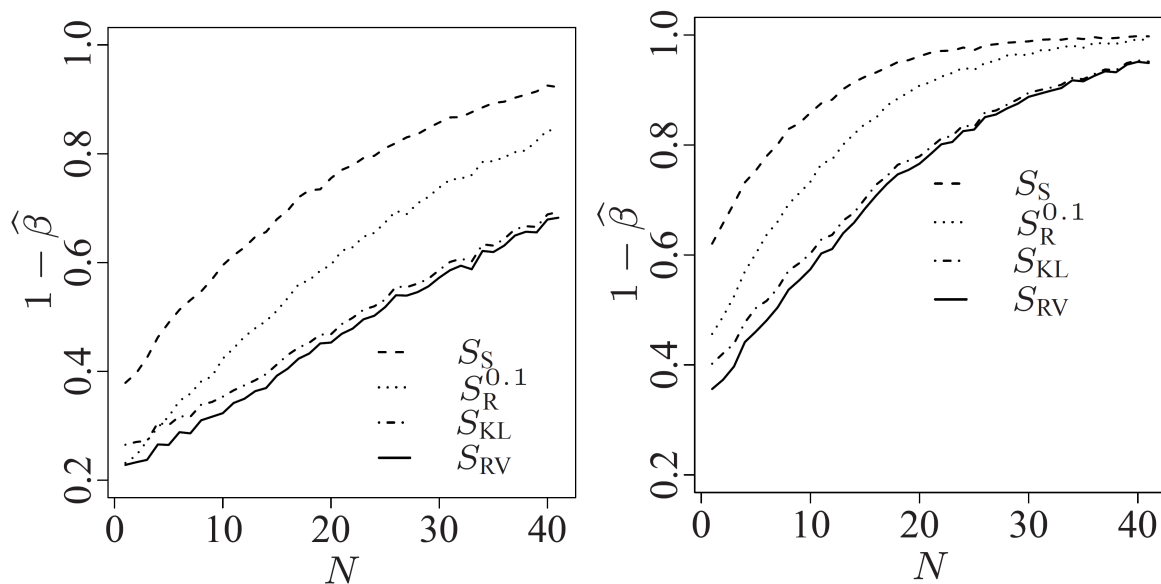
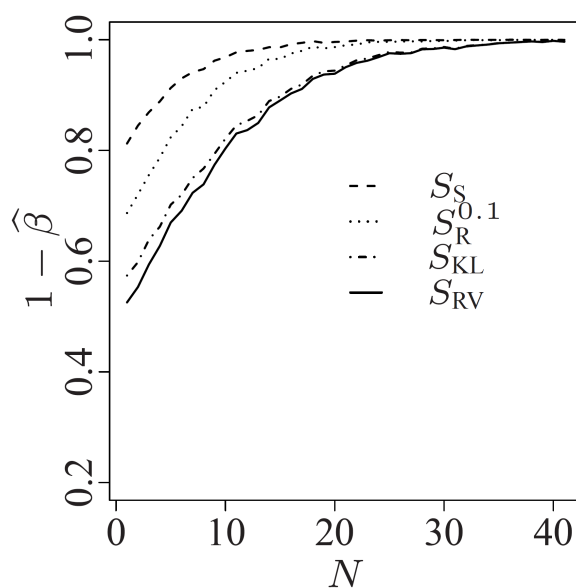
Seguindo a mesma estrutura de aplicação da Seção 5.5 com a restrição da hipótese nula a  $\mathcal{H}_0 : \Sigma_1 = \Sigma_2$ , nós estimamos os tamanhos dos testes para os dados reais. Tal restrição parece razoável a este conjunto de dados, pois as estimativas de  $L$  são próximas entre si. Adicionalmente, como as estimativas são próximas do número nominal de looks,  $L = 4$ , as curvas das distribuições  $\mathcal{W}_R$  e  $\mathcal{W}$  não devem destoar muito. A fim de checar este fato, as Figuras 5.29(a)-5.29(c) mostram os ajustes de duas distribuições Wishart complexa escalonadas:  $\mathcal{W}_R(\hat{\Sigma}, \hat{L})$  (curva na cor preta) e  $\mathcal{W}(\hat{\Sigma}, 4)$  (cinza). Note que os ajustes em ambos os casos não diferem muito para as amostras utilizadas, o que sugere como razoável a comparação apenas das matrizes de covariâncias para este banco de dados.

A fim de quantificar a performance dos teste para os dados da Figura 5.26 os valores de  $\hat{\alpha}$  foram estimados. Os resultados em termos dos tamanhos de amostra e das regiões selecionadas são apresentados na Figura 5.30. Neste caso, verifica-se que a desigualdade entre as estimativas dos tamanhos dos testes em dados sintéticos se repete em dados reais. Os resultados mostraram evidência de que o teste  $S_R^{0.1}$  supera os outros testes. Este teste apresentou bons resultados até mesmo para tamanhos pequenos de amostra de dados PolSAR. Assim, esta medida é sugerida como uma ferramenta pertinente a análise em pequenas janelas, tais como em técnicas de filtragem em dados PolSAR.

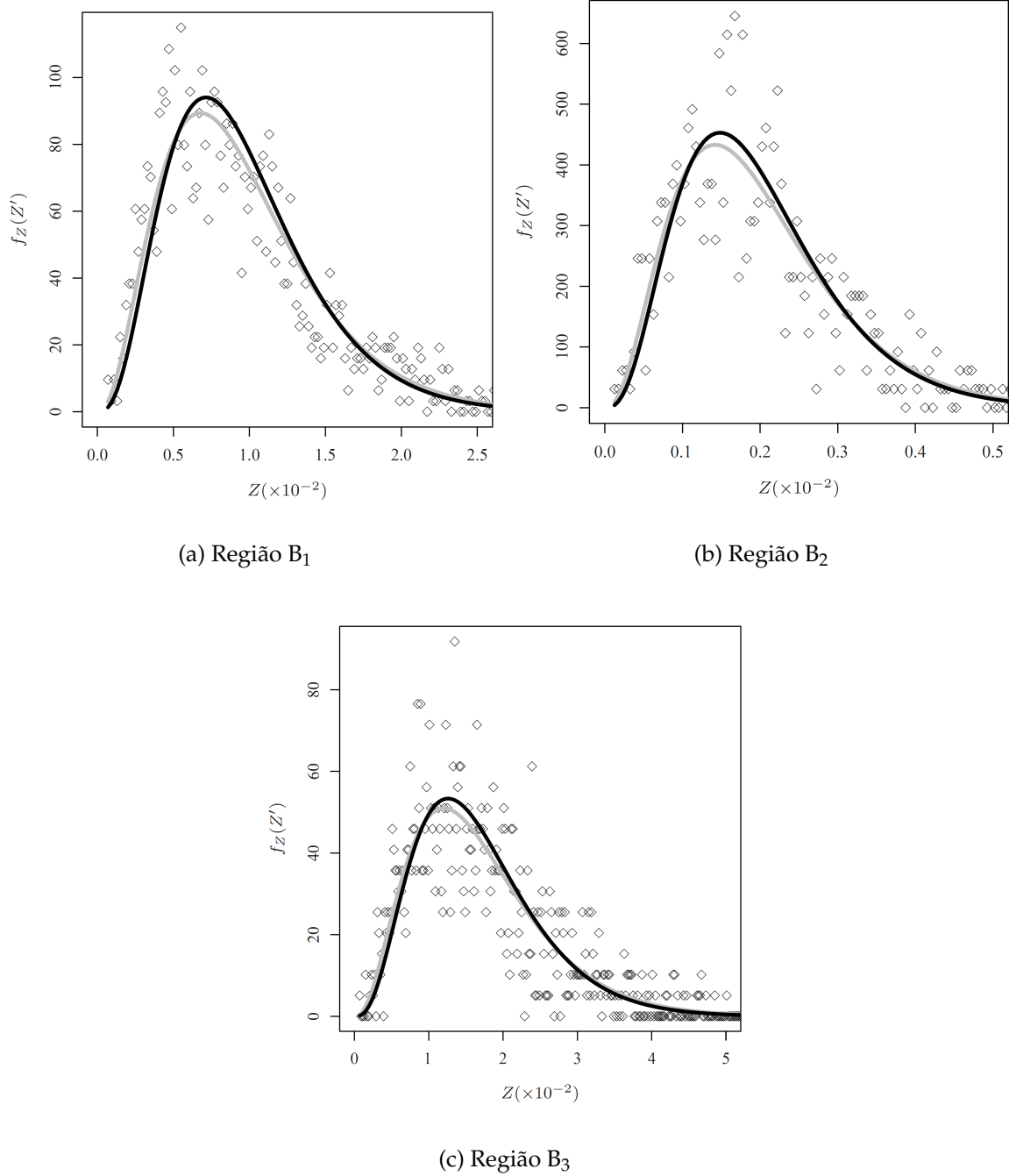
### 5.6.3 Considerações da Seção 5.6

Nesta seção, nós estimamos a performance de quatros testes de homogeneidade ((a) razão de verossimilhança,  $S_{RV}$ , (b) Entropia de Shannon,  $S_S$ , (c) Entropia de Rényi,  $S_R^{0.1}$ , (d) Distância de Kullback-Leibler,  $S_{KL}$ ) para dados de imagens sintéticas e reais. Para este fim, as estimativas de seus tamanhos e poderes foram utilizadas como critério de seleção.

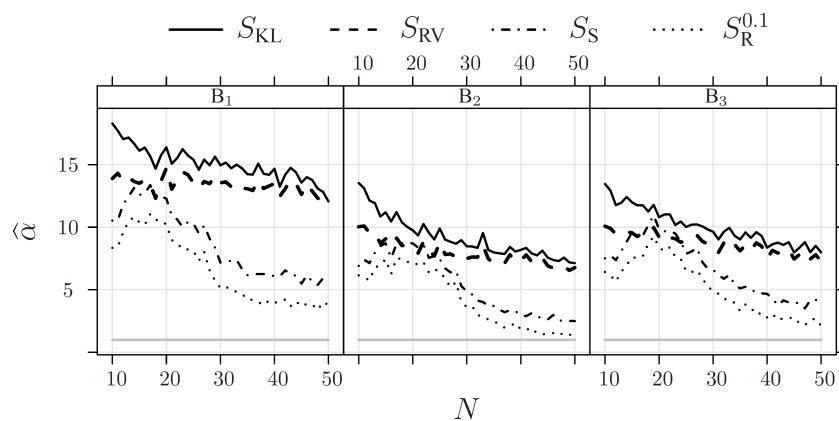
Cenários sintéticos foram gerados segundo experimentos de Monte Carlo a exemplo da Seção 5.5. Em termos dos tamanhos dos testes estimados, os testes da razão

(a)  $(\Sigma, 4)$  vs.  $(\Sigma_B \cdot (1 + 0.2), 4)$ (b)  $(\Sigma, 4)$  vs.  $(\Sigma_B \cdot (1 + 0.3), 4)$ (c)  $(\Sigma, 4)$  vs.  $(\Sigma_B \cdot (1 + 0.4), 4)$ 

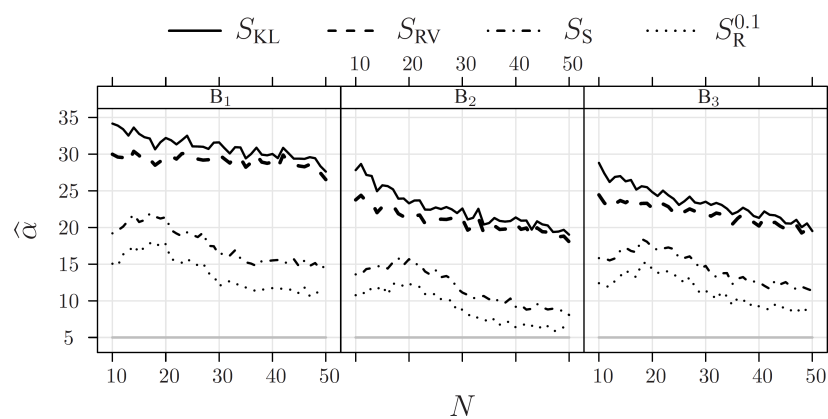
**Figura 5.28** Poderes empíricos dos testes baseados em dados sintético para diferentes cenários no nível 1%.



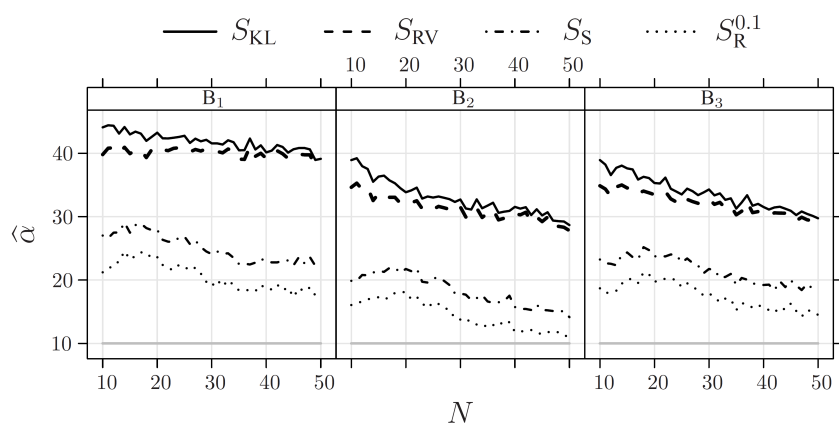
**Figura 5.29** Histogramas e curvas de densidades ajustadas: com número de looks fixado (cinza) e estimável (preta).



(a) 1%



(b) 5%



(c) 10%

**Figura 5.30** Tamanhos empíricos do testes para dados reais em três diferentes regiões e nos níveis 1%, 5% e 10%.

de verossimilhança e aquele baseado na entropia de Shannon apresentaram os melhores desempenhos. Em particular,  $S_S$  se mostrou mais preciso em tamanhos amostra menores. Adicionalmente, o seguinte comportamento foi observado: os testes como estatísticas  $S_{RV}$  e  $S_{KL}$  tendem a sobrestimar o nível nominal adotado, enquanto que aqueles baseados nas entropias tendem a subestimar. Quanto aos poderes estimados, o teste baseado na entropia de Shannon apresentou o melhor resultado em todos os cenários considerados.

Como os tamanhos dos testes se mostraram bastante competitivos em imagens sintéticas, uma aplicação a dados reais foi realizada avaliando as estimativas deste critério. Para todas as regiões selecionadas, o teste baseado na entropia de Rényi com ordem 0.1 apresentou os melhores resultados. Note que esta entropia converge para a de Shannon quando sua ordem tende para um. Assim, parece sugestivo que o melhor teste é baseado na entropia de Rényi de tal forma que possamos obter o melhor  $\beta$  a partir dos dados.

#### 5.6.4 Trabalhos futuros

Como complemento desta seção, objetiva-se

- ✎ Desenvolver métodos de estimação da ordem  $\beta$  da entropia de Rényi a partir de dados PolSAR.
- ✎ aplicar as medidas propostas baseadas em entropias a cenários de filtragem.
- ✎ explorar este mesmo estudo em outros cenários com dados PolSAR reais.

## Capítulo 6

# Conclusões

Esta tese propôs métodos de inferência estatística baseados em Teoria da Informação para dados polarimétricos, em particular, para aplicações de radar de abertura sintética. Além da abordagem estatística, as ferramentas derivadas são incorporadas a contextos de processamento de imagens PolSAR. Desta forma, a tese pode ser dividida em duas partes. Inicialmente, discute-se toda a parte de métodos envolvendo conceitos de (i) radar polarimétrico, (ii) modelagens de imagens polarimétricas e (iii) teoria estatística de informação. Em segundo lugar, apresentam-se os principais resultados desta tese na forma de seis seções, que são resumidas abaixo.

**Seção 5.1:** Considerando que os pixels de uma imagem SAR são representados por uma variável aleatória escalar com distribuição  $\mathcal{G}_I^0$ , nós comparamos os desempenhos de testes baseados na classe de distância  $(h, \phi)$  como o teste não paramétrico de Kolmogorov-Smirnov. Este estudo foi baseado em experimentos Monte Carlo, tendo como critérios de comparação os poderes e tamanhos estimados dos testes. Adicionalmente, os resultados foram quantificados na presença de um modelo paramétrico de contaminação que representa a presença dos fenômenos de “reflexão de canto” e “double bounce” (típicos de regiões de cidades e florestas densas). Os resultados mostraram evidências de que os testes baseados nas distâncias triangular e aritmética geométrica apresentaram, em média, melhores resultados do que os demais (tais como o não paramétrico de Kolmogorov Smirnov e aquele baseado na distância de Kulback-Leibler).

**Seção 5.2:** Imagens PolSAR são afetadas por um tipo particular de ruído speckle multidimensional, tal como discutido por López-Martínez et al. (2005). Neste caso, os pixels podem ser entendidos como realizações de uma distribuição matricial Wishart complexa escalonada. Um importante parâmetro desta distribuição é o número de looks, que controla a influência deste ruído sobre a qualidade de imagens PolSAR. Assim, a proposta de metodologias de estimação melhorada para o número de looks é uma contribuição importante no contexto de processamento de imagens PolSAR. Nesta seção, nós propomos duas contribuições neste sentido: (i) expressão para o viés de segunda ordem do estimador de máxima verossimilhança com base na fórmula proposta por Cox & Snell (1968) e (ii) expressão para a verossimilhança perfilada modificada por Barndorff-Nielsen (1994).

Como consequência direta, derivaram-se dois novos estimadores com vieses de ordens  $O(N^{-2})$  e  $O(N^{-3/2})$ .

