Machine learning and image processing methods for acoustic seafloor mapping and classification

Knut Landmark

February 12, 2016

Preface

This thesis has been submitted to the Faculty of Mathematics and Natural Sciences at the University of Oslo in partial fulfillment of the requirements for the degree *Philosophiae Doctor (Ph.D.)*. The work was carried out at the Department of Informatics, University of Oslo, in the digital signal processing and image analysis (DSB) research group. The thesis has been supervised by Professor Anne H. Schistad Solberg, Associate Professor Andreas Austeng, and Associate Professor Roy Edgar Hansen. During this period I have been on a partial leave of absence from the Norwegian Defence Research Establishment (FFI). The work was funded by the Research Council of Norway (NFR) through a Ph.D. grant from the project Acoustic Image Formation and Interpretation (ACIM).

Acknowledgements

First, I would like to thank my supervisors at the University of Oslo (UiO), Anne H. S. Solberg, Andreas Austeng, and Roy Edgar Hansen, for their guidance and encouragements throughout this Ph.D. project. With their great knowledge and open attitude, the faculty of the DSB group have created a fertile, cross-disciplinary environment in which I have thrived very much. Special thanks to Anne, my main supervisor, who has patiently trusted me to work in a somewhat long-winded, meandering fashion. I would also like to thank my two other co-authors, professor Fritz Albregtsen, who contributed important insights about texture operators and unbiased classifier evaluation in Paper III, and senior scientist Ellen Johanne Eidem of FFI, with whom I enjoyed a very good collaboration on Paper I. It was Ellen, along with Per Eidsvig and Dr. Stein Nilsen, who pointed me in the direction of the Ph.D. position and generally encouraged me to do research. I am also grateful to Dr. Atle Ommundsen, who suggested I remain at FFI in a part-time position, which made it easier for me to engage in Ph.D. studies at this point in life. Atle has also helped me by reviewing manuscripts.

The data for this project were provided by the hydrography group of the FFI Maritime systems division; in particular, I would like to thank Arnfinn Karlsen, Bård Henriksen, Kjersti Hovemoen, and Christian Ugelstad. Also, on several occasions, I have had the chance to discuss echo sounder technology with scientists and engineers of Kongsberg Maritime AS (Horten, Norway), and in particular I am grateful to Håvard Nes, Ragnar Eckhoff, and Berit Horvei. I have enjoyed many fruitful discussions with fellow Ph.D. students and post-docs, especially my office mates Angélique Berthelot, Andreas Kleppe, Vikash Pandey, and Wei Zhang. Are Fritz Charles Jensen provided Matlab code for the function approximation method used in Paper II, and Carl-Inge Colombo Nilsen shared with us his deep knowledge of array signal processing.

I think it is also appropriate to acknowledge all those whose scientific work I have built on, and I hope the bibliography properly reflects that. Here I would just like to point out one particularly useful work, namely the excellent monograph by Jackson and Richardson [1]. My sincere thanks also to Prof. Dr. Matthias Troyer and the Atelier Berann (Lans, Austria), for the kind permission to reproduce H. C. Berann's World Ocean Map in Ch.1. Finally, my heartfelt thanks to my parents, Christin and Arne, and my wife, Elina, for their invaluable help, support, care, and encouragements.

Abstract

The nature of marine sediments reflects past and present physical processes on the seafloor, and it is a critical factor for benthic ecology, shallow water acoustics, and oceanic geotechnical engineering. The aim of this thesis is an accurate, high-resolution acoustic remote sensing technique for seabed classification. For this purpose, single-beam and multibeam echo sounder (SBES and MBES) datasets from the northern North Sea were analyzed along with supporting seismo-acoustic and sedimentological data. The research is presented in three scientific papers.

A reference point is provided in Paper I, which presents a new regional surficial seabed sediment map of the study area. The map was obtained with an industry-strength statistical seabed classification method, QTC, applied to SBES data. The theoretical basis of the QTC algorithm is examined, and the classification results found to be largely consistent with supporting data. Of particular note are two erosional zones: 1) the western slope of the Norwegian Channel has sandy sediments with elongated pockmarks resulting from fluid seepage and ocean current erosion; 2) the eastern near-shore area around the Sognefjorden and Fensfjorden inlets has a heterogeneous mix of sediments as a result of glacial erosion, with a surface imprint revealing the Weichselian drainage of ice from the fjords into the Norwegian Channel Ice Stream.

For Paper II, MBES signal processing was implemented for bathymetric mapping and estimation of the incidence angle-dependent scattering cross section, $\sigma(\theta)$, whose characteristics depend on seabed composition. The σ-function is represented by an adaptive, piecewise constant function approximation, obtained after a triple averaging of samples in the time, space, and angular domains. The result is a random feature vector whose probability density function (PDF) is approximately multivariate Gaussian. The PDF is estimated using training data near sediment sampling sites, and Bayesian decision theory is invoked to classify seabed types. A statistical model is developed to compute theoretical error bounds as a function of sediment scattering characteristics and experimental parameters. Theoretical estimates and experimental cross validation results show that good classification accuracy and spatial resolution are possible even with limited training data.

The geomorphology of the seafloor is a result of erosional and depositional processes. Acoustical classification methods may therefore be extended by considering terrain features. A set of invariant features, based on local distributions of intrinsic geometric properties, is proposed and tested in Paper III. The features are sensitive to noise in MBES data, and a noise filter is therefore developed. The filter is based on a composite image transform consisting of a derivative operator plus a fast, discrete Radon transform (DRT). The derivative operator is implemented by circular convolution with a Laplacian point spread function modified to make the operator invertible. An iterative inverse DRT is proposed and shown numerically to converge faster than, e.g., solutions based on the normal equations. The DRT-based filter significantly improved image quality and terrain classification accuracy.

Further results derived from Paper III include local DRT filters applied to fault line detection in seismic images, invertible gradient operators, and enhancement of synthetic aperture radar images. Natural extensions of Paper II include non-linear Bayesian inversion, non-supervised learning, improvements in accuracy and spatial resolution, and improved estimation of $\sigma(\theta)$.

Publications

I. Ellen Johanne Eidem and Knut Landmark, "Acoustic seabed classification using QTC IMPACT on single-beam echo sounder data from the Norwegian Channel, northern North Sea", *Continental Shelf Research*, vol. 68, pp. 1–14, Oct. 2013

II. Knut Landmark, Anne H. Schistad Solberg, Andreas Austeng, and Roy Edgar Hansen, "Bayesian seabed classification using angle-dependent backscatter data from multibeam echo sounders", *IEEE Journal of Oceanic Engineering*, vol. 39, no. 4, pp. 724–739, Oct. 2014

III. Knut Landmark, Anne H. Schistad Solberg, Fritz Albregtsen, Andreas Austeng, and Roy Edgar Hansen, "A Radon-transform-based image noise filter — with applications to multibeam bathymetry", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 11, pp. 6252–6273, Nov. 2015

Related publications

I. Knut Landmark, Anne H. Schistad Solberg, Carl-Inge Colombo Nilsen, Andreas Austeng, and Roy Edgar Hansen, "Bayesian seabed classification using angle-dependent backscatter data from multibeam sonar", *11th European Conference on Underwater Acoustics (ECUA) 2012*, vol. 1, pp. 417–424, Institute of Acoustics (IOA), 2012

II. Knut Landmark and Anne H. Schistad Solberg, "Reducing scalloping in synthetic aperture radar images using a composite image transform", Proc. SPIE 9643, Image and Signal Processing for Remote Sensing XXI, 96431B (October 15, 2015); doi:10.1117/12.219 4952, http://dx.doi.org/10.1117/12.2194952

Contents

List of Figures

List of Tables

Notation and conventions

Symbol	Units	Description
$*$	–	convolution (binary operator)
$\oplus$	–	direct sum (binary operator)
$\lVert \cdot \rVert$	–	vector or matrix norm
$\star$	–	element-wise multiplication (binary operator)
$./$	–	element-wise division (binary operator)
$\bar{z}$	–	complex conjugate of z
$\mathbf{1}_A$	–	indicator function (characteristic function) for the set A
$\langle \cdot \rangle$	–	average
α	dB/m	attenuation coefficient
$\partial\Omega$	–	the boundary of set Ω
$\zeta(\boldsymbol{r})$	m	relief function
η	s/m	ray constant
γ	–	Euler-Mascheroni constant
$\gamma_{\boldsymbol{u}}$	–	acoustic ray (parametric curve)
Γ	–	resolution cell, ensonified area
$\delta(x)$	–	Dirac delta function; for $\boldsymbol{x} \in \mathbb{R}^n$, $\delta(\boldsymbol{x}) \equiv \delta(x_1)\cdots\delta(x_n)$
Δ	–	Laplace operator
$\mathrm{E}[\cdot]$	–	expectation value
θ	rad	angle of incidence
λ	m	wavelength
$\lambda(\omega_i \mid \omega_j)$	–	loss function
$\boldsymbol{\mu}$	–	class mean vector
μ_M	–	Mth order statistical moment
ν	Hz	frequency
Ξ	–	scintillation index
$\Pi(\boldsymbol{r}, t)$	Pa	complex acoustic pressure field
Π_s, Π_i	Pa	complex pressure of scattered and incident fields
ρ	kg/m^3	mass density, bulk density
ρ	m	horizontal range, radius in cylindrical coordinates
$\rho(\mathcal{A})$	–	spectral radius of linear operator $\mathcal{A}$
σ	–	scattering cross section
$\boldsymbol{\Sigma}$	–	covariance matrix
τ	s	pulse length, travel time
$\chi^2(K)$	–	χ^2 distribution with K degrees of freedom
Ψ_{RX}	rad	3 dB beamwidth, receiver array
Ψ_{TX}	rad	3 dB beamwidth, transmitter array
ϕ	–	Krumbein ϕ-scale
$\varphi_{\boldsymbol{X}}$	–	characteristic function for random vector X
ω	Hz	angular frequency
ω_j	–	class label
$\Omega(\boldsymbol{r}, \boldsymbol{r}_0)$	–	ensonification function for array phase centre $\boldsymbol{r}_0$

$\mathcal{B}$	—	approximate GDB inverse transform
$B^D(\boldsymbol{x}; r)$	—	D-dimensional ball, $\left\{ \boldsymbol{y} \in \mathbb{R}^D \mid \|\boldsymbol{y} - \boldsymbol{x}\| \le r \right\}$
c	—	number of classes
c	m/s	speed of sound
$C^k(\cdot)$	—	the set of functions that are k times differentiable
C^n	—	n-dimensional unit cylinder
d	m	lattice spacing (array)
D	m	particle diameter, water depth
Df	—	differential map of $f : \mathbb{R}^n \to \mathbb{R}^m$
$D_n(h, s)$	—	graph with n nodes, intercept h, and rise parameter s (GDB transform)
D_{ref}	m	reference depth
$\mathcal{F}[\cdot]$	—	Fourier transform (discrete and continuous)
D_c	—	Dunn index for c clusters
DB_c	—	Davies-Bouldin index for c clusters
$\boldsymbol{G}$	—	gray level co-occurrence matrix (GLCM)
$\Im f$	—	imaginary part of f
$g_i(\boldsymbol{y})$	—	discriminant function for class ω_i
h	—	intercept parameter (GDB transform)
H	nat	entropy (information theoretic)
H	m^{-1}	mean curvature
$\mathcal{H}[\cdot]$	—	Hilbert transform
$H(\boldsymbol{\theta}, s)$	—	the hyperplane perpendicular to $\boldsymbol{\theta}$, with distance s to the origin
I^α	—	Riesz potential
$\mathcal{I}(\cdot)$	—	range (image) of operator
ICV	—	inverse coefficient of variation
$\boldsymbol{J}$	—	Jacobian matrix
$\boldsymbol{k}$	m^{-1}	wave vector
k_r	m^{-1}	radial wave number
K	m^{-2}	Gaussian curvature
$\mathcal{K}_k(\mathcal{A}, b)$	—	the Krylov space, $\mathcal{K}_k(\mathcal{A}, b) = \text{span}\left\{ b, \mathcal{A}b, \dots, \mathcal{A}^{k-1}b \right\}$
$\mathcal{L}$	—	modified Laplace operator
$L(\boldsymbol{m})$	—	likelihood function for model parameter vector $\boldsymbol{m}$
$\boldsymbol{m}$	—	model parameter vector
$\mathcal{M}$	—	median filter
MA	—	mapping accuracy
M_z	—	(graphic) mean grain size (ϕ-scale)
$\boldsymbol{n}$	—	unit surface normal vector
N	—	image size, i.e., $N \times N$ pixels
$\mathcal{N}(\cdot)$	—	null space (kernel) of operator
$N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	—	normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix Σ
$\mathcal{P}$	—	Cauchy principal value
$p_Y(y),\ f_Y(y)$	—	probability density function for random variable Y
$p(\boldsymbol{d} \mid \omega)$	—	class-conditional probability density function for data $\boldsymbol{d}$
$p(\boldsymbol{r}, t)$	Pa	acoustic pressure field

$P(\omega)$	–	prior probability of class ω
$P(\omega \mid \boldsymbol{d})$	–	posterior probability for class ω given data $\boldsymbol{d}$
$Q(\boldsymbol{r}, \boldsymbol{r}_0)$	m^{-2}	propagation factor
$\boldsymbol{r}$	m	position vector
R	m	slant range
$\mathcal{R}$	–	Radon transform
$\mathcal{R}^i$	–	GDB (DRT) transform transform for ith quadrant
$\mathcal{R}^+$	–	adjoint Radon transform (backprojection operator)
$\mathbb{R}^+$	–	positive real numbers
$\Re f$	–	real part of f
s	–	average scattering cross section, rise parameter (GDB transform)
S	ppt	salinity
s_0	$\mathrm{m} \cdot \mathrm{Pa}$	source strength
S^n	–	the unit n-sphere, $\{\, \boldsymbol{x} \in \mathbb{R}^{n+1} \mid \|\boldsymbol{x}\| = 1 \,\}$
S_b	dB	seabed scattering strength
S_c	–	silhouette index for c clusters
s_r	V/Pa	receiver sensitivity
t	s	time
T	–	transmission coefficient (water-seabed interface)
T	°C	temperature
$\mathbb{T}^d$	–	the torus $\mathbb{R}^d / \mathbb{Z}^d$
$T_k(x)$	–	Chebyshev polynomial of the first kind on the interval $[-1, 1]$
$\mathrm{Vol}\,(\cdot)$	–	volume
V	–	reflection coefficient (water-seabed interface)
$\mathrm{Var}[\cdot]$	–	variance
$w(\boldsymbol{r})$	–	aperture function
$W(\boldsymbol{k})$	m^4	2D roughness spectrum
$W(\boldsymbol{k})$	m^3	aperture smoothing function
y	Pa^2	square modulus of complex pressure, $y = |\Pi|^2$
z	m	depth
z_0	m	acoustic source depth

The preferred convention for the the Fourier transform of $f : \mathbb{R}^n \to \mathbb{C}$ is here

$$\hat{f}(\boldsymbol{\xi}) \equiv \mathcal{F}\,[f](\boldsymbol{\xi}) = \int_{\mathbb{R}^n} f(\boldsymbol{x}) e^{-2\pi i \boldsymbol{x} \cdot \boldsymbol{\xi}}\,\mathrm{d}\boldsymbol{x}$$

$$f(\boldsymbol{x}) = \int_{\mathbb{R}^n} \hat{f}(\boldsymbol{\xi}) e^{2\pi i \boldsymbol{x} \cdot \boldsymbol{\xi}}\,\mathrm{d}\boldsymbol{\xi}.$$

However, in some instances where it is natural and very common to use angular frequency and wave numbers (Appendix A, in particular), we have used the alternative convention

$$\hat{f}(\boldsymbol{\omega}) \equiv \mathcal{F}\,[f](\boldsymbol{\omega}) = \int_{\mathbb{R}^n} f(\boldsymbol{x}) e^{-i\boldsymbol{\omega} \cdot \boldsymbol{x}}\,\mathrm{d}\boldsymbol{x}$$

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{f}(\boldsymbol{\omega}) e^{i\boldsymbol{\omega} \cdot \boldsymbol{x}}\,\mathrm{d}\boldsymbol{\omega}.$$

Acronyms

ANN artificial neural network

ARA angular range analysis

AVO amplitude versus offset

ASC acoustic seabed classification

ACIM Acoustic Image Formation and Interpretation

BIC Bayesian information criterion

CDF cumulative distribution function

CLT central limit theorem

CTD conductivity, temperature, and depth

DAS delay-and-sum (beamforming)

DFT discrete Fourier transform

DRT discrete Radon transform

DTM digital terrain model

ECDF empirical cumulative distribution function

EM expectation-maximization

ESA European Space Agency

FFI Norwegian Defence Research Establishment

GDB Götz, Druckmüller, and Brady (transform)

GLCM gray level co-occurrence matrix

GMRES Generalized Minimum Residual (iterative method)

GSAB Generic Seafloor Acoustic Backscatter

HRP heave, roll, and pitch

KSAT Kongsberg Satellite Services

LBP local binary pattern

LDC linear discriminant classifier

LHS left hand side (of equation)

LTI linear time-invariant

MAP maximum a posteriori estimate

MRA main response axis

MB multibeam bathymetry

MBES multibeam echo sounder

ML maximum likelihood

MSP mean square pressure

NC Norwegian Channel (Norwegian Trench)

NGU Geological Survey of Norway

PCA principal component analysis

PDF probability density function

PPD posterior probability density

PSF point spread function

QDC quadratic discriminant classifier

QTC Quester Tangent Corporation

RHS right hand side (of equation)

RMS root mean square

RX receiver (array)

SAR synthetic aperture radar

SBES single-beam echo sounder

SI scintillation index

SOM self-organizing map

SSP sound speed profile

SVM support vector machine

TX transmitter (array)

UDC uncorrelated discriminant classifier

UiB University of Bergen

UiO University of Oslo

WKBJ Wentzel-Kramers-Brillouin-Jeffreys

Chapter 1

Introduction

1.1   Motivation

To the modern marine science professional, a pause to look at the big picture brings consolation and relief (Fig. 1.1). The mapping of the topography (bathymetry) and geology of the world's ocean floors was a crucial step towards the modern understanding of plate tectonics and continental drift [2]. Data about the ocean floors yield important insights about Earth's geological history as well as present marine biological and physical processes. This is no less true at the regional and local scales, in particular on the continental shelves, which are subject to political dispute and the conflicting interests of resource development and nature conservation. Seabed mapping is important for marine resource exploration, for understanding benthic marine life, and for monitoring the impact of human activity. The physical character of the seabed largely determines what kinds of organisms can exist there, and consequently what kinds of fish species will feed there. Knowledge about the physical character of the seabed is also needed in numerous practical applications, ranging from offshore construction to fishing and military use of sonar. It seems likely that seabed mapping and monitoring will remain important activities, both commercial, political, and scientific.

Since the publication of the World ocean map (Fig. 1.1), Earth science has been transformed by the new technologies of digital information processing. This multi-faceted technological revolution involves new remote sensing systems, high-performance computing, and theoretical advances in signal processing and data analysis. Remote sensing techniques in earth science include the imaging of the Earth's surface, oceans, and crust by electromagnetic and seismo-acoustic waves. The image formation process is followed interpretation, to extract the useful information from the data. Observational science is increasingly data intensive because instrumentation evolves towards higher spatial resolution and sampling rates. A consequence is that interpretation and classification of data must to a greater extent be done automatically using computer algorithms. New methods are being developed for extracting more information from observations. In particular, machine learning [3] and inverse theory [4] provide quite general theoretical frameworks and algorithms for making predictions and parameter estimates based on observed data.

Presently, a large volume of high-resolution, continuous-coverage acoustic data is being generated by modern swath sonars, including multibeam echo sounders (MBESs) [5, Ch. 8],
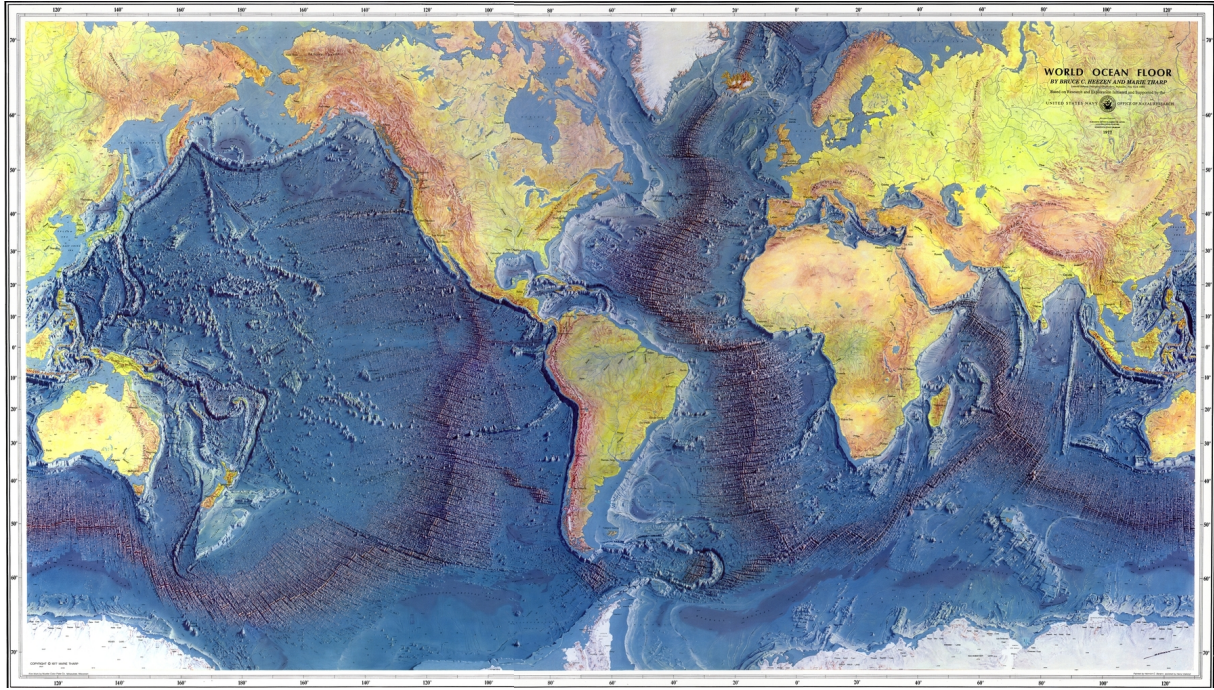
Fig. 1.1: World Ocean Floor relief map (1977), painted by Austrian artist and cartographer H. C. Berann, in collaboration with American geologists M. Tharp and B. Heezen of the Lamont Geological Laboratory in New York. The discovery of the volcanically active mid-ocean ridges provided an important piece of evidence for the theory of plate tectonics and continental drift, which by the 1960s was emerging as a unified explanation of a range of phenomena observed in new data, including paleomagnetics, gravimetry, sedimentology, and bathymetry [2]. Reproduced with the permission of Atelier Berann, Lans, Austria.

for both commercial, governmental, and scientific purposes. Echo sounding data have primarily been used for making navigation charts and digital terrain (water depth) models, as well as derived attribute images that are useful for geological interpretation. Sediment maps have mainly been based on sparse sedimentological data, i.e., from physical samples. The purpose of acoustic seabed classification (ASC) and characterization is to extract from sonar data additional information about the physical composition and geoacoustic properties of the seabed. To the extent that this approach is successful, it provides a faster, more cost-effective alternative or supplement to classical seabed mapping of extended areas, with a spatial resolution that is closer to the natural spatial variability of the seabed.