**Seção 5.3:** Testes de hipóteses e técnicas de processamento de imagens com base em medidas de discriminação para dados PolSAR são duas linhas que seguem um papel importante em análise de dados PolSAR. Conradsen et al. (2003) propuseram uma metodologia baseada sobre o teste da razão de verossimilhanças objetivando discriminar duas variáveis Wishart distribuídas. Por outro lado, Goudail & Réfrégier (2004) aplicaram medidas de discriminação para caracterizar performance em algoritmos de detecção de mudanças e segmentação em processamento de imagens PolSAR. Esta seção avança no sentido de unir as duas linhas através da proposta de testes de hipótese baseados em medidas de distâncias estocásticas e aplicar estas medidas a análise de agrupamento de dados PolSAR. Neste sentido, foram propostas oito estatísticas de teste, sendo quatro delas baseadas na matriz de covariância e as outras no coeficiente correlação complexo. Estas estatísticas foram baseadas nas distâncias de Kullback-Leibler, Rényi, Bhattacharyya e Hellinger. Em geral, com base em estudos de simulação Monte Carlo e aplicações a dados reais, obtiveram-se as seguintes conclusões: os testes baseados nas distâncias de Rényi e Bhattacharyya apresentaram melhores desempenhos, enquanto que a análise de agrupamento em função da distância de Rényi se apresentou mais flexível do que as demais.

**Seção 5.4:** Imagens PolSAR são corrompidas pelo ruído speckle, dificultando o processamento de imagem, tal como em detecção de bordas. Esta seção compara quatro abordagens alternativas baseadas em distâncias estocásticas com uma outra proposta por Gambini et al. (2008). Experimento com imagem sintética apresentou evidências de que as metodologias baseadas em distâncias estocásticas são aptas a encontrar pontos de bordas em imagens PolSAR. Aplicações a dados reais mostraram que os métodos baseados sobre as distâncias de Rényi e Bhattacharyya apresentaram melhores desempenhos.

**Seção 5.5:** Analisar dados de imagens PolSAR é equivalente a lidar com um ruído multidimensional resultante de seus processamentos. Este fato exige que uma modelagem apropriada seja definida *a priori*, tal como a distribuição Wishart complexa escalonada. No contexto desta distribuição, a proposta de medidas de variabilidade escalares que carregam simultaneamente as informações da matriz de covariância e do número de looks é bastante importante. Morio et al. (2008) analisaram a entropia de Shannon para caracterização de imagens polari-métricas. Nesta seção, nós derivamos as expressões analíticas das entropias de Shannon, Rényi e Tsallis para a distribuição Wishart complexa escalonada. Expressões para suas variâncias assintóticas são derivadas e incorporadas na classe de entropias  $(h, \phi)$ , resultando na proposta de dois testes de hipóteses baseados em medidas de entropias. Estudos de simulação bem como aplicações a dados reais sugeriram a entropia de Shannon como a melhor fonte de discriminação e

a entropia de Rényi com ordem 0.1 como a mais apta a reconhecer homogeneidade.

**Seção 5.6:** A matriz de covariância tem um papel importante em análise de dados PolSAR, pois guarda tanto as intensidades de cada canal de polarização quanto as suas covariâncias. Assim a proposta de testes de hipótese baseados sobre tais matrizes são procurados. Conradsen et al. (2003) propuseram uma metodologia baseada na razão de verossimilhanças objetivando discriminar duas regiões distribuídas conforme uma lei Wishart complexa original (ou não escalonada). Por outro lado, considerando ferramentas da Teoria da Informação, testes de hipóteses para duas amostras de dados PolSAR foram propostos e discutidos por Nascimento et al. (2010) num sentido univariado e por Frery, Nascimento & Cintra (2011) em termos polarimétricos. Através de experimentos Monte Carlo, o desempenho de quatro metodologias paramétricas para avaliar homogeneidade baseadas na matriz de covariância foi analisado: dois dos métodos consideraram as entropias de Shannon e Rényi, um outro a distância de Kullback-Leibler e outro a estatística da razão de verossimilhança. O resultados mostraram evidências de que o melhor teste é aquele baseado na entropia de Rényi de tal forma que possamos obter o melhor  $\beta$  a partir dos dados.



## Fórum PolSARpro: Seu uso na captura de imagens polarimétricas

A fim de obter os conjuntos de dados reais utilizados nesta tese, as seguintes etapas foram empregadas:

1. Inicialmente, é necessário baixar os arquivos (\*.stk, \*.co) dos dados para, a partir deles, obter suas informações "ENVI", [http://geol.hu/data/online\\_help/-ENVI\\_Header\\_Format.html](http://geol.hu/data/online_help/-ENVI_Header_Format.html), com ponteiros em extensões \*.bin e \*.hdr para os elementos da matriz de coerência de cada pixel. Esta etapa foi realizada através do sítio do "PolSARproESA" (<http://earth.eo.esa.int/polsarpro/>). Nesta página, dentre as informações importantes, tais como os links para fazer download do programa PolSARpro e de cursos sobre processamentos de imagens PolSAR, o caminho

[Airborne Data Sources](#) – [Download sample data](#)

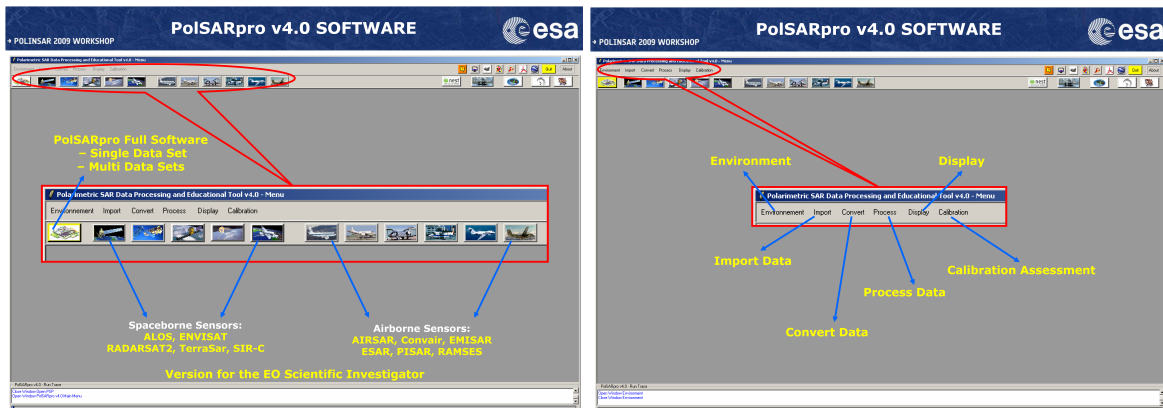
fornece as informações dos bancos de dados referentes a vários sistemas aero-transportados. Em particular, esta tese trabalhará com três destes sistemas, discutidos posteriormente: AIRSAR, EMISAR e E-SAR. Quanto ao sensor EMISAR, suas informações ENVI podem ser importadas diretamente do sítio do PolSARproESA.

2. Uma vez instalado o software PolSARpro, o seu menu principal terá as seguintes informações destacadas nas Figuras A.1(a) e A.1(b). Estas são as funções principais para a captura de dados (menus spaceborne e airborne, Enviroment, Import Data e Convert Data), o processamento de imagens PolSAR (Process Data e Calibration Assessment) e a visualização (Display). Dentro disto, considere inicialmente o seguinte caminho para mudança de diretório:

[AISAR](#) – [Enviroment](#)

3. Subsequentemente, para extrair as informações dos canais das imagens PolSAR a partir dos arquivos obtidos do sítio PolSARESA, considere:

[AISAR](#) – [Import](#) – [Input Data](#)



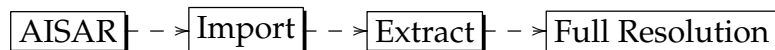
(a) Menu I

(b) Menu II

Figura A.1 Software PolSARpro.

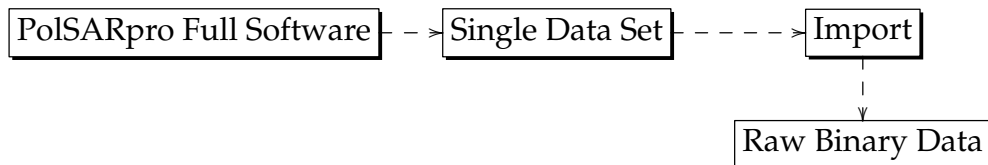
No estágio do “Input Data” algumas informações sobre os dados são solicitadas (tais como (i) a versão o processador do sensor, (ii) se os dados são mono ou multipolarizados e o (iii) número de linhas e colunas dos dados). Estas informações estão todas no arquivo \*.PDF obtido pelo sítio. Além disso os arquivos com os ponteiros aos dados são requeridos. Como suas informações ENVI são obtidas do sítio PolSARESA, este e o próximo passo podem ser desconsiderados para o sensor EMISAR.

#### 4. Usando



as informações ENVI, \*.bin e \*.hdr, para os elementos da matriz de coerência são salvas no diretório pré-especificado no passo 2.

#### 5. Para converter os elementos da matriz de coerência em elementos de covariância, considere o caminho



e no estágio “Raw Binary Data” é preciso alimentar o software com as informações \*.bin dos dados obtidos no item anterior. Feito isto, a opção “Convert” estará disponível e poderá ser utilizada no processo de conversão dos dados.

#### 6. Uma vez obtidas as informações ENVI da matriz de covariância, o seguinte código, produzido pela Profa. Dra. Michelle Horta, foi utilizado para carregar uma imagem PolSAR no software R:

---

```

# # Função ``le_arqbin_envi``
# ==> Le arquivo binário de acordo com o header do formato ENVI
# Entrada:
#   I -> nome do arquivo
#   C -> samples
#   F -> lines
#   HO -> header offset
#   DT -> data type (tipos 4-float ou 5-double)
#   BO -> byte order
# Saída:
#   Z -> Image float: NumeroLinhas x NumeroColunas

le_arqbin_envi <- function(I,C,F,HO=0,DT=4,BO=0){
  f <- 0
  conexioParaLeer <- file(I,open="r+b")
  Z <- matrix(nrow=F,ncol=C)
  # HEADER OFFSET
  if (HO!=0)
  {
    readBin(conexioParaLeer, numeric(), size=HO)
  }
  # DATA TYPE 4->32bit_float 8->64bit_double
  if (DT == 4) SZ=4 else SZ=8

  # BYTE ORDER
  if (BO == 0) BO1="little" else BO1="big"

  while (f!=F){
    c <- 0
    while (c!=C){
      ##Reading the complex data
      Z[f+1,c+1]<-readBin(conexioParaLeer,
        numeric(),size=SZ,signed=FALSE,endian=BO1, n=1)
      c <- c+1
    }
    f <- f+1
  }
  close(conexioParaLeer,rw="read")
  Z}

```

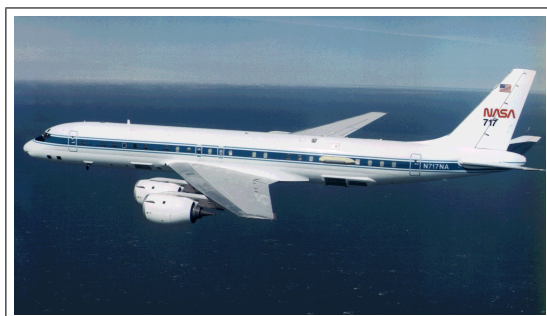
---

A seguir, nós apresentamos uma breve discussão a respeito dos sensores aerotransportados utilizados nesta tese.

## A.1 Sistemas SAR aerotransportados

### A.1.1 Sensor AIRSAR (*Airborne Synthetic Aperture Radar* - AIRSAR) e dados utilizados

Projetado e construído pelo laboratório de propulsão (*Jet Propulsion Laboratory* - NASA-JPL) em 1980, a primeira versão do sistema AIRSAR foi acoplada à aeronave CV-990. Entretanto, a CV-990 foi envolvida em um acidente em julho de 1985, em que esta versão foi completamente destruída. Após o incidente, uma nova versão foi construída pelo JPL, a qual ficou conhecida por AIRSAR, cuja imagem é apresentada na Figura A.2. Esta versão operou em modo polarimétrico, considerando as bandas P-(0.45



**Figura A.2** Imagem do sistema AIRSAR.

GHz), L-(1.26 GHz) e C-(5.31 GHz) simultaneamente. O AIRSAR fez o seu primeiro voo em 1987 e, desde então, realiza, no mínimo, uma atividade por ano.

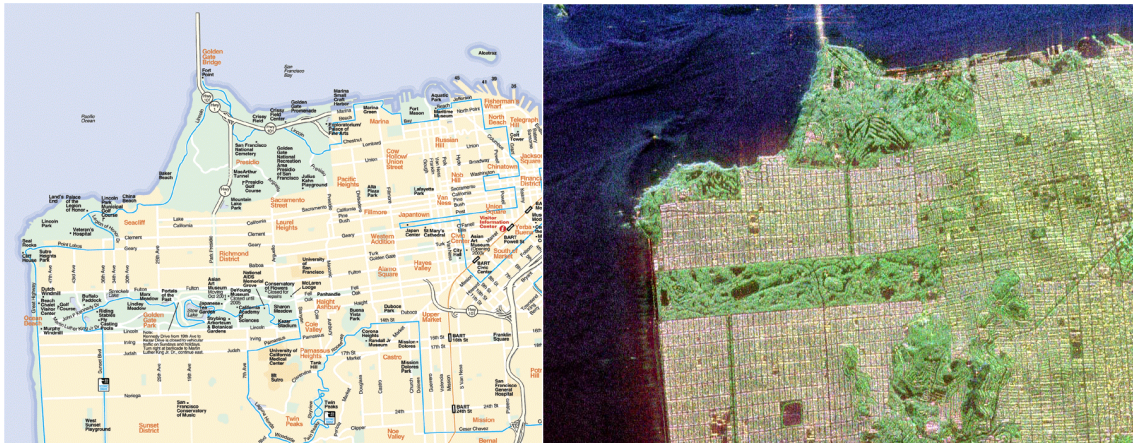
Esta tese utilizou duas imagens obtidas pelo sensor AIRSAR. Uma breve descrição delas é considerada na seguinte discussão.

A Figura A.3 (Fonte: [http://earth.eo.esa.int/polsarpro/Manuals/021\\_polarimetric\\_airborne\\_sensors.pdf](http://earth.eo.esa.int/polsarpro/Manuals/021_polarimetric_airborne_sensors.pdf)) mostra uma imagem de São Francisco com coloração segundo a codificação de Pauli <sup>1</sup> capturada adotando um número de looks quatro. Esta imagem PolSAR é bastante utilizada em vários artigos pois apresenta três regiões bastante definidas: oceano, floresta e cidades. Estas regiões têm graus de texturas bem definidos possibilitando que um método utilizado seja submetido a várias situações. Adicionalmente, um mapa especificando as regiões é apresentado na Figura A.3.

A Figura A.4 mostra uma imagem de Flevoland estudada por Pottier, Lee & Ferro-Famil (2006) bem como um mapa das áreas de cultivo presentes. Esta imagem foi

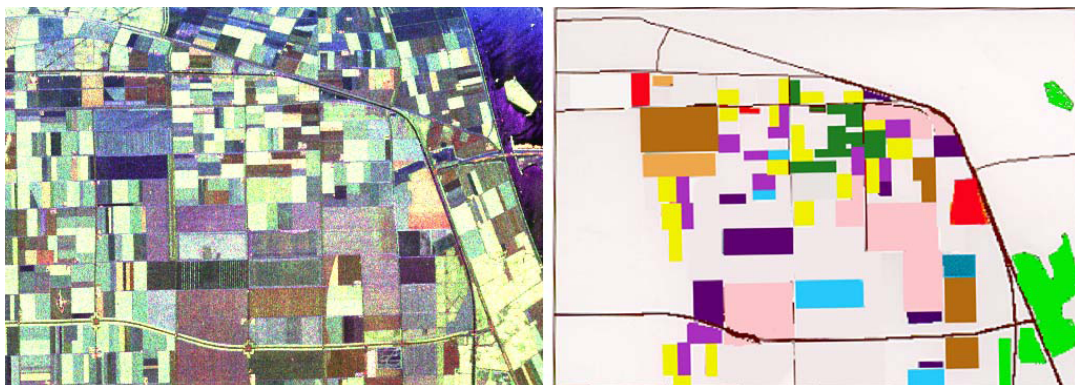
---

<sup>1</sup> É possível acrescentar coloração a uma imagem PolSAR permitindo que algumas de suas características importantes sejam realçadas. Dentre elas a *representação de Pauli* tem recebido considerável atenção. Neste caso, ao invés de construir o esboço da imagem sobre os canais HH, HV e VV, faz-se isto sobre as combinações HH+VV, HH-VV, HV associando às cores azul, vermelho e verde.



**Figura A.3** Mapa de São Francisco (USA) e sua imagem AIRSAR codificada pela coloração Pauli.

obtida utilizando um número de looks oito e representa regiões de pastos, florestas e oceano.



**Figura A.4** Imagem AIRSAR codificada pela coloração Pauli de Flevoland, Holanda, (esquerda) e mapa de áreas de cultivo [feijão (vermelho), floresta (verde claro), batata (amarelo), alfafa (azul claro), trigo (rosa), apenas solo (marrom claro), beterraba (roxo claro), colza (marrom escuro), ervilhas (roxo escuro), grama (verde escuro), água (azul escuro)] (direita).