1.2 Objectives

The principal aim of this thesis is to develop an accurate, high-resolution acoustical remote sensing technique for seabed classification. The emphasis has been on machine learning (statistical classification) methods applied to echo sounder data. The research was carried out as part of the project Acoustic Image Formation and Interpretation (ACIM), whose main aim has been to develop new techniques for seismic and sonar imaging and image analysis. The approach is cross-disciplinary, and the project should demonstrate the benefits of combining theory and methods from, e.g., statistical classification theory, seismo-acoustic signal processing, and im-

age analysis.

The thesis is based on three research papers, referred to as Paper I, II, and III. The dual objectives of Paper I were 1) to perform a theoretical and experimental evaluation of a state-of-the-art ASC algorithm, and 2) apply it to produce a new regional seabed sediment map of a part of the northern North Sea (Sec. 1.4, Fig. 1.2). The purpose was also to present new geophysical, sedimentological, and bathymetric data of the study area, and the analysis of these datasets is the basis for the interpretation and evaluation of the ASC results. Furthermore, these data provide context and ground truth for Paper II, where a new supervised statistical seabed classification method for MBES data is presented. This method is our proposed resolution of the thesis' main objective. The purpose of Paper II is also to examine, theoretically and experimentally, the method's practical applicability, i.e., how survey parameters influence the performance (classification accuracy) of the algorithm. Paper III was conceived as an extension of Paper II, by considering additional features for seabed classification. The emphasis is, however, on image processing. In particular, the main objective of Paper III is to develop a denoising algorithm for improving data quality and thereby classification accuracy. The effect of noise reduction on classification accuracy is considered. The theory developed in Paper III is not specific to the ASC problem, and some possible other applications are discussed in Sec. 5.1.

1.3   The classification problem

The acoustic seabed classification (machine learning) problem can be briefly stated as follows. Sound waves generated by the echo sounder propagate through the water column and interact with the seabed. The interaction creates incoherent scattered and coherent reflected fields that are sampled in space and time by the echo sounder receiver array. Acoustic scattering and reflection depends on the physical properties of the seabed. A digitized received signal is, mathematically, a finite-dimensional vector, $x \in \mathbb{R}^D$, containing samples in space or time. Each signal can be mapped to a region on the seafloor. The signals are normalized by applying corrections that compensate for extraneous factors that influence signals independently of seabed properties. Such factors include propagation loss, sonar characteristics, and measurement geometry. A key step is feature extraction, which may be viewed as a transformation $\varphi : \mathbb{R}^D \to \mathbb{R}^d$ such that $\phi(x)$, the feature vector, contains the relevant and non-redundant information in x, and where, usually, $d < D$ and as small as possible to ensure statistical robustness. The choice of feature extraction function ϕ should yield a high classification accuracy. This effectively means that different seabed types should form separable clusters in the feature space $\mathbb{R}^d$. Moreover, feature extraction should preferably not be tailored to a specific dataset or sonar system, but rather be based on basic physical principles (generalizability). A range of available mathematical algorithms, known as classifiers, may subsequently be applied to form a partition of the feature space $\mathbb{R}^d$. In supervised learning, the partition is computed based on training data, i.e., data where the class is known (labeled data). This is the resolution of the classification problem. Any new observation x is assigned to the class corresponding to the partition in which the feature vector $\varphi(x)$ lies.
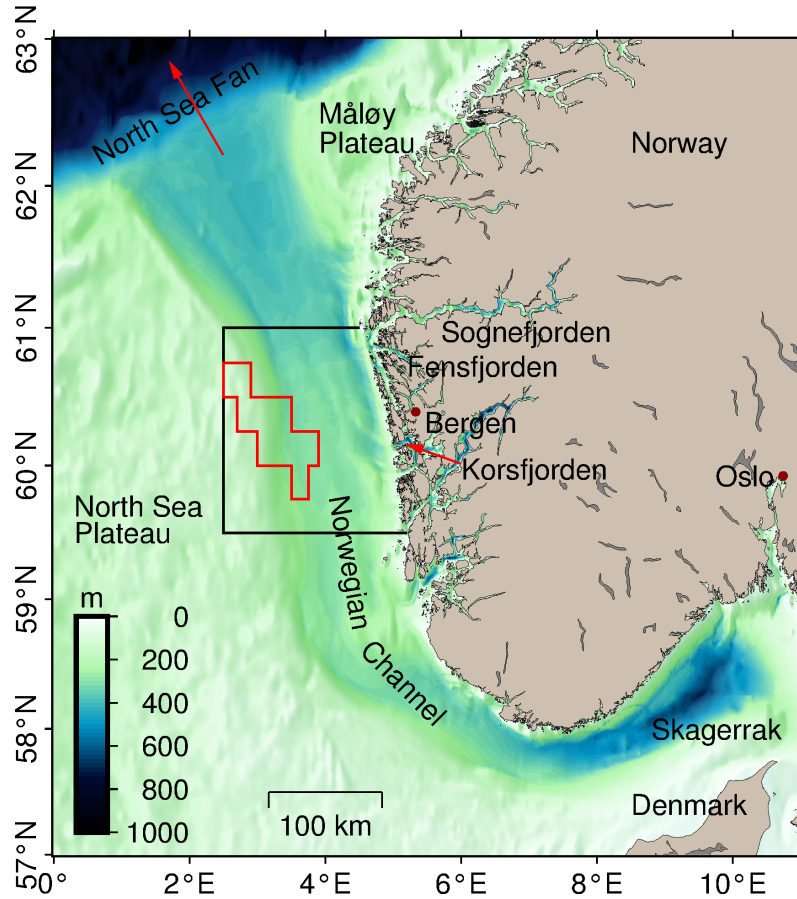
Fig. 1.2: Overview and bathymetry of the Norwegian Channel in the North Sea. The black rectangle outlines the study area of the SBES-based seabed classification work described in Paper I [10]. The red polygon contained in the rectangle shows the extent of the MBES dataset on which Papers II and III were based. The MBES dataset contains data from two different instruments (Kongsberg Maritime EM 1002 and EM 710).

1.4   Study area

To work on the seabed classification problem, MBES data from the northern North Sea, with supporting data, were obtained from the Norwegian Defence Research Establishment (FFI) (Kjeller/Horten, Norway). The study area covers a part of the Norwegian Channel (Norwegian Trench) off the coast of western Norway, and a part of the eastern margin of the shallow North Sea Plateau. The Norwegian Channel (NC), a major topographic feature of the North Sea, is a depression in the continental shelf along the coast of southern Norway, with depths of about 300 m in the study area (Fig. 1.2) [6]. The western slope of the NC is one of the major pathways for inflowing North Atlantic water, whereas outflowing mixed oceanic and fresh water in the eastern and central parts of the channel makes up the northbound Norwegian Coastal Current [7]. The topography of the NC therefore influences the circulation pattern of the North Sea, and, conversely, the NC is an important trap for water-borne sediments in the North Sea [8]. The shallow geology of the NC is a result of a series of alternating glaciations and inundations during the Quaternary, and both marine, glacio-marine, and glacigenic sediments are found in the area [9], with basement outcrops near the Norwegian shore.

4

1.5 Structure of the thesis

The thesis is organized in two parts: the three papers are reprinted in the second part, while the first part, Ch. 1–6, is intended to introduce the work in the papers and their common themes, providing motivation, foundational material, and overview. Chapter 2 explains some basic theory and concepts from echo sounding, statistical classification theory, and digital image analysis. The exposition is limited to material on which the papers are directly based, foundations for further work described in Ch. 5, and material used in the data preparation workflow. The latter is generally not described in detail in the papers. Chapter 3 is a brief introduction to seabed classification: physical sediment characterization, different types of features and algorithms, and current work in the field. Chapter 4 introduces the three papers and summarizes the main results and developments. Several related topics that were explored to varying degrees in the course of the project are briefly discussed in Ch. 5, where some directions for further work are pointed out. Chapter 6 is the conclusion of this work.

Chapter 2

Foundations

This chapter, with Appendix A–C, provides background material on theory, measurement principles, and data preparation for the algorithms considered in Paper I–III. The chapter is also intended as a reference for further work outlined in Ch. 5. The material is standard for the most part, although the derivations are our own (except Appendix A, which is textbook based).

Section 2.1 reviews basic definitions, physics, and signal processing principles of echo sounding, with emphasis on 1) data preparation for classification algorithms and 2) the random nature of MBES data. Section 2.1.1–2.1.3 plus appendices describe the basic principles of multibeam sonars, and the corrections required to estimate seabed scattering strength from recorded acoustic pressure data. This is key to the method proposed in Paper II, and also for further work on parameter estimation (Sec. 5.3). The basis for the statistical model of MBES data used in Paper II is described in Sec. 2.1.4. As suggested in Sec. 5.3, this theory could also be employed in geoacoustic parameter estimation. Section 2.2.1 and Sec. 2.2.2 review the classification theory used in Paper II and Paper I, respectively. Section 2.3 presents the basis for the forward and inverse composite image transforms developed in Paper III as part of a data enhancement algorithm.

2.1 Echo sounding

Echo sounding follows the same basic principles as electromagnetic imaging and remote sensing. A wavefield is created (emitted) by an array of controlled sources, the waves propagate through a medium (water) and interact with an object of interest (the seabed). The interaction creates scattered and reflected fields that are sampled in time and space by an array of receivers. Array signal processing [11] is applied to the recorded signals (echoes) to locate or image the objects of interest. The received signals depend on the physical properties of the objects, and ASC is about extracting this information.

An echo time series depends on intrinsic properties of the seabed (roughness and volume scattering strength, reflection coefficient, sound attenuation), but also extraneous factors (source strength, receiver sensitivity, pulse length, frequency, beamwidth, propagation loss, water depth, and seabed slope). For seabed classification and characterization, the influence of extraneous factors should be reduced as much as possible.

2.1.1 Radiation and directivity

Whereas point-like or spherical acoustic sources radiate isotropically (in an unbounded medium), the radiation pattern of a transducer element may be highly directive depending on its shape, size, and emitted wavelengths. For instance, the energy flux from a vibrating piston element in a plane baffle will peak around the axis perpendicular to the plane. The larger the element surface, the more directive the emitter is. A large emitter or receiver can be formed by combining several transducer elements into an array (antenna). This introduces extra degrees of freedom, as the phase and source/receiver strength of each element can be adjusted individually so as to alter the array's main response axis (MRA) and beam pattern (beam steering). The greater the physical extent of the array, the more directive it is. However, in analogy with digital sampling of time-varying signals, the element spacing, D, must be sufficiently small so as to avoid spatial aliasing, i.e., $D < \lambda/2$ for linear arrays [11], where λ is the wavelength.

For a receiver array, from the signal processing viewpoint, there is an unknown field $p(\boldsymbol{x}, t)$ that one wishes to obtain information about by sampling the field in space and time. The geometry and response of the array is encoded in the aperture function $w \colon \mathbb{R}^3 \to [0, 1]$, which is defined such that the field observed at position $\boldsymbol{r}$ is [11, Sec. 3.1]

$$f(\boldsymbol{r}, t) = w(\boldsymbol{r})p(\boldsymbol{r}, t). \tag{2.1}$$

The aperture function w is zero where there is no sensor present. Letting w take values in the range $[0, 1]$ allows for a relative weighting of different parts of the aperture. Let $\hat{p}(\boldsymbol{\xi}, t)$ denote the 3D Fourier transform

$$\hat{p}(\boldsymbol{\xi}, t) = \int_{\mathbb{R}^3} p(\boldsymbol{r}, t)e^{-2\pi i \boldsymbol{\xi} \cdot \boldsymbol{r}} \, \mathrm{d}\boldsymbol{r}. \tag{2.2}$$

The observed signal integrated over space is

$$\begin{aligned}
s(t) &= \int_{\mathbb{R}^3} w(\boldsymbol{r})p(\boldsymbol{r}, t) \, \mathrm{d}\boldsymbol{r} \\
&= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} w(\boldsymbol{r})\hat{p}(\boldsymbol{\xi}, t)e^{2\pi i \boldsymbol{\xi} \cdot \boldsymbol{r}} \, \mathrm{d}\boldsymbol{\xi} \, \mathrm{d}\boldsymbol{r}.
\end{aligned}$$

This may be rewritten as the scalar product

$$s(t) = \langle \hat{p}(\boldsymbol{\xi}, t), \hat{w}(\boldsymbol{\xi}) \rangle = \int_{\mathbb{R}^3} \hat{p}(\boldsymbol{\xi}, t)\overline{\hat{w}}(\boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi} = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \hat{p}\left(\boldsymbol{k}/2\pi, t\right) W(\boldsymbol{k}) \, \mathrm{d}\boldsymbol{k}, \tag{2.3}$$

where $\boldsymbol{k} = 2\pi\boldsymbol{\xi}$ is the wave vector, $\overline{W}$ denotes the complex conjugate, and

$$W(\boldsymbol{k}) \equiv \overline{\hat{w}}(\boldsymbol{k}/2\pi) = \int_{\mathbb{R}^3} w(\boldsymbol{r})e^{i\boldsymbol{k} \cdot \boldsymbol{r}} \, \mathrm{d}\boldsymbol{r} \tag{2.4}$$

is the aperture smoothing function. The integrated signal is seen to be a weighted average of $\hat{p}$ over wavenumber space (2.3) with respect to $W(\boldsymbol{k})$, also termed the array pattern. $W(\boldsymbol{k})$ determines the directivity of the array. For a uniform rectangular transducer element in the x-y plane, with area $L_x \times L_y$, the aperture function is

$$w(\boldsymbol{r}) = h(x, y)\delta(z) \tag{2.5a}$$

$$h(x, y) = \begin{cases} 1 & \text{if } -L_x/2 \le x \le L_x/2 \text{ and } -L_y/2 \le y \le L_y/2 \\ 0 & \text{otherwise.} \end{cases} \tag{2.5b}$$

The well-known array pattern for this aperture is the product of two sinc functions since

$$W(\boldsymbol{k}) = \int_{\mathbb{R}^3} h(x, y)\delta(z)e^{i(k_x x + k_y y + k_z z)} \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z \tag{2.6}$$

$$= \int_{-L_x/2}^{L_x/2} e^{ik_x x} \, \mathrm{d}x \int_{-L_y/2}^{L_y/2} e^{ik_y y} \, \mathrm{d}y = \left[\frac{\sin k_x L_x/2}{k_x/2} \right] \left[\frac{\sin k_y L_y/2}{k_y/2} \right]. \tag{2.7}$$

A composite array formed by M non-overlapping, translated copies of a basic aperture element, $w_0(\boldsymbol{r})$, has aperture function

$$w(\boldsymbol{r}) = \sum_{n=0}^{M-1} v_n w_0(\boldsymbol{r} - \boldsymbol{r}_n), \tag{2.8}$$

with weights $v_n \in [0, 1]$, $n = 1, \dots, M-1$. A time-domain, linear time-invariant (LTI) filter may be applied to each channel (aperture element), so that (2.1) becomes

$$f(t) = \sum_{n=0}^{M-1} v_n w_0(\boldsymbol{r} - \boldsymbol{r}_n) h_n(t) * p(\boldsymbol{r}, t) \tag{2.9}$$

where $*$ denotes convolution with respect to t. The far-field of an acoustic source may often, locally, be approximated by a plane wave. A plane wave is on the form

$$p(\boldsymbol{r}, t) = y(t - (\boldsymbol{u}/c) \cdot \boldsymbol{r}) \tag{2.10}$$

where $\boldsymbol{u}$ is a unit vector, c is the phase speed, and $y : \mathbb{R} \to \mathbb{R}$ is the pulse form. For the plane wave, $\hat{p}(\boldsymbol{\xi}, \nu) = \hat{y}(-\nu)\delta(\xi + (\nu/c)\boldsymbol{u})$. The signal observed with the composite array then becomes

$$s(t) = \int_{-\infty}^{\infty} \left[\sum_{n=0}^{M-1} \hat{h}_n(-\nu)v_n e^{2\pi i(\nu/c)\boldsymbol{u} \cdot \boldsymbol{r}_n} \times \int_{\mathbb{R}^3} w_0(\boldsymbol{\xi})e^{2\pi i(\nu/c)\boldsymbol{u} \cdot \boldsymbol{\xi}} \, \mathrm{d}\boldsymbol{\xi} \right] \hat{y}(\nu)e^{-2\pi i\nu t} \, \mathrm{d}\nu$$

$$= \int_{-\infty}^{\infty} [W_a(\boldsymbol{k}_0)W_0(\boldsymbol{k}_0)] \, \hat{y}(\nu)e^{-2\pi i\nu t} \, \mathrm{d}\nu, \tag{2.11}$$

where $W_a(\boldsymbol{k}_0) = \sum_{n=0}^{M-1} \hat{h}_n(-ck_0/2\pi)v_n \exp(i\boldsymbol{k}_0 \cdot \boldsymbol{r}_n)$, and $\boldsymbol{k_0} = (2\pi\nu/c)\boldsymbol{u}$ is the wave number of the plane wave (2.10). If we choose

$$\hat{h}_n(\nu) = |\hat{h}_n(\nu)|e^{2\pi i\nu\boldsymbol{u} \cdot \boldsymbol{r}_n/c}, \tag{2.12}$$

then W_a attains its maximum value, i.e., $W_a(\boldsymbol{k}_0) = \sum_{n=0}^{M-1} |\hat{h}_n|v_n$. The exponential factor in (2.12) corresponds to a time shift $\Delta t_n = \boldsymbol{u} \cdot \boldsymbol{r}_n/c$; it causes the wavefronts across the aperture to add constructively. This is delay-and-sum (DAS) beamforming, which makes it possible to electronically control the array's direction of maximum response. This is a central principle of MBES systems (Sec. 2.1.3). If the array elements form a regular, plane lattice with positions

$\{\,(md, nd) \mid m, n \in \mathbb{Z}\,\}$ (with zero-padding), the array smoothing function (setting $\hat{h}_n = 1$) becomes

$$W(\boldsymbol{k}) = W_0(\boldsymbol{k})W_a(k_x d, k_y d) = W_0(\boldsymbol{k}) \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} v_{mn} e^{id(k_x m + k_y n)},$$

which is a conjugate discrete-space Fourier transform modulated by the factor W_0, which represents the intrinsic directivity of each element. The weights v_{mn} may be chosen so as to control the sidelobe levels of the array. For example, one may take v_{mn} as samples of the product of two Dolph-Chebyshev window functions, in analogy with filter design in discrete-time signal processing.

To make a link with physics, it is necessary to compute the (far-) field emitted by an acoustic array. Anticipating the example of the EM 710, one of two MBES systems used in the present study, a plane array will be considered here. The motion of each transducer element is perpendicular to the plane; if the array surface lies in the x-y plane, the velocity of the array surface is given by a function on the form $\boldsymbol{u} = \begin{bmatrix} 0 & 0 & u(x, y, t) \end{bmatrix}^T$. It will be assumed that the plane surface on which the elements are mounted extends sufficiently far beyond the elements (i.e. several wavelengths), so that in calculations the plane surface can be taken as infinite (with the array enclosed in a bounded part of it) without incurring significant errors. With this simplification, the far field in a homogeneous medium with density ρ can be computed analytically using Green's functions [12, Sec. 7.4]. The result is (Appendix A)

$$\hat{p}(\boldsymbol{r}, \omega) \simeq -ik\rho c \frac{e^{ikr}}{2\pi r} \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{u}(x_0, y_0, \omega) e^{-i(k_x x_0 + k_y y_0)} \, \mathrm{d}x_0 \, \mathrm{d}y_0, \tag{2.13}$$

where $\hat{u}(x, y, \omega)$ is the temporal Fourier transform of $u(x, y, t)$ for angular frequency ω. Equation (2.13) should be compared with (2.1), (2.3), and (2.4): the array directivity is given by the spatial Fourier transform of either $\hat{u}(x, y, \omega)$ or the weight function w over the aperture. The Fourier transform of w is the array pattern $W(\boldsymbol{k})$ [$W(k_x, k_y)$ for a plane array].

2.1.2 Refraction and attenuation

Equation (2.13) explains the directivity of a plane array, and it is a formula for computing the acoustic far-field pressure in an iso-speed medium. For echo sounding, the speed of sound c is variable, but may be assumed to be a function of depth z only, for two reasons. The first, practical reason is that only the vertical sound speed profile (SSP) is known from measurements. SSPs are obtained from conductivity, temperature, and depth (CTD) probes intermittently lowered from the survey ship through the water column. The second reason is that, across the distances that echo sounder waves propagate, the ocean is approximately horizontally layered; in open sea areas away from oceanographic fronts, the ratio between the vertical and horizontal sound speed gradients is typically about 10^3 [13, Sec. 1.1]. For stratified media with $c = c(z)$, the solution of the Helmholtz equation may be written as a closed-form integral expression by means of the Hankel transform and 1D Green's function techniques [14, Sec. 2.4 & Ch. 4], [15, Ch. 4 & 6]. However, given the high frequency range and array beam processing of the MBES, it is more practical to use ray theory [14, Ch. 3], [16, App. E] to solve the homogeneous Helmholtz

equation (A.2). Subsequently, the iso-speed solution may be used to provide initial values for the ray equations, to properly account for directivity and source strength.

In first order ray theory (Appendix B), the solution for the pressure due to a point source is on the form

$$\hat{p}(\boldsymbol{r}, \omega) = A_0(\boldsymbol{r}) e^{i\omega\tau(\boldsymbol{r})}. \tag{2.14}$$

A solution to the ray equations (Eqn. B.9) defines a parameterized curve (ray) $\gamma_{\boldsymbol{u}} : \mathbb{R}^+ \to \mathbb{R}^3$, where $\boldsymbol{u} \in S^2$ (the unit sphere) specifies the initial direction of the ray. In MBES data processing, $\boldsymbol{u}$ is given by the MRA of a beam (Sec. 2.1.3). In the stratified, axially symmetric model, $\boldsymbol{u}$ may be specified by a single angle, e.g. ϑ, the initial ray angle with respect to the horizontal plane. Using cylindrical coordinates for γ, the ray equations may be written in the form we have used for numerical integration (Appendix B):

$$\frac{\mathrm{d}\rho}{\mathrm{d}z} = \frac{\eta c(z)}{\sqrt{1 - \eta^2 c(z)^2}} \tag{2.15a}$$

$$\frac{\mathrm{d}\tau}{\mathrm{d}z} = \frac{1}{c(z)\sqrt{1 - \eta^2 c^2(z)}} \tag{2.15b}$$

$$\frac{\mathrm{d}s}{\mathrm{d}\tau} = c(z) \tag{2.16}$$

where ρ is horizontal range, s is arc length, $\eta = \cos\vartheta / c(z_0)$, and z_0 is the source depth. The ray amplitude is (Eqn. B.19)

$$A_0(s) = A_0(0) \sqrt{\frac{c}{c(z_0)} \frac{\det \boldsymbol{J}_s(0)}{\det \boldsymbol{J}_s}} \tag{2.17}$$

where, in the stratified model,

$$\det \boldsymbol{J}_s = \rho \sqrt{\left(\frac{\partial z}{\partial \vartheta}\right)^2 + \left(\frac{\partial \rho}{\partial \vartheta}\right)^2}. \tag{2.18}$$

Equations (2.17)–(2.18) are the formulas used for computing the effect of geometrical spreading or focusing on the amplitude.