#### **A.1.2 Sensor E-SAR (*Experimental Airborne SAR Sensor - E-SAR*) e dados utilizados**

De propriedade do centro aeroespacial alemão (*German Aerospace Center - DLR*) e projetado pelo Instituto de Radar e Microondas (*Microwaves and Radar Institute - DLR-HR*) em conjunto com o Instituto de Instalações de Pesquisas de Vôos (*Research Flight Facilities - DLR-FB*), o E-SAR é uma sistema polarimétrico de multifrequência. A Figura A.5 exibe uma imagem deste sistema. Este sensor opera em quatro bandas de frequência,

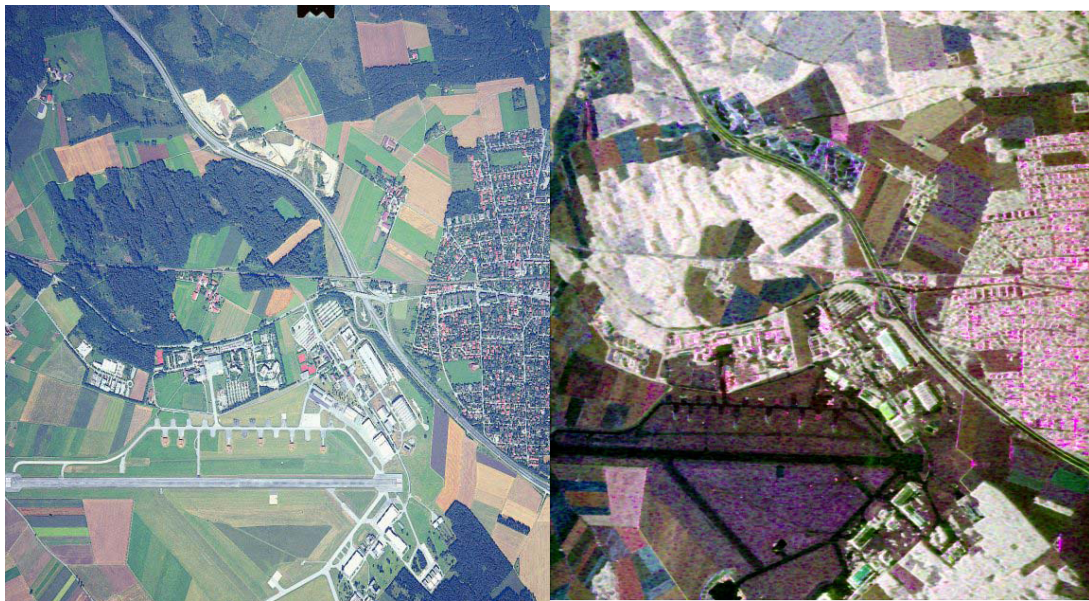




**Figura A.5** Imagem do sistema E-SAR.

X-(9.6 GHz), C-(5.3 GHz), L-(1.3 GHz) e P-(360 MHz), e seu perfil de medição vai de operações com um canal de polarização até modos PolSAR, InSAR e Pol-InSAR.

A imagem Weßling (Munique, Alemanha) mostrada na Figura A.6 foi utilizada em experimentos nesta tese. Esta imagem foi obtida pelo sensor E-SAR considerando um número de looks três e apresenta regiões de pastos, florestas e cidades.



**Figura A.6** Imagem óptica e E-SAR codificada pela coloração Pauli de Munique (Alemanha).

### **A.1.3 Sensor EMISAR (*Electromagnetics Institute SAR* - EMISAR) e dados utilizados**

Desenvolvido pelo Instituto de Eletromagnética da Universidade Técnica da Dinamarca (Lyngby), o EMISAR foi inicialmente um sistema aerotransportado SAR C-band

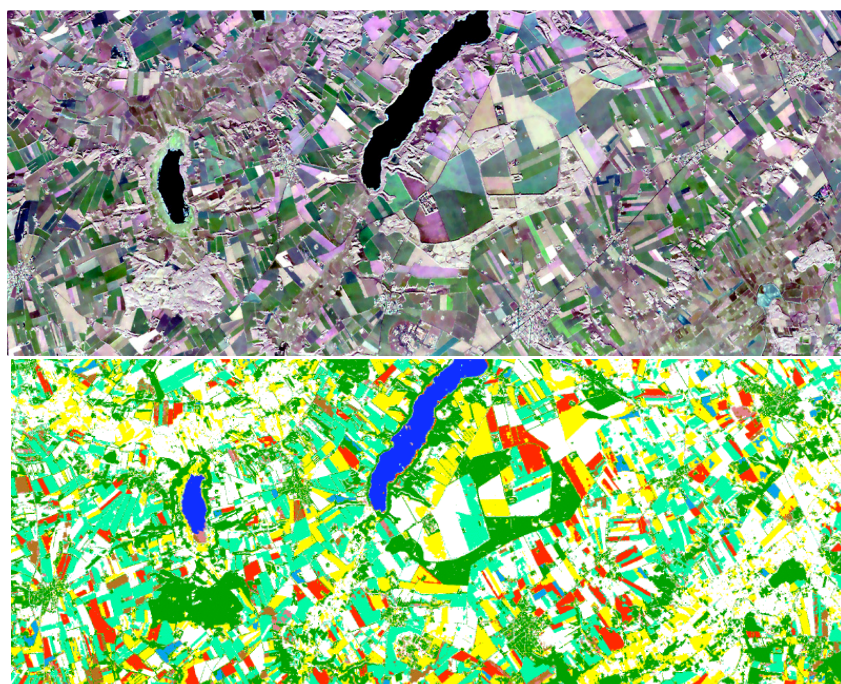
(5.3 GHz). Seu desenvolvimento teve como objetivo (i) a produção de imagens com alta resolução e (ii) garantir flexibilidade significativa na geometria de aquisição tal que pudesse servir como instrumentos de voo para sistemas SAR de satélite da ESA. A Figura A.7 apresenta uma imagem do EMISAR.



**Figura A.7** Imagem do sistema EMISAR.

Posteriormente, o sistema C-band foi estendido numa versão com capacidade polarimétrica: um sistema adicional L-band (1.25 GHz) foi testado em 1995. Atualmente, tal sistema é utilizado para coletar dados SAR de alta qualidade para programas científicos internacionais e para pesquisas em sensoriamento remoto no *Centro de Danish para sensoriamento Remoto*.

Esta tese utiliza recortes da imagem de regiões agrícolas de Folloum, apresentada na Figura A.8 (Fonte: Skriver, Dall, Le Toan, Quegan, Ferro-Famil, Pottier, Lumsdon & Moshammer 2005), obtida pelo sensor EMISAR. Tal imagem é referenciada por *imagem EMISAR de Follum* e foi obtida considerando um número de looks nominal quatro (como descrito em Skriver et al. 2005). A Figura A.8 também apresenta um mapa estimado das regiões agrícolas via método de árvores de decisão hierárquica.



**Figura A.8** Imagem EMISAR codificada pela coloração Pauli de áreas agrícolas de Follum, Dinamarca, (topo) e mapa de áreas cultivadas [ervilhas (vermelho), cevada de primavera (azul claro), beterraba (verde médio), batata (roxo), cevada de inverno (azul médio), colza de inverno (amarelo) e trigo de inverno (branca)] (base).



## Apêndice B

### Expressões para as integrais da Seção 5.1

---


$$\int_0^\infty \frac{x^a}{(b+nx)^c} dx = \frac{b^{-c} \left(\frac{n}{b}\right)^{-1-a} \Gamma(1+a) \Gamma(-1-a+c)}{\Gamma(c)},$$

$$\int_0^\infty \frac{\log(b+nx)x^a}{(b+nx)^c} dx = \frac{1}{\Gamma(c)} \left( b^{-c} \left(\frac{n}{b}\right)^{-1-a} \Gamma(1+a) \Gamma(-1-a+c) (\log(b) + \psi^{(0)}(c) - \psi^{(0)}(-1-a-c)) \right),$$

$$\begin{aligned} \int_0^\infty \frac{\log(b_1+n_1x)x^a}{(b_2+n_2x)^c} dx = & \frac{b_1 b_2^{-c} \pi \csc(a\pi)}{n_1} \left( \frac{n_2^2 (b_1/n_1)^{2-a} c {}_2F_1(2+a, 1+c; 3+a; \frac{b_1 n_2}{b_2 n_1})}{2+3a+a^2} - \right. \\ & \frac{2+a}{2+3a+a^2} \left( \frac{b_2 n_1 - b_1 n_2}{b_2 n_1} \right)^{-c} + \frac{(n_2/b_2)^{-a} \Gamma(-1-a+c) {}_3F_2(1, 1, -a+c; 2, 1-a; \frac{b_1 n_2}{b_2 n_1})}{\Gamma(1-a) \Gamma(c)} + \\ & \left. \frac{n_1 (b_2/n_2)^{1+a}}{b_1} \frac{\Gamma(-1-a+c)}{\Gamma(-a) \Gamma(c)} \left( \log\left(\frac{n_2}{n_1 b_2}\right) + \psi^{(0)}(-1-a+c) - \psi^{(0)}(-a) + \pi \cot(a\pi) \right) \right), \end{aligned}$$

em que  ${}_2F_1(\cdot, \cdot; \cdot)$  e  ${}_3F_2(\cdot, \cdot, \cdot; \cdot, \cdot)$  são funções hipergeométricas,

$$\int_0^\infty \frac{\log(x)x^a}{(b+nx)^c} dx = -\frac{1}{\Gamma(c)} \left[ b^{-c} (n/b)^{-1-a} \Gamma(1+a) \Gamma(-1-a+c) (-H_{[a+1]} + H_{[-1-a+c]} + \log(n/b)) \right],$$

em que  $H_n$  é a  $n$ -ésimo número harmônico, e

$$\begin{aligned} \int_0^\infty x^a (b_1+n_1x)^{c_1} (b_2+n_2x)^{c_2} dx = & b_1^{c_1} b_2^{c_2} \pi \csc((c_2+a)\pi) \\ & \left[ -\frac{\left(\frac{n_1}{b_1}\right)^{-1-c_2-a} \left(\frac{n_2}{b_2}\right)^{c_2} \Gamma(-1-c_1-c_2-a) {}_2F_1(-c_2, -1-c_1-c_2-a; -c_2-a; \frac{b_2 n_1}{b_1 n_2})}{\Gamma(-c_1) \Gamma(-c_2-a)} + \right. \\ & \left. \frac{\left(\frac{n_2}{b_2}\right)^{-1-a} \Gamma(1+a) {}_2F_1(-c_1, 1+a; 2+c_2+a; \frac{b_2 n_1}{b_1 n_2})}{\Gamma(-c_2) \Gamma(2+c_2+a)} \right]. \end{aligned}$$


---

**Figura B.1** Identities de integrais para suporte das distâncias para o modelo  $\mathcal{G}^0$ . A notação paramétrica usada nesta figura independe do restante do texto.

## Apêndice C

# Expressões para as contribuições teóricas da Seção 5.2: Propriedades estatísticas da distribuição Wishart complexa

Este apêndice objetiva apresentar as demonstrações dos principais resultados teóricos desenvolvidos na Seção 5.2. Vamos apresentar inicialmente alguns resultados de análise matriciais.

Considere três matrizes quadradas  $A$ ,  $B$  e  $C$  de ordem  $m$ . De acordo com Brewer (1978), as seguintes propriedades se verificam

$$(P1) \quad [B \otimes A] = U_{m \times m} [A \otimes B] U_{m \times m};$$

$$(P2) \quad [A \otimes B]^{-1} = A^{-1} \otimes B^{-1};$$

$$(P3) \quad [A \otimes B][D \otimes G] = A D \otimes B G;$$

$$(P4) \quad [A \otimes B] \otimes C = A \otimes [B \otimes C];$$

$$(P5) \quad [A + C] \otimes [B + D] = A \otimes B + A \otimes D + C \otimes B + C \otimes D;$$

$$(P6) \quad \partial[A C] / \partial B = [\partial A / \partial B] (I_m \otimes C) + (I_m \otimes A) [\partial C / \partial B];$$

$$(P7) \quad \partial[A \otimes C] / \partial B = [\partial A / \partial B] \otimes C + A \otimes [\partial C / \partial B].$$

em que  $U_{m \times m}$  é uma matriz de permutação.

A seguir, é apresentado um esquete para as derivações da expressão do viés de segunda ordem do estimador MV segundo Cox & Snell (1968). Consideremos primeiro um resumo com as principais definições.

### C.1 Correção de Viés

Seja  $\{Z_1, Z_2, \dots, Z_N\}$  uma amostra iid de uma matriz aleatória populacional  $Z$  com densidade dada por  $f_Z(Z'; \theta)$ , em que  $\dim(\theta) = p$  e  $\dim(\cdot)$  representa a dimensão

de um vetor. Esta seção resume a obtenção de estimadores corrigidos baseados na expressão de viés de segunda ordem proposta por Cox & Snell (1968).

Este método se baseia na derivação dos valores esperados das derivadas de segunda e terceira ordem da função de log-verossimilhança total sob  $N$  observações:

$$\ell(\boldsymbol{\theta}) = \sum_{k=1}^N \log f_{\mathbf{Z}}(\mathbf{Z}_k; \boldsymbol{\theta}).$$

Para este fim nós introduzimos, a notação que representa a notação padrão em pesquisas de Teoria Assintótica (para maiores detalhes, consultar Cordeiro 2011).

Seja  $U_i$  (for  $i = 1, 2, \dots, p$ ) a derivada de  $\ell(\boldsymbol{\theta})$  no  $i$ -ésimo elemento do vetor  $\boldsymbol{\theta}$  (denominado por  $\theta_i$ ). Nós definimos as derivadas de segunda e terceira ordem por

$$U_{ij} = \partial^2 \ell(\boldsymbol{\theta}) / \partial \theta_i \partial \theta_j \text{ e } U_{ijk} = \partial^3 \ell(\boldsymbol{\theta}) / \partial \theta_i \partial \theta_j \partial \theta_k,$$

para  $i, j, k = 1, 2, \dots, p$ .

A partir destes termos, as expressões para os cumulantes de  $\ell(\boldsymbol{\theta})$  são definidas por:  $\kappa_{ij} = E(U_{ij})$ ,  $\kappa_{i,j} = E(U_i U_j)$ ,  $\kappa_{ijk} = E(U_{ijk})$ ,  $\kappa_{i,jk} = E(U_i U_{jk})$ , e assim por diante. Além disso, as derivadas dos cumulantes sobre os parâmetros são representadas por

$$\kappa_{ij}^{(k)} = \partial \kappa_{ij} / \partial \theta_k \quad \text{e} \quad \kappa_{ijk}^{(h)} = \partial \kappa_{ijk} / \partial \theta_h,$$

para  $i, j, k, h = 1, 2, \dots, p$  representando índices no espaço paramétrico. Esta notação para os cumulantes é conhecida como *notação tensorial* e foi introduzida por Lawley (1956), sendo utilizada até os dias de hoje.

Note que o elemento  $\kappa_{i,j} = -\kappa_{ij}$  (tal equação é conhecida como uma das *identidades de Bartlett*, Barndorff-Nielsen & Cox 1989), em que  $\kappa_{ij}$  representa a entrada  $(i, j)$  da Matriz Informação de Fisher esperada ( $\mathcal{K}(\boldsymbol{\theta})$ ). Assim,  $\kappa^{i,j} = -\kappa^{ij}$ , em que  $\kappa^{ij}$  representa a entrada  $(i, j)$  de  $\mathcal{K}(\boldsymbol{\theta})^{-1}$ .

Considerando as definições acima, Cox & Snell (1968) propuseram uma fórmula geral para o viés de segunda ordem do estimador MV:

$$B(\hat{\theta}_i) = \sum_{r,s,t} \kappa^{i,r} \kappa^{s,t} \left( \kappa_{rs,t} + \frac{1}{2} \kappa_{rst} \right) = \sum_{r,s,t} \kappa^{i,r} \kappa^{s,t} \left( \kappa_{rs}^{(t)} - \frac{1}{2} \kappa_{rst} \right), \quad (\text{C.1})$$

em que  $\hat{\theta}_i$  é o estimador relativo ao  $i$ -ésimo termo de  $\boldsymbol{\theta}$ , para  $i = 1, 2, \dots, p$ . A segunda igualdade da expressão (C.1) é justificável, pois o termo  $\kappa_{rs,t} + \frac{1}{2} \kappa_{rst}$  pode ser substituído por  $\kappa_{rs}^{(t)} - \frac{1}{2} \kappa_{rst}$  como consequência das identidades de Bartlett. Ainda com relação aos cumulantes,  $\kappa$ 's, adotaremos o seguinte abuso de notação em termos da notação usual de Lawley, a fim de ser explícito quanto aos parâmetros da distribuição Wishart complexa escalonada:

-  $\kappa_{\theta_i\theta_j} = \kappa_{ij}$ ,  $\kappa_{\theta_i\theta_j} = \kappa_{i,j}$ ,  $\kappa^{\theta_i\theta_j} = \kappa^{i,j}$ ,  $\kappa_{\theta_i\theta_j\theta_k} = \kappa_{ijk}$  e, conseqüentemente, a equação (C.1) reescrita como

$$B(\hat{\theta}_i) = \sum_{r,s,t} \kappa^{\theta_i\theta_r} \kappa^{\theta_s\theta_t} \left( \kappa_{\theta_r\theta_s}^{(\theta_t)} - \frac{1}{2} \kappa_{\theta_r\theta_s\theta_t} \right).$$

Por fim, o estimador corrigido é definido por  $\tilde{\theta} = \hat{\theta} - \hat{B}(\hat{\theta})$ , em que  $\hat{B}(\hat{\theta})$  representa o viés  $B(\hat{\theta})$  avaliado em  $\hat{\theta}$ . É conhecido que o viés de  $\tilde{\theta}$  é de ordem  $O(N^{-2})$ , isto é,  $E(\hat{\theta}) = \theta + O(N^{-2})$  enquanto que  $\hat{\theta}$  é de ordem  $O(N^{-1})$ .

## C.2 Derivadas de primeira e segunda ordem

Seja  $\{Z_1, Z_2, \dots, Z_N\}$  uma amostra iid de  $Z \sim \mathcal{W}(\theta)$  tal que  $\theta = [L, \text{vec}(\Sigma)^t]^t$ , em que

$$\text{vec}(\Sigma) = [\Sigma_{11} \Sigma_{21} \cdots \Sigma_{m1} \cdots \Sigma_{1m} \Sigma_{2m} \cdots \Sigma_{mm}]^t = (\sigma_i)_{i=1,2,\dots,m^2}.$$

Diretamente da densidade da Wishart complexa escalonada, na equação (3.7), decorre-se que

$$\begin{aligned} \ell(\theta) = & mLN \log L + (L - m) \sum_{i=1}^N \log |Z_i| - LN \log |\Sigma| \\ & - N \frac{m(m-1)}{2} \log \pi - N \sum_{k=0}^{m-1} \log \Gamma(L - k) - LN \text{tr}(\Sigma^{-1} \bar{Z}). \end{aligned} \quad (\text{C.2})$$

Além disso, considerando os resultados (a)  $\partial \ell(\theta) / \partial \text{vec}(\Sigma) = \text{vec}(\partial \ell(\theta) / \partial \Sigma)$  (dado em Hjorungnes & Gesbert 2007), (b)  $\partial \text{tr}[\Sigma^{-1} Z] / \partial \Sigma = -\Sigma^{-1} Z \Sigma^{-1}$  e (c)  $\partial \log |\Sigma| / \partial \Sigma = \Sigma^{-1}$ , as funções escores correspondentes são dadas por

$$U_L = \frac{\partial \ell(\theta)}{\partial L} = mN(\log L + 1) + \sum_{i=1}^N \log |Z_i| - N \log |\Sigma| - N \psi_m^{(0)}(L) - N \text{tr}(\Sigma^{-1} \bar{Z})$$

e

$$U_{\text{vec}(\Sigma)} = \frac{\partial \ell(\theta)}{\partial \text{vec}(\Sigma)} = LN \text{vec}(\Sigma^{-1} \bar{Z} \Sigma^{-1} - \Sigma^{-1}).$$

Note que o termo  $U_{\text{vec}(\Sigma)}$  está generalizado com relação a seção anterior. Para estabelecer uma ligação com aquela seção, note que  $U_{\sigma_i}$  é o  $i$ -ésimo elemento do vetor  $m^2$  dimensional  $U_{\text{vec}(\Sigma)}$ .

Baseando-se sobre estes resultados, temos que os elementos da matriz informação de Fisher observada (ou matriz Hessiana) são dados por:

(a)

$$U_{LL} = \frac{\partial[U_L]}{\partial L} = \frac{Nm}{L} - N\psi_m^{(1)}(L), \quad (C.3)$$

em que  $\psi_m^{(0)}(\cdot)$  representa o termo de ordem zero da função poligamma multivariada de  $v$ -ésima ordem dada por  $\psi_m^{(v)}(L) = \sum_{i=0}^{m-1} \psi^{(v)}(L-i)$ ,  $\psi^{(v)}(\cdot)$  é a função poligamma ordinária dada por  $\psi^{(v)}(L) = \partial^{v+1} \log \Gamma(L) / \partial L^{v+1}$ , para  $v \geq 0$  (neste caso,  $\psi^{(0)}(\cdot)$  é conhecida como a função digamma).

(b)

$$U_{L \text{vec}(\Sigma)} = \frac{\partial[U_L]}{\partial \text{vec}(\Sigma)} = N \text{vec}(\Sigma^{-1} \bar{Z} \Sigma^{-1} - \Sigma^{-1})^t \quad (C.4)$$

e

$$U_{\text{vec}(\Sigma)^* L} = \frac{\partial[U_{\text{vec}(\Sigma)^*}]}{\partial L} = N \text{vec}(\Sigma^{-1} \bar{Z} \Sigma^{-1} - \Sigma^{-1})^*. \quad (C.5)$$

Similarmente à proposta de Anfinson, Doulgeris & Eltoft (2009), considere a versão complexa e restrita à matrizes quadradas da definição da página 127 de Kollo & Rosen (2005).

**Definição 6.** Dadas duas matrizes positivas definidas Hermitianas  $X$  e  $Y$ , tem-se

$$\frac{\partial Y}{\partial X} = \frac{\partial}{\partial \text{vec}(X)^*} \text{vec}(Y)^t = \text{vec}(Y)^t \otimes \frac{\partial}{\partial \text{vec}(X)^*},$$

em que

$$\frac{\partial}{\partial \text{vec}(X)^*} = \left[ \frac{\partial}{\partial x_{11}^*}, \dots, \frac{\partial}{\partial x_{m1}^*}, \frac{\partial}{\partial x_{12}^*}, \dots, \frac{\partial}{\partial x_{m2}^*}, \dots, \frac{\partial}{\partial x_{1m}^*}, \dots, \frac{\partial}{\partial x_{mm}^*} \right]^t.$$

Utilizando um argumento de indução similar ao de Kollo & Rosen (2005), uma versão complexa do teorema 1.4.4 página 138 de Kollo & Rosen é dada por

**Teorema 1.** A  $k$ -ésima derivada  $\partial^k Y / \partial X^k$  pode ser reescrita como

$$\frac{\partial^k Y}{\partial X^k} = \frac{\partial}{\partial \text{vec}(X)^*} \text{vec}(Y)^t \otimes \underbrace{\frac{\partial}{\partial \text{vec}(X)^H} \cdots \frac{\partial}{\partial \text{vec}(X)^H}}_{k-1 \text{ termos}}.$$

Este resultado estende a definição anterior e nos ajuda a entender que as derivações de altas ordem em termos do operador  $\text{vec}(\cdot)$  podem ser escritas em termos de derivadas matriciais.

A seguinte discussão considera derivações de ordem 2 para log-verossimilhança da distribuição Wishart complexa.

(c) Da definição 6, tem-se

$$U_{\text{vec}(\Sigma) \text{vec}(\Sigma)^*} = \frac{\partial}{\partial \text{vec}(\Sigma)^*} \text{vec} \left( \frac{\partial \ell(\theta)}{\partial \Sigma} \right)^t = \frac{\partial^2 \ell(\theta)}{\partial \Sigma^2} = NL \left\{ \underbrace{\frac{\partial(\Sigma^{-1} \bar{Z} \Sigma^{-1})}{\partial \Sigma}}_{T_1} - \underbrace{\frac{\partial \Sigma^{-1}}{\partial \Sigma}}_{T_2} \right\}.$$

Anfinsen, Doulgeris & Eltoft (2009) mostraram que

(P8)

$$T_1 = -\Sigma^{-1} \otimes \Sigma^{-1} \bar{Z} \Sigma^{-1} - \Sigma^{-1} \otimes \Sigma^{-1} \Sigma^{-1} \bar{Z}.$$

Além disso, é conhecido de Hjorungnes & Gesbert (2007) o seguinte resultado:

(P9)

$$T_2 = -\Sigma^{-1} \otimes \Sigma^{-1}.$$

Assim, nós temos

$$\begin{aligned} U_{\text{vec}(\Sigma) \text{vec}(\Sigma)^*} &= \frac{\partial [U_{\text{vec}(\Sigma)}^t]}{\partial \text{vec}(\Sigma)^*} = L(T_1 - T_2) = \\ &= NL(\Sigma^{-1} \otimes \Sigma^{-1} - \Sigma^{-1} \otimes \Sigma^{-1} \bar{Z} \Sigma^{-1} - \Sigma^{-1} \otimes \Sigma^{-1} \Sigma^{-1} \bar{Z}). \end{aligned} \quad (\text{C.6})$$

O resumo dos termos derivados até então segue nas identidades matriciais abaixo.

$$\begin{aligned} \begin{bmatrix} U_{LL} & U_{L \text{vec}(\Sigma)}^t \\ U_{\text{vec}(\Sigma)^* L} & U_{\text{vec}(\Sigma) \text{vec}(\Sigma)^*} \end{bmatrix} &= \begin{bmatrix} \frac{mN}{L} - N\psi_m^{(1)}(L) & N \text{vec}(\Sigma^{-1} \bar{Z} \Sigma^{-1} - \Sigma^{-1})^t \\ N \text{vec}(\Sigma^{-1} \bar{Z} \Sigma^{-1} - \Sigma^{-1})^* & U_{\text{vec}(\Sigma) \text{vec}(\Sigma)^*} \end{bmatrix}, \\ \begin{bmatrix} \kappa_{L,L} & \kappa_{L, \text{vec}(\Sigma)}^t \\ \kappa_{\text{vec}(\Sigma)^*, L} & \kappa_{\text{vec}(\Sigma), \text{vec}(\Sigma)^*} \end{bmatrix} &= \begin{bmatrix} N\psi_m^{(1)}(L) - \frac{Nm}{L} & \text{vec}(\mathbf{0}_m)^t \\ \text{vec}(\mathbf{0}_m) & LN \Sigma^{-1} \otimes \Sigma^{-1} \end{bmatrix} \\ \text{e} \\ \begin{bmatrix} \kappa_{L,L}^L & \kappa_{L, \text{vec}(\Sigma)}^{L,t} \\ \kappa_{\text{vec}(\Sigma)^*, L}^L & \kappa_{\text{vec}(\Sigma), \text{vec}(\Sigma)^*}^L \end{bmatrix} &= \begin{bmatrix} [N\psi_m^{(1)}(L) - \frac{mN}{L}]^{-1} & \text{vec}(\mathbf{0}_m)^t \\ \text{vec}(\mathbf{0}_m) & (NL)^{-1} \Sigma \otimes \Sigma \end{bmatrix}, \end{aligned}$$

em que  $\mathbf{0}_m$  é a matriz nula de ordem  $m$ .

Fazendo uma ligação à seção anterior, note que:

- O termos  $U_{L\sigma_i}$ ,  $\kappa_{L,\sigma_i}$  e  $\kappa_{L,\sigma_i}^{L,\sigma_i}$  são dados pelos  $i$ -ésimos elementos dos  $m^2$  vetores  $U_{L \text{vec}(\Sigma)}$ ,  $\kappa_{L, \text{vec}(\Sigma)}$  e  $\kappa_{L, \text{vec}(\Sigma)}^{L, \text{vec}(\Sigma)}$ , respectivamente.
- os elementos  $U_{\sigma_i \sigma_j}$ ,  $\kappa_{\sigma_i, \sigma_j}$  e  $\kappa_{\sigma_i, \sigma_j}^{\sigma_i, \sigma_j}$  representam as entradas  $(i, j)$  das matrizes  $U_{\text{vec}(\Sigma) \text{vec}(\Sigma)^*}$ ,  $\kappa_{\text{vec}(\Sigma), \text{vec}(\Sigma)^*}$  e  $\kappa^{\text{vec}(\Sigma), \text{vec}(\Sigma)^*}$ , respectivamente.

Baseando-se sobre esses elementos, tem-se os seguintes resultados para as derivadas dos cumulantes:

- Como  $\kappa_{L\sigma_i} = \kappa_{\sigma_i L} = 0$ , então  $\kappa_{L\sigma_i}^{(\sigma_j)} = \kappa_{\sigma_i L}^{(\sigma_j)} = 0$ ;
- Como  $\kappa_{LL}$  independe de  $\sigma_i$ ,  $\kappa_{LL}^{(\sigma_j)} = 0$ .
- Note ainda que o resultado

$$\kappa_{\text{vec}(\Sigma) \text{vec}(\Sigma)^*}^{(L)} = \frac{\partial}{\partial L} [\kappa_{\text{vec}(\Sigma) \text{vec}(\Sigma)^*}] = -N\Sigma^{-1} \otimes \Sigma^{-1}$$

é facilmente obtido de  $\kappa_{\text{vec}(\Sigma) \text{vec}(\Sigma)^*}$ . Neste caso, o elemento  $\kappa_{\sigma_i \sigma_j}^{(L)}$  representa a entrada  $(i, j)$  da matriz  $\kappa_{\text{vec}(\Sigma) \text{vec}(\Sigma)^*}^{(L)}$ .

- Considere a seguinte demonstração para o termo

$$\kappa_{\text{vec}(\Sigma) \text{vec}(\Sigma)^*}^{(\text{vec}(\Sigma)^*)} = \frac{\partial}{\partial \text{vec}(\Sigma)^*} [\kappa_{\text{vec}(\Sigma) \text{vec}(\Sigma)^*}] = \frac{\partial}{\partial \Sigma} [-NL \underbrace{\Sigma^{-1} \otimes \Sigma^{-1}}_{A_1}] :$$

$$\begin{aligned} \frac{\partial}{\partial \Sigma} A_1 &\stackrel{(p7)}{=} \frac{\partial \Sigma^{-1}}{\partial \Sigma} \otimes \Sigma^{-1} + \Sigma^{-1} \otimes \frac{\partial \Sigma^{-1}}{\partial \Sigma} \stackrel{(p9)}{=} -[\Sigma^{-1} \otimes \Sigma^{-1}] \otimes \Sigma^{-1} - \Sigma^{-1} \otimes [\Sigma^{-1} \otimes \Sigma^{-1}] \\ &\stackrel{(p6)}{=} -2[\Sigma^{-1} \otimes \Sigma^{-1}] \otimes \Sigma^{-1}. \end{aligned}$$

Para este caso, o elemento  $\kappa_{\sigma_i \sigma_j}^{(\sigma_k)}$  é representado pela entrada  $(k + (i - 1)m^2, k + (j - 1)m^2)$  da matriz  $\kappa_{\text{vec}(\Sigma) \text{vec}(\Sigma)^*}^{(\text{vec}(\Sigma)^*)} = 2LN[\Sigma^{-1} \otimes \Sigma^{-1}] \otimes \Sigma^{-1}$ .

### C.3 Derivadas de terceira ordem

(a) Cálculo para  $U_{\text{vec}(\Sigma) \text{vec}(\Sigma)^* \text{vec}(\Sigma)^*} = \partial^3 \ell(\theta) / \partial \Sigma^3$ :

Da equação (C.6), tem-se

$$\frac{\partial^3 \ell(\theta)}{\partial \Sigma^3} = LN \frac{\partial}{\partial \Sigma} \left\{ \underbrace{\Sigma^{-1} \otimes \Sigma^{-1}}_{A_1} - \underbrace{\Sigma^{-1} \otimes \Sigma^{-1} \bar{Z} \Sigma^{-1}}_{A_2} - \underbrace{\Sigma^{-1} \otimes \Sigma^{-1} \Sigma^{-1} \bar{Z}}_{A_3} \right\}.$$

O elemento  $A_1 = -2[\Sigma^{-1} \otimes \Sigma^{-1}] \otimes \Sigma^{-1}$  foi obtido anteriormente.

Para  $A_2$ ,

$$\begin{aligned}\frac{\partial}{\partial \Sigma} A_2 &\stackrel{(p7)}{=} \frac{\partial \Sigma^{-1}}{\partial \Sigma} \otimes [\Sigma^{-1} \bar{Z} \Sigma^{-1}] + \Sigma^{-1} \otimes T_1 \stackrel{(p9)}{=} -[\Sigma^{-1} \otimes \Sigma^{-1}] \otimes [\Sigma^{-1} \bar{Z} \Sigma^{-1}] + \Sigma^{-1} \otimes T_1 \\ &= -[\Sigma^{-1} \otimes \Sigma^{-1}] \otimes [\Sigma^{-1} \bar{Z} \Sigma^{-1}] - \Sigma^{-1} \otimes [\Sigma^{-1} \otimes \Sigma^{-1} \bar{Z} \Sigma^{-1}] \\ &\quad - \Sigma^{-1} \otimes [\Sigma^{-1} \otimes \Sigma^{-1} \Sigma^{-1} \bar{Z}] \\ &= -2[\Sigma^{-1} \otimes \Sigma^{-1}] \otimes [\Sigma^{-1} \bar{Z} \Sigma^{-1}] - [\Sigma^{-1} \otimes \Sigma^{-1}] \otimes [\Sigma^{-1} \Sigma^{-1} \bar{Z}].\end{aligned}$$

Para  $A_3$ , seguindo (P8)

$$\frac{\partial}{\partial \Sigma} A_3 \stackrel{(p7)}{=} \frac{\partial \Sigma^{-1}}{\partial \Sigma} \otimes [\Sigma^{-1} \Sigma^{-1} \bar{Z}] + \Sigma^{-1} \otimes \underbrace{\frac{\partial \Sigma^{-1} \Sigma^{-1} \bar{Z}}{\partial \Sigma}}_{A_{31}}.$$

Neste caso, o termo  $A_{31}$  é derivado por

$$\begin{aligned}A_{31} &\stackrel{(p7)}{=} \frac{\partial \Sigma^{-1}}{\partial \Sigma} [I_m \otimes \Sigma^{-1} \bar{Z}] + [I_m \otimes \Sigma^{-1}] \frac{\partial [\Sigma^{-1} \bar{Z}]}{\partial \Sigma} \\ &\stackrel{(P6) \text{ e } (P9)}{=} [-\Sigma^{-1} \otimes \Sigma^{-1}] [I_m \otimes \Sigma^{-1} \bar{Z}] + [I_m \otimes \Sigma^{-1}] [-\Sigma^{-1} \otimes \Sigma^{-1}] [I_m \otimes \bar{Z}] \\ &= -2[\Sigma^{-1} \otimes \Sigma^{-1} \Sigma^{-1} \bar{Z}]\end{aligned}$$

Assim, temos que

$$\begin{aligned}\frac{\partial}{\partial \Sigma} A_3 &= -[\Sigma^{-1} \otimes \Sigma^{-1}] \otimes [\Sigma^{-1} \Sigma^{-1} \bar{Z}] - 2\Sigma^{-1} \otimes [\Sigma^{-1} \otimes \Sigma^{-1} \Sigma^{-1} \bar{Z}] \\ &= -3[\Sigma^{-1} \otimes \Sigma^{-1}] \otimes [\Sigma^{-1} \Sigma^{-1} \bar{Z}]\end{aligned}$$