Volume attenuation due to viscous absorption and relaxation processes in the water column is accounted for by including an exponential damping factor

$$A_0(s) \to A_0(s) e^{-\int_0^s \alpha(s')/(20\log_{10} e)\,\mathrm{d}s'}, \tag{2.19}$$

where α is the attenuation coefficient in units of $\mathrm{dB\,m}^{-1}$. This attenuation coefficient may be computed using the empirical formulas of Francois and Garrison [17], [1, App. B], as a function of frequency ν [Hz], depth z [m], salinity S [ppt], temperature T [°C], and seawater pH:

$$\alpha = \frac{A_1 P_1 \nu_1 \nu^2}{\nu^2 + \nu_1^2} + \frac{A_2 P_2 \nu_2 \nu^2}{\nu^2 + \nu_2^2} + A_3 P_3 \nu^2 \tag{2.20}$$

11

where
$$\nu_1 = 2.8 \left[\frac{S}{35}\right]^{0.5} \times 10^{7-1245/(273+T)} \text{ and } \nu_2 = \frac{8.17 \times 10^{11-1990/(273+T)}}{1 + 0.0018(S - 35)}$$

are the relaxation frequencies of magnesium sulphate $MgSO_4$ and boric acid H_3BO_3;

$$P_2 = 1 - 1.37 \times 10^{-4}z + 6.2 \times 10^{-9}z^2 \text{ and } P_3 = 1 - 3.83 \times 10^{-5}z + 4.9 \times 10^{-10}z^2$$

account for the pressure dependence of absorption by $MgSO_4$ and viscosity, respectively;

$$A_1 = \frac{8.86}{c}10^{0.78\text{pH}-5}, \ A_2 = 21.44\frac{S}{c}(1 + 0.025T), \text{ and}$$

$$A_3 = \begin{cases} 4.937 \times 10^{-4} - 2.59 \times 10^{-5}T + 9.11 \times 10^{-7}T^2 - 1.50 \times 10^{-8}T^3 & \text{for } T \leq 20°\text{C} \\ 3.964 \times 10^{-4} - 1.146 \times 10^{-5}T + 1.45 \times^1 0^{-7}T^2 - 6.5 \times 10^{-10}T^3 & \text{for } T > 20°\text{C} \end{cases}$$

account for the temperature dependence of absorption by $MgSO_4$, H_3BO_3, and viscosity, respectively, and

$$c = 1412 + 3.21T + 1.19S + 0.0167z$$

is the calculated sound speed. The Francois-Garrison formula (2.20) is expected to be accurate in open sea conditions with $S \sim 30$ to 40 ppt. The typical pH range for sea water is 7.8 to 8.3. When pH is not available from measurements, we have used the constant value pH $= 8.0$.

2.1.3 Sonar configuration

Figure 2.1 shows a typical array configuration for an MBES system. In this example there are two separate rectangular transducer arrays, one for transmission (TX) and one for reception (RX). Both arrays are highly directive, with one dimension being much longer than the wavelength $\lambda = c/\nu$, where ν is the center frequency and c is the speed of sound in water. The Kongsberg Maritime EM 710, one of the MBES systems providing data for the present study, has a 1940×224 mm TX array and a 970×224 mm RX array, with center frequency 70–100 kHz, which corresponds to $\lambda \approx 15$–20 mm [18]. Fig. 2.1 shows one possible TX beam pattern, in the x-y half-plane, for the simplified limiting case of a line array ($L_y \to 0$) with isotropic (non-directive) elements. The line array pattern is rotationally symmetric about the x-axis. Hence the mainlobe emission of the TX array forms a narrow vertical fan perpendicular to the vessel's longitudinal axis (x-axis). This fan may be tilted a few degrees forward and backward (rotated about the y-axis) in quick succession by electronic beam steering; thereby, two close strips of seafloor may be ensonified within the recording time of a single ping.

Similarly, the RX array is sensitive in a narrow vertical fan parallel to the vessel's longitudinal axis (x-axis). By electronic beam steering, this fan may be tilted away from the vertical as shown in Fig. 2.1. For each ping the received data are processed a number of times at different steering angles to cover a wide angular sector athwartships, up to $140°$ and 400 angles for the EM 710, and $150°$ and 111 angles for the EM 1002. The intersection of the RX and TX mainlobe fans defines a beam, and taking into account travel time, sound refraction (ray bending) and the vessel's attitude (heave, roll, and pitch), the samples of a received echo are mapped to the seafloor, referred to a hull-fixed cartesian coordinate system. The soundings are
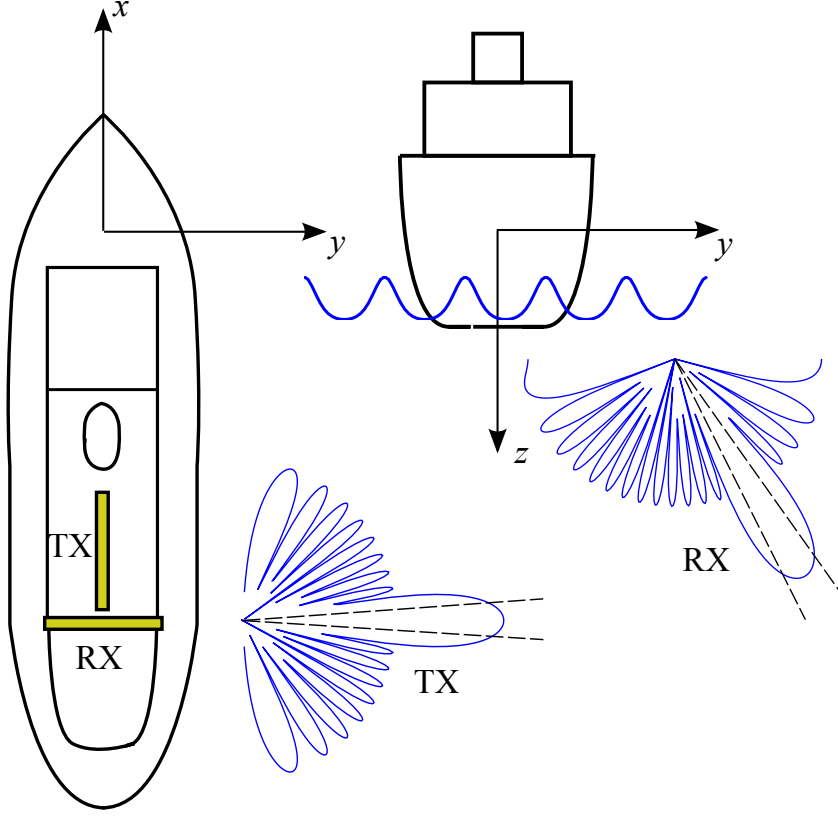
Fig. 2.1: Sketch of a multibeam echo sounder (MBES) system with a T-shaped array with separate rectangular transmit (TX) and receive (RX) parts. Excluding sidelobes, the TX array emits sound in a narrow vertical fan in the y-z-plane perpendicular to the vessel track. The TX beampattern is shown projected onto the x-y half-plane. The unsteered RX array is sensitive in a similar narrow fan in the x-z-plane. The RX array is steered at a range of angles away from the vertical. The intersection of the RX and TX fans forms a beam. The beampatterns shown here both have uniform -20 dB sidelobe levels and 3 dB beamwidth $\Psi_{RX} = \Psi_{TX} = 10°$; these angles are shown with dashed lines. The actual beamwidths of the EM 1002 are $\Psi_{RX} = \Psi_{TX} = 2.0°$, and the beamwidths of the EM 710 are $\Psi_{RX} = 1.0°$ and $\Psi_{TX} = 0.5°$, but we do not know the exact beampatterns.

subsequently positioned in an Earth-fixed spatial reference system using accurate differential GPS position data.

Figure 2.2 shows the MBES acquisition geometry in the vertical across-track plane. The MBES is downward-looking, highly directive, and short-pulsed, so multipath propagation is not considered. The position (phase center) of the array, $\boldsymbol{r}_0$, and a position on the seafloor, $\boldsymbol{r}$, define the direct eigenray connecting them. This ray defines a direction from the array, given, e.g., by the spherical angles ϑ and φ (Sec. 2.1.1). Using complex signal representation (Appendix C), the pressure field is

$$\Pi(\boldsymbol{r}, t) = p_a(\boldsymbol{r}, t)e^{-i\omega_c t}, \tag{2.21}$$

where ω_c is the carrier (angular) frequency of the signal, and p_a is the analytic signal (C.2), whose temporal Fourier transform is

$$\hat{p}_a(\boldsymbol{r}, \omega) = \hat{p}(\boldsymbol{r}, \omega) + \text{sgn}(\omega)\hat{p}(\boldsymbol{r}, \omega). \tag{2.22}$$
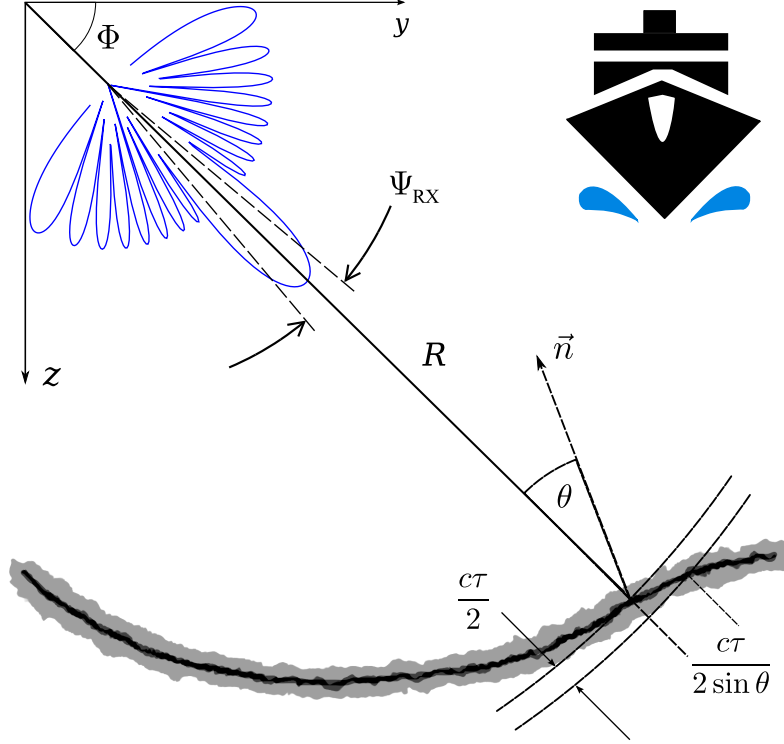
Fig. 2.2: MBES geometry in the across-track plane (from Paper II). The vessel moves along the x-axis. A single ping with pulse length τ ensonifies a strip of seafloor which is wide in y and narrow in x. The receiver array steers many beams at different depression angles, Φ, yielding backscatter response at different incidence angles θ. A resolution cell has a length of $c\tau/2\sin\theta$ (for gated CW pulses); the beam footprint length is approximated as $R\Psi_{\mathrm{RX}}$. The beampattern shown here, with $3\,\mathrm{dB}$ beamwidth $\Psi_{\mathrm{RX}} = 10°$ and uniform $-20\,\mathrm{dB}$ sidelobe levels, is only for illustration purposes; it is not the actual beampattern of the EM 1002 or EM 710.

From (2.13), the envelope of the incident complex pressure, Π_i, satisfies

$$|\Pi_i(\boldsymbol{r}, \omega; \boldsymbol{r}_0)| \propto W_{\mathrm{TX}}[\boldsymbol{k}(\boldsymbol{r}; \boldsymbol{r}_0, \omega)] \tag{2.23}$$

for $\omega > 0$, where $W_{\mathrm{TX}}(\boldsymbol{k})$ is the TX aperture smoothing function, conventionally normalized to unity on the MRA. The ensonified area, or resolution cell $\Gamma(t)$, is the patch of seafloor from which scattered acoustic energy reaches the receiver at time t. For a stationary signal, Γ is the beam footprint: the region where $W_{\mathrm{TX}}[\boldsymbol{k}(\boldsymbol{r}; \boldsymbol{r}_0, \omega)] \neq 0$. We write $y \equiv |\Pi|^2$ and define the time-averaged (mean) square pressure (MSP)

$$\langle y(\boldsymbol{r}) \rangle_T = \frac{1}{T} \int_0^T |\Pi(\boldsymbol{r}, t)|^2 \, \mathrm{d}t. \tag{2.24}$$

The source factor s_0^2 [19, Ch. 3] is defined such that

$$\langle y(\boldsymbol{r}) \rangle_T |\boldsymbol{r} - \boldsymbol{r}_0|^2 = s_0^2 |W_{\mathrm{TX}}[\boldsymbol{k}(\boldsymbol{r}, \boldsymbol{r}_0)]|^2 \tag{2.25}$$

in the far field under iso-speed (constant c) conditions. Transmission loss, including geometric

spreading and absorption, is represented by the propagation factor $Q(\boldsymbol{r}, \boldsymbol{r}_\text{ref})$ defined by

$$\langle y(\boldsymbol{r})\rangle_T = \langle y(\boldsymbol{r}_\text{ref})\rangle_T \left|\boldsymbol{r} - \boldsymbol{r}_\text{ref}^2\right|^2 Q(\boldsymbol{r}, \boldsymbol{r}_\text{ref}). \tag{2.26}$$

In practice $\boldsymbol{r}_\text{ref}$ is taken to be close to the source, e.g. 1 m, where absorption and refraction is negligible. However, one should bear in mind that the simple relationship (2.23) is a far-field approximation. Substituting (2.25) with $\boldsymbol{r}_0 = \boldsymbol{r}_\text{ref}$ into (2.26) yields

$$\langle y(\boldsymbol{r})\rangle_T = s_0^2 \left|W_\text{TX}[\boldsymbol{k}(\boldsymbol{r}, \boldsymbol{r}_0)]\right|^2 Q(\boldsymbol{r}, \boldsymbol{r}_0). \tag{2.27}$$

By the reciprocity principle, the received mean-square *incoherent* (scattered) pressure for co-located RX and TX arrays is

$$\langle y_s\rangle = \int_\Gamma \sigma(\theta) \left|\Pi_i(\boldsymbol{r}; \boldsymbol{r}_0)\right|^2 Q(\boldsymbol{r}, \boldsymbol{r}_0) \left|W[\boldsymbol{k}(\boldsymbol{r}, \boldsymbol{r}_0)]\right|^2 \mathrm{d}^2\boldsymbol{r}$$
$$= s_0^2 \int_\Gamma \sigma(\theta) Q(\boldsymbol{r}, \boldsymbol{r}_0)^2 \left|W_\text{TX}[\boldsymbol{k}(\boldsymbol{r}, \boldsymbol{r}_0)]\right|^2 \left|W_\text{RX}[\boldsymbol{k}(\boldsymbol{r}, \boldsymbol{r}_0)]\right|^2 \mathrm{d}^2\boldsymbol{r}. \tag{2.28}$$

This equation introduces the dimensionless scattering cross section $\sigma(\theta)$, which describes the incidence-angle-dependent incoherent scattering properties of the seabed (Sec. 2.1.4). In formal scattering theories, the average $\langle\cdots\rangle$ in (2.28) is over a hypothetical ensemble of data points obtained with identical measurement geometries over different realizations of a random medium. In practice we will average over data obtained from distinct ensonified areas all assumed to lie within a seabed region with stationary statistical properties (Sec. 2.1.4). Thus the average is over both time and consecutive pings, at a fixed incidence angle and corrected for propagation loss.

For a transient, gated CW pulse with length τ, the ensonified area Γ is limited to $c\tau/2\sin\theta$ in the across-track (y) direction, except near vertical incidence (Fig. 2.2). For typical short-pulse, narrow-beam MBES data we may as a good approximation consider θ and Q constant across Γ. The ensonification function Ω may be defined as

$$\Omega(\boldsymbol{r}, \boldsymbol{r}_0) = \mathbf{1}_\Gamma(\boldsymbol{r}) \left|W_\text{TX}[k(\boldsymbol{r}, \boldsymbol{r}_0)]\right|^2 \left|W_\text{RX}[\boldsymbol{k}(\boldsymbol{r}, \boldsymbol{r}_0)]\right|^2, \tag{2.29}$$

where $\mathbf{1}_\Gamma(\boldsymbol{r})$ is the characteristic (indicator) function of $\Gamma \subset \mathbb{R}^2$. Equation (2.28) may then be compactly written

$$\langle y_s\rangle = s_0^2 Q^2 \sigma(\theta) \int_{\mathbb{R}^2} \Omega(\boldsymbol{r}, \boldsymbol{r}_0) \mathrm{d}^2\boldsymbol{r}. \tag{2.30}$$

This equation is the basis for estimating the seabed scattering cross section, or the seabed scattering strength $S_b = 10\log_{10}\sigma$. The propagation factor is found using the ray methods briefly outlined in Sec. 2.1.2 and Appendix B. A simple approximation to the integral in (2.30) is

$$\int_{\mathbb{R}^2} \Omega(\boldsymbol{r}, \boldsymbol{r}_0) \mathrm{d}^2\boldsymbol{r} \approx R\Psi_t \min\left[R\Psi_r, c\tau/2\sin\theta\right], \tag{2.31}$$

where Ψ_TX and Ψ_RX are the TX and RX 3 dB beamwidths respectively. A better approximation may be obtained by assuming $\left|W_\text{TX}[k(\boldsymbol{r}, \boldsymbol{r}_0)]\right|^2 \left|W_\text{RX}[\boldsymbol{k}(\boldsymbol{r}, \boldsymbol{r}_0)]\right|^2$ is separable into a product of functions of a depression angle and an azimuthal angle respectively [1, App. G.2]. Yet another

approach is noted in Sec. 5.2, but not developed in detail in this thesis.

2.1.4   Scattering statistics

A high-resolution MBES emits ultrasound with wavelengths on the order of 1 cm. At this length scale, seabed sediments are heterogeneous; both the surface structure (roughness) and volume structure are irregular and practically impossible to measure or characterize in full detail. High-frequency seafloor acoustics therefore employs statistical measures to characterize sediment structure, and, likewise, the scattered acoustic field. The complex scattered acoustic field, Π_s, may be defined as the deviation from the complex mean field, i.e. [1, Sec. 2.3],

$$\Pi = \langle \Pi \rangle + \Pi_s. \tag{2.32}$$

The mean field $\langle \Pi \rangle$ is then attributed to the incident and coherently reflected fields. This interpretation of the terms in the decomposition (2.32) is meaningful if 1) measurements are taken in a fixed geometry and 2) the statistical moments of Π are position independent: their asymptotic values in the large ensemble limit are the same for any patch of seafloor. If 2) holds for any fixed geometry, we consider the seafloor to be statistically homogeneous and assume it can be meaningfully labeled a specific "seabed class".

The main quantity to study is the scattered MSP, i.e.,

$$\langle y_s \rangle = \left\langle |\Pi|^2 \right\rangle - |\langle \Pi \rangle|^2 \tag{2.33}$$

where $y_s \equiv |\Pi_s|^2$. Another useful statistical quantity is the scintillation index (SI)

$$\Xi = \frac{\langle y_s^2 \rangle}{\langle y_s \rangle^2} - 1. \tag{2.34}$$

The SI measures the level of fluctuation in a probability distribution or ensemble of observed values. For the monostatic MBES configuration we may disregard the coherently reflected field, hence the second term in (2.33), except near vertical incidence. We therefore drop the subscript s. Following [1] we will also write $x \equiv |\Pi_s|$. For an MBES, the geometry is given by the incidence angle θ and slant range R (Fig. 2.2), so the average $\langle \cdots \rangle$ is taken at fixed incidence angle. We consider y the realization of a θ-dependent random variable Y with distribution $f_{Y|\Theta}(y \mid \theta)$. The range dependence is absorbed in the two factors Ω and Q in (2.30).

If the true (generally unknown) seafloor surface is given by a depth (or elevation) function $z(\boldsymbol{r})$, where $\boldsymbol{r}$ represent horizontal position, then two statistical measures relevant for scattering [1, Sec. 6.1] are the RMS roughness

$$h^2 = \left\langle \zeta(\boldsymbol{r})^2 \right\rangle, \tag{2.35}$$

where $\zeta(\boldsymbol{r}) \equiv z(\boldsymbol{r}) - \langle z \rangle$, and the spatial covariance

$$B(\boldsymbol{r}) = \left\langle \zeta(\boldsymbol{r}_0 + \boldsymbol{r})\zeta(\boldsymbol{r}_0) \right\rangle. \tag{2.36}$$

For (2.36) to be a covariance independent of $\boldsymbol{r}_0$, it must be assumed that the relief function

$\zeta(\boldsymbol{r})$ is stationary, so that $\langle\zeta(\boldsymbol{r}_0)\zeta(\boldsymbol{r}_1)\rangle = B(\boldsymbol{r}_1 - \boldsymbol{r}_0)$ depends only on the difference $\boldsymbol{r}_1 - \boldsymbol{r}_0$. Mathematical theories of scattering generally treat ζ (or related quantities) as a random field, and the averaging operation $\langle\cdots\rangle$ is formally the expectation value over all realizations of the field. In particular, some scattering theories assume that ζ is a Gaussian random field; such a random field is completely defined by its mean $\langle\zeta\rangle$ and covariance function $B(\boldsymbol{r})$ (2.36), or equivalently the roughness spectrum, which is the Fourier transform of $\boldsymbol{B}$,

$$W(\boldsymbol{k}) = \int_{\mathbb{R}^2} B(\boldsymbol{r}) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \, \mathrm{d}^2\boldsymbol{r}. \tag{2.37}$$

In small-roughness perturbation theory, the monostatic scattering cross section may be written

$$\sigma(\theta) = \left(\frac{\omega}{c}\right)^4 |A(\boldsymbol{k}_r)|^2 W(-2\boldsymbol{k}_r), \tag{2.38}$$

where $\boldsymbol{k}_r$ is the radial (tangent plane) component of the incident acoustic wave vector in water, with $|\boldsymbol{k}_r| = (\omega/c)\sin\theta$ [1, Sec. 13.1, App. K]. The coefficient $A(\boldsymbol{k}_r)$ depends on the acoustic properties of the seabed, e.g., for homogeneous fluid [20], homogeneous elastic [21], or homogeneous poro-elastic conditions [22]. Similar forms for σ are found in the small-slope approximation and Kirchhoff scattering theories.