Portanto, temos que

$$\begin{aligned}\frac{\partial^3 \ell(\theta)}{\partial \Sigma^3} &= LN \left\{ 4[\Sigma^{-1} \otimes \Sigma^{-1}] \otimes [\Sigma^{-1} \Sigma^{-1} \bar{Z}] + 2[\Sigma^{-1} \otimes \Sigma^{-1}] \otimes [\Sigma^{-1} \bar{Z} \Sigma^{-1}] \right. \\ &\quad \left. - 2[\Sigma^{-1} \otimes \Sigma^{-1}] \otimes \Sigma^{-1} \right\}\end{aligned}$$

E o seu valor esperado é dado por

$$\kappa_{\text{vec}(\Sigma) \text{ vec}(\Sigma)^* \text{ vec}(\Sigma)^*} = E \left[ \frac{\partial^3 \ell(\theta)}{\partial \Sigma^3} \right] = 4NL[\Sigma^{-1} \otimes \Sigma^{-1}] \otimes \Sigma^{-1}.$$

(b) Demais derivadas de terceira ordem:

$$\begin{aligned}\bullet \quad \kappa_{LLL} &= \kappa_{LL}^{(L)} = -N \left[ \frac{m}{L^2} + \psi_m^{(2)}(L) \right]; \quad \kappa_{LL \text{ vec}(\Sigma)} = \kappa_{L \text{ vec}(\Sigma) L} = \kappa_{\text{vec}(\Sigma) LL} = 0; \\ \kappa_{\text{vec}(\Sigma) \text{ vec}(\Sigma)^* L} &= \kappa_{\text{vec}(\Sigma) L \text{ vec}(\Sigma)^*} = \kappa_{L \text{ vec}(\Sigma) \text{ vec}(\Sigma)^*} = -N \Sigma^{-1} \otimes \Sigma^{-1}\end{aligned}$$



### C.4 Expressão para o viés de segunda ordem

Devido a ortogonalidade entre os parâmetros  $L$  e  $\sigma_i$  para  $i = 1, 2, \dots, m^2$  (justificada pelo resultado  $\kappa^{L\sigma_i} = 0$ ), a equação (C.1) é degenerada na seguinte fórmula para este caso:

$$B(\hat{L}) = \kappa^{L,L} \kappa^{L,L} \left( \kappa_{LL}^{(L)} - \frac{1}{2} \kappa_{LLL} \right) + \kappa^{L,L} \underbrace{\sum_{i,j=1}^{m^2} \kappa^{\sigma_i, \sigma_j} \left( \kappa_{L\sigma_i}^{(\sigma_i)} - \frac{1}{2} \kappa_{L\sigma_i \sigma_j} \right)}_{Q_1}.$$

Note que a derivação da fórmula fechada para  $B(\hat{L})$  depende apenas da expressão para  $Q_1$ . Como  $\kappa_{L\sigma_i}^{(\sigma_i)} = 0$ , o termo  $Q_1$  é simplificado em

$$Q_1 = -\frac{1}{2} \sum_{i,j=1}^{m^2} \kappa^{\sigma_i, \sigma_j} \kappa_{L\sigma_i \sigma_j}.$$

Para a derivação deste termo, considere a proposta do seguinte teorema.

**Teorema 2.** Se  $\kappa^{\sigma_i, \sigma_j}$  e  $\kappa_{L\sigma_i \sigma_j}$  são as entradas  $(i, j)$  das matrizes  $\kappa^{\text{vec}(\Sigma), \text{vec}(\Sigma)^*} = (NL)^{-1} \Sigma \otimes \Sigma$  e  $\kappa_{L \text{vec}(\Sigma) \text{vec}(\Sigma)^*} = -N(\Sigma \otimes \Sigma)^{-1}$ , respectivamente, então o seguinte somatório se verifica:

$$-\frac{1}{2} \sum_{i,j=1}^{m^2} \kappa^{\sigma_i, \sigma_j} \kappa_{L\sigma_i \sigma_j} = (2L)^{-1} m^2.$$

---

**Prova:** Sejam  $\kappa^{\sigma_i, \sigma_j}$  e  $\kappa_{L\sigma_i \sigma_j}$  as entradas  $(i, j)$  das matrizes  $\kappa^{\text{vec}(\Sigma), \text{vec}(\Sigma)^*} = (NL)^{-1} \Sigma \otimes \Sigma$  e  $\kappa_{L \text{vec}(\Sigma) \text{vec}(\Sigma)^*} = -N(\Sigma \otimes \Sigma)^{-1}$ , respectivamente. A seguinte argumentação considera o cálculo do termo

$$H_1 = -\frac{1}{2} \sum_{i,j=1}^{m^2} \kappa^{\sigma_i, \sigma_j} \kappa_{L\sigma_i \sigma_j}.$$

Definamos ainda as matrizes  $B = \Sigma \otimes \Sigma$  e  $C = [\Sigma \otimes \Sigma]^{-1}$ . Neste caso, os termos  $\kappa^{\sigma_i, \sigma_j}$  e  $\kappa_{L\sigma_i \sigma_j}$  podem ser reescrito como  $\kappa^{\sigma_i, \sigma_j} = (NL)^{-1} b_{ij}$  e  $\kappa_{L\sigma_i \sigma_j} = -N c_{ij}$ , em que  $b_{ij}$  e  $c_{ij}$  são as entradas  $(i, j)$  das matrizes  $B$  e  $C$ , respectivamente. Assim o termo  $H_1$  pode ser reescrito como

$$H_1 = \frac{1}{2L} \sum_{i,j=1}^{m^2} b_{ij} c_{ij}.$$

De Johnson & Wichern (1988, p. 82), a seguinte relação entre  $b_{ij}$  e  $c_{ij}$  se verifica:  $c_{ij} = |B_{j,i}| |B|^{-1} (-1)^{i+j}$ , em que  $B_{j,i}$  representa a matriz de cofatores obtida desconsiderando a  $j$ -ésima linha e a  $i$ -ésima coluna da matriz  $B$ . Note ainda que, como  $B$  é

uma matriz Hermitiana,  $c_{ij} = |B_{i,j}||\mathbf{B}|^{-1}(-1)^{i+j}$ . Combinando este resultado ao fato que  $|\mathbf{B}| = \sum_{j=1}^{m^2} b_{ij} |B_{i,j}|(-1)^{i+j}$  para qualquer  $i = 1, 2, \dots, m^2$  fixado (como proposto em Johnson & Wichern 1988, p. 79), tem-se que

$$\sum_{j=1}^{m^2} c_{ij} b_{ij} = 1,$$

para qualquer  $i$ .

Portanto, temos que

$$H_1 = \frac{1}{2L} \sum_{i=1}^{m^2} \left[ \sum_{j=1}^{m^2} b_{ij} c_{ij} \right] = (2L)^{-1} m^2.$$

A expressão para o viés de segunda ordem do estimador MV é, então, dada por

$$B(\hat{L}) = -\frac{1}{2N} \frac{\frac{m}{2L} + \sum_{i=0}^{m-1} \psi^{(2)}(L-i)}{[\psi_m^{(1)}(L) - \frac{m}{L}]^2} + \frac{m^2}{2NL[\psi_m^{(1)}(L) - \frac{m}{L}]}. \quad (\text{C.7})$$

## Apêndice D

# Resultados para distribuição $\mathcal{G}_m^{0(L)}$

**Teorema 3.** Se  $\mathbf{Z} \sim \mathcal{G}_m^{0(L)}(\boldsymbol{\Sigma}, \alpha, L)$  com função densidade de probabilidade dada na Tabela 3.5, então os seguintes momentos são válidos:

$$\mathbb{E} \left[ \frac{\text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z})}{L \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (\alpha - 1)} \right] = \frac{mL}{mL - \alpha}, \quad (\text{D.1})$$

$$\mathbb{E}[\log(L \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (-\alpha - 1))] = \psi^{(0)}(mL - \alpha) + \log(-\alpha - 1) - \psi^{(0)}(-\alpha), \quad (\text{D.2})$$

$$\mathbb{E} \left[ \frac{1}{L \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (\alpha - 1)} \right] = \frac{\alpha}{(\alpha - mL)(-\alpha - 1)}. \quad (\text{D.3})$$

**Prova:** Seja  $\mathbf{Z} \sim \mathcal{G}_p^0(\boldsymbol{\theta})$  tal que  $\boldsymbol{\theta} = [\alpha, L, \text{vec}(\boldsymbol{\Sigma})^t]^t$ , em que

$$\text{vec}(\boldsymbol{\Sigma}) = [\Sigma_{11} \ \Sigma_{21} \ \cdots \ \Sigma_{m1} \ \cdots \ \Sigma_{1m} \ \Sigma_{2m} \ \cdots \ \Sigma_{mm}]^t = (\sigma)_{i=1,2,\dots,m^2}.$$

Diretamente da densidade da  $\mathcal{G}_m^{0(L)}$ , na Tabela 3.5, sua log-verossimilhança é dada por

$$\begin{aligned} \ell(\boldsymbol{\theta}) = \log f_{\mathcal{G}^p}(\mathbf{Z}) &= mL \log L + \log \Gamma(mL - \alpha) - L \log |\boldsymbol{\Sigma}| - \log \Gamma(-\alpha) \\ &\quad - \alpha \log(-\alpha - 1) - \frac{m(m-1)}{2} \log \pi - \sum_{i=0}^{m-1} \log \Gamma(L - i) + (L - m) \log |\mathbf{Z}| \\ &\quad + (\alpha - mL) \log(L \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (-\alpha - 1)). \end{aligned}$$

Considerando a mesma notação utilizada no Apêndice C, suas funções escore rela-

tivas são dadas por

$$\begin{aligned}
 U_\alpha &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} = -\psi^{(0)}(mL - \alpha) + \psi^{(0)}(-\alpha) - \log(-\alpha - 1) + \frac{\alpha}{-\alpha - 1} \\
 &\quad + \log(L \operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (-\alpha - 1)) - \frac{\alpha - mL}{L \operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (-\alpha - 1)}, \\
 U_L &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial L} = m(\log L + 1) + m\psi^{(0)}(mL - \alpha) - \log |\boldsymbol{\Sigma}| + \log |\mathbf{Z}| - \psi_m^{(0)}(L) \\
 &\quad - m \log(L \operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (-\alpha - 1)) + (\alpha - mL) \frac{\operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z})}{L \operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (-\alpha - 1)}, \\
 U_{\operatorname{vec}(\boldsymbol{\Sigma})} &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \operatorname{vec}(\boldsymbol{\Sigma})} = -L \operatorname{vec} \left( \boldsymbol{\Sigma}^{-1} + L(\alpha - mL) \frac{\boldsymbol{\Sigma}^{-1} \mathbf{Z} \boldsymbol{\Sigma}^{-1}}{L \operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (-\alpha - 1)} \right).
 \end{aligned}$$

Cosideremos, inicialmente, a derivação do resultado (D.1). Da equivalência  $E(U_{\operatorname{vec}(\boldsymbol{\Sigma})}) = 0$ , tem-se que

$$E \left[ \frac{\boldsymbol{\Sigma}^{-1} \mathbf{Z} \boldsymbol{\Sigma}^{-1}}{L \operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (-\alpha - 1)} \right] = \frac{\boldsymbol{\Sigma}^{-1}}{mL - \alpha}.$$

Aplicando a operação  $\boldsymbol{\Sigma}[\mathbf{R}]\boldsymbol{\Sigma}$ , para  $\mathbf{R}$  como um argumento matricial compatível, em ambos os lados da equivalência acima, temos que

$$E \left[ \frac{\mathbf{Z}}{L \operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (-\alpha - 1)} \right] = \frac{\boldsymbol{\Sigma}}{mL - \alpha}.$$

Utilizando a operação  $\boldsymbol{\Sigma}^{-1}[\mathbf{R}]$  em ambos os lados da igualdade acima, então

$$E \left[ \frac{\boldsymbol{\Sigma}^{-1} \mathbf{Z}}{L \operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (-\alpha - 1)} \right] = \frac{I_m}{mL - \alpha},$$

em que  $I_m$  é a matriz de identidade de ordem  $m$ . Por fim, tomando os traços de ambos os lados, decorre-se que

$$E \left[ \frac{\operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z})}{L \operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (-\alpha - 1)} \right] = \frac{m}{mL - \alpha}.$$

A seguinte discussão se destina a provar equação (D.2). É decorrente de  $E(U_L) = 0$ ,

$$\begin{aligned}
 -m E[\log(L \operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (-\alpha - 1))] &= -(\alpha - mL) E \left[ \frac{\operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z})}{L \operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (-\alpha - 1)} \right] \\
 + \log |\boldsymbol{\Sigma}| - m(\log L + 1) - m\psi^{(0)}(mL - \alpha) &- E(\log |\mathbf{Z}|) + \psi_m^{(0)}(L).
 \end{aligned}$$

Neste caso, tem-se que  $\mathbf{Z} = \mathbf{Y} \cdot X$  em que  $X \sim \Gamma^{-1}(-\alpha, -\alpha - 1)$  e  $\mathbf{Y} \sim \mathcal{W}(\boldsymbol{\Sigma}, L)$ . Assim, de (i)  $E(\log X) = \log(-\alpha - 1) - \psi^{(0)}(-\alpha)$  e (ii)  $E(\log |\mathbf{Y}|) = \log |\boldsymbol{\Sigma}| + \psi_m^{(0)}(L) - m \log L$ , derivado por Anfinsen, Douglgeris & Eltoft (2009), decorre-se que

$$\begin{aligned} E(\log |\mathbf{Z}|) &= E(\log |\mathbf{Y}|) + m E(\log X) \\ &= \log |\boldsymbol{\Sigma}| + \psi_m^{(0)}(L) - m \log L + m[\log(-\alpha - 1) - \psi^{(0)}(-\alpha)]. \end{aligned} \quad (\text{D.4})$$

Finalmente, combinando esta última expressão com o resultado (D.1), tem-se

$$E[\log(L \operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (-\alpha - 1))] = \psi^{(0)}(mL - \alpha) + \log(-\alpha - 1) - \psi^{(0)}(-\alpha).$$

Esta expressão é crucial para a obtenção da entropia de Shannon da distribuição polarimétrica  $\mathcal{G}_m^{0(L)}$  (veja Seção 4.1).

Por fim, consideremos a derivação do resultado (D.3). De  $E(U_\alpha) = 0$  combinado ao resultado equação (D.2), menores manipulações resultam em

$$E \left[ \frac{1}{L \operatorname{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{Z}) + (\alpha - 1)} \right] = \frac{\alpha}{(\alpha - mL)(-\alpha - 1)}.$$

## Apêndice E

# Expressões para as contribuições teóricas da Seção 5.3

Nesta seção, nós consideramos a derivação das distâncias estocásticas para as distribuições polarimétricas e depois particularizamos para o contexto da distribuição Wishart complexa escalonada.

Baseando-se na equação (3.29), seja duas matrizes aleatórias  $\mathbf{Z}_1$  e  $\mathbf{Z}_2$  com densidades

$$f_{\mathbf{Z}_i}(\mathbf{Z}'; \boldsymbol{\theta}_i, L_i) = \underbrace{\frac{h_{\mathcal{M}}(\boldsymbol{\theta}_i, L_i)}{\Gamma_m(L_i)}}_{T_i} |\mathbf{Z}'|^{L_i-m} g(\text{tr}(\boldsymbol{\Sigma}_i^{-1} \mathbf{Z}'); \boldsymbol{\theta}_i, L_i), \quad (\text{E.1})$$

em que as funções  $g(\text{tr}(\boldsymbol{\Sigma}_i^{-1} \mathbf{Z}'); \boldsymbol{\theta}_i, L_i)$  e  $h_{\mathcal{M}}(\boldsymbol{\theta}_i, L_i)$  são dadas na Tabela 3.5. As derivações subsequentes estarão baseadas neste contexto.

**Para a distância de Kullback-Leibler:** Para consideração geral desta medida de distância nas distribuições polarimétricas, o seguinte resultado decorre da densidade (E.1):

$$\log \frac{f_{\mathbf{Z}_1}(\mathbf{Z}'; \cdot)}{f_{\mathbf{Z}_2}(\mathbf{Z}'; \cdot)} = \log \frac{T_1}{T_2} + (L_1 - L_2) \log |\mathbf{Z}| + \log \frac{[g(\text{tr}(\boldsymbol{\Sigma}_1^{-1} \mathbf{Z}'); \boldsymbol{\theta}_1, L_1)]}{[g(\text{tr}(\boldsymbol{\Sigma}_2^{-1} \mathbf{Z}'); \boldsymbol{\theta}_2, L_2)]}$$

Desta expressão, obtêm-se que

$$\begin{aligned} 2d_{\text{KL}}(\mathbf{Z}_1, \mathbf{Z}_2) &= \int_{\mathcal{A}} [f_{\mathbf{Z}_1}(\mathbf{Z}'; \cdot) - f_{\mathbf{Z}_2}(\mathbf{Z}'; \cdot)] \log \frac{f_{\mathbf{Z}_1}(\mathbf{Z}'; \cdot)}{f_{\mathbf{Z}_2}(\mathbf{Z}'; \cdot)} d\mathbf{Z}' \\ &= (L_1 - L_2) [\mathbb{E}(\log |\mathbf{Z}_1|) - \mathbb{E}(\log |\mathbf{Z}_2|)] \\ &\quad + \mathbb{E}(\log [g(\text{tr}(\boldsymbol{\Sigma}_1^{-1} \mathbf{Z}_1); \boldsymbol{\theta}_1, L_1)]) + \mathbb{E}(\log [g(\text{tr}(\boldsymbol{\Sigma}_2^{-1} \mathbf{Z}_2); \boldsymbol{\theta}_2, L_2)]) \\ &\quad - \mathbb{E}(\log [g(\text{tr}(\boldsymbol{\Sigma}_1^{-1} \mathbf{Z}_2); \boldsymbol{\theta}_1, L_1)]) - \mathbb{E}(\log [g(\text{tr}(\boldsymbol{\Sigma}_2^{-1} \mathbf{Z}_1); \boldsymbol{\theta}_2, L_2)]). \end{aligned} \quad (\text{E.2})$$

Nesta expressão, note que os log momentos para  $\mathbf{Z}_i = \mathbf{X}_i \cdot \mathbf{Y}_i$ ,  $\mathbb{E}(\log |\mathbf{Z}_i|)$  para  $i = 1, 2$ , em que  $\mathbf{Y}_i \sim \mathcal{W}(\boldsymbol{\Sigma}_i, L_i)$  e  $\mathbf{X}_i$  é descrito por uma das distribuições elencadas

na Seção 3.3.2, são definidos pela expressão geral:

$$\begin{aligned} E(\log |\mathbf{Z}_i|) &= E(\log |\mathbf{Y}_i \cdot X_i|) = E(\log |\mathbf{Y}_i|) + m E(\log X_i) \\ &= \log |\boldsymbol{\Sigma}| + \psi_m^{(0)}(L) - m \log L + m E(\log X_i). \end{aligned}$$

em que  $m$  é a ordem da matriz  $\mathbf{Z}_i$ . A seguir consideramos uma discussão para duas distribuições polarimétricas.

- Distribuição Wishart: Para  $\Pr(X = 1) = 1$ , temos que  $\mathbf{Z}_i \sim \mathcal{W}(\boldsymbol{\Sigma}_i, L_i)$ . Para esta distribuição menores manipulações resultam em

$$E(\log[g(\text{tr}(\boldsymbol{\Sigma}_i^{-1} \mathbf{Z}_j); \boldsymbol{\theta}_i, L_i)]) = \begin{cases} -mL_i, & \text{se } i = j \\ -L_i \text{tr}(\boldsymbol{\Sigma}_i^{-1} \boldsymbol{\Sigma}_j), & \text{se } i \neq j. \end{cases}$$

Aplicando os estes resultados na equação (E.2), obtêm-se

$$\begin{aligned} d_{\text{KL}}(\mathbf{Z}_1, \mathbf{Z}_2) &= \frac{L_1 - L_2}{2} \underbrace{\left\{ \log \frac{|\boldsymbol{\Sigma}_1|}{|\boldsymbol{\Sigma}_2|} - m \log \frac{L_1}{L_2} + [\psi_m^{(0)}(L_1) - \psi_m^{(0)}(L_2)] \right\}}_{T^{\mathcal{W}}} \\ &\quad + \frac{\text{tr}(L_2 \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 + L_1 \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_2)}{2} - \frac{m(L_1 + L_2)}{2}. \end{aligned}$$

- Distribuição  $\mathcal{G}^0$ : Para  $X \sim \Gamma^{-1}(-\alpha, (-\alpha - 1))$ , temos que  $\mathbf{Z}_i \sim \mathcal{G}_p^0(\boldsymbol{\Sigma}_i, \alpha_i, L_i)$ . Para este caso, o termo  $(L_1 - L_2)[E(\log |\mathbf{Z}_1|) - E(\log |\mathbf{Z}_2|)]$  da expressão (E.2) é

$$T_1^{\mathcal{G}^0} = (L_1 - L_2) \left\{ T^{\mathcal{W}} + m \log \frac{(-\alpha_1 - 1)}{(-\alpha_2 - 1)} - m[\psi^{(0)}(-\alpha_1) - \psi^{(0)}(-\alpha_2)] \right\}.$$

Além disso, pelo resultado do Apêndice D, a soma  $E(\log[g(\text{tr}(\boldsymbol{\Sigma}_1^{-1} \mathbf{Z}_1); \boldsymbol{\theta}_1, L_1)]) + E(\log[g(\text{tr}(\boldsymbol{\Sigma}_2^{-1} \mathbf{Z}_2); \boldsymbol{\theta}_2, L_2)])$  é dada por

$$T_2^{\mathcal{G}^0} = \sum_{i=1}^2 (\alpha_i - mL_i) [\psi^{(0)}(mL_i - \alpha_i) + \log(-\alpha_i - 1) - \psi^{(0)}(-\alpha_i)].$$

Os termos cruzados  $(\alpha_i - mL_i) E\{\log[(-\alpha_i - 1) + L_i \text{tr}(\boldsymbol{\Sigma}_i^{-1} \mathbf{Z}_j)]\}$  se mostraram analiticamente intratáveis. Entretanto, a seguinte discussão elabora um limite inferior para esta quantidade. Partindo do resultado  $-\log(X) \geq 1 - X$ , para todo  $X$  real positivo (em que a igualdade é satisfeita se somente se  $X=1$ ), então se considerarmos a variável estocástica  $X = (-\alpha_i - 1) + L_i \text{tr}(\boldsymbol{\Sigma}_i^{-1} \mathbf{Z}_j)$ , o seguinte resultado é válido:

$$E\{-\log[(-\alpha_i - 1) + L_i \text{tr}(\boldsymbol{\Sigma}_i^{-1} \mathbf{Z}_j)]\} \geq 2 + \alpha_i - L_i \text{tr}(\boldsymbol{\Sigma}_i^{-1} \boldsymbol{\Sigma}_j),$$

valendo a igualdade se só se  $\text{tr}(\Sigma_i^{-1} \mathbf{Z}_j) = (2 + \alpha_i)/L_i$ .

Por hora, então, parece sugestivo utilizar o seguinte limite inferior para a distância de Kullback-Leibler, quando se quer contrastar elementos da família de distribuições  $\mathcal{G}_m^{0(L)}$ :

$$d_{\text{KL}}(\mathbf{Z}_1, \mathbf{Z}_2) = T_1^{\mathcal{G}^0} + T_2^{\mathcal{G}^0} + (\alpha_1 - mL_1)[2 + \alpha_1 - L_1 \text{tr}(\Sigma_1^{-1} \Sigma_2)] \\ + (\alpha_2 - mL_2)[2 + \alpha_2 - L_2 \text{tr}(\Sigma_2^{-1} \Sigma_1)].$$

**Para a distância de Bhattacharyya:** Agora consideremos a derivação da distância de Bhattacharyya para distribuições polarimétricas com densidades definidas por (E.1). De (E.1), é válido o resultado

$$d_B(\mathbf{Z}_1, \mathbf{Z}_2) = -\log \left[ \int_{\mathcal{A}} \sqrt{f_{\mathbf{Z}_1}(\mathbf{Z}'; \cdot) f_{\mathbf{Z}_2}(\mathbf{Z}'; \cdot)} d\mathbf{Z}' \right] = -\frac{\log T_1 + \log T_2}{2} \\ - \underbrace{\log \left[ \int_{\mathcal{A}} |\mathbf{Z}'|^{\frac{L_1+L_2}{2}-m} \sqrt{g(\text{tr}(\Sigma_1^{-1} \mathbf{Z}'); \theta_1, L_1) g(\text{tr}(\Sigma_2^{-1} \mathbf{Z}'); \theta_2, L_2)} d\mathbf{Z}' \right]}_{T_B}.$$

- Distribuição Wishart: Para este caso, temos

$$\sqrt{g(\text{tr}(\Sigma_1^{-1} \mathbf{Z}'); \theta_1, L_1) g(\text{tr}(\Sigma_2^{-1} \mathbf{Z}'); \theta_2, L_2)} \\ = \exp\left\{-\text{tr} \underbrace{[2^{-1}(L_1 \Sigma_1^{-1} + L_2 \Sigma_2^{-1}) \mathbf{Z}']}_{\Sigma_c^{-1}}\right\} \quad (\text{E.3})$$

Baseando-se no artigo de Maiwald & Kraus (2000) (que demonstrou propriedades importantes para a Wishart não escalonada bem com propuseram a densidade para a inversa da distribuição Wishart complexa que está definida em um suporte positivo definido), temos que a matriz  $\Sigma_i^{-1}$  é Hermitiana positiva definida e, como  $L_1, L_2 \geq m$ , as matrizes  $\Sigma_c^{-1}$  e  $\Sigma_c$  também o são. Desta forma, considerando  $\Sigma_c$  e  $L_c = (L_1 + L_2)/2$  como os parâmetros de uma distribuição Wishart não escalonada (vide equação (3.8)), o resultado seguinte decorre de sua densidade.

$$\int_{\mathcal{A}} |\mathbf{Z}'|^{L_c-m} \exp\{-\text{tr}[\Sigma_c^{-1} \mathbf{Z}']\} d\mathbf{Z}' = |\Sigma_c|^{L_c} \Gamma_m(L_c).$$

A partir dele, temos a quantidade  $T_B$  para a distribuição Wishart complexa escalonada é

$$T_B = -\frac{L_1 + L_2}{2} \log \left| \left( \frac{L_1 \Sigma_1^{-1} + L_2 \Sigma_2^{-1}}{2} \right)^{-1} \right| - \log \Gamma_m \left( \frac{L_1 + L_2}{2} \right)$$



Considerando o termo  $T_i = L_i^{mL_i} |\Sigma_i|^{-L_i}$ , para  $i = 1, 2$ , obtidos da Tabela 3.5 combinado ao resultado acima, a distância de Bhattacharyya entre elementos da distribuição Wishart escalonada é

$$d_B(\mathbf{Z}_1, \mathbf{Z}_2) = \quad (E.4)$$

$$\sum_{k=0}^{m-1} \log \frac{\sqrt{\Gamma(L_1 - k)\Gamma(L_2 - k)}}{\Gamma(\frac{L_1 + L_2}{2} - k)} + \frac{L_1 \log |\Sigma_1|}{2} + \frac{L_2 \log |\Sigma_2|}{2} - \frac{m}{2} (L_1 \log L_1 + L_2 \log L_2) - \frac{L_1 + L_2}{2} \log \left| \left( \frac{L_1 \Sigma_1^{-1} + L_2 \Sigma_2^{-1}}{2} \right)^{-1} \right|. \quad (E.5)$$

A distância de Hellinger é obtida da transformação

$$d_H(\mathbf{Z}_1, \mathbf{Z}_2) = 1 - \exp\{-d_B(\mathbf{Z}_1, \mathbf{Z}_2)\}.$$

- Notas sobre as distriuições  $\mathcal{G}^0$  e KummerU: A partir da derivação da Wishart, outras distâncias podem ser obtidas usando a mesma metodologia. Por exemplo, as distribuições  $\mathcal{G}_p^0$  e KummerU apresentam distâncias que dependem fundamentalmente das seguintes integrais, respectivamente:

$$\int_{\mathcal{A}} |\mathbf{Z}'|^{\frac{L_1 + L_2}{2} - m} [L_1 \text{tr}(\Sigma_1^{-1} \mathbf{Z}') + (-\alpha_1 - 1)]^{\frac{\alpha_1 - mL_1}{2}} [L_2 \text{tr}(\Sigma_2^{-1} \mathbf{Z}') + (-\alpha_2 - 1)]^{\frac{\alpha_2 - mL_2}{2}} d\mathbf{Z}'$$

e

$$\int_{\mathcal{A}} |\mathbf{Z}'|^{\frac{L_1 + L_2}{2} - m} U\left(L_1 m - \alpha, 1 + L_1 m - \mathcal{L}_1, \frac{L \text{tr}(\Sigma_1^{-1} \mathbf{Z}') \mathcal{L}_1}{(-\alpha_1 - 1)}\right)^{1/2} U\left(L_2 m - \alpha, 1 + L_2 m - \mathcal{L}_2, \frac{L \text{tr}(\Sigma_2^{-1} \mathbf{Z}') \mathcal{L}_2}{(-\alpha_2 - 1)}\right)^{1/2} d\mathbf{Z}'.$$

**Proposta a distância de Rényi:** A seguinte discussão versa sobre cálculo da distância de Rényi para distribuições polarimétricas com densidades definidas por (E.1).

De (E.1), é valido o resultado

$$\begin{aligned} d_R^\beta(\mathbf{Z}_1, \mathbf{Z}_2) &= \frac{1}{\beta - 1} \log \left[ \frac{\int_{\mathcal{A}} f_{\mathbf{Z}_1}^\beta(\mathbf{Z}'; \cdot) f_{\mathbf{Z}_2}^{1-\beta}(\mathbf{Z}'; \cdot) d\mathbf{Z}' + \int_{\mathcal{A}} f_{\mathbf{Z}_1}^{1-\beta}(\mathbf{Z}'; \cdot) f_{\mathbf{Z}_2}^\beta(\mathbf{Z}'; \cdot) d\mathbf{Z}'}{2} \right] \\ &= \frac{\log 2}{1 - \beta} + \frac{1}{\beta - 1} \log \left[ T_1^\beta T_2^{1-\beta} \int_{\mathcal{A}} |\mathbf{Z}'|^{\beta(L_1 - m) + (1-\beta)(L_2 - m)} g(\text{tr}(\Sigma_1^{-1} \mathbf{Z}'); \theta_1, L_1)^\beta g(\text{tr}(\Sigma_2^{-1} \mathbf{Z}'); \theta_2, L_2)^{1-\beta} d\mathbf{Z}' \right. \\ &\quad \left. + \int_{\mathcal{A}} T_1^{1-\beta} T_2^\beta |\mathbf{Z}'|^{\beta(L_2 - m) + (1-\beta)(L_1 - m)} g(\text{tr}(\Sigma_1^{-1} \mathbf{Z}'); \theta_1, L_1)^{1-\beta} g(\text{tr}(\Sigma_2^{-1} \mathbf{Z}'); \theta_2, L_2)^\beta d\mathbf{Z}' \right] \end{aligned}$$

- Distribuição Wishart: Para este caso, temos

$$\begin{aligned}
d_R^\beta(\mathbf{Z}_1, \mathbf{Z}_2) = & \\
= & \frac{\log 2}{1-\beta} + \frac{1}{\beta-1} \log \left[ T_1^\beta T_2^{1-\beta} \int_{\mathcal{A}} |\mathbf{Z}'| \overbrace{\beta L_1 + (1-\beta)L_2}^{L_{12}} - m \right. \\
& \left. \exp(-\text{tr}[\underbrace{\beta L_1 \Sigma_1^{-1} + (1-\beta)L_2 \Sigma_2^{-1}}_{\Sigma_{12}^{-1}}] \mathbf{Z}') d\mathbf{Z}' \right. \\
& + T_1^{1-\beta} T_2^\beta \int_{\mathcal{A}} |\mathbf{Z}'| \overbrace{\beta L_2 + (1-\beta)L_1}^{L_{21}} - m \\
& \left. \exp(-\text{tr}[\underbrace{\beta L_2 \Sigma_2^{-1} + (1-\beta)L_1 \Sigma_1^{-1}}_{\Sigma_{21}^{-1}}] \mathbf{Z}') d\mathbf{Z}' \right] \quad (\text{E.6})
\end{aligned}$$

Similarmente à distancia de Bhattacharyya, considerando  $\Sigma_{ij}$  e  $L_{ij}$ , para  $(i, j) = (1, 2), (2, 1)$ , como os parâmetros de uma distribuição Wishart não escalonada, tem-se que

$$\int_{\mathcal{A}} |\mathbf{Z}'|^{L_{ij}-m} \exp\{-\text{tr}[\Sigma_{ij}^{-1} \mathbf{Z}']\} d\mathbf{Z}' = |\Sigma_{ij}|^{L_{ij}} \Gamma_m(L_{ij}).$$

A partir dele, (E.6) se reduz a

$$\begin{aligned}
d_R^\beta(\mathbf{Z}_1, \mathbf{Z}_2) = & \\
= & \frac{\log 2}{1-\beta} + \frac{1}{\beta-1} \log \left[ T_1^\beta T_2^{1-\beta} |\Sigma_{12}|^{L_{12}} \Gamma_m(L_{12}) + T_1^{1-\beta} T_2^\beta |\Sigma_{21}|^{L_{21}} \Gamma_m(L_{21}) \right]. \quad (\text{E.7})
\end{aligned}$$

Considerando o termo  $T_i = L_i^{mL_i} |\Sigma_i|^{-L_i}$ , para  $i = 1, 2$ , obtidos da Tabela 3.5 combinado ao resultado acima, a distância de Rényi entre elementos da distribuição Wishart escalonada é

$$d_R^\beta(\theta_1, \theta_2) = \frac{1}{\beta-1} \log \frac{I(\theta_1, \theta_2)}{2}, \quad (\text{E.8})$$

tal que

$$\begin{aligned}
I(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) &= \left[ \Gamma(L_1 - m + 1)^m |\boldsymbol{\Sigma}_1|^{L_1} L_1^{-mL_1} \prod_{i=1}^{m-1} (L_1 - i)^i \right]^{-\beta} |\boldsymbol{\Sigma}_{12}|^{E_1} \\
&\times \left[ \Gamma(L_2 - m + 1)^m |\boldsymbol{\Sigma}_2|^{L_2} L_2^{-mL_2} \prod_{i=1}^{m-1} (L_2 - i)^i \right]^{\beta-1} \\
&\times \Gamma(E_1 - m + 1)^m \prod_{i=1}^{m-1} (E_1 - i)^i \\
&+ \left[ (\Gamma(L_1 - m + 1)^m |\boldsymbol{\Sigma}_1|^{L_1} L_1^{-pL_1} \prod_{i=1}^{m-1} (L_1 - i)^i) \right]^{\beta-1} |\boldsymbol{\Sigma}_{21}|^{E_2} \\
&\times \left[ \Gamma(L_2 - m + 1)^m |\boldsymbol{\Sigma}_2|^{L_2} L_2^{-mL_2} \prod_{i=1}^{m-1} (L_2 - i)^i \right]^{-\beta} \\
&\times \Gamma(E_2 - m + 1)^m \prod_{i=1}^{m-1} (E_2 - i)^i,
\end{aligned}$$

em que  $E_1 = \beta L_1 + (1 - \beta)L_2$ ,  $E_2 = \beta L_2 + (1 - \beta)L_1$ ,  $0 < \beta < 1$ , e  $\boldsymbol{\Sigma}_{ij} = |(L_i \beta \boldsymbol{\Sigma}_i^{-1} + L_j (1 - \beta) \boldsymbol{\Sigma}_j^{-1})^{-1}|$  para  $(i, j) = (1, 2), (2, 1)$ .