Assuming ζ is stationary Gaussian, McDaniel [23] found using small-roughness perturbation theory that $\Xi = 1$ when the ensonified area is sufficiently large compared to the acoustic wavelength. (An additional assumption in [23] is that $W(\boldsymbol{k})$ varies slowly near the Bragg wave vector $\pm 2\boldsymbol{k}_r$. Moreover, a Gaussian function is used as an analytically tractable approximation to the ensonification function Ω.) The result $\Xi = 1$ is consistent with a Rayleigh PDF for the pressure amplitude x, i.e. (suppressing the conditional parameter θ),

$$f_X(x) = \frac{2x}{\langle x^2\rangle} e^{-x^2/\langle x^2\rangle}. \tag{2.39}$$

The general formula for the PDF of the product of two random variables, $Y = SX$, is

$$p_Y(y) = \int_{-\infty}^{\infty} \frac{1}{|s|} p_{S,X}(s, \frac{y}{s}) \, \mathrm{d}s. \tag{2.40}$$

Evaluating (2.40) using the joint distribution $p_{X,X}(s, x) = \delta(s - x)p_X(x)$, with $S = X$ and p_X a Rayleigh distribution, shows that the distribution for the squared pressure envelope y is exponential, i.e.,

$$p_Y(y) = \frac{1}{\langle x^2\rangle} e^{-y/\langle x^2\rangle}, \; y \geq 0 \; (0 \text{ otherwise}). \tag{2.41}$$

This PDF yields $\Xi = 1$ in (2.34). Although we are not aware of a theoretical development similar to [23] for estimating Ξ from volume scattering, it is still expected that the pressure amplitude $x = |\Pi_s|$ will be approximately Rayleigh distributed when the ensonified area is sufficiently large. This follows from considering the scattered field $\Pi_s(t)$ as a sum of contributions from N discrete scatterers within the ensonified area $\Gamma(t)$. Such discrete scatterers may for example be surface facets or volume inhomogeneities with high acoustic impedance contrast, such as pebbles, shells, or gas bubbles. Let B denote the quantity $B \equiv s_0^2 Q^2 \int_{\mathbb{R}^2} \Omega(\boldsymbol{r}, \boldsymbol{r_0}) \, \mathrm{d}^2\boldsymbol{r}$,

so that (2.30)

$$\langle y_s \rangle \equiv \langle |\Pi_s|^2 \rangle = \sigma(\theta)B. \tag{2.42}$$

We may then write

$$\Pi_s(\boldsymbol{r}, t) = \sqrt{B} \sum_{i=1}^{N} \frac{\kappa}{\sqrt{N}} a_i(\boldsymbol{r}, t) e^{i\phi_i(\boldsymbol{r}, t)}, \tag{2.43}$$

where the factor a_i ($i = 1, \ldots, N$) is a real random variable that, scaled by $\kappa/\sqrt{N}$, determines the scattering amplitude of the ith scatterer, and the constant κ must be determined so that (2.42) is maintained. The phase factor ϕ_i ($i = 1, \ldots, N$) is a random variable that depends on the travel time, and if Γ is large compared to the wavelength, ϕ_i may be assumed to be uniformly distributed over $[0, 2\pi]$, i.e., with PDF $p_\phi(\phi) = \mathbf{1}_{[0,2\pi]}/2\pi$ [24]. On this condition, and with the assumption that the random variables $a_1, \ldots, a_N, \phi_1, \ldots, \phi_N$ are statistically independent, it follows that in the limit $N \to \infty$, the pressure envelope is Rayleigh distributed independent of the statistics of the amplitude factors.

The characteristic function $\varphi_{\boldsymbol{X}} : \mathbb{R}^2 \to \mathbb{C}$ for the random vector $\boldsymbol{X} = [\Re\Pi_s, \Im\Pi_s]$ is

$$\varphi_{\boldsymbol{X}}(\lambda_1, \lambda_2) \equiv \mathrm{E}[e^{i\boldsymbol{\lambda} \cdot \boldsymbol{X}}] = \left\langle \prod_{i=1}^{N} \int_0^{2\pi} \frac{\mathrm{d}\phi_i}{2\pi} e^{i\kappa\sqrt{B}(a_i/\sqrt{N})\lambda \cos(\psi - \phi_i)} \right\rangle_{\boldsymbol{a}} \tag{2.44}$$

where $\langle \cdots \rangle_{\boldsymbol{a}}$ denotes the expectation value with respect to the random vector $\boldsymbol{a} = [a_1, \ldots, a_N]$, and $\boldsymbol{\lambda}$ has been expressed in polar coordinates: $\lambda_1 = \lambda \cos\psi$ and $\lambda_2 = \lambda \sin\psi$. With the change of variables $\tau = \phi_i - \psi - \pi/2$, the integral in (2.44) is seen to be an integral representation of the Bessel function $J_0(\kappa\lambda\sqrt{B}a_i/\sqrt{N})$. Since the a_i factors are independent, expanding $J_0(x)$ in powers of x ([25, Eqn. 9.1.12]) and taking the limit $N \to \infty$ gives

$$\begin{aligned} \varphi_{\boldsymbol{X}}(\boldsymbol{\lambda}) &= \left\langle J_0\left(\kappa\lambda\sqrt{B}\frac{a}{\sqrt{N}}\right)\right\rangle_a^N = \left(1 - \frac{1}{4}\lambda^2 B\kappa^2 \langle a^2 \rangle_a \frac{1}{N} + \mathcal{O}(\frac{1}{N^2})\right)^N \\ &\xrightarrow{N\to\infty} e^{-\lambda^2 B\kappa^2 \langle a^2 \rangle_a/4}. \end{aligned} \tag{2.45}$$

Equation (2.45) is seen to be the characteristic function of a bivariate Gaussian distribution with mean $\boldsymbol{\mu} = 0$ and covariance matrix $\boldsymbol{\Sigma} = \xi^2 \mathbf{1}_{2\times 2}$ where $\xi^2 = (1/2)B\kappa^2 \langle a^2 \rangle_a$. Hence $\Re\Pi_s$ and $\Im\Pi_s$ are independent normal random variables with equal variance ξ^2 and zero mean, so $x \equiv |\Pi_s| = |X|$ is Rayleigh distributed with mean $\langle x \rangle = \xi\sqrt{\pi/2}$. Consequently, $\langle y_s \rangle = 2\xi^2$, and from (2.42) we determine that

$$\kappa = \sqrt{\frac{\sigma(\theta)}{\langle a^2 \rangle}}. \tag{2.46}$$

When the ensonified area is not large, which may well happen with a high-resolution MBES, it is expected that there will be larger intensity fluctuations, i.e., with $\Xi > 1$. Several heuristic theories have been developed to explain non-Rayleigh fluctuations in randomly scattered wavefields, e.g., by imposing a particular distribution for the scattering amplitudes or the number of scatterers or the scattering cross sections [24, 26–31]. McDaniel [23], on the other hand, accounts for non-Rayleigh scattering statistics by allowing the relief function ζ to be non-stationary. The roughness spectrum (2.37) then becomes position dependent, $W = W(\boldsymbol{k}, \boldsymbol{\rho})$,

where $\boldsymbol{\rho}$ is a horizontal position vector. Thus $W(\boldsymbol{k})$ completely defines the statistics of a stationary Gaussian relief, while $W(\boldsymbol{k}, \boldsymbol{\rho})$ yields a particular form of non-stationary relief statistics. The final expressions for the (frequency-space) moments in (2.34) depend on the mean spectrum $\overline{W}_B(\boldsymbol{\rho})$ and the two-point function

$$C(\boldsymbol{\rho}, \boldsymbol{\rho}') \equiv \overline{W_B(\boldsymbol{\rho})W_B(\boldsymbol{\rho}')} = W_B^2\left[1 + \gamma C_W(\boldsymbol{\rho} - \boldsymbol{\rho}')\right], \tag{2.47}$$

where $\gamma = \mathrm{Var}[W_B]/\langle W_B\rangle^2$, $W_B(\boldsymbol{\rho})$ is a weighted average of $W(\boldsymbol{k}, \boldsymbol{\rho})$ over wavenumber space, and C_W is a correlation function with $C_W(0) = 1$, assumed to be exponential or Gaussian. The bar denotes averaging over the ensemble of different non-stationary roughness spectra $W(\boldsymbol{k}, \boldsymbol{\rho})$. This second form of ensemble average is formal, not computational, as $\overline{W}_B(\boldsymbol{\rho})$ and $C(\boldsymbol{\rho}, \boldsymbol{\rho}')$ (2.47) are simply given by assumption. In particular, it is assumed that $\overline{W}_B$ and C are position independent, with $C(\boldsymbol{\rho}, \boldsymbol{\rho}') = C(\boldsymbol{\rho} - \boldsymbol{\rho}')$, so that averaging over non-stationary roughness spectra yields stationary statistics. The result is that $\Xi > 1$ and increases with decreasing ensonified area,

$$\Xi = 1 + \frac{2\gamma}{\sqrt{[1 + (\Delta x/L)^2][1 + (\Delta y/L)^2]}}, \tag{2.48}$$

when the area is $\sim \Delta x \Delta y$ [23].

This result can be reproduced and generalized [1, Sec. 16.3] by letting the scattering cross section be position dependent, replacing (2.47) with

$$\langle \sigma(\boldsymbol{r})\sigma(\boldsymbol{r}')\rangle = \langle \sigma\rangle^2\left[1 + \gamma C_\sigma(\boldsymbol{r} - \boldsymbol{r}')\right], \tag{2.49}$$

where now $\gamma = \mathrm{Var}[\sigma]/\langle \sigma\rangle^2$. It is then not possible to take σ outside the integral in (2.28)–(2.30), and one defines instead the average scattering cross section

$$s = \int_{\mathbb{R}^2} \sigma(\boldsymbol{r})\Omega(\boldsymbol{r}, \boldsymbol{r}_0)\,\mathrm{d}^2\boldsymbol{r}. \tag{2.50}$$

To obtain (2.48), Jackson and Richardson [1] propose that $x = |\Pi_s|$ is Rayleigh distributed (2.39), as in the stationary case, for an ensemble of measurements with a fixed value of s. This implies that $y = |\Pi_s|^2$ is conditionally exponentially distributed, and leads to the following compound distribution for y

$$p_Y(y) = \int_0^\infty \frac{1}{s} e^{-y/s} p_S(s)\,\mathrm{d}s, \tag{2.51}$$

where $p_S(s)$ is the PDF for the average scattering cross section. This PDF is the basis for the theoretical analysis of the statistical seabed classification algorithm of Paper II.

Suppose that S and Z are two independent random variables with $S > 0$ and

$$p_Z(z) = e^{-z}. \tag{2.52}$$

Then evaluation of the general formula for the product of two random variables (2.40) shows that $p_{SX}(y) = p_Y(y)$ as given by (2.51). This means that Y may be interpreted as the product of a random scattering strength variable and an exponentially distributed variable with unit

variance. Such a product model, with a specific distribution for S, the gamma distribution, has been used as a model for radar clutter [32, Ch. 5.4], [30, 31, 33] and sonar reverberation [34]. This point will be revisited in Sec. 3.3.2.

2.2 Statistical classification

2.2.1 Bayesian supervised classification

Section 1.3 introduced the concept of the feature vector, $\boldsymbol{y} = \varphi(\boldsymbol{x})$, where $\boldsymbol{x} \in \mathbb{R}^D$ is the observed signal, and the map $\varphi : \mathbb{R}^D \to \mathbb{R}^d$, known as feature extraction, is construed such that, ideally, signals representing different states of nature are mapped to different, separable regions of the feature space $\mathbb{R}^d$ while keeping d as low as possible. Here we will assume that there are a finite number of states (classes) to discriminate between, denoted $\omega_1, \ldots, \omega_c$.

Bayesian decision theory [3, Ch. 2] is founded on the following formula (Bayes' theorem) for the posterior $P(\omega_j \mid \boldsymbol{y})$:

$$P(\omega_j \mid \boldsymbol{y}) = \frac{p(\boldsymbol{y} \mid \omega_j)P(\omega_j)}{p(\boldsymbol{y})}. \tag{2.53}$$

The posterior $P(\omega_j \mid \boldsymbol{y})$ is the probability that an observation $\boldsymbol{y}$ represents class ω_j. The prior probability $P(\omega_j)$ is the base rate, the fraction of observations expected to belong to class ω_j. If nothing is known a priori about the base rate, a natural choice is to assign equal prior probabilities, i.e. $P(\omega_j) = 1/c$ for all $j = 1, \ldots, c$. The likelihood $p(\boldsymbol{y} \mid \omega_j)$, viewed as a function of $\boldsymbol{y}$, is the probability density distribution of observations from class ω_j (class-conditional density). The evidence factor $p(\boldsymbol{y})$ may be decomposed as

$$p(\boldsymbol{y}) = \sum_{j=1}^{c} p(\boldsymbol{y} \mid \omega_j)P(\omega_j), \tag{2.54}$$

from which the normalization $\sum_{j=1}^{c} P(\omega_j \mid \boldsymbol{y}) = 1$ also follows. The factor $p(\boldsymbol{y})$ is class independent and therefore usually disregarded. However, suppose that contrary to assumption an observation does not fit into any class, so that $p(\boldsymbol{y} \mid \omega) \ll 1$ for all $j = 1, \ldots, c$. Then $p(\boldsymbol{y})$ will be small, and normalization ensures that at least one $P(\omega_j \mid \boldsymbol{y})$ will be greater than or equal to $1/c$. This is a potential problem in seabed classification because it may not be strictly known in advance how many classes are represented in a dataset.

The mathematical (conditional) risk associated with assigning an observation $\boldsymbol{y}$ to a particular class ω_j is defined as the expected loss

$$R(\omega_i \mid \boldsymbol{y}) = \sum_{i=1}^{c} \lambda(\omega_i \mid \omega_j)P(\omega_j \mid \boldsymbol{y}), \tag{2.55}$$

where $\lambda(\omega_i \mid \omega_j)$, the loss function, quantifies the loss incurred by assigning an observation from class ω_j to class ω_i. The Bayes decision rule states that one should always choose the

class for which $R(\omega_i \mid \boldsymbol{y})$ is minimum. In the present work the following loss function is used:

$$\lambda(\omega_i \mid \omega_j) = \begin{cases} 0 & i = j \\ 1 & i \neq j \end{cases} \qquad i, j = 1, \ldots, c. \tag{2.56}$$

It then follows from (2.55) that

$$R(\omega_i \mid \boldsymbol{y}) = \sum_{j \neq i} P(\omega_j \mid \boldsymbol{y}) = 1 - P(\omega_i \mid \boldsymbol{y}). \tag{2.57}$$

The Bayes decision rule then implies that $\boldsymbol{y}$ should be assigned to the class ω_i with the largest posterior $P(\omega_i \mid \boldsymbol{y})$. This decision rule is optimum in the sense that it minimizes the probability of choosing a wrong class.

From (2.53) it follows that to apply the Bayes decision rule one must compare the quantities $p(\boldsymbol{y} \mid \omega_j)P(\omega_j)$, $j = 1, \ldots, c$, which presupposes that the class-conditional PDFs $p(\boldsymbol{y} \mid \omega_j)$ are known. In supervised classification, a set of observations whose class is known, the training dataset, is used to estimate the class-conditional PDFs. In Paper II the feature vectors are constructed such that, it is argued, $p(\boldsymbol{y} \mid \omega_i)$ is well approximated by a multivariate normal (Gaussian) PDF,

$$p(\boldsymbol{y} \mid \omega_i) = \frac{(2\pi)^{-d/2}}{|\boldsymbol{\Sigma}_i|^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1} (\boldsymbol{y} - \boldsymbol{\mu}_i)\right]. \tag{2.58}$$

The problem of estimating $p(\boldsymbol{y} \mid \omega_i)$ then reduces to computing the sample mean and covariance matrix

$$\hat{\boldsymbol{\mu}}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} \boldsymbol{y}_k^i \tag{2.59a}$$

$$\hat{\boldsymbol{\Sigma}}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} (\boldsymbol{y}_k^i - \hat{\boldsymbol{\mu}}_i)(\boldsymbol{y}_k^i - \hat{\boldsymbol{\mu}}_i)^T, \tag{2.59b}$$

where the training samples from class ω_i are denoted $\boldsymbol{y}_k^i$, $k = 1, \ldots, n_i$. In seabed classification, acoustic training data may be obtained from areas with well understood seabed conditions, or from small neighborhoods of sediment sampling sites. With scarce training data it is particularly important to keep the feature space dimension (d) low to get robust estimates of the PDFs. One general approach to achieve this is described in Sec. 2.2.3 below and applied in Paper II.

In terms of the discriminant function

$$\begin{aligned} g_i(\boldsymbol{y}) &= \log p(\boldsymbol{y} \mid \omega_i) + \log P(\omega_i) \\ &= -\frac{1}{2}(\boldsymbol{y} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(\boldsymbol{y} - \boldsymbol{\mu}_i) - \frac{d}{2}\log 2\pi - \frac{1}{2}\log |\boldsymbol{\Sigma}_i| + \log P(\omega_i), \end{aligned} \tag{2.60}$$

which increases monotonously with increasing posterior $P(\omega_i \mid \boldsymbol{y})$, Bayes decision rule may be formulated as follows: Select a class ω_i for which

$$g_i(\boldsymbol{y}) \geq g_j(\boldsymbol{y}) \text{ for all } j. \tag{2.61}$$
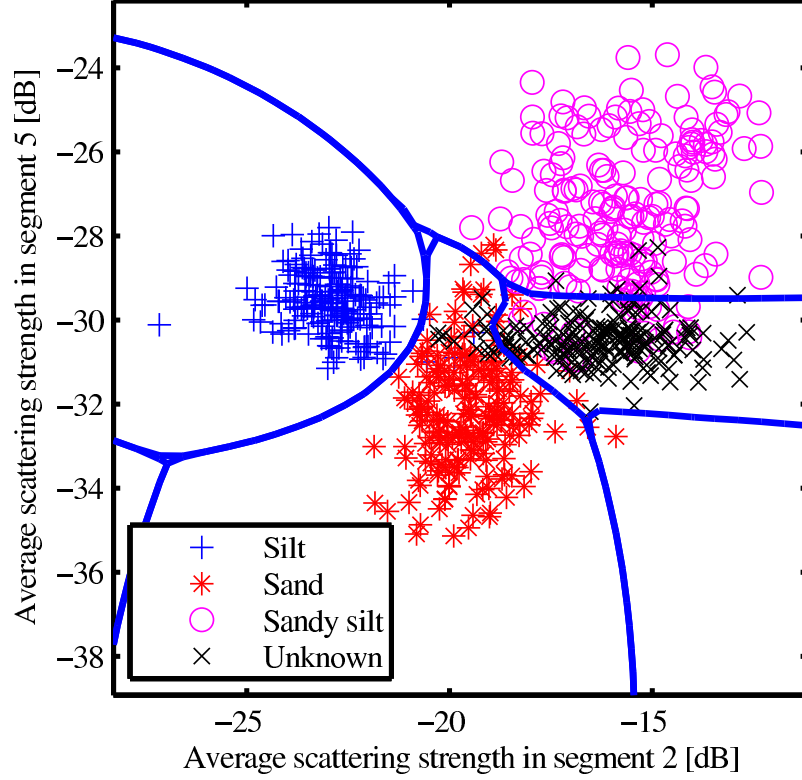
Fig. 2.3: Scatterplot and decision boundaries for a two-dimensional feature vector (from Paper II). The features represent seabed scattering strength at low (5° to 12°) and medium (32° to 50°) incidence angles, respectively. The decision boundaries were computed using the quadratic discriminant classifier (QDC). The plot contains 800 observations from four different seabed regions (200 observations per class).

The set of quadratic polynomial equations $g_i(\boldsymbol{y}) = g_j(\boldsymbol{y}),\ i \neq j$, determines the boundaries ("decision boundaries") of a partition of feature space, with each (interior) region associated with one of the classes $\omega_1, \ldots, \omega_c$ (Fig. 2.3). Using a parametric model for the class-conditional PDFs, as in (2.58), reduces the risk of overfitting the model to the data. When the estimated covariance matrix $\boldsymbol{\Sigma}_i$ is near-singular, e.g., if there are few and closely correlated training samples, a more robust classifier may be obtained by assuming either that the covariance matrices are all equal, or that each covariance matrix is diagonal. If the covariance matrices are equal, the equations $g_i(\boldsymbol{y}) = g_j(\boldsymbol{y}),\ i \neq j$ are linear and the decision boundaries are hyperplanes. This classifier is referred to as the linear discriminant classifier (LDC). The classifier with diagonal covariance matrices is termed the uncorrelated discriminant classifier (UDC), while the general case is the quadratic discriminant classifier (QDC).

2.2.2 Unsupervised learning and mixture models

In supervised classification, a classifier is trained on a limited set of labeled samples (i.e., with known class). In seabed classification problems, the amount of acoustic data that can be reliably used as labeled training data may be small, as there is often a scarcity of physical sediment samples. In addition, there may be acoustic data from sediment classes for which there are

no ground truth observations at all. Another challenge is that the acoustic signature of the seabed may change over time, for example if ocean waves or currents alter the micro-topography and surface roughness. An unsupervised learning approach may then be appropriate, in which the classifier is trained on a large amount of unlabeled data, and ground truth, if available, is introduced at the end to try and label groupings found in the data [3, Ch. 10].

Suppose that samples can be drawn from c classes $\omega_1, \ldots, \omega_c$, and assume that the class-conditional PDFs $p(\boldsymbol{y} \mid \omega_j, \boldsymbol{q}_j)$ have identical parametric form with unknown parameters $\boldsymbol{q}_j$, $j = 1, \ldots, c$. The PDF for the samples is

$$p(\boldsymbol{y}; \boldsymbol{q}, \boldsymbol{w}) = \sum_{j=1}^{c} p(\boldsymbol{y} \mid \omega_j; \boldsymbol{q}_j) P(\omega_j), \tag{2.62}$$

where $\boldsymbol{q} = \begin{bmatrix} \boldsymbol{q}_1, \ldots, \boldsymbol{q}_c \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} P(\omega_1), \ldots, P(\omega_{c-1}) \end{bmatrix}$ are unknown. For a training dataset consisting of n unlabeled samples, $\mathcal{D} = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\}$, independently drawn from the distribution (2.62), the joint distribution is

$$p(\mathcal{D}; \boldsymbol{q}, \boldsymbol{w}) = \prod_{k=1}^{n} p(\boldsymbol{y}_k; \boldsymbol{q}, \boldsymbol{w}), \tag{2.63}$$

and $p(\boldsymbol{y}_k; \boldsymbol{q}, \boldsymbol{w})$ is related to the posterior $P(\omega_j \mid \boldsymbol{y}_k; \boldsymbol{q}, \boldsymbol{w})$ by Bayes' rule:

$$P(\omega_j \mid \boldsymbol{y}_k; \boldsymbol{q}, \boldsymbol{w}) = \frac{p(\boldsymbol{y}_k \mid \omega_j; \boldsymbol{q}_j) P(\omega_j)}{p(\boldsymbol{y}_k; \boldsymbol{q}, \boldsymbol{w})}. \tag{2.64}$$

If the vectors $\boldsymbol{q}$ and $\boldsymbol{w}$ can be estimated from the unlabeled training data, then the Bayes decision rule (2.61) may be applied as for supervised classification. This approach is taken in the two algorithms described below, both of which were applied for echo classification in Paper I, with Gaussian mixtures. This means that (2.62) is assumed, with $p(\boldsymbol{y} \mid \omega_j; \boldsymbol{q}_j)$ given by (2.58), where the components of $\boldsymbol{q}_j$ are the unknown parameters of $\boldsymbol{\mu}_j$ and $\boldsymbol{\Sigma}_j$.

Expectation-maximization clustering

The expectation-maximization (EM) algorithm treats the class (cluster) labels $\{\,\omega_j\,\}$ as missing data over which the likelihood function is marginalized. The joint distribution for observing $\boldsymbol{y}$, belonging to ω_j, is

$$p(\boldsymbol{y}, \omega_j; \boldsymbol{q}) = p(\boldsymbol{y} \mid \omega_j; \boldsymbol{q}) P(\omega_j). \tag{2.65}$$

The corresponding log-likelihood function for observing the dataset $\mathcal{D}$ with associated class labels $\omega_{j_1}, \ldots, \omega_{j_n}$ is

$$L(\boldsymbol{q}) = \log \prod_{k=1}^{n} p(\boldsymbol{y}_k, \omega_{j_k}; \boldsymbol{q}) = \sum_{k=1}^{n} \log \left[p(\boldsymbol{y}_k \mid \omega_{j_k}; \boldsymbol{q}) P(\omega_{j_k}) \right]. \tag{2.66}$$

The EM algorithm is iterative. At each iteration t, new estimates for the unknown parameters, $\boldsymbol{q}_t, \boldsymbol{w}_t$, are computed. First the expected value of $L(\boldsymbol{q})$ is computed with respect to the joint

conditional distribution

$$W(j_1, \ldots, j_n) \equiv P(\omega_{j_1}, \ldots, \omega_{j_n} \mid \mathcal{D}; \boldsymbol{q}_t, \boldsymbol{w}_t) = \prod_{k=1}^{n} P(\omega_{j_k} \mid \boldsymbol{y}_k; \boldsymbol{q}_t, \boldsymbol{w}_t).$$

The result (the "E-step") is

$$
\begin{aligned}
Q(\boldsymbol{q}, \boldsymbol{w} \mid \boldsymbol{q}_t, \boldsymbol{w}_t) &= \mathrm{E}_W \left[\sum_{k=1}^{n} \log \left(p(\boldsymbol{y}_k \mid \omega_{j_k}; \boldsymbol{q}) w_{j_k} \right) \right] \\
&= \sum_{k=1}^{n} \mathrm{E}_W \left[\log \left(p(\boldsymbol{y}_k \mid \omega_{j_k}; \boldsymbol{q}) w_{j_k} \right) \right] \\
&= \sum_{k=1}^{n} \sum_{j_1=1}^{c} \cdots \sum_{j_n=1}^{c} \log \left(p(\boldsymbol{y}_k \mid \omega_{j_k}; \boldsymbol{q}) w_{j_k} \right) \prod_{l=1}^{n} P(\omega_{j_l} \mid \boldsymbol{y}_l; \boldsymbol{q}_t, \boldsymbol{w}_t) \\
&= \sum_{k=1}^{n} \sum_{j=1}^{n} P(\omega_j \mid \boldsymbol{y}_k; \boldsymbol{q}_t, \boldsymbol{w}_t) \log \left(p(\boldsymbol{y}_k \mid \omega_j; \boldsymbol{q}) w_j \right). \quad (2.67)
\end{aligned}
$$

The probability $P(\omega_j \mid \boldsymbol{y}_k; \boldsymbol{q}_t, \boldsymbol{w}_t)$ is evaluated using (2.64) and (2.62) with the parameter estimates $\boldsymbol{q}_t$ and $\boldsymbol{w}_t$. For Gaussian mixtures, the second factor is on the form (2.60). Next, the quantity $Q(\boldsymbol{q}, \boldsymbol{w} \mid \boldsymbol{q}_t, \boldsymbol{w}_t)$ is maximized with respect to $\boldsymbol{q}$ and $\boldsymbol{w}$, subject to the constraint $\sum_{j=0}^{c} w_j = 1$, which yields closed-form expressions for the updated estimates $\boldsymbol{q}_{t+1}$ and $\boldsymbol{w}_{t+1}$ (the "M-step") [35, Sec. 14.2]. The procedure is repeated until a convergence criterion is met, e.g., $\|\boldsymbol{q}(t+1) - \boldsymbol{q}(t)\| + \|\boldsymbol{w}(t+1) - \boldsymbol{w}(t)\| < \epsilon$ for some small positive tolerance ϵ. The initial ($t = 0$) estimates $\boldsymbol{q}_0$ and $\boldsymbol{w}_0$ may be chosen arbitrarily, subject only to the constraint $\sum_{j=0}^{c} w_j = 1$.

The QTC algorithm

The QTC algorithm [36–38] is a proprietary algorithm which is also based on the Gaussian mixture assumption (2.62), and which also estimates the unknown parameters $\boldsymbol{q} = \left[\boldsymbol{q}_1, \ldots, \boldsymbol{q}_c \right]$ and $\boldsymbol{w} = \left[P(\omega_1), \ldots, P(\omega_{c-1}) \right]$, so that a sample can be classified using Bayes' decision rule. For a fixed number of classes c, QTC applies stochastic optimization on the unlabeled training dataset to minimize a cost function E with respect to $\boldsymbol{q}$ and $\boldsymbol{w}$. The process is iterative; at each step the samples are assigned to one of c clusters. If we let $\boldsymbol{y}_{ki}(t)$ denote the ith member of the kth cluster at step t, the cost function may be expressed as follows in terms of the discriminant functions of (2.60):

$$E = -2 \sum_{k=1}^{c} \sum_{i=1}^{n_k} g_k(\boldsymbol{y}_{ki}). \quad (2.68)$$

Here n_k is the number of observations assigned to the kth class at step t, so the total number of training samples is $n = \sum_{k=1}^{c} n_k$. Consider a small perturbation to the cluster structure at step t, e.g., moving a single observation $\boldsymbol{y}$ from cluster k to cluster j. The induced change in the cost function is $E' - E = -2 \left[g_j'(\boldsymbol{y}) - g_k(\boldsymbol{y}) \right]$. If n_k and n_j are large, then $g_j' \approx g_j$. The classification rule (2.61) then implies that the perturbation rule should be accepted if $E' < E$. QTC has a scheme for generating random perturbations, and sampling a large number of configurations.

However, to avoid finding a local minimum of E instead of the global minimum, perturbations with $E' > E$ are sometimes accepted, i.e., a form of simulated annealing.

At every step t, maximum likelihood (ML) estimates $\hat{q}_j^{\mathrm{ML}}(t)$ and $\hat{w}_j^{\mathrm{ML}}(t)$ are computed using the samples assigned to class ω_j at that stage. These estimates are accurate when the classes are well separated. The number of classes is determined by repeating this process for different values of c to find the value that minimizes the quantity $E + M \log n$, where $M = \dim \boldsymbol{q} + \dim \boldsymbol{w}$ is the number of free parameters. The term $M \log n$ represents a penalty for adding terms to the series in (2.62), and hence counters the tendency to overfit the data by choosing very flexible models. The quantity $E + M \log n$ is related to the Bayesian information criterion (BIC) [39],

$$\mathrm{BIC} = -2L(\boldsymbol{q}^{\mathrm{ML}}, \boldsymbol{w}^{\mathrm{ML}}) + M \log n, \tag{2.69}$$

where L is the log-likelihood function evaluated at the ML values of the unknown parameters $\boldsymbol{q}$ and $\boldsymbol{w}$. By requiring that $\nabla_{\boldsymbol{\theta}_i} L = 0$, $i = 1, \ldots, M$, where $\boldsymbol{\theta}^T = \begin{bmatrix} \boldsymbol{q}^T & \boldsymbol{w}^T \end{bmatrix}$, it can be shown that the parameter estimates $\hat{\boldsymbol{\mu}}_i$, $\hat{\Sigma}_i$, and $\hat{P}(\omega_i)$ at a finite, local maximum of the likelihood function must satisfy the following set of equations [3, Sec. 10.4]:

$$\hat{P}(\omega_i) = \frac{1}{n} \sum_{k=1}^{n} \hat{P}(\omega_i \mid \boldsymbol{y}_k, \hat{\boldsymbol{\theta}}) \tag{2.70}$$

$$\hat{\boldsymbol{\mu}}_i = \frac{\sum_{k=1}^{n} \hat{P}(\omega_i \mid \boldsymbol{y}_k, \hat{\boldsymbol{\theta}}) \boldsymbol{y}_k}{\sum_{k=1}^{n} \hat{P}(\omega_i \mid \boldsymbol{y}_k, \hat{\boldsymbol{\theta}})} \tag{2.71}$$

$$\hat{\Sigma}_i = \frac{\sum_{k=1}^{n} \hat{P}(\omega_i \mid \boldsymbol{y}_k, \hat{\boldsymbol{\theta}})(\boldsymbol{y}_k - \hat{\boldsymbol{\mu}}_i)(\boldsymbol{y}_k - \hat{\boldsymbol{\mu}}_i)^T}{\sum_{k=1}^{n} \hat{P}(\omega_i \mid \boldsymbol{y}_k, \hat{\boldsymbol{\theta}})}, \tag{2.72}$$

where

$$\hat{P}(\omega_i \mid \boldsymbol{x}_k, \hat{\boldsymbol{\theta}}) = \frac{p(\boldsymbol{y}_k \mid \omega_i, \hat{\boldsymbol{\theta}}_i) \hat{P}(\omega_i)}{\sum_{j=1}^{c} p(\boldsymbol{y}_k \mid \omega_j, \hat{\boldsymbol{\theta}}_j) P(\omega_j)}. \tag{2.73}$$

Assuming that, after many iterations ($t \gg 1$),

$$\hat{P}(\omega_i \mid \boldsymbol{y}_k, \boldsymbol{\theta}) \simeq \begin{cases} 1 & \text{if } \boldsymbol{y}_k \text{ is from class } \omega_i \\ 0 & \text{otherwise}, \end{cases} \tag{2.74}$$

it follows from (2.70)–(2.72) that $\hat{P}(\omega_i) \simeq n_i/n$ and $\hat{\boldsymbol{\mu}}_i$ and $\hat{\Sigma}_i$ are given by (2.59). Consequently,

$$\min E \approx E(t \gg 1) \approx -2 \log p(\mathcal{D} \mid \hat{\boldsymbol{q}}^{\mathrm{ML}}, \hat{\boldsymbol{w}}^{\mathrm{ML}}; c), \tag{2.75}$$

hence

$$\mathrm{BIC} \approx \min E + M \log n. \tag{2.76}$$

Treating c, the number of terms (classes or clusters) in the mixture model (2.62), as an unknown parameter, Bayes' rule implies that

$$\begin{aligned} P(c \mid \mathcal{D}) &\propto W(c) P(\mathcal{D} \mid c) \\ &\propto W(c) \int_{\mathbb{R}^M} p(\mathcal{D} \mid \boldsymbol{q}, \boldsymbol{w}; c) p(\boldsymbol{q}, \boldsymbol{w} \mid c) \, \mathrm{d}\boldsymbol{q} \, \mathrm{d}\boldsymbol{w}, \end{aligned} \tag{2.77}$$

where $W(c)$ is the prior probability that the unknown mixture model has c terms. Assuming there is no prior information about c or $\boldsymbol{q}, \boldsymbol{w}$, so that W and $P(\boldsymbol{q}, \boldsymbol{w} \mid c)$ are uniform, equation (2.77) can be evaluated using Laplace's method for approximating integrals (e.g., [40]), with the result that

$$-2 \log P(\mathcal{D} \mid c) \approx \min E + M \log n. \tag{2.78}$$

The approximation to the integral in (2.77) is accurate in the large n limit if $L(\boldsymbol{q}, \boldsymbol{w})$ decays rapidly around the ML point. Subject to these assumptions, as well as (2.74), minimizing the right hand side of (2.78) as in the QTC algorithm maximizes the likelihood of the observed data $\mathcal{D}$. QTC software has been extensively used in ASC work, including this thesis (Paper I).

2.2.3 Feature reduction

The volume of a D-dimensional ball $B^D(\boldsymbol{x}; r) = \{ \boldsymbol{x}' \in \mathbb{R}^D \mid \|\boldsymbol{x}' - \boldsymbol{x}\| \leq r \}$ is (e.g. [41, Ch. 10])

$$\text{Vol}\left(B^D\right) = \begin{cases} \frac{\pi^m}{m!} r^D & \text{for } D = 2m \text{ an even integer} \\ \frac{2^{2m+1} m! \pi^m}{(2m+1)!} r^D & \text{for } D = 2m + 1 \text{ an odd integer.} \end{cases} \tag{2.79}$$

Applying Stirling's approximation $m! \approx \sqrt{2\pi m}\, (m/e)^m$ in (2.79) shows that $\log \text{Vol}\left(B^D\right) \to -\infty$ for $D \to \infty$, which means that $\text{Vol}\left(B^D\right) \to 0$. In a hypercube of side length L, large enough to contain all the training samples, the average number of samples per unit volume is n/L^D. Since the number of training samples n is finite, the average number density is finite (and tends to zero for $D \to \infty$ if $L > 1$). Therefore, the expected number of training samples in a B^D-neighborhood of an arbitrary point $\boldsymbol{x}$ becomes small when D is large; hence it becomes difficult to estimate the class-conditional densities $p(\boldsymbol{x} \mid \omega_i)$.

This sparsity of training data in high dimensions is one reason for reducing the number of features with respect to the number of components in the observed signal. It is pertinent in seabed classification, where there will typically be a limited number of sediment samples available. Training observations (acoustic data) must be acquired in the vicinity of the sampling sites so that they can reasonably be assumed to represent the class of the sediment sample. The observed signal may often be viewed as (samples of) some function of an independent variable, such as complex pressure envelope versus time (echoes), power spectral density versus frequency, or as in Paper II, scattering strength versus incidence angle. For such signals, function approximation is an effective means of feature reduction. For example, computing the least squares error fit of a discrete time series $\boldsymbol{x} = [p(t_1), \ldots, p(t_D)] \in \mathbb{R}^D$ to the polynomial model $p(t) = a_0 + a_1 t + \ldots a_{d-1} a^{d-1}$ produces the feature vector $\boldsymbol{y} = [a_0, \ldots, a_{d-1}] \in \mathbb{R}^d$. An even simpler approximation method is the bin smoother [42, Sec. 2.1], where the time (independent variable) domain is partitioned into d disjoint bins, and the signal is represented by the set of mean data values in each bin. With this method each feature has a simple physical interpretation. In Paper II we apply a more sophisticated variant of the bin smoother, in which the bin sizes are determined from the data such that the approximation error is minimized for the data set as a whole [43].

2.3 Image restoration

The standard end product of MBES seabed mapping is a digital terrain model (DTM), a gridded or triangulated model of water depth. Many attribute images can be derived from a DTM, including gradient, curvature, texture, and shaded relief (hill-shade) images. Such attributes are important interpretative tools as means for visualizing terrain forms and seabed geomorphology, and may also be used for automated statistical classification of terrain type, geological provinces, or even sediment type [44–46]. It should be possible to improve the seabed classification accuracy by combining acoustic features (e.g. echo shape) with terrain attributes. We therefore investigated attributes with useful invariant properties that turned out to be sensitive to noise or artifacts in the bathymetric data. This is the background for Paper III, on terrain classification and, primarily, noise reduction in terrain models and attribute images.

In analogy with filtering in the frequency (Fourier) domain, the main idea is to suppress noise in a transformed representation of the image (DTM) and computing the inverse transform to obtain the restored image. The transform consists of a modified Laplace operator composed with a discrete Radon transform. It is chosen such that the signature of the transformed noise becomes easier to suppress. The Laplace operator accentuates fine-scale intensity changes; the DRT focuses linear features because each DRT value is the sum of pixel intensities (depth values) along a straight line. However, the standard Laplace operator is not invertible, and the inverse Radon transform will in general be a pseudo-inverse. The following subsections provide some background on the key elements of the algorithm.

2.3.1 The Laplace operator

The Laplace operator (or Laplacian) $\Delta \equiv \nabla^2$ is the second-order differential operator appearing in the Helmholtz equation (A.2). In Euclidean space $\mathbb{R}^d$ it is defined as

$$\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \ldots + \frac{\partial^2 f}{\partial x_d^2}, \tag{2.80}$$

using Cartesian coordinates. The Laplace operator is related to neighborhood averages. If f is a twice-differentiable real function on an open set $\Omega \subseteq \mathbb{R}^d$, then for any $\boldsymbol{x} \in \Omega$ [47, Sec. 1.4]

$$\lim_{r \to 0} \frac{1}{r^2 \operatorname{Vol}(B^d(\boldsymbol{x}; r))} \int_{B^d(\boldsymbol{x}; \boldsymbol{r})} (f(\boldsymbol{y}) - f(\boldsymbol{x})) \, \mathrm{d}\boldsymbol{y} = \frac{1}{2(d+2)} \Delta f(\boldsymbol{x}), \tag{2.81}$$

where B^d is the d-dimensional ball (Sec. 2.2.3). Equation (2.81) shows that Δ is an isotropic detector for boundaries between distinct regions. For continuous images ($d = 2$) the Laplace operator accentuates edges, and the local value $\Delta f(\boldsymbol{x})$ is independent of the orientation of the structures it detects. A digital (discrete) image $\boldsymbol{f} = [f(u, v)]_{u,v=1,\ldots,N}$ with $N \times N$ pixels may be regarded as regular samples of the periodic band-limited function given by the Fourier series

$$f : \mathbb{T}^2 \to \mathbb{R}, \quad f(x, y) = \sum_{m=-N/2}^{N/2-1} \sum_{n=-N/2}^{N/2-1} c_{mn} e^{2\pi i(nx+my)}. \tag{2.82}$$

27

The torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ is obtained by identifying any point $(x, y) \in \mathbb{R}^2$ with the corresponding unique point $(x', y') \in [0, 1) \times [0, 1)$ such that $(x, y) - (x', y') \in \mathbb{Z}^2$. (Geometrically $\mathbb{T}^2$ is obtained from the unit square $[0, 1] \times [0, 1]$ by gluing together the two pairs of opposing edges.) By letting the first row in the image ($m = 1$) correspond to $y = 0$, the coefficients c_{mn} in (2.82) are related to the 2D discrete Fourier transform $\mathcal{F}$ by

$$c_{mn} = \frac{1}{N^2} \mathcal{F}[\boldsymbol{f}](m + N \bmod N, n + N \bmod N), \tag{2.83}$$

$$\mathcal{F}[\boldsymbol{f}](m, n) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} f(u, v) e^{-2\pi i(mu+nv)/N}. \tag{2.84}$$

The square-integrable functions $L^2(\mathbb{T}^2) = \{\, f : \mathbb{T}^2 \to \mathbb{R} \mid \int_{\mathbb{T}^2} |f(\boldsymbol{x})|^2 \, \mathrm{d}\boldsymbol{x} < \infty \,\}$ may be identified with their associated set of Fourier coefficients

$$f \longleftrightarrow c_{mn} = \int_0^1 \int_0^1 f(x, y) e^{2\pi i(mx+ny)} \, \mathrm{d}x \, \mathrm{d}y, \ (m, n) \in \mathbb{Z}^2. \tag{2.85}$$

If $f \in L^2(\mathbb{T}^2) \cap C^2(\mathbb{T}^2)$ is twice differentiable (we do not consider weak derivatives), the action of Δ is

$$\Delta : c_{mn} \longmapsto -4\pi^2(m^2 + n^2)c_{mn}, \ (m, n) \in \mathbb{Z}^2. \tag{2.86}$$

In terms of the wave vector $\boldsymbol{k} = 2\pi(m, n)$, the transfer function for Δ is

$$\Delta(\boldsymbol{k}) = -\boldsymbol{k} \cdot \boldsymbol{k}, \tag{2.87}$$

which is another manifestation of the rotation invariance of the Laplace operator. The kernel (null space) of Δ is the set of constant functions, for which $c_{mn} = 0$ for $(m, n) \neq 0$. An inverse Laplace operator exists for the restricted set of functions

$$L_0^2(\mathbb{T}^2) = \left\{\, f \in L^2(\mathbb{T}^2) \,\bigg|\, \int_{\mathbb{T}^2} f \, \mathrm{d}x = 0 \,\right\}, \tag{2.88}$$

for which $c_{00} \equiv 0$. The inverse Δ^{-1} may then be defined by

$$\Delta^{-1} : c_{mn} \longmapsto \begin{cases} \frac{1}{-4\pi^2(m^2+n^2)} c_{mn} & \text{for } (m, n) \in \mathbb{Z}^2 \setminus \{0\} \\ 0 & \text{for } m = n = 0. \end{cases} \tag{2.89}$$

This means that the inverse Laplacian is only defined up to a constant function. Moreover, the periodic function f in (2.82) is smooth (infinitely differentiable) for $(x, y) \in (0, 1) \times (0, 1)$ but possibly discontinuous on the boundary of the image. It is therefore in general not possible to use the natural, frequency-domain definition of the Laplace operator on images, and discrete versions of Δ in the spatial domain are also singular; indeed any spatial derivative filter must suppress the mean value [48, Sec. 12.3] and therefore has a non-trivial kernel.

In the spatial domain the Laplace operator may be approximated by a convolutional operator given by a discrete point spread function $\boldsymbol{h}$, a matrix with dimension $(2n + 1) \times (2n + 1)$ and

$n < N/2$, i.e.

$$(\Delta \boldsymbol{f})(i, j) = (\boldsymbol{h} * \boldsymbol{f})(i, j) = \sum_{u=-n}^{n} \sum_{v=-n}^{n} \boldsymbol{h}(u, v) \boldsymbol{f}(i - u, j - v). \tag{2.90}$$

The transfer function $\mathcal{F}[\boldsymbol{h}]$ is obtained from (2.84) after zero-padding $\boldsymbol{h}$ to size $N \times N$. There is no universal choice for $\boldsymbol{h}$; two possible choices are

$$\boldsymbol{h}_a = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{2.91}$$

and

$$\boldsymbol{h}_b = \frac{1}{2} \begin{bmatrix} 1/2 & 1 & 1/2 \\ 1 & -6 & 1 \\ 1/2 & 1 & 1/2 \end{bmatrix} \tag{2.92}$$

The transfer functions are

$$\mathcal{F}[\boldsymbol{h}_a](\hat{\boldsymbol{k}}) = -4 \left[\sin^2(\hat{k}_x/2) + \sin^2(\hat{k}_y/2) \right] \tag{2.93}$$

and

$$\mathcal{F}[\boldsymbol{h}_b](\hat{\boldsymbol{k}}) = 4 \left[\cos^2(\hat{k}_x/2) \cos^2(\hat{k}_y/2) - 1 \right], \tag{2.94}$$

respectively, where $\hat{\boldsymbol{k}} = 2\pi(m, n)/N$ is a scaled wave vector. The following series expansions [48, Sec. 12.5], using polar coordinates $(\hat{k}, \phi)$ for $\hat{\boldsymbol{k}}$, express the anisotropy of the discrete approximations and should be compared with the ideal isotropic continuous Laplace operator (2.87):

$$\mathcal{F}[\boldsymbol{h}_a](\hat{\boldsymbol{k}}) \approx -\hat{k}^2 + \frac{3}{48}\hat{k}^4 + \frac{1}{48}\cos(4\phi)\hat{k}^4 + \mathcal{O}(\hat{k}^6) \tag{2.95}$$

and

$$\mathcal{F}[\boldsymbol{h}_b](\hat{\boldsymbol{k}}) \approx -\hat{k}^2 + \frac{3}{32}\hat{k}^4 - \frac{1}{96}\cos(4\phi)\hat{k}^4 + \mathcal{O}(\hat{k}^6). \tag{2.96}$$

Equations (2.95) and (2.96) show that the discrete Laplace operators are isotropic for small wave numbers (long wavelengths). They also show that $\boldsymbol{h}_b$ (2.92) is more isotropic than $\boldsymbol{h}_a$ (2.91). Rotation invariance is desirable in the applications discussed in Paper III. We use $\boldsymbol{h}_b$ with circular convolution, i.e. periodic boundary conditions, as the starting point for constructing an invertible spatial Laplace operator in Paper III. The operator is applied to detrended images such that discontinuities along the boundaries are insignificant.