**Distância de Rényi:** Uma segunda simetrização da divergência de Rényi é dada por

$$\begin{aligned}
d_{\text{R}}^{\beta}(\mathbf{Z}_1, \mathbf{Z}_2) &= [2(\beta - 1)]^{-1} \left\{ \log \left[ \int_{\mathcal{A}} f_{\mathbf{Z}_1}^{\beta}(\mathbf{Z}'; \cdot) f_{\mathbf{Z}_2}^{1-\beta}(\mathbf{Z}'; \cdot) d\mathbf{Z}' \right] \right. \\
&\quad \left. + \log \left[ \int_{\mathcal{A}} f_{\mathbf{Z}_1}^{1-\beta}(\mathbf{Z}'; \cdot) f_{\mathbf{Z}_2}^{\beta}(\mathbf{Z}'; \cdot) d\mathbf{Z}' \right] \right\}
\end{aligned}$$

Considerando os mesmos argumentos anteriores, temos a seguinte distância mais simplificada:

$$\begin{aligned}
d_{\text{R}2}^{\beta}(\mathbf{Z}_1, \mathbf{Z}_2) &= [2(1 - \beta)]^{-1} \left[ \sum_{k=0}^{m-1} \log \frac{\Gamma(L_1 - k) \Gamma(L_2 - k)}{\Gamma(E_1 - k) \Gamma(E_2 - k)} + L_1 \log |\boldsymbol{\Sigma}_1| + L_2 \log |\boldsymbol{\Sigma}_2| \right. \\
&\quad \left. - E_1 \log |\boldsymbol{\Sigma}_{12}| - E_2 \log |\boldsymbol{\Sigma}_{21}| - m(L_1 \log L_1 + L_2 \log L_2) \right],
\end{aligned}$$

em que  $E_i$ , para  $i = 1, 2$ , e  $\boldsymbol{\Sigma}_{ij}$  foram definidos no item anterior.

## Referências Bibliográficas

- Abbasnejad, M. & Arghami, N. R. (2011), 'Rényi entropy properties of order statistics', *Communications in Statistics - Theory and Methods* **40**(1), 40–52.
- Abramowitz, M. & Stegun, I. A. (1964), *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, Dover.
- Ainsworth, T. L., Schuler, D. L. & Lee, J. S. (2008), 'Polarimetric SAR characterization of man-made structures in urban areas using normalized circular-pol correlation coefficients', *Remote Sensing of Environment* **112**(6), 2876–2885.
- Ali, S. M. & Silvey, S. D. (1996), 'A general class of coefficients of divergence of one distribution from another', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **26**, 131–142.
- Allende, H., Frery, A. C., Galbiati, J. & Pizarro, L. (2006), 'M-estimators with asymmetric influence functions: The GA0 distribution case', *The Journal of Statistical Computation and Simulation*. **76**(11), 941–956.
- Andai, A. (2009), 'On the geometry of generalized Gaussian distributions', *Journal of Multivariate Analysis* **100**(4), 777–793.
- Anfinsen, S. N., Doulgeris, A. P. & 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, S. N., Doulgeris, A. P. & Eltoft, T. (2011), 'Goodness-of-fit tests for multi-look polarimetric radar data based on the Mellin transform', *IEEE Transactions on Geoscience and Remote Sensing* **49**(7), 2764 –2781.
- Anfinsen, S. N., Eltoft, T. & Doulgeris, A. P. (2009), A relaxed Wishart model for polarimetric SAR data, in 'Proceedings of the 4th International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLinSAR'2009)', Vol. ESA SP-668, Frascati, Italy.
- Anfinsen, S. N., Jenssen, R. & Eltoft, T. (2007), Spectral clustering of polarimetric SAR data with Wishart-derived distance measures, in 'Proceedings of the 3rd Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLinSAR'2007)', Frascati, Italy.

- Arimoto, S. (1971), 'Information-theoretical considerations on estimation problems', *Information and Control* **19**, 181–194.
- Baratpour, S., Ahmadi, J. & Arghami, N. R. (2008), 'Characterizations based on Rényi entropy of order statistics and record values', *Journal of Statistical Planning and Inference* **138**(8), 2544 – 2551.
- Barndorff-Nielsen, O. E. (1994), 'Adjusted versions of profile likelihood and directed likelihood, and extended likelihood', *Journal of the Royal Statistical Society. Series B (Methodological)* **56**(1), 125–140.
- Barndorff-Nielsen, O. E. & Cox, D. R. (1989), *Asymptotic Techniques for use in Statistics*, Chapman & Hall, London.
- Basseville, M. (1989), 'Distance measures for signal processing and pattern recognition', *Signal Processing* **18**, 349–369.
- Beauchemin, M., Thomson, K. & Edwards, G. (1998), 'On nonparametric edge detection in multilook SAR images', *IEEE Transactions on Geoscience and Remote Sensing* **36**(5), 1826–1829.
- Beaulieu, J. M. & Touzi, R. (2004), 'Segmentation of textured polarimetric SAR scenes by likelihood approximation', *IEEE Transactions on Geoscience and Remote Sensing* **42**(10), 2063–2072.
- Blahut, R. E. (1972), *An hypothesis testing approach to information theory*, by Richard E. Blahut, Cornell University, August.
- Blake, A. & Isard, M. (1998), *Active Contours*, Springer, Berlin.
- Blatt, D. & Hero, A. O. (2007), 'On tests for global maximum of the log-likelihood function', *IEEE Transactions on Information Theory* **53**, 2510–2525.
- Boerner, W. M. & Lee, J.-S. (2007), Review of existing monographs and books on radar polarimetry and polarimetric SAR with the aim of justifying the need of updates, in 'IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2007)', pp. 180–183.
- Bombrun, L., Anfinson, S. N. & Harant, O. (2011), A complete coverage of log-cumulant space in terms of distributions for polarimetric SAR data, in 'Proceedings of the 5th International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLinSAR'2011)', Frascati, Italy.
- Bombrun, L. & Beaulieu, J. M. (2008), 'Fisher distribution for texture modeling of polarimetric SAR data', *IEEE Geoscience and Remote Sensing Letters* **5**(3), 512–516.

- Bombrun, L., Vasile, G., Gay, M., Pascal, F. & Ovarlez, J.-P. (2010), KummerU clutter model for PolSAR data: Application to segmentation and classification, in '2nd SONDRRA Workshop', Cargèse France. Département Images et Signal.
- Bombrun, L., Vasile, G., Gay, M. & Totir, F. (2011), 'Hierarchical segmentation of polarimetric SAR images using heterogeneous clutter models', *IEEE Transactions on Geoscience and Remote Sensing* **49**(2), 726–737.
- Bregman, L. M. (1967), 'The Relaxation Method of Finding the Common Point of Convex Sets and Its Application to the Solution of Problems in Convex Programming', *USSR Computational Mathematics and Mathematical Physics* **7**, 200–217.
- Brewer, J. (1978), 'Kronecker products and matrix calculus in system theory', *IEEE Transactions on Circuits and Systems* **25**(9), 772–781.
- Broniatowski, M. & Keziou, A. (2009), 'Parametric estimation and tests through divergences and the duality technique', *Journal of Multivariate Analysis* **100**(1), 16–36.
- Burbea, J. (1984), 'The Bose-Einstein entropy of degree  $\alpha$  and its Jensen difference', *Utilitas Mathematica* **25**, 225–240.
- Burbea, J. & Rao, C. (1982a), 'Entropy differential metric, distance and divergence measures in probability spaces: a unified approach', *Journal of Multivariate Analysis* **12**, 575–596.
- Burbea, J. & Rao, C. (1982b), 'On the convexity of higher order Jensen differences based on entropy functions', *IEEE Transactions on Information Theory* **28**(6), 961–963.
- Burbea, J. & Rao, C. R. (1982c), 'On the convexity of some divergence measures based on entropy functions', *IEEE Transactions on Information Theory* **IT-28**, 489–495.
- Bustos, O. H., Lucini, M. M. & Frery, A. C. (2002), 'M-estimators of roughness and scale for GA0-modelled SAR imagery', *EURASIP Journal on Applied Signal Processing* **2002**(1), 105–114.
- Cao, F., Hong, W., Wu, Y. & Pottier, E. (2007), 'An unsupervised segmentation with an adaptive number of clusters using the SPAN/H/ $\alpha$ /A space and the complex Wishart clustering for fully polarimetric SAR data analysis', *IEEE Transactions on Geoscience and Remote Sensing* **45**(11), 3454–3467.
- Casella, G. & Berger, R. L. (2002), *Statistical Inference*, Duxbury Press.
- Chambers, J. M. (2008), *Software for data analysis: programming with R*, Statistics and computing, Springer.
- Chernoff, H. (1952), 'A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations', *The Annals of Mathematical Statistics* **23**(4), 493–507.

- Cintra, R. J., Frery, A. C. & Nascimento, A. D. C. (2011), 'Parametric and nonparametric tests for speckled imagery', *Pattern Analysis and Applications* pp. 1–21.
- Cloude, S. R. & Pottier, E. (1997), 'An entropy based classification scheme for land application of polarimetric SAR', *IEEE Transactions on Geoscience and Remote Sensing* **35**, 68–78.
- Cochran, W. G. (1934), 'The distribution of quadratic forms in a normal system, with applications to the analysis of covariance', *Mathematical Proceedings of the Cambridge Philosophical Society* **30**(2), 178–191.
- Conradsen, K., Nielsen, A. A., Schou, J. & 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.
- Cordeiro, G. M. (2011), Bias correction, in M. Lovric, ed., 'International Encyclopedia of Statistical Science', Springer Berlin Heidelberg, pp. 148–152.
- Cover, T. M. & Thomas, J. A. (1991), *Elements of Information Theory*, Wiley-Interscience, New York.
- Cox, D. R. & Reid, N. (1987), 'Parameter orthogonality and approximate conditional inference', *Journal of the Royal Statistical Society. Series B (Methodological)* **49**(1), 1–39.
- Cox, D. R. & Reid, N. (1992), 'A note on the difference between profile and modified profile likelihood', *Biometrika* **79**(2), 408–411.
- Cox, D. R. & Snell, E. J. (1968), 'A general definition of residuals', *Journal of the Royal Statistical Society. Series B (Methodological)* **30**(2), 248–275.
- Cressie, N. & Read, T. R. C. (1984), 'Multinomial goodness-of-fit tests', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **46**, 440–464.
- Cribari-Neto, F. & Cordeiro, G. M. (1996), 'On Bartlett and Bartlett-type corrections', *Econometric Reviews* **15**(4), 339–367.
- Cribari-Neto, F., Frery, A. C. & Silva, M. F. (2002), 'Improved estimation of clutter properties in speckled imagery', *Computational Statistics & Data Analysis* **40**(4), 801–824.
- Cribari-Neto, F. & Zarkos, S. G. (1999), 'R: yet another econometric programming environment', *Journal of Applied Econometrics* **14**, 319–329.
- Csiszár, I. (1967), 'Information type measures of difference of probability distributions and indirect observations', *Studia Scientiarum Mathematicarum Hungarica* **2**, 299–318.

- Delignon, Y. & Pieczynski, W. (2002), 'Modeling non-Rayleigh speckle distribution in SAR images', *IEEE Transactions on Geoscience and Remote Sensing* **40**(6), 1430–1435.
- Deza, M. M. & Deza, E. (2009), *Encyclopedia of Distances*, 1 edn, Springer.
- Dharmawansa, P. & McKay, M. R. (2009), 'Diagonal distribution of a complex non-central Wishart matrix: A new trivariate non-central chi-squared density', *Journal of Multivariate Analysis* **100**(4), 561–580.
- Donoho, D. L., Vetterli, M., DeVore, R. A. & Daubechies, I. (1998), 'Data compression and harmonic analysis', *IEEE Transactions on Information Theory* **44**(6), 2435–2476.
- Doornik, J. A. (1999), *Object-Oriented Matrix Programming Using Ox*, 3 edn, Timberlake Consultants Press.
- Doulgeris, A. P., Anfinson, S. N. & Eltoft, T. (2011), 'Automated non-gaussian clustering of polarimetric synthetic aperture radar images', *IEEE Transactions on Geoscience and Remote Sensing* **49**(10), 3665–3676.
- Eguchi, S. & Copas, J. (2006), 'Interpreting Kullback-Leibler divergence with the Neyman-Pearson lemma', *Journal of Multivariate Analysis* **97**, 2034–2040.
- Epstein, B. (1948), 'Some applications of the Mellin transform in statistics', *The Annals of Mathematical Statistics* **19**(3), 370–379.
- Ersahin, K., Cumming, I. & Ward, R. K. (2010), 'Segmentation and classification of polarimetric SAR data using spectral graph partitioning', *IEEE Transactions on Geoscience and Remote Sensing* **48**(1), 164–174.
- European Space Agency (2009), *Polarimetric SAR Data Processing and Educational Tool (POLSARPRO)*, 4.0 edn, European Space Agency.
- Fannjiang, A. C., Solna, K. & Yan, P. (2009), 'Synthetic aperture imaging of multiple point targets in rician fading media', *SIAM Journal on Imaging Sciences* **2**(2), 344–366.
- Ferguson, T. S. (1967), *Mathematical Statistics: A Decision Theoretic Approach*, Probability and mathematical statistics, Academic Press.
- Ferreri, C. (1980), 'Hypoentropy and related heterogeneity, divergence and information measures', *Statistica* **2**, 155–167.
- Filho, C. R. S. & Crósta, P. A. (2003), 'Geotecnologia aplicadas à geologia', *Revista Brasileira de Geociência* **33**(2), 1–4.
- Freeman, A. & Durden, S. L. (1998), 'A three-component scattering model for polarimetric SAR data', *IEEE Transactions on Geoscience and Remote Sensing* **36**(3), 963–973.



- Freitas, C. C., Frery, A. C. & Correia, A. H. (2005), 'The polarimetric G distribution for SAR data analysis', *Environmetrics* **16**(1), 13–31.
- Freitas, C. C., Sant'Anna, S. J. S., Rennó, C. D. & Correia, A. H. (2003), Utilização de imagens de radar de abertura sintética na classificação do uso e ocupação do solo, in 'Embrapa (Org.). Geomática: Técnicas e Aplicações Avançadas'.
- Frery, A. C., Cintra, R. J. & Nascimento, A. D. C. (2011), Hypothesis test in complex Wishart distributions, in 'Proceedings of the 5th International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLINAR'2011)', Frascati, Italy.
- Frery, A. C., Cintra, R. J. & Nascimento, A. D. C. (2012), 'Entropy-based statistical analysis of PolSAR data', *Submitted to IEEE Geoscience and Remote Sensing*.
- Frery, A. C., Correia, A. H. & Freitas, C. C. (2007), 'Classifying multifrequency fully polarimetric imagery with multiple sources of statistical evidence and contextual information', *IEEE Transactions on Geoscience and Remote Sensing* **45**, 3098–3109.
- Frery, A. C., Cribari-Neto, F. & Souza, M. O. (2004), 'Analysis of minute features in speckled imagery with maximum likelihood estimation', *EURASIP Journal on Applied Signal Processing* **2004**(16), 2476–2491.
- Frery, A. C., Jacobo-Berlles, J., Gambini, J. & Mejail, M. (2010), 'Polarimetric SAR image segmentation with B-splines and a new statistical model', *Multidimensional Systems and Signal Processing* **21**, 319–342.
- Frery, A. C., Muller, H. J., Yanasse, C. C. F. & Sant'Anna, S. J. S. (1997), 'A model for extremely heterogeneous clutter', *IEEE Transactions on Geoscience and Remote Sensing* **35**(3), 648–659.
- Frery, A. C., Nascimento, A. D. C. & Cintra, R. J. (2010a), Analytic expressions for stochastic distances between relaxed complex Wishart distributions. Submitted to IEEE Geoscience and Remote Sensing.
- Frery, A. C., Nascimento, A. D. C. & Cintra, R. J. (2010b), Contrast in speckled imagery with stochastic distances, in 'International Conference on Image Processing (ICIP)', Hong Kong, pp. 26–29.
- Frery, A. C., Nascimento, A. D. C. & Cintra, R. J. (2011), 'Information theory and image understanding: An application to polarimetric SAR imagery', *Chilean Journal of Statistics* **2**(2), 81–100.
- Gambini, J., Mejail, M., Jacobo-Berlles, J. & Frery, A. C. (2006), 'Feature extraction in speckled imagery using dynamic B-spline deformable contours under the G0 model', *International Journal of Remote Sensing* **27**(22), 5037–5059.