2.3.2 The Radon transform

The Radon transform of a function $w : \mathbb{R}^n \to \mathbb{R}$ integrates w over hyperplanes in $\mathbb{R}^n$. A hyperplane $H \subset \mathbb{R}^n$ may be specified as a tangent plane to the unit sphere S^{n-1}, with a signed distance s to the origin:

$$H(\boldsymbol{\theta}, s) = \left\{ \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{\theta} \in S^{n-1}, \ s \in \mathbb{R}, \ \boldsymbol{x} \cdot \boldsymbol{\theta} = s \right\}. \tag{2.97}$$
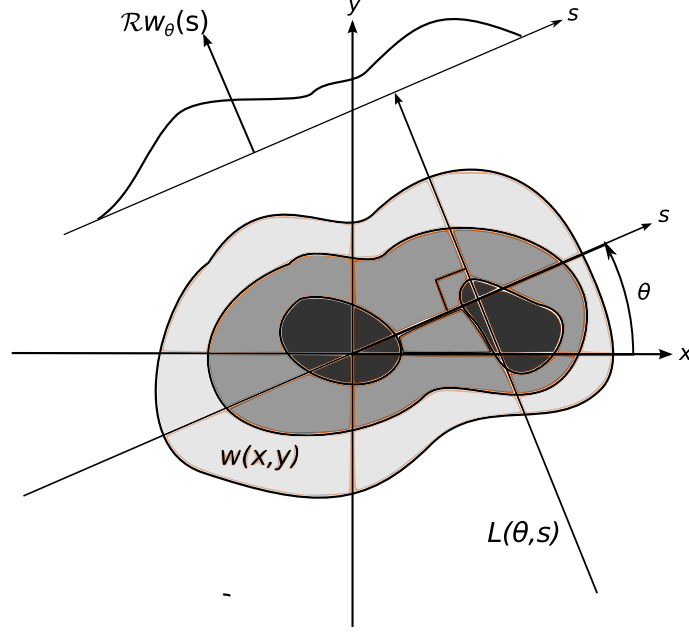
Fig. 2.4: The continuous Radon transform $\mathcal{R}w$ of a function $w : \mathbb{R}^2 \to \mathbb{R}$ is formed by integrating w along the set of straight lines $L(\theta, s)$, $\theta \in [0, 2\pi)$, $s \in \mathbb{R}$. The line $L(\theta, s)$ has a signed distance s from the origin and is perpendicular to the axis with unit vector $[\cos\theta \ \sin\theta]$. The function w is pictured here as a contour map. For a fixed angle θ one obtains the *projection* $s \longmapsto \mathcal{R}w_\theta(s) = \mathcal{R}w(\theta, s)$.

For $n = 2$, H is a straight line and $\boldsymbol{\theta} \in S^1$, the unit circle. The Radon transform of w is

$$(\mathcal{R}w)(\boldsymbol{\theta}, s) = \int_{H(\boldsymbol{\theta}, s)} w(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}. \tag{2.98}$$

This integral is referred to as a *projection* (Fig. 2.4), and in 2D computerized tomography it is a model for the cumulative attenuation of a linear beam passing through the object being imaged. The domain of $\mathcal{R}w$ is the set of all hyperplanes, and may be identified with the unit cylinder

$$C^n = \left\{ (\boldsymbol{\theta}, s) \mid \boldsymbol{\theta} \in S^{n-1}, s \in \mathbb{R} \right\}. \tag{2.99}$$

By definition, $H(-\boldsymbol{\theta}, -s) = H(\boldsymbol{\theta}, s)$ and $\mathcal{R}w(-\boldsymbol{\theta}, -s) = \mathcal{R}w(\boldsymbol{\theta}, s)$. The backprojection (adjoint) operator, applied to $g : C^n \to \mathbb{R}$, is a function $\mathcal{R}^+g : \mathbb{R}^n \to \mathbb{R}$ defined by

$$(\mathcal{R}^+g)(\boldsymbol{x}) = \int_{S^{n-1}} g(\boldsymbol{\theta}, \boldsymbol{x} \cdot \boldsymbol{\theta}) \, \mathrm{d}\boldsymbol{\theta}. \tag{2.100}$$

This definition implies that $(\mathcal{R}^+\mathcal{R}w)(\boldsymbol{x})$ is the average (up to a constant) of the projections of w along hyperplanes that intersect $\boldsymbol{x}$.

The GDB transform

Götz and Druckmüller [49] and Brady [50] (GDB) have proposed a discrete approximation to the 2D Radon transform, in which the line integrals are replaced by sums of pixel intensities along certain graphs that include exactly one pixel from each column in the zero-padded image
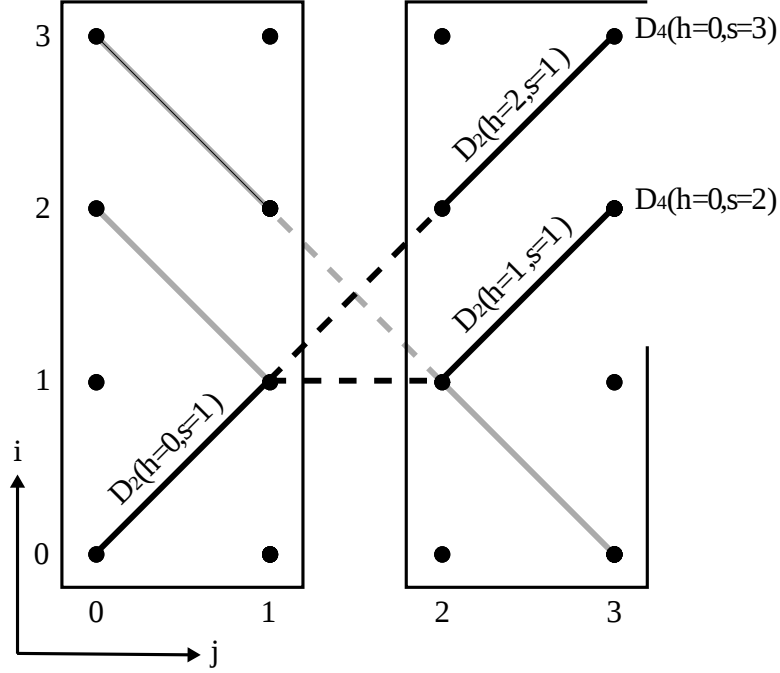
Fig. 2.5: Recursive construction of graphs in a 4×4 pixels image. The graphs are formed by joining graphs in the two 4×2 pixels subimages according to the rule in Eqn. (2.101). The rise parameter s ranges from 0 to 1 in 4×2 subimages, and from 0 to 3 in the 4×4 image. When the image is flipped upside down, the graphs are mapped to lines with slopes in the range $-45°$ to $0°$ (faint lines). (Illustration from Paper III.)

(matrix). The algorithm is recursive, with $4N^2 \log_2 N$ arithmetic operations (additions) for $N \times N$ pixels images. This is faster than conventional $\mathcal{O}(N^3)$ algorithms. Specifically, let $f_{i,j} \in \mathbb{R}^{N \times N}$ denote the pixel intensity in row i and column j, where $f_{i,j} = 0$ for i,j outside the range $0 \leq i,j < N-1$, and the lower left corner pixel is $(0,0)$. The image size N must be a power of 2. In the first quadrant, i.e., for slopes from $0°$ to $45°$, a graph representing a straight line starts at pixel $(h,0)$ and ends at $(h+s, N-1)$, where h is the intercept and s is the slope; this graph is denoted $D_N(h,s)$. The same notation may be applied to the subgraphs obtained when the image is recursively split in vertical halves to produce subimages of size $N \times N/2$, $N \times N/4$, ..., $N \times 1$. Graphs in subimages of size $N \times 1$ contain a single pixel. The full graph is constructed recursively by combining subgraphs in pairs of adjacent subimages, according to the rule [49–51]

$$D_n(h, 2s) = D_{n/2}^{(L)}(h, s) \cup D_{n/2}^{(R)}(h+s, s) \tag{2.101a}$$

$$D_n(h, 2s+1) = D_{n/2}^{(L)}(h, s) \cup D_{n/2}^{(R)}(h+s+1, s). \tag{2.101b}$$

Here, the right-hand side of the union refers to the right (R) half of the image, and correspondingly for the left-hand (L) side. The procedure is illustrated in Fig. 2.5. Slopes in the other quadrants are obtained by transposing or flipping the image. The full DRT is the disjoint union

of four parts (quadrants):

$$(\mathcal{R}^1 f)(h,s) = \sum_{(i,j) \in D_N(h,s)} f_{i,j} \qquad (0° \text{ to } 45°) \tag{2.102a}$$

$$(\mathcal{R}^2 f)(h,s) = \sum_{(i,j) \in D_N(h,s)} f_{j,i} \qquad (45° \text{ to } 90°) \tag{2.102b}$$

$$(\mathcal{R}^3 f)(h,s) = \sum_{(i,j) \in D_N(h,s)} f_{j,N-1-i} \qquad (-90° \text{ to } -45°) \tag{2.102c}$$

$$(\mathcal{R}^4 f)(h,s) = \sum_{(i,j) \in D_N(h,s)} f_{N-1-i,j} \qquad (-45° \text{ to } 0°). \tag{2.102d}$$

The number of non-zero elements in each quadrant is at most $N^2 + N(N-1)/2$, and the minimum value for h is $-N+1$, for which the associated graph intersects the image only at a corner pixel. The range $\mathcal{R}\left[\mathbb{R}^{N \times N}\right]$ therefore has dimension $P = 4(N^2 + N(N-1)/2) = 6N^2 - 2N$, and $\mathcal{R} : \mathbb{R}^Q \to \mathbb{R}^P$ is a linear operator between vector spaces, with $Q = N^2$. There is also a recursive, linear adjoint algorithm, $\mathcal{R}^+ : \mathbb{R}^P \to \mathbb{R}^Q$ [50], whose matrix representation is the transpose of the matrix representation of $\mathcal{R}$ [52]. If $g = \mathcal{R}^+ \mathcal{R} f$, then $g_{i,j}$ is the sum of f along all graphs containing pixel (i,j), i.e., g is the backprojection of $\mathcal{R}f$.

2.3.3 Krylov methods

Krylov subspaces

In the discrete setting, computing the inverse Radon transform means solving the linear system

$$\mathcal{R}g = d \tag{2.103}$$

with respect to $g \in \mathbb{R}^{N \times N}$, where $d \in \mathbb{R}^{6N^2 - 2N}$ is given. In Paper III, d is the output of a filter in the DRT domain, and is generally not contained in the image of $\mathcal{R}$. A particular pseudo-inverse is defined as the solution to the linear system

$$\mathcal{A}g = b, \tag{2.104}$$

where

$$\mathcal{A} \equiv \mathcal{B}\mathcal{R} \text{ and } b \equiv \mathcal{B}d, \tag{2.105}$$

and $\mathcal{B}$ is an approximate inverse DRT algorithm due to Press [51]; thus $\mathcal{B}$ acts as a preconditioner for (2.103). Unless the image size, N, is small, this system must be solved using only matrix-vector multiplications implemented by the two algorithms $\mathcal{R}$ and $\mathcal{B}$. This may be accomplished with Krylov methods, which encompass several important algorithms for solving linear systems.

Krylov methods are iterative, and find approximate solutions in the sequence of Krylov subspaces defined by

$$\mathcal{K}_{k+1}(\mathcal{A}, b) \equiv \text{span}\left\{ b, \mathcal{A}b, \ldots, \mathcal{A}^k b \right\}, \; k = 1, 2, \ldots \tag{2.106}$$

If $\mathcal{A}$ is invertible, the solution $g = \mathcal{A}^{-1}b$ lies in a Krylov subspace: Suppose that $\boldsymbol{A}$ is a matrix representation of $\mathcal{A}$ with inverse $\boldsymbol{A}^{-1}$, and that

$$q(\lambda) = c_0 + c_1\lambda + \ldots + c_m\lambda^m, \; c_m \neq 0 \tag{2.107}$$

is a polynomial of minimal degree such that $q(\boldsymbol{A}) = 0$. Then $c_0 \neq 0$, otherwise

$$c_1\boldsymbol{I} + c_2\boldsymbol{A} + \ldots + c_m\boldsymbol{A}^{m-1} = 0$$

contrary to assumption. Hence,

$$0 = \frac{1}{c_0}q(\boldsymbol{A}) = \boldsymbol{I} + \frac{1}{c_0}\boldsymbol{A}\left[c_1 + c_2\boldsymbol{A} + \ldots + c_m\boldsymbol{A}^{m-1}\right]$$

and consequently

$$g = \boldsymbol{A}^{-1}\boldsymbol{b} = -\frac{1}{c_0}\left[\sum_{j=0}^{m-1} c_{j+1}\boldsymbol{A}^j\right]\boldsymbol{b} \in \mathcal{K}_m(\boldsymbol{A}, \boldsymbol{b}). \tag{2.108}$$

The existence of a polynomial such as (2.107) is ensured by the Cayley-Hamilton theorem, which states that the characteristic polynomial

$$p(\lambda) \equiv \det(\lambda\boldsymbol{I} - \boldsymbol{A}) = \lambda^n - a_1\lambda^{n-1} + \ldots + (-1)^n \det\boldsymbol{A} \tag{2.109}$$

is such that $p(\boldsymbol{A}) = 0$ (here, $n = N^2$).

Generalized Minimum Residual

In Paper III, we use in particular the generalized minimum residual (GMRES) algorithm [53, 54], which produces a series of approximations, g_k, defined as the solutions to the least squares problems

$$\min_{g \in \mathcal{K}_k(\mathcal{A},b)} \|b - \mathcal{A}g\|, \; k = 1, 2, \ldots \tag{2.110}$$

in the Euclidean norm. Unlike many other methods, GMRES can be applied to arbitrary (non-singular) matrices, and it is furthermore not necessary to have an implementation of the adjoint operator (matrix transpose). Both of these points are important with respect to inverting the DRT. The GMRES algorithm proceeds by recursively constructing an orthonormal basis for $\mathcal{K}_{k+1}(\boldsymbol{A}, b)$: Suppose the columns of the matrix $\boldsymbol{V}_k = \begin{bmatrix} \boldsymbol{v}_1 & \ldots & \boldsymbol{v}_k \end{bmatrix}$ form an orthonormal basis for $\mathcal{K}_k(\boldsymbol{A}, b)$, with $\boldsymbol{v}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|$. The orthogonal projection of $\boldsymbol{A}\boldsymbol{v}_k$ onto $\mathcal{K}_k(\boldsymbol{A}, b)$ is $\sum_{i=1}^k (\bar{\boldsymbol{v}}_i^T \boldsymbol{A}\boldsymbol{v}_k)\boldsymbol{v}_i$. Therefore, the new basis vector

$$\boldsymbol{v}_{k+1} = \hat{\boldsymbol{v}}_{k+1}/\|\hat{\boldsymbol{v}}_{k+1}\|, \tag{2.111}$$

where

$$\hat{\boldsymbol{v}}_{k+1} = \boldsymbol{A}\boldsymbol{v}_k - \sum_{i=1}^k (\bar{\boldsymbol{v}}_i^T \boldsymbol{A}\boldsymbol{v}_k)\boldsymbol{v}_i = \boldsymbol{A}\boldsymbol{v}_k - \boldsymbol{V}_k\boldsymbol{h}_k \tag{2.112}$$

and $\boldsymbol{h}_k = \overline{\boldsymbol{V}}_k^T \boldsymbol{A} \boldsymbol{v}_k$, is orthogonal to $\mathcal{K}_k(\boldsymbol{A}, b)$, and by construction, $\boldsymbol{v}_{k+1} \in \mathcal{K}_{k+1}(\boldsymbol{A}, b)$. This orthogonalization scheme, known as Arnoldi's method, may be written as [55]

$$\boldsymbol{A}\boldsymbol{V}_k = \boldsymbol{V}_{k+1}\boldsymbol{H}_k, \tag{2.113}$$

where

$$\boldsymbol{H}_k = \begin{bmatrix} \boldsymbol{H}_{k-1} & \boldsymbol{h}_k \\ 0 & \|\hat{\boldsymbol{v}}_{k+1}\| \end{bmatrix} \in \mathbb{R}^{(k+1)\times k}, \ \boldsymbol{H}_1 = \begin{bmatrix} \boldsymbol{h}_1 \\ \|\hat{\boldsymbol{v}}_2\| \end{bmatrix}. \tag{2.114}$$

Any $\boldsymbol{g} \in \mathcal{K}_k(\boldsymbol{A}, b)$ may be decomposed as $\boldsymbol{g} = \boldsymbol{V}_k \boldsymbol{y}$ for some $\boldsymbol{y} \in \mathbb{R}^k$, and $\boldsymbol{v}_1 = \boldsymbol{V}_{k+1}\boldsymbol{e}_1$ where $\boldsymbol{e}_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}^T \in \mathbb{R}^{k+1}$. The norm in (2.110) may therefore be rewritten as

$$\|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{g}\| = \|b\boldsymbol{V}_{k+1}\boldsymbol{e}_1 - \boldsymbol{V}_{k+1}\boldsymbol{H}_k\boldsymbol{y}\| = \|\boldsymbol{V}_{k+1}\| \, \|b\boldsymbol{e}_1 - \boldsymbol{H}_k\boldsymbol{y}\|$$

where $b = \|\boldsymbol{b}\|$. The GMRES approximate solution is $\boldsymbol{g}_k = \boldsymbol{V}_k\boldsymbol{y}_k$, where

$$\boldsymbol{y}_k = \min_{\boldsymbol{y}} \|b\boldsymbol{e}_1 - \boldsymbol{H}_k\boldsymbol{y}\|. \tag{2.115}$$

Eigenvalue estimation

Suppose that $(\boldsymbol{y}, \lambda)$ is an eigenpair for the $k \times k$ matrix $\hat{\boldsymbol{H}}_k = \begin{bmatrix} \boldsymbol{H}_{k-1} & \boldsymbol{h}_k \end{bmatrix}$, so that $\hat{\boldsymbol{H}}_k\boldsymbol{y} = \lambda\boldsymbol{y}$. Then

$$\boldsymbol{V}_{k+1}\boldsymbol{H}_k\boldsymbol{y} = \begin{bmatrix} \boldsymbol{V}_k & \boldsymbol{v}_{k+1} \end{bmatrix} \begin{bmatrix} \lambda\boldsymbol{y} \\ \|\hat{\boldsymbol{v}}_{k+1}\| \, y^k \end{bmatrix} = \lambda\boldsymbol{V}_k\boldsymbol{y} + \boldsymbol{v}_{k+1} \|\hat{\boldsymbol{v}}_{k+1}\| \, y^k,$$

so that by (2.113),

$$\|\boldsymbol{A}\boldsymbol{V}_k\boldsymbol{y} - \lambda\boldsymbol{V}_k\boldsymbol{y}\| = \|\hat{\boldsymbol{v}}_{k+1}\| \, \left|y^k\right| \tag{2.116}$$

where y^k is the last component of $\boldsymbol{y} \in \mathbb{R}^k$. If $\left|y^k\right|$ or $\|\hat{\boldsymbol{v}}_{k+1}\|$ is small, (2.116) implies that $(\boldsymbol{V}_k\boldsymbol{y}, \lambda)$ is an approximate eigenpair for $\boldsymbol{A}$. This observation is the basis for efficient methods to estimate the largest eigenvalues of $\boldsymbol{A}$ using Arnoldi iteration [56,57]. This is used in Paper III to estimate the largest eigenvalues of the operator $\mathcal{E} = I - \mathcal{A}$, because the alternative iterative solution to (2.103) given by [51]

$$g_0 = b \tag{2.116a}$$

$$g_{k+1} = \left[\sum_{j=0}^{k+1} \mathcal{E}^j \right] b = g_k + \mathcal{B}(d - \mathcal{R}g_k) \tag{2.116b}$$

exists provided all eigenvalues of $\mathcal{E}$ have magnitude less than unity.

Chapter 3

Seabed classification

This chapter provides a brief introduction to the field of seabed classification, including a review (by no means comprehensive) of some relevant literature. Other introductions may be found in [58] and [59]. The section on acoustic parameters (Sec. 3.2.2) is largely based on [1].

3.1 Overview

The purpose of acoustic seabed classification (ASC) and characterization is to map the physical composition, distribution, and acoustic properties of surficial seabed sediments. ASC is mainly applied to data from high-frequency surveying sonars, including single or multibeam echo sounders [5] and sidescan sonars [60]. The basic premise is that differences in sediment composition often imply differences in acoustic response. Provided the recorded echoes have been corrected for extraneous influences, such as water depth or instrument characteristics (see Sec. 2.1.3), spatial variations in seabed properties may be resolved with higher resolution than is practically possible by only taking sediment samples. Information about seabed composition is useful in a range of problems, including the development of marine renewable energy [61], sonar performance prediction, benthic habitat mapping and marine resource management [59, 62–67], environmental monitoring [68, 69], and geotechnical engineering [70]. A scientific understanding of the physical processes that form the seabed is also a goal in itself, and new acoustic techniques for seabed mapping are instrumental to achieve this.

There appears to be three main approaches to seabed sediment mapping with high-frequency acoustic data. First, the direct interpretation of geo-registered sidescan imagery in conjunction with bathymetry. A sidescan image mosaic shows the backscattered intensity of each geo-located resolution cell, corrected for transmission loss, source directivity, and the incidence angle-dependence of the scattering cross section, $\sigma(\theta)$. The latter correction may be obtained by assuming that scattering from the seabed is perfectly diffuse (isotropic). In this instance, the backscattered as well as the incident intensities are proportional to the geometrical cross section, and σ follows Lambert's law, $\sigma(\theta) = \mu \cos^2 \theta$, where μ is a constant. A sidescan image may therefore show the spatial variability of μ [71, 72]. Attribute images derived from DTMs, especially gradient, curvature, and hill-shade images, reveal geomorphological features that reflect erosional and depositional processes on the seabed. Taken together, images of seabed scattering strength and geomorphology provide a powerful interpretational tool for experts on

sedimentology and geology, and accordingly this approach is much used in the marine geology literature.

The second approach is physics-based parameter estimation (inversion). Formally, one assumes a known functional relation, $\boldsymbol{d} = F(\boldsymbol{m})$, between a feature vector $\boldsymbol{d}$ (measured data) and a set of model parameters $\boldsymbol{m}$ that represent the acoustic properties of the seabed (sound speed, density, attenuation coefficient, roughness spectrum coefficients, etc.). The feature vector commonly represents the echo envelope [73–75] or the (estimated) variation of σ with incidence angle [76–78]. As argued by Tarantola [4] and others, the complete solution to the inverse problem should be given as a posterior probability distribution over the parameter space, $p(\boldsymbol{m} \mid \boldsymbol{d}) \propto P(\boldsymbol{m})L(\boldsymbol{m})$, where $P(\boldsymbol{m})$ is the prior probability that represents a priori information about the model, and $L(\boldsymbol{m})$ is the likelihood function that measures how well a model $\boldsymbol{m}$ fits the data. However, this probabilistic approach does not appear to be widespread in high-frequency ASC work. Instead, point estimates for the model parameters are given, in particular the least squares best fit $\operatorname{argmin}_{\boldsymbol{m}} \|d - F(\boldsymbol{m})\|$, which is the maximum likelihood estimate $\boldsymbol{m}_{\mathrm{ML}}$ if one assumes that the observational errors are Gaussian distributed as $p(\boldsymbol{d} \mid \boldsymbol{m}) \sim N(F(\boldsymbol{m}), \sigma_D^2 \boldsymbol{I})$ for σ_D^2 a constant variance. Finding the best fit is an optimization problem that may be solved with several standard iterative techniques. However, point estimates are less useful if $p(\boldsymbol{m} \mid \boldsymbol{d})$ is multi-modal or flat. It is critical that the forward model, F, sufficiently accurately represents the physics of the process that generates the data, in particular the acoustic seabed interaction. This concern must be balanced against the practical concern that the forward model should be computationally fast and the number of model parameters sufficiently low so that efficient random sampling of the model space becomes feasible. Two models that have been used in practice are the echo time domain model of [73] and the scattering strength model of [79]. The latter accounts for both surface and volume scattering in a statistically homogeneous seabed. Surface scattering is computed using the Kirchhoff approximation near vertical incidence, and the composite roughness approximation for oblique angles (see [1]). The roughness spectrum (2.37) is assumed to be isotropic and on the form

$$W(\boldsymbol{k}) = \frac{w}{k^\gamma}, \tag{3.1}$$

i.e., a power law with two free parameters w and γ. Recalling equation (2.38) (and similar expressions), it is straightforward to substitute different theories for wave propagation in the sediment, e.g., fluid, elastic, or poro-elastic models. One model that has been used in practice for inversion of MBES data is the effective density fluid model of Williams [80], which is an approximation to the full Biot poro-elastic model, suitable for porous fluid-saturated granular materials (sands). A recent model that reportedly is promising for inversion of MBES data was presented in [81]. This model, however, uses only the Lambertian approximation for random scattering.

The third approach to ASC is the machine learning approach, which includes supervised and unsupervised statistical learning. Machine learning algorithms use data to make predictions on feature vectors. In principle, a prediction could be a regression for a continuous physical parameter such as density or mean grain size. However, probably as a result of the scarcity of training data, learning algorithms for sonar data are generally used for classification, i.e., the assignment of a feature vector to one of a discrete set of classes (seabed types). Training data are

feature vectors (measured data) obtained near sediment sampling sites or in other areas where the seabed character is known with reasonable confidence. The many examples of machine learning ASC include [38, 64, 82–84]. The most common types of features represent either the echo time evolution, the scattering strength versus angle relation, or the local area statistics of scattering strength in sidescan imagery.

Section 2.2 described the basic principles of statistical classification in the Bayesian framework. There are several other general classification methods that are being applied in ASC and that require no knowledge of underlying probability distributions. They include simple nearest-neighbor methods [85, Ch. 13], linear discriminant functions and their generalization, the support vector machines (SVMs) [3, Ch. 5], [85, Ch. 12], and the non-linear methods of artificial neural networks (ANNs) [85, Ch. 11], [3, Ch. 5]. Nevertheless, the classification problem is the same, namely to obtain a partition of the feature space. Arguably, the main distinction between various acoustic seabed classification (ASC) methods is what types of features are employed. Therefore, we concentrate on some of the characteristics of seabed sediments and the features that are commonly derived from seabed-interacting acoustic signals.

3.2   Physical characterization

3.2.1   Grain size

Granular materials such as unconsolidated or clastic sediments can be characterized by grain size, given as the diameter D or as $\phi = -\log_2(D/1\,\text{mm})$ (the Krumbein ϕ–scale). Table 3.1 shows the size terms according to the dyadic scale of Wentworth [86]. The principal classes silt, sand, and gravel are divided into subclasses with the qualifiers *very coarse* (excl. silt), *coarse*, *medium*, *fine*, and *very fine*. The principal classes have overall different modes of transport by running water, from traction (gravel) to colloidal suspension (clay) [87]. There is also a rough correspondence between the principal classes and main forms of rock erosion: breakage (gravel), abrasion and granular disintegration (sand and coarse silt), and chemical decay (clay) [88]. The grains of seabed sediment deposits are however not uniformly sized, and samples are appropriately characterized by the grain size distribution: the cumulative distribution of mass fraction as a function of increasing particle size. In a laboratory, the grain size distribution may be determined by sieving a sample through progressively finer meshes down to $0.063\,\text{mm}$; the distribution of finer grains may be determined, e.g., by measuring X-ray attenuation through a suspension. The grain size distribution is sometimes, particularly in acoustical work [89], represented by a single summary statistic, the ϕ-scale mean grain size M_z, which is often estimated as [90]

$$M_z \approx \frac{\phi_{16} + \phi_{50} + \phi_{84}}{3} \ \text{ or } \ M_z \approx \frac{3\phi_{10} + 4\phi_{50} + 3\phi_{90}}{10}, \tag{3.2}$$

where ϕ_p denotes the particle size $\phi(D)$ at the pth percentile of the cumulative grain size distribution. Mixed sediments can be distinguished by whether they are matrix supported or grain supported [1, Sec. 3.2]. In matrix-supported sediments, the coarse particles are suspended without contact in predominantly fine-grained material. In grain-supported sediments, the coarse particles are in physical contact, forming a porous structure that may be filled with either fluid or clay-sized particles.

Table 3.1: Definition of sediment types

Type	Grain size (mm)	ϕ [a]
Clay	< 0.002	> 9
Silt	0.002–0.0625	9 to 4
Very fine sand	0.0625–0.125	4 to 3
Fine sand	0.125–0.25	3 to 2
Medium sand	0.25–0.5	2 to 1
Coarse sand	0.5–1.0	1 to 0
Very coarse sand	1.0–2.0	0 to -1
Very fine gravel	2.0–4.0	-1 to -2
Fine gravel	4.0–8.0	-2 to -3
Medium gravel	8.0–16	-3 to -4
Coarse gravel	16–32	-4 to -5
Very coarse gravel	32–64	-5 to -6
Cobble	64–256	-6 to -8
Boulder	> 256	< -8

[a] $\phi = -\log_2(D/1\,\mathrm{mm})$, where D is particle diameter.

A simplified classification system for mixed sediments is provided by the Folk triangle (Fig. 3.1) [91], which has ten classes defined according to the proportions of sand, silt, and clay. Figure 3.2 shows a modified version of Folk's system, where the principal classes are sand, gravel, and mud (clay and silt combined) [92]. Both systems combined were applied to classify the sediment samples used as supporting data in the acoustic classification experiments of Papers I and II [10].

3.2.2 Acoustic parameters

Recalling Eqn. (2.4) (p. 17), the monostatic seabed scattering cross section for small-roughness surface scattering may be expressed on the form $\sigma_{\mathrm{surf.}} \propto |A|^2 W$, where W is the roughness spectrum encoding the statistics of the surface relief. The factor A depends on the acoustic properties of the seabed and water, including the flat-interface (no roughness), plane wave reflection coefficient V. For surface scattering from a homogeneous fluid sediment, the following expression may be derived [1, App. K,M]:

$$A = \frac{1}{2}\left[1 + V(\theta, \rho_b/\rho, c_b/c)\right]^2 G(\theta, \rho_b/\rho, c_b/c),\qquad(3.3)$$

where G, like V, is a function of incidence angle, the seabed-water mass density ratio, ρ_b/ρ, and the seabed-water sound speed ratio, c_b/c. Here, the seabed sound speed is a complex number,
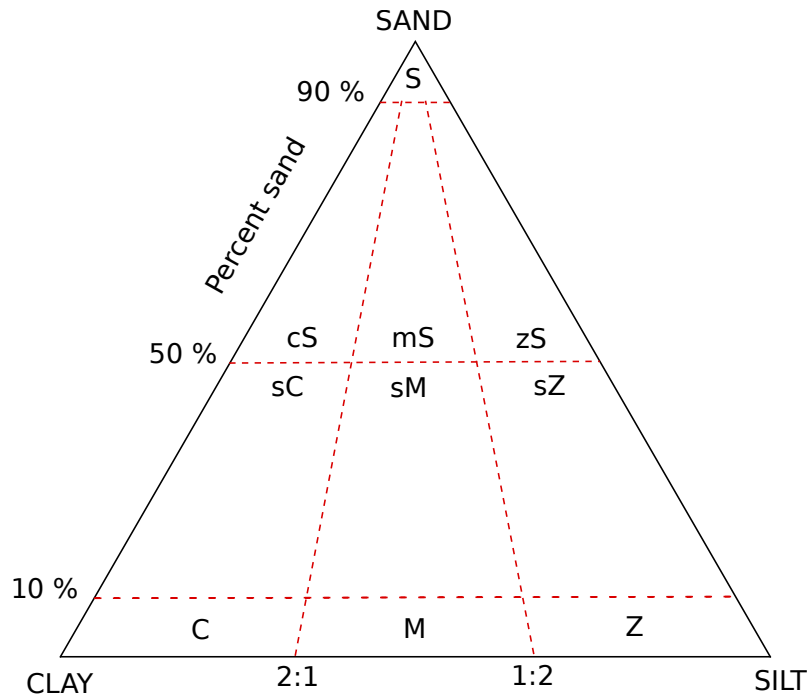
Fig. 3.1: Folk's sediment classification system [91]. The main classes are sand (S), silt (Z), and clay (C), cf. Table 3.1. A mix of clay and silt is termed mud (M). Subclasses with the qualifiers sandy (s), muddy (m), clayey (c), or silty (z) are mixtures with proportions as indicated in the triangle.
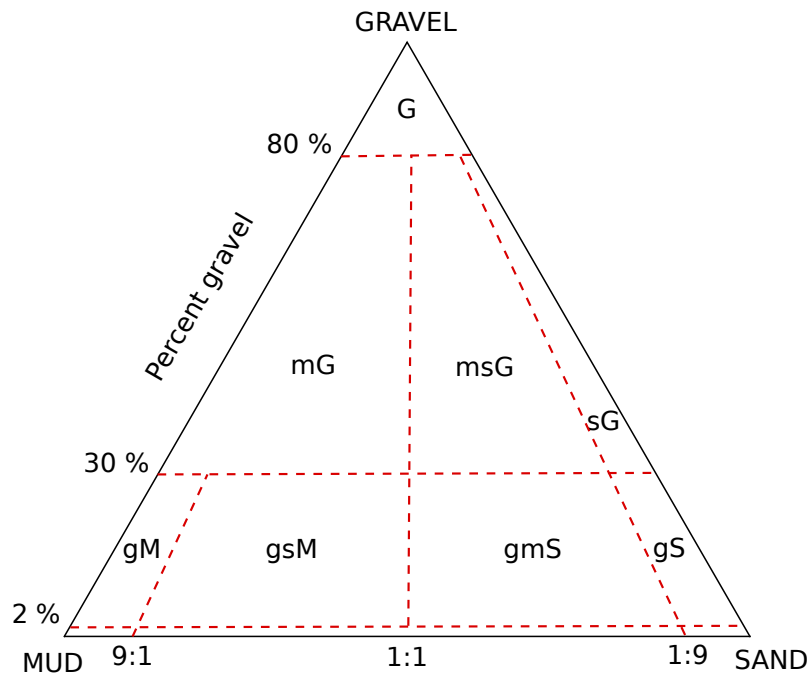

Fig. 3.2: An alternative version of Folk's classification system, by the Geological Survey of Norway (NGU) [92]. The main classes are sand (S), mud (M), and gravel (G), cf. Table 3.1. Subclasses with the qualifiers sandy (s), muddy (m), or gravelly (g) are mixtures with proportions as indicated in the triangle. A sediment is gravelly if it contains more than 2 % gravel. Note that the original gM class in [92] has been split into gM and gsM, as in [10].

to account for attenuation. Equivalently, the wave number is complex, i.e.,

$$k_b = \frac{\omega}{c_b} = \Re k_b + i \Im k_b, \tag{3.4}$$

and the sound speed ratio may be expressed as

$$\frac{c_b}{c} = \frac{v_b/c}{1 + i\delta_b} \tag{3.5}$$

where $v_b = \omega/\Re k_b$ is the (real) phase speed in the seabed and $\delta_b = \Im k_b / \Re k_b$. In terms of these parameters, the seabed attenuation coefficient in units of $\mathrm{dB/m}$ [cf. Eqn. (2.19)] becomes

$$\alpha_b = \frac{20}{\log 10} \frac{\omega \delta_b}{v_b}. \tag{3.6}$$

Similarly, in the simplest model of high-frequency volume scattering from a fluid sediment, the volume contribution to the scattering cross section is on the form $\sigma_{\mathrm{vol.}} \propto |T(\theta, \rho_b/\rho, c_b/c)|^4 \, \sigma_v$, where T is the water-sediment transmission coefficient, and σ_v is the local volume scattering cross section, assumed to be constant [1, Sec. 14.1]. A minimal six-parameter model for monostatic scattering, $\sigma(\theta) = \sigma_{\mathrm{surf.}}(\theta) + \sigma_{\mathrm{vol.}}(\theta)$, has been specified in [79, Ch. 4] using the following parameterization: ρ_b/ρ, v_b/c, δ_b, σ_v/α_b, γ, and w, where the two latter specify the isotropic roughness spectrum in (3.1).

There are many simplifications in this model. Seabed sediments are normally fluid-saturated porous media (two-phase media), and will behave acoustically like a fluid only when the porosity, β, defined as the ratio of pore fluid volume to total volume, is high (critical porosity [93]). The simplest parameters to obtain experimentally from sediment samples are porosity, density, and grain size, and there are strong correlations between v_b/c and, respectively, β, ρ_b, and M_z [1, Ch. 5]. However, the correlation with attenuation is weaker, and it seems difficult to relate roughness parameters (γ, w, or RMS roughness h^2) to physical properties like grain size [94]. Moreover, the roughness spectrum can change over time, sometimes in a matter of hours, due to hydrodynamic forcing or biological processes. This aspect, too, must be kept in mind when interpreting ASC results or inversion results. Considering in particular that σ is sensitive to the parameter w and σ_v/α_b [79, Ch. 4], this is arguably a reason to treat σ itself as a random variable, as in Paper II.

3.3 Features

3.3.1 Intensity versus angle

Equation (2.30) relates the measured (incoherent) mean square pressure to the angle-dependent scattering cross section $\sigma(\theta)$. In this and the following sections, we will assume that the measured squared pressure envelope y, or the proportional quantity intensity, $y/\rho c$, has been corrected for source strength, propagation loss, and array directivity. Using notation from Sec. 2.1.4, we define

$$y_{\mathrm{corr}} = \frac{y}{s_0^2 Q^2 \int_{\mathbb{R}^2} \Omega(\boldsymbol{r}, \boldsymbol{r}_0)}. \tag{3.7}$$

By (2.30), the average of y_{corr} is an estimate for the scattering cross section σ in a fixed ensonification geometry. Since virtually all ASC work involves a corrected intensity, we generally drop the subscript on y_{corr} and write simply y. Applied to each MBES beam separately, the combined result yields an estimate of $\sigma(\theta)$ across the angular range of the MBES. This functional relationship is the basis for the classification method developed in Paper II.

In seismology, it has been recognized that reflected seismic wave amplitudes vary with the horizontal distance (offset) between source and receiver. This amplitude variation may be used to determine geoacoustic properties, in particular wave speed contrasts, at reflecting subsurface interfaces. This observation has led to the amplitude versus offset (AVO) technique in seismic data analysis. In analogy to AVO, Fonseca and Mayer [76] have developed angular range analysis (ARA), a parameter estimation method for MBES data. The feature vector, the ARA parameters, includes the slope and intercept of the $\sigma(\theta)$ curve. Moreover, the curve is divided into three separate angular ranges, a near-nadir range $[0°, 25°)$, a far range $[25°, 55°)$, and an outer range $[55°, 85°)$. The model parameters include water-to-sediment density and sound speed ratios, five other geoacoustic parameters, and the two parameters of the power law roughness spectrum (3.1). From estimates of w and γ, the RMS roughness $h^2 = B(0)$ may be obtained [cf. (2.35)–(2.37)]. A key assumption is that certain features correlate with certain model parameters; in particular, the slope is strongly correlated with roughness, while the intercept is correlated with impedance contrast. The inversion algorithm therefore proceeds in steps, where some model parameters are adjusted to fit an ARA-parameter while other model parameters are held fixed. Another assumption is that the model parameters are constant along the whole swath. One procedure advocated in [95] is to visually inspect sidescan images to ensure that the ARA method is applied to data from an acoustically homogeneous area.

Model fitting results for observed $\sigma(\theta)$ curves are also reported in [77, 78, 96], where the power law spectrum (3.1) is assumed. The parameter estimates of [77] are consistent with independent estimates made by inversion of a time-domain (echo shape) model applied to single beam echo sounder data in [97]. However, there is generally a lack of experimental data on the important roughness (and volume) scattering parameters, which makes it difficult to verify the inversion results. There also appears to be little discussion of model ambiguity, confidence intervals for the parameter estimates, or the significance of theory errors. Machine learning methods, including Bayesian classifiers, neural networks, and cluster algorithms, have also been applied to $\sigma(\theta)$ data in several works [98–102]. Simulation results in [99] showed best performance with a Bayesian classifier, albeit on data with known statistical distributions. In a comparison of classifiers in [102], the Bayesian classifier performed worse than, e.g., a SVM classifier, in a high-dimensional problem (71 features). However, it is difficult, by comparing results on different datasets, to draw any conclusions on general classifier performance for ASC problems.

Figure 3.3 illustrates how the angular dependence of S can be used to distinguish between sediment types, and shows a way of forming feature vectors with constant function approximations. This method is used in Paper II. An alternative to either this or the ARA feature vector may be obtained with the Generic Seafloor Acoustic Backscatter (GSAB) model [103, 104]. This is a six-parameter heuristic model for the observed angular response

$$y(\theta) = Ae^{-\theta^2/2B^2} + C\cos^D\theta + Ee^{-\theta^2/2F^2}. \tag{3.8}$$
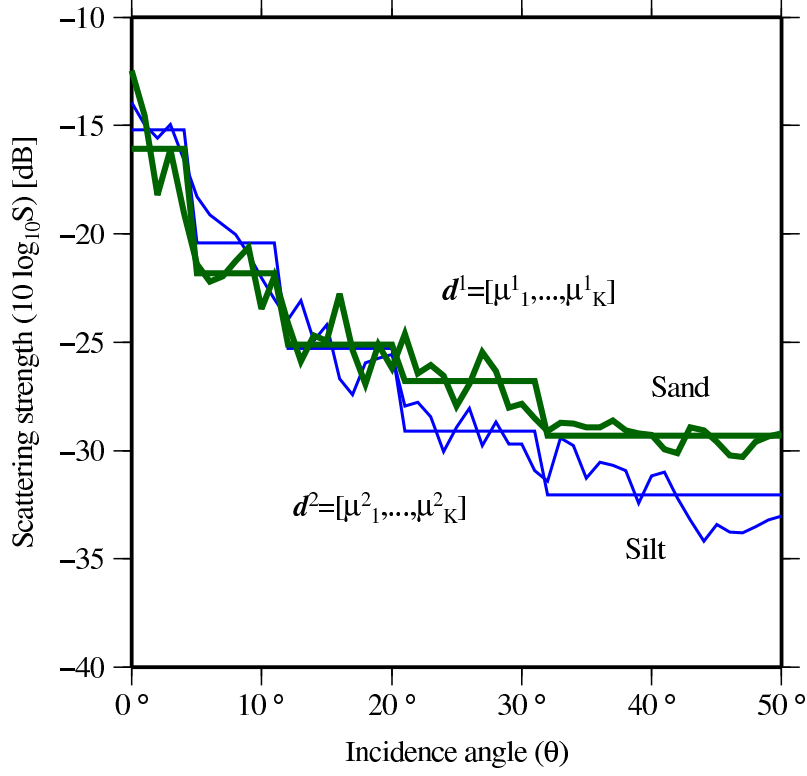
41

Fig. 3.3: Two observations of scattering strength versus incidence angle, $S(\theta)$ (irregular curves), formed by averaging data from 10 consecutive pings (from Paper II). The angular dependence of S makes it possible to distinguish between sediment types: sand (thick, green lines) and silt (thin, blue lines) in this example. In Paper II, observations are represented by piecewise constant approximations. A K-dimensional feature vector, $\boldsymbol{d}$, comprises the mean values in K angular segments (here, $K = 5$).

Here the first Gaussian function models the response in the near-nadir, specular region, while the second term is a modified Lambert law for scattering at oblique angles, and the last Gaussian functions accounts for a transition between the to angular regimes. The six parameters A to F, obtained by fitting observed data to (3.8), form a feature vector that may be used with machine learning or inverse methods, or as useful descriptors for direct interpretation of observed data [104].

Finally, the $\sigma(\theta)$ curve has also been used as an aid for more direct geological interpretation and identification of sediment facies, see, e.g., [105] and [106].

3.3.2 Intensity distributions

In Sec. 2.1.4 it was noted that the observed, corrected pressure envelope, x, should be Rayleigh distributed when the ensonified area is sufficiently large, and then, consequently, the squared envelope y should be exponentially distributed. However, deviations from Rayleigh statistics have been observed in real data. Equation (2.51) in Sec. 2.1.4 is a more general distribution that depends on an unknown distribution for the average scattering cross section. If a parametric form for $p_Y(y)$ or $p_X(x)$ is assumed, the parameter estimates may be used as features. Helle-quin et al [34] present histograms of measured corrected intensity Y that show deviations from

Rayleigh statistics, especially on gravelly and rocky (rough) seafloors. A product model for Y (see page 19) is introduced, based on [33], which gives a K-distribution for Y, i.e.,

$$p_Y(y) = \frac{2}{\Gamma(\alpha_{\text{eff}})y}(\beta y)^{(\alpha_{\text{eff}}+1)/2}K_{\alpha_{\text{eff}}-1}\left(2\sqrt{\beta y}\right). \tag{3.9}$$

Here K_n is the modified Bessel function of the second kind of order n. The two parameters are $\beta = \alpha_{\text{eff}}/\langle y \rangle$ and α_{eff}, the effective shape parameter, which depends on the relative size of the ensonified area and correlation length of the local, position-dependent scattering cross section σ. There is a close similarity between this statistical model and the models of McDaniel [23] and Jackson and Richardson [1, Sec. 16.3] discussed in Sec. 2.1.4. However, the latter do not assume a specific distribution for the (average) scattering cross section. It is demonstrated that the K-distribution can be fitted to the observed histograms. Moreover, the shape factor α_{eff} is incidence angle-dependent and is shown to be a good feature for discriminating between seabed types. These aspects are explored further in [107] and [108].

The interest in the K-distribution as a statistical model for clutter and reverberation seems to originate in work by Jakeman and Pusey [24, 26, 27]. They consider the discrete scattering model for the received electric field (cf. discussion on page 18),

$$E(r,t) = e^{it}\sum_{i=1}^{N} a_i(\boldsymbol{r},t)e^{i\varphi_i(\boldsymbol{r},t)} = e^{it}A(\boldsymbol{r},t)e^{i\Phi(\boldsymbol{r},t)}, \tag{3.10}$$

where a_i and φ_i ($i = 1,\ldots,N$) are independent random variables and the phases φ_i are assumed to be uniformly distributed over $[0,2\pi)$. It is shown that for finite N, the distribution for A is a K-distribution if the amplitudes a_i are themselves K-distributed. Furthermore, if the number of scatterers N is allowed to vary according to a negative binomial distribution, A will be K-distributed, independent of the amplitude (a_i) statistics, though only in the large N limit. The physical interpretation of the negative binomial distribution is that scatterers tend to occur in bunches, as from ripples on top of larger swells. This is, however, only a heuristic model, although Jakeman and Pusey show that the K-distribution fits data very well in various scattering experiments (microwaves on sea surface, light through turbulent fluid layers, starlight scintillation).