- Gambini, J., Mejail, M., Jacobo-Berlles, J. & Frery, A. C. (2008), 'Accuracy of edge detection methods with local information in speckled imagery', *Statistics and Computing* **18**(1), 15–26.
- Gao, G. (2010), 'Statistical modeling of SAR images: A survey', *Sensors* **10**, 775–795.
- Gentle, J. E. (2002), *Elements of Computational Statistics*, Statistics and computing, Springer.
- Gierull, C. H. & Sikaneta, I. C. (2002), 'Estimating the effective number of looks in interferometric SAR data', *IEEE Transactions on Geoscience and Remote Sensing* **40**(8), 1733–1742.
- Giet, L. & Lubrano, M. (2008), 'A minimum Hellinger distance estimator for stochastic differential equations: An application to statistical inference for continuous time interest rate models', *Computational Statistics & Data Analysis* **52**(6), 2945–2965.
- Goodman, J. W. (1985), *Statistical Optics*, Wiley Series in Pure and Applied Optics, New York.
- Goodman, N. R. (1963a), 'The distribution of determinant of a complex Wishart distributed matrix', *The Annals of Mathematical Statistics* **34**, 178–180.
- Goodman, N. R. (1963b), 'Statistical analysis based on a certain complex Gaussian distribution (an introduction)', *The Annals of Mathematical Statistics* **34**, 152–177.
- Goudail, F. & Réfrégier, P. (2004), 'Contrast definition for optical coherent polarimetric images', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**(7), 947–951.
- Goudail, F., Réfrégier, P. & Delyon, G. (2004), 'Bhattacharyya distance as a contrast parameter for statistical processing of noisy optical images', *Journal of the Optical Society of America A* **21**(7), 1231–1240.
- Hagedorn, M., Smith, P. J., Bones, P. J., Millane, R. P. & Pairman, D. (2006), 'A trivariate chi-squared distribution derived from the complex Wishart distribution', *Journal of Multivariate Analysis* **97**(3), 655–674.
- Harant, O., Bombrun, L., Gay, M., Fallourd, R., Trouvã, E. & Tupin, F. (2009), Segmentation and classification of polarimetric SAR data based on the KummerU distribution, in 'Proceedings of the 4th International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLinSAR'2009)', Vol. ESA SP-668.
- Hartley, R. V. L. (1928), 'Transmission of information', *Bell System Technical Journal* **7**, 535–563.
- Havrda, J. & Chalvat, F. (1967), 'Concept of structural  $\alpha$ -entropy', *Kybernetika* **3**, 30–35.

- Hjorungnes, A. & Gesbert, D. (2007), 'Complex-valued matrix differentiation: Techniques and key results', *IEEE Transactions on Signal Processing* **55**(6), 2740–2746.
- Horta, M. M. (2009), Modelos de Mistura de Distribuições na Segmentação de Imagens SAR Polarimétricas Multi-look, PhD thesis, Instituto de Física de São Carlos.
- Hu, W. (2005), Calibration of Multivariate Generalized Hyperbolic Distributions Using the EM Algorithm, with Applications in Risk Management, Portfolio Optimization, and Portfolio Credit Risk, PhD thesis, Florida State University.
- Ihaka, R. & Gentleman, R. (1996), 'R: A language for data analysis and graphics', *Journal of Computational and Graphical Statistics* **5**(3), 299–314.
- Inglada, J. & 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.
- Johnson, R. A. & Wichern, D. W. (1988), *Applied Multivariate Statistical Analysis*, Prentice-Hall series in statistics, Prentice-Hall.
- Kailath, T. (1967), 'The divergence and Bhattacharyya distance measures in signal selection', *IEEE Transactions on Communication Technology* **15**, 52–60.
- Kapur, J. N. (1972), 'Measures of uncertainty, mathematical programming and physics.', *Journal of the Indian Society of Agriculture and Statistics* **24**, 47–66.
- Kersten, P. R., Lee, J. S. & Ainsworth, T. L. (2005), '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.
- Kollo, T. & Rosen, D. (2005), *Advanced Multivariate Statistics with Matrices*, Mathematics and its applications, Springer.
- Kolmogorov, A. N. (1963), 'On the approximation of distributions of sums of independent summands by infinitely divisible distributions', *The Indian Journal of Statistics, Series A* **25**(2), 159–174.
- Kullback, S. (1978), *Information Theory and Statistics*, Dover.
- Kullback, S. & Leibler, R. A. (1951), 'On information and sufficiency', *The Annals of Mathematical Statistics* **22**(1), 79–86.
- Lawley, D. N. (1956), 'A general method for approximating to the distribution of likelihood ratio criteria', *Biometrika* **43**(4), 295–303.
- Lee, J. S., Grunes, M. R. & Kwok, R. (1994), 'Classification of multi-look polarimetric SAR imagery based on complex Wishart distribution', *International Journal of Remote Sensing* **15**(11), 2299–2311.

- Lee, J. S., Hoppel, K. W., Mango, S. A. & Miller, A. R. (1994), 'Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery', *IEEE Transactions on Geoscience and Remote Sensing* **32**(5), 1017–1028.
- Lee, J.-S., Miller, A. R. & Hoppel, K. W. (1994), 'Statistics of phase difference and product magnitude of multi-look processed Gaussian signals', *Waves in Random Media* **4**(3), 307–319.
- Lee, J. S. & Pottier, E. (2009), *Polarimetric Radar Imaging: From Basics to Applications*, CRC, Boca Raton.
- Lee, J. S., Schuler, D. L., Lang, R. H. & Ranson, K. J. (1994), *k*-distribution for multi-look processed polarimetric sar imagery, in 'Geoscience and Remote Sensing Symposium, 1994 (IGARSS '94)', Vol. 4, pp. 2179–2181.
- Lefebvre, E. (2007), *Advances and challenges in multisensor data and information processing*, NATO security through science series: Information and communication security, IOS Press.
- Liese, F. & Vajda, I. (2006), 'On divergences and informations in statistics and information theory', *IEEE Transactions on Information Theory* **52**(10), 4394–4412.
- Liuni, M., Robel, A., Romito, M. & Rodet, X. (2011), Rényi information measures for spectral change detection, in 'IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'2011)', pp. 3824–3827.
- Lopes, A. & 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.
- López-Martínez, C. & Fabregas, X. (2003), 'Polarimetric SAR speckle noise model', *IEEE Transactions on Geoscience and Remote Sensing* **41**(10), 2232–2242.
- López-Martínez, C., Fábregas, X. & Pottier, E. (2005), 'Multidimensional speckle noise model', *EURASIP Journal on Applied Signal Processing* **2005**(20), 3259–3271.
- López-Martínez, C. & Pottier, E. (2007), 'On the extension of multidimensional speckle noise model from single-look to multilook SAR imagery', *IEEE Transactions on Geoscience and Remote Sensing* **45**(2), 305–320.
- Mahalanobis, P. C. (1936), 'On the generalized distance in statistics', *National Institute of Science* **12**, 49–55.
- Maiwald, D. & Kraus, D. (2000), 'Calculation of moments of complex Wishart and complex inverse Wishart distributed matrices', *IEE Proceedings - Radar, Sonar and Navigation* **147**(4), 162–168.

- Mathai, A. M. (1997), *Jacobians of Matrix Transformations and Functions of Matrix Arguments*, World Scientific, New York, USA.
- Maybank, S. J. (2007), 'Application of the Fisher-Rao metric to ellipse detection', *International Journal of Computer Vision* **72**(3), 287–307.
- Mejail, M. E., Frery, A. C., Jacobo-Berlles, J. & Bustos, O. H. (2001), 'Approximation of distributions for SAR images: Proposal, evaluation and practical consequences', *Latin American Applied Research* **31**, 83–92.
- Mejail, M. E., Jacobo-Berlles, J., Frery, A. C. & Bustos, O. H. (2003), 'Classification of SAR images using a general and tractable multiplicative model', *International Journal of Remote Sensing* **24**(18), 3565–3582.
- Menéndez, M. L., Pardo, J. A. & Pardo, L. (2007), 'The moment-corrected phi-divergence test statistics for symmetry', *Journal of Computational and Applied Mathematics* **202**(2), 315–327.
- Menéndez, M., Morales, D., Pardo, L. & Salicrú, M. (1995), 'Asymptotic behaviour and statistical applications of divergence measures in multinomial populations: a unified study', *Statistical Papers* **36**, 1–29. 10.1007/BF02926015.
- Morio, J., Réfrégier, P., Goudail, F., Dubois-Fernandez, P. C. & Dupuis, X. (2008), 'Information theory-based approach for contrast analysis in polarimetric and/or interferometric SAR images', *IEEE Transactions on Geoscience and Remote Sensing* **46**(8), 2185–2196.
- Morio, J., Réfrégier, P., Goudail, F., Dubois Fernandez, P. C. & Dupuis, X. (2009), 'A characterization of Shannon entropy and Bhattacharyya measure of contrast in polarimetric and interferometric SAR image', *Proceedings of the IEEE (PIEEE)* **97**(6), 1097–1108.
- Nadarajah, S. & Zografos, K. (2003), 'Formulas for Rényi information and related measures for univariate distributions', *Information Science* **155**, 119–138.
- Nadarajah, S. & Zografos, K. (2005), 'Expressions for Rényi and Shannon entropies for bivariate distributions', *Information Sciences* **170**(2-4), 173–189.
- Nascimento, A. D. C., Cintra, R. J. & Frery, A. C. (2009a), Measuring synthetic aperture radar target differences with stochastic distances, in 'TIC-STH Symposium on Advances in Systems and Sensors', Toronto, Canada, pp. 587–592.
- Nascimento, A. D. C., Cintra, R. J. & Frery, A. C. (2009b), Stochastic distances and hypothesis testing in speckled data, in 'XIV Brazilian Remote Sensing Symposium (SBSR)', Vol. 14, Natal, RN, pp. 7353–7360.

- Nascimento, A. D. C., Cintra, R. J. & Frery, A. C. (2010), 'Hypothesis testing in speckled data with stochastic distances', *IEEE Transactions on Geoscience and Remote Sensing* **48**(1), 373–385.
- Nascimento, A. D. C., Cintra, R. J. & Frery, A. C. (2012), 'Modified profile likelihood and bias correction in Wishart complex distribution'.
- Nascimento, A. D. C., Frery, A. C. & Cintra, R. J. (2009), Analysis of speckled imagery with parametric and nonparametric tests, in 'Sibgrapi 2009 - Theses and Dissertations'.
- Nascimento, A. D. C., Frery, A. C. & Cintra, R. J. (2012a), 'An extension for the family of  $\mathcal{G}_{\mathcal{I}}$  distributions'.
- Nascimento, A. D. C., Frery, A. C. & Cintra, R. J. (2012b), 'Parametric hypothesis tests in PolSAR imagery: Information theory vs. likelihood ratio statistics'.
- Nascimento, A. D. C., Horta, M. M. & Frery, A. C. (2011), Boundary detection with stochastic distance in polarimetric imagery, in 'Proceedings of the 5th International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLinSAR'2011)', Frascati, Italy.
- Neuman, E. (1990), 'Inequalities involving multivariate convex functions II', *Proceedings of the American Mathematical Society* **109**(4), 965–974.
- Novak, L. M. & Burl, M. C. (1990), 'Optimal speckle reduction in polarimetric SAR imagery', *IEEE Transactions on Aerospace and Electronic Systems* **26**(2), 293–305.
- Nyquist, H. (1924), 'Certain factors affecting telegraph speed', *Bell System Technical Journal* **3**(2), 324–346.
- Nyquist, H. (1928), 'Certain topics in telegraph transmission theory', *Transaction of the American Institute of Electrical Engineers transaction* **47**, 617–644.
- Oliver, C. & Quegan, S. (1998), *Understanding Synthetic Aperture Radar Images*, The Sci-Tech radar and defense series, SciTech Publishing.
- Pardo, L., Morales, D., Salicrú, M. & Menéndez, M. L. (1997), 'Large sample behavior of entropy measures when parameters are estimated', *Communications in Statistics - Theory and Methods* **26**(2), 483–501.
- Piessens, R., de Doncker-Kapenga, E., Uberhuber, C. W. & Kahaner, D. K. (1983), *QUADPACK: A Subroutine Package for Automatic Integration*, Springer-Verlag, New York.
- Pottier, E., Lee, J. S. & Ferro-Famil, L. (2006), Advanced concepts in polarimetric SAR image analysis: A tutorial review, in 'EuSAR 2006 Tutorial, EuSAR 2006 Conference', Dresden, Germany.

- Prudnikov, A. P., Brychkov, Y. A. & Maichev, I. . (1986), *Integrals and Series*, Vol. 2, New York.
- Rao, C. (1990), 'On the use and interpretation of distance functions in statistics', *Bulletin of the International Statistical Institute* **34**.
- Rao, C. R. (1984), 'Convexity properties of entropy functions and analysis of diversity', *Lecture Notes-Monograph Series* **5**, pp. 68–77.
- Rényi, A. (1961), On measures of entropy and information, in '4th Berkeley Symposium on Mathematical Statistics and Probability', Vol. 1, pp. 547–561.
- Rizki Akbar, P., Tetuko S. S., J. & Kuze, H. (2010), 'A novel circularly polarized synthetic aperture radar (cp-sar) system onboard a spaceborne platform', *International Journal of Remote Sensing* **31**, 1053–1060.
- Rosso, O. A., De-Micco, L., Larrondo, H. A., Martín, M. T. & Plastino, A. (2010), 'Generalized statistical complexity measure', *International Journal of Bifurcation and Chaos* **20**(3), 775–785.
- Rubinstein, R. Y. & Kroese, D. P. (2008), *Simulation and the Monte Carlo Method*, Wiley Series in Probability and Statistics, John Wiley & Sons.
- Salicrú, M., Mendéndez, M. L. & Pardo, L. (1993), 'Asymptotic distribution of  $(h, \phi)$ -entropy', *Communications in Statistics - Theory Methods* **22**(7), 2015–2031.
- Salicrú, M., Menéndez, M. L., Pardo, L. & Morales, D. (1994), 'On the applications of divergence type measures in testing statistical hypothesis', *Journal of Multivariate Analysis* **51**, 372–391.
- Salicrú, M. & Vives, Ocaña, S. J. (2005), 'Testing the homogeneity of diversity measures: a general framework', *Journal of Statistical Planning and Inference* **132**, 117–129.
- Sarabandi, K. (1992), 'Derivation of phase statistics from the Mueller matrix', *Radio Science* **27**, 553–560.
- Schou, J., Skriver, H., Nielsen, A. H. & Conradsen, K. (2003), 'CFAR edge detector for polarimetric SAR images', *IEEE Transactions on Geoscience and Remote Sensing* **41**(1), 20–32.
- Seber, G. A. F. (2007), *A Matrix Handbook for Statisticians*, 1st edn, Wiley-Interscience, New York, NY, USA.
- Seghouane, A. K. & Amari, S. I. (2007), 'The AIC criterion and symmetrizing the Kullback-Leibler divergence', *IEEE Transactions on Neural Networks* **18**(1), 97–106.
- Severini, T. A. (2000), *Likelihood Methods in Statistics*, Oxford: Oxford university Press.

- Shannon, C. E. (1948), 'A mathematical theory of communication', *Bell System Technical Journal* **27**, 379–423.
- Sharma, B. D. & Mittal, D. P. (1977), 'New non-additive measures of relative information.', *Journal of Combinatory Information & Systems Science* **2**, 122–133.
- Silva, M., Cribari-Neto, F. & Frery, A. C. (2008), 'Improved likelihood inference for the roughness parameter of the GA0 distribution', *Environmetrics* **19**(4), 347–368.
- Skriver, H., Dall, J., Le Toan, T., Quegan, S., Ferro-Famil, L., Pottier, E., Lumsdon, P. & Moshhammer, R. (2005), Agriculture classification using polsar data, in 'Proceedings of the 2nd International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLinSAR'2005)', Vol. ESA SP-586, Frascati, Italy.
- Taneja, I. J. (2006), 'Bounds on triangular discrimination, harmonic mean and symmetric chi-square divergences', *Journal of Concrete and Applicable Mathematics* **4**, 91–111.
- Ulabay, F. T. & Elachi, C. (1990), *Radar Polarimetry for Geoscience Applications*, Artech House, Norwood.
- Ulabay, F. T., Moore, R. K. & Fung, A. K. (1986), *Microwave Remote Sensing Active and Passive: Radar Remote Sensing and Surface Scattering and Emission Theory*, Artech House, Inc, Norwood, MA.
- Ullah, A. (1996), 'Entropy, divergence and distance measures with econometric applications', *Journal of Statistical Planning and Inference* **49**(1), 137–162.
- Van den Bos, A. (1994), 'A Cramer-Rao lower bound for complex parameters', *IEEE Transactions on Signal Processing* **42**(10), 2859.
- Varma, R. S. (1966), 'Generalizations of Rényi's entropy of order  $\alpha$ ', *Journal of Mathematical Sciences*, **1**, 34–48.
- Vasconcellos, K. L. P., Frery, A. C. & Silva, L. B. (2005), 'Improving estimation in speckled imagery', *Computational Statistics* **20**(3), 503–519.
- Vorobyov, A. V., Gershman, A. B. & Wong, K. M. (2005), 'Maximum likelihood direction-of-arrival estimation in unknown noise fields using sparse sensor arrays', *IEEE Transactions on Signal Processing* **53**(1), 34–43.
- Woelder, K. & Granholm, J. (2010), 'Cross-polarization and sidelobe suppression in dual linear polarization antenna arrays', *IEEE Transactions on Antennas and Propagation* **45**, 1727–1740.



- Yan, W., Yang, W., Liu, Y. & Sun, H. (2010), Unsupervised classification of PolinSAR image based on Shannon entropy characterization, in 'Proceedings of the 2010 IEEE International Conference on Signal Processing (ICSP 2010)', pp. 2192–2195.
- Yao, K. (1973), 'A representation theorem and its applications to spherically-invariant random processes', *IEEE Transactions on Information Theory* **19**(5), 600–608.
- Yueh, S. H., Kong, J. A., Jao, J. K., Shin, R. T. & Novak, L. M. (1989), 'K-distribution and polarimetric terrain radar clutter', *Journal of Electromagnetic Waves and Applications* **3**(8), 747–768.
- Zografos, K. & Nadarajah, S. (2005), 'Expressions for Rényi and Shannon entropies for multivariate distributions', *Statistics & Probability Letters* **71**(1), 71–84.