Another model that predicts a K-distribution for reverberation envelope statistics was developed in [28, 29]. The seabed is considered as a finite set of scattering objects (patches) of varying size. The patch size is assumed to be exponentially distributed and proportional to the scattering strength. The authors show that this assumption, also, leads to a K-distribution for the envelope $|A|$ assuming an expression similar to (3.10). This conclusion holds for finite N and the shape factor is $\alpha = N/2$. Thus there is a link between the parameters of the K-distribution and physical properties of the seafloor (number of scatterers, average scatter size). The choice of the exponential distribution is heuristic, but apparently justified by observation of size distributions of geological structures. Note that the K-distributions obtained in [28] and [24] are for the modulus of the complex (pressure) envelope, while the K-distribution in (3.9) is for the corrected, squared envelope.

A different approach is taken by Simons and Snellen [109]. They consider the statistical

distribution of the quantity

$$\bar{y} = \frac{1}{N_s} \sum_{i=1}^{N_s} 10 \log_{10} y \qquad (3.11)$$

for a single fixed beam; N_s is the number of resolution cells in the beam footprint. If N_s is large and each resolution cell is an independent secondary source, with random scattering strength drawn from identical distributions, it is expected by the central limit theorem (CLT) that the distribution for $\bar{y}$ approaches a normal distribution. Simons and Snellen consider the histogram for $\bar{y}$ for the whole survey dataset (for a fixed beam) and assume that each present seabed type contributes a normal component. Thus the proposed statistical model is a one-dimensional normal mixture distribution

$$p_Y(y \mid \boldsymbol{w}, \boldsymbol{\mu}, \boldsymbol{\xi}) = \sum_{k=1}^{c} w_k e^{-(y-\mu_k)/2\xi_k^2}. \qquad (3.12)$$

There are $3c + 1$ unknown parameters: $\boldsymbol{w}, \boldsymbol{\mu}, \boldsymbol{\xi}$, and the number of seabed classes c. These parameters are determined by a curve fitting procedure for the histogram. With all parameters determined, each component in (3.12) is interpreted as the class-conditional PDF for a single seabed class, and the Bayesian decision rule is subsequently be applied to classify each observed value $\bar{y}$. This is therefore, effectively, a Bayesian unsupervised classification method (Sec. 2.2.2) in a one-dimensional feature space. This method has the significant advantage (along with QTC) that the classes can be determined from the data. One the other hand, the use of a single feature and a single beam may be a limitation.

3.3.3 Echo descriptors

Echo descriptors have been mainly used for single-beam ASC. Echoes must be 1) aligned at the first bottom-reflected (backscattered) time sample, 2) compensated for propagation loss, 3) compensated for the relative stretching or compression in time that occurs due to variable water depth, and 4) padded or truncated to a fixed length of some N time samples. Such a normalized echo may be viewed as a feature vector $\boldsymbol{x} \in \mathbb{R}^N$. However, as noted in Sec. 2.2.3, it is usually desirable to employ a smaller set of features that captures the essential information in the data. A single feature f_i may in principle be any function (algorithm) $f_i : \mathbb{R}^N \to \mathbb{R}$; a feature vector with D components is then simply $\begin{bmatrix} f_1(\boldsymbol{x}) & \dots & f_D(\boldsymbol{x}) \end{bmatrix}$.

The most important feature appears to be the total energy

$$f(\boldsymbol{x}) = E = \sum_{k=1}^{N} x_k^2. \qquad (3.13)$$

The total energy depends on the reflection coefficient by $R \propto \sqrt{E}$. With calibration, and combined with regression relations for sound speed and density as functions of mean grain size [110], this has been used to obtain maps of grain size distribution [111]. Some general, efficient feature reduction methods that may be applied to echoes are principal component analysis (PCA) and thresholding in the frequency or wavelet transform domains. van Walree et al. [84] employ five features, besides energy, that describe the time spread, echo asymmetry,

fractal nature, and power spectral skewness of the echo signal. Several other authors also use fractal descriptors, e.g., the Hausdorff dimension.

Regarding the echo amplitude as a random variable, the histogram of an echo $\boldsymbol{x}$, denoted $p(v)$, represents the probability (relative frequency) of observing amplitude $x = v$. Suppose that $\boldsymbol{v} = \left[v_1, \ldots, v_L\right]$ represents the bins on the amplitude scale with L levels. The mth order statistical moment is defined as [112, Sec. 11.2.4]

$$\mu_M(\boldsymbol{v}) = \sum_{k=1}^{L} (v_k - \bar{v})^m p(v_k) \tag{3.14}$$

where

$$\bar{v} = \sum_{k=1}^{L} v_k p(v_k). \tag{3.15}$$

The first few statistical moments are used as shape descriptors in image analysis, and should be equally applicable for echo classification. Another possibility we have not seen in the sonar literature is the bin smoother discussed in Sec. 2.2.3.

3.3.4 Texture and geomorphology

Texture measures [112, Sec. 11.3.3] are used for image segmentation and classification in a broad range of applications. In ASC, texture measures have mostly been applied to backscatter image mosaics from side-scan sonars or MBESs. Two important groups of texture measures are based on first and second order statistics, respectively. The simplest approach derives features from the local pixel intensity distribution (histogram) $p(v)$: mean, variance, higher moments, uniformity $U(\boldsymbol{v}) = \sum_{k=1}^{L} p^2(v_k)$, entropy $H(\boldsymbol{v}) = -\sum_{k=1}^{L} p(v_k) \log_2 p(v_i)$. Second order statistics may be obtained with gray level co-occurrence matrices (GLCMs). A GLCM, $\boldsymbol{G}$, describes the statistics of pixel pairs with a fixed relative position, defined by a row and column displacement $\boldsymbol{\rho} = (\Delta x, \Delta y)$. With L pixel intensity bins $\boldsymbol{v}$, $\boldsymbol{G}$ is an $L \times L$ matrix where G_{ij} is the probability that two pixels with separation $\boldsymbol{\rho}$ have intensities v_i and v_j, respectively. Certain attributes, derived from GLCMs in [113], are widely used as texture measures. Some examples from ASC can be found in [38, 114–116].

Another form of texture measure, that does not appear to be widespread in ASC, is the local binary pattern (LBP) [117]. The basic LBP is formed by thresholding the pixel values of an 8-neighborhood with respect to the center pixel. This produces a string of eight zeros and ones, which is interpreted as a binary number in the range 0–255. The distribution (histogram) of LBPs in an image block of size $S \times S$ pixels may be used as a measure of the texture in that block [118]; the number of features equals the number of bins in the histogram. Since this gives very high-dimensional feature vectors (256 components), it is common to consider only the so-called uniform patterns corresponding to the binary strings with at most two transitions between 0 and 1. There are 58 such patterns for the 8-neighborhood [119]. LBPs are invariant with respect to uniform changes in image intensity. Furthermore, by identifying all patterns that are equal modulo a circular bit shift, one obtains a set of rotation-invariant distributions with fewer features (36 general and nine uniform) [120].

In Paper III, LBPs are applied to a residual terrain model to classify geomorphology, an-

ticipating that geomorphological patterns will contribute additional valuable information about seabed geology. More generally, features derived from terrain analysis and geostatistics should complement acoustic signal features. Some examples of terrain analysis applied to ASC can be found in [45] and [46].

Chapter 4

Summary of research

4.1 Paper I

This paper presents a new map of surficial seabed sediments in a 23.000 km^2 area of the Norwegian Channel (Norwegian Trench) (NC) and North Sea plateau. The map was obtained by applying the QTC Impact ASC system to single-beam echo sounder data and interpreting the segmentation results through analysis of supporting data. The supporting data included sediment samples, parametric sonar profiles, shallow seismics, and multibeam bathymetry. A part of the paper is devoted to an analysis and experimental evaluation of the QTC algorithm, which is a widely used industrial-strength method for acoustic sediment mapping.

The QTC algorithm has three main steps: 1) data normalization with respect to extraneous factors, including the effect of variable water depth, 2) echo shape analysis, feature extraction, and feature reduction, and 3) a stochastic clustering algorithm based on the assumption that observations from a sediment class are normally distributed. The acoustic classification results were highly sensitive to the sand and gravel content in the top layer, generally consistent at track line intersections and for different pulse lengths, and in good agreement with supporting data and earlier mapping work in an adjacent area. The normal distribution assumption (see Sec. 2.2.2, p. 24) did not hold. The EM clustering algorithm (Sec. 2.2.2, p. 23), which is fast and simple to implement, gave results very similar to QTC when applied to the same set of feature vectors. In contrast, the k-means and SOM clustering algorithms [35, 121] did not produce equally consistent results. A cluster validity analysis showed that there was no clear cluster structure in the data. Since the QTC algorithm tries to fit a normal mixture distribution to the observations, where the number of classes, the class means, and the class covariance matrices are all unknown, this lack of structure implies a degree of arbitrariness in the classification results. In particular, the number of classes had to be fixed by hand using physical considerations and supporting data.

In the study area, the NC forms a roughly 110 km wide depression (water depth 300 m) in the continental shelf, bounded by the shallow North Sea plateau to the west and the Norwegian coastline to the east. The central NC is a trap for water-borne sediments, and the top sediment layer is a Holocene marine deposition that is a few meters thick as observed in parametric sonar data. Data from the eastern part of the area, especially multibeam bathymetry, show traces of the ice dynamics of the last, Weichselian glaciation. Several large moraines with ploughmarks and

parallel lineaments are seen to extend north-eastwards from the coast. Massive glacial erosion has formed a flow-like pattern of parallel, large-scale depressions outside the Sognefjorden and Fensfjorden inlets, which is evidence that ice was drained from the mainland ice sheet in this area and deflected northwards into the NC Weichselian ice stream. The western side of the NC is marked by two erosional processes in combination: water current erosion and the seepage of vertically migrating fluid (gas). These processes have formed fluted patterns of elongated depressions (pockmarks), similar to those observed by Bøe et al. in the Skagerrak [122].

Based on the ASC results, the area can be divided into five main parts with the following sedimentological classes:

1. soft, fine-grained marine sediments (silt and mud) in the deep, central part of the NC;

2. a heterogeneous mix of clay, silt, and sand in the north-eastern coastal zone and on the lower western slope of the NC;

3. sandy mud with gravel, mainly on the moraines and upper western slope;

4. sand with varying amounts of gravel on the North Sea plateau;

5. clay and sandy clay in the fjords of western Norway.

The effect of glacial erosion and presence of glacigenic sediments, in combination with a variable but generally low marine sedimentation rate, explains the heterogeneous nature of the sediments on the eastern, near-shore side of the NC. The ASC results furthermore indicate that, closer to the western slope of the NC, sand is mixed with the fine-grained Holocene marine unit. We think a plausible explanation for this observation can be found in the work of Hovland [123], who based on analysis of shallow seismics from the western slope proposed that vertically migrating gas is lifting sediments into suspension above the elongated pockmarks; the suspended sediments are subsequently transported away by water currents.

The paper concludes that, overall, the QTC algorithm gave plausible and consistent results as applied to a large data set from a physically diverse area, but points out some possible ways to further develop the algorithm, see also Sec. 5.3 (p. 64) for more details.

4.2   Paper II

This paper develops a supervised statistical seabed classification method for multibeam echo sounders. A main objective is to evaluate the method's theoretical and experimental performance and, consequently, its practical applicability.

The first step is to estimate from MBES data the seabed scattering strength as a function of incidence angle, $S(\theta)$. The shape of $S(\theta)$ depends on the seabed reflection coefficient, the roughness spectrum, and the strength of volume scattering [1, Ch. 13–14]. For this reason, an estimate of the function $S(\theta)$ ($\theta_{\min} \leq \theta \leq \theta_{\max}$), subsequently referred to as an observation, is used as the input signal to the classification algorithm. The second step, feature reduction, consists of finding a piecewise constant function approximation for $S(\theta)$, so that $S(\theta)$ can be replaced by a vector $\boldsymbol{d} = \begin{bmatrix} d_1 & \ldots & d_K \end{bmatrix}^T$, where d_k is the average scattering strength in a segment $\theta_{k-1} \leq \theta < \theta_k$. The number of components in $\boldsymbol{d}$ represents a trade-off between

approximation accuracy and robustness, and the segment limits are determined such that the approximation error is minimum in the least squares sense for the dataset as a whole (including all classes). The mean values, d_k $(k = 1, \ldots, K)$, may be computed using measurements from several consecutive pings (spatial averaging). In addition, averaging takes place in the angular and time domains. It is argued using the multidimensional CLT [124, Ch. 11] that the class-dependent PDFs, $p(\boldsymbol{d} \mid \omega_i)$, can be approximated by multivariate normal distributions. The third step is therefore to estimate each $p(\boldsymbol{d} \mid \omega_i)$ by computing the sample mean and covariance matrix from a set of training observations, typically obtained from data in the vicinity of one or more sediment samples (ground truth). Finally, the Bayesian minimum risk decision rule (Sec. 2.2.1) is applied to classify new observations.

To obtain further quantitative predictions, it is assumed that the measured complex pressure, $\Pi(t)$, scattered at a fixed incidence angle θ from a random surface with a constant scattering cross section, s, is a Gaussian random process with independent, identically distributed real and imaginary parts (Sec. 2.1.4). The implication is that the squared envelope $Y = |\Pi(t)|^2$ (or intensity $Y/\rho c$) will have an exponential PDF, i.e., $f_{Y|S}(y \mid S = s) = (1/s) \exp\left(-y/s\right)$, with $s > 0$, $y > 0$. A seabed class is characterized by an (unknown) angle-dependent distribution for s, $f_{S|\Theta}(s \mid \theta)$, so that the compound distribution for the observed, corrected intensity becomes

$$f_{Y|\Theta}(y \mid \theta) = \int_0^\infty \frac{1}{s} e^{-y/s} f_{S|\Theta}(s \mid \theta) \, \mathrm{d}s. \tag{4.1}$$

From this starting point, error bounds for two-class classification problems are calculated for both linear intensity and log-intensity. To compare classifiers based on intensity and log-intensity, it is necessary to compare the means and variances for s and $\log s$ without knowledge of the underlying scattering cross section distribution $f_{S|\Theta}(s \mid \theta)$. To proceed further, the maximum entropy principle is applied by constraining the mean and variance of $\log s$. The results may be used to predict classifier performance in terms of the intra-class variance for s and the degree of spatial averaging, given the modeled or measured scattering cross sections as a function of θ. One conclusion is that applying the classifier to log-intensity yields the better classification accuracy.

To assess the performance and spatial resolution of the proposed method experimentally, a set of randomized tests on real data with four classes was devised to see how the classification accuracy depended on:

- the dimension K of the observation vector $\boldsymbol{d}$;

- the amount of training data;

- the amount of spatial averaging;

- the total range of incidence angles over which data were collected.

The best accuracy (95 %) was obtained for log-intensity data using only 3–5 components in the observation vectors, and the amount of training data could be limited to the equivalent of about 500 pings per class. The spatial averaging, which limits along-track spatial resolution, could be limited to 10–20 pings. Similar across-track resolution was attained by dividing the full swath into separate, independent angular sectors, but only at reduced accuracy (87 %). Furthermore,

it was verified that the empirical class-wise distributions $p(\boldsymbol{d} \mid \omega_i)$ were largely consistent with the Gaussian (normal) assumption in the sense that, by assumption, $p(\boldsymbol{d} \mid \omega_i) \propto \exp\left(-r^2/2\right)$, where

$$r^2 = (\boldsymbol{d} - \hat{\boldsymbol{\mu}}_i)^T \hat{\boldsymbol{\Sigma}}_i^{-1} (\boldsymbol{d} - \hat{\boldsymbol{\mu}}_i) \tag{4.2}$$

was found to be distributed as for a true multivariate normal distribution.

The method was also used to make a sediment map of an area stretching from the eastern margin of the North Sea plateau and into the central part of the Norwegian Channel. This area is contained in the larger study area of Paper I. The map shows a transition from sand to fine-grained sediments along the west-to-east axis, very much like the map from Paper I. A part of the area corresponding the heterogeneous class 2 of Paper I was classified as sand. This result may reflect the transport phenomenon mentioned in Sec. 4.1, whereby sand from the upper slope may accumulate at the foot of the slope. There was no available ground truth in this area, but two nearby sediment samples contained a large sand fraction in the top layer. Note that the difference in operating frequency (95 kHz for the MBES and 38 kHz for the SBES) implies that the MBES signals have lower penetration and are consequently less sensitive to the sub-surface layers.

4.3 Paper III

The origin of this work was an attempt to extend and improve the classification method of Paper II by including invariant terrain, or geomorphological, features in the algorithm, since morphological patterns reflect the erosional and depositional processes on the seabed. Terrain features are derived from the bathymetry data, e.g., from a triangulated surface or regular grid digital terrain model (DTM). By invariant features we mean intrinsic features that are independent of coordinate translations and rotations, absolute depth, or low-frequency trends in the DTM. The features are also required to not depend on externally imposed parameters such as the spatial relationship of pixel pairs in second-order statistics, or the direction of derivatives. To this end we investigated primarily local distributions of intrinsic geometrical properties related to the mean and Gaussian curvatures [125, 126], as well as rotation-invariant local binary patterns [119, 120]. These features are sensitive to noise or artifacts in the MBES bathymetry data, in particular motion-induced errors and track line artifacts. A denoising algorithm, the main theme of the paper, was therefore developed prior to the classification tests. The denoising results presented in the paper are from applications to gridded multibeam bathymetry, but tests on synthetic aperture radar (SAR) images (Sec. 5.1 and [127]) suggest that the algorithm may be useful for correcting other types of remote sensing imagery as well.

The algorithm is image transform based, and analogous to applying a band-stop filter in the Fourier transform (frequency) domain. As for frequency domain filters, a noisy image is transformed to a representation where the noise is easier to suppress. Here, the transform is the composite operator $\mathcal{R} \circ \mathcal{L}$, where $\mathcal{L}$ is an edge-sensitive operator and $\mathcal{R}$ denotes the GDB discrete Radon transform (Sec. 2.3.2). The operator $\mathcal{L}$ accentuates fine-scale intensity changes, and $\mathcal{R}$ focuses linear features because each DRT component is the sum of pixel intensities along a linear graph. The inverse transform is subsequently applied to obtain the corrected image. The

operator $\mathcal{L}$ is defined by

$$\mathcal{L}\boldsymbol{f} = \boldsymbol{f} * \boldsymbol{h}_\epsilon, \tag{4.3}$$

where $*$ denotes circular convolution and $\boldsymbol{h}_\epsilon$ satisfies

$$\lim_{\epsilon \to 0} \mathcal{F}[\boldsymbol{h}] = \mathcal{F}[\boldsymbol{h}_b]. \tag{4.4}$$

Here, $\boldsymbol{h}_b$ is the point spread function (PSF) of the discrete Laplace operator (2.92) (see Sec. 2.3.1). Provided $\epsilon \ll 1$, the operator $\mathcal{L}$ therefore has the same desirable properties as the standard discrete Laplacian, i.e., it is linear, rotation invariant, edge sensitive, noise sensitive, and a measure of curvature. Moreover, the inverse, $\mathcal{L}^{-1}$, exists and can be computed by direct deconvolution with respect to $\boldsymbol{h}_\epsilon$.

Inverting the DRT is equivalent to solving the over-determined linear system

$$\boldsymbol{R}\boldsymbol{f} = \boldsymbol{d}, \tag{4.5}$$

where $\boldsymbol{f} \in \mathbb{R}^{N \times N}$ (digital image or DTM), $\boldsymbol{R} \in \mathbb{R}^{P \times Q}$ is the matrix representation of the linear, recursive GDB algorithm with $P = 6N^2 - 2N$ and $Q = N^2$, and $\boldsymbol{d} \in \mathbb{R}^P$ is the image representation in the DRT domain. Three iterative methods for solving (4.5) were considered. First, noticing that the recursive GDB backprojection algorithm is equivalent to the transpose matrix $\boldsymbol{R}^T$, the normal equations $\boldsymbol{R}^T\boldsymbol{R}\boldsymbol{f} = \boldsymbol{R}^T\boldsymbol{d}$ were solved using the LSQR algorithm [128]. Second, we used the iterative algorithm of [51], which is expressed in terms of a Neumann series and an approximate inverse algorithm for the GDB transform, denoted $\mathcal{B}$ (with matrix representation $\boldsymbol{B}$). A numerical analysis showed that the initial rate of convergence of the Neumann series was higher than iterative solutions of the normal equations. However, the analysis also indicated that a necessary criterion for convergence, namely that the spectral radius of the matrix $\boldsymbol{1}_Q - \boldsymbol{B}\boldsymbol{R}$ is less than unity, only holds for for $N < 512$. Since the approximate solution $\boldsymbol{f}_k$ ($k = 0, 1, \ldots$) obtained with the Neumann series lies in the Krylov subspace (Sec. 2.3.3)

$$\mathcal{K}_{k+1}(\boldsymbol{A}, \boldsymbol{b}) \equiv \text{span}\{\boldsymbol{b}, \boldsymbol{A}\boldsymbol{b}, \ldots, \boldsymbol{A}^k\boldsymbol{b}\}, \tag{4.6}$$

where $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{R}$ and $\boldsymbol{b} = \boldsymbol{B}\boldsymbol{d}$, we tried instead to solve the least squares problems

$$\min_{\boldsymbol{x} \in \mathcal{K}_k(\boldsymbol{A}, \boldsymbol{b})} \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|, \ k = 1, 2, \ldots. \tag{4.7}$$

This is accomplished with the GMRES (generalized minimum residual) algorithm, which showed the fastest convergence, also for $N \geq 512$.

In the DRT domain, linear artifacts are focused to high-intensity spots that may be suppressed by resetting the matrix elements to zero. When the linear artifacts are parallel, which is typical for track line artifacts, the high-intensity spots appear in a single narrow vertical band in the DRT. It is, however, not necessary to assume that the lines are parallel, periodic, or have a particular orientation. In MBES data, a second noise signature is conjectured to be linked with motion-induced errors because of its orientation (mostly concentrated at angles perpendicular to the track lines). This noise may be suppressed with a small median filter, $\mathcal{M}$. The composite operator $\mathcal{R} \circ \mathcal{L}$ is applied to the detrended image $\boldsymbol{f} - \mathcal{T}[\boldsymbol{f}]$, where $\mathcal{T}[\boldsymbol{f}] \sim \boldsymbol{f}$ is an approxima-

51

tion that does not preserve noise or artifacts. If this step is omitted, large values of $\mathcal{L}[\boldsymbol{f}]$ along the boundary of the image may be altered by the DRT domain filter, causing distortions in the restored image. The approximation operator is implemented as a least squares approximation to a tensor product basis of Chebyshev polynomials. The full algorithm can be summarized in seven steps:

Algorithm 1

Precondition: $\boldsymbol{f} \in \mathbb{R}^{N \times N}$ is a square image of size $N \times N$ pixels, where N is a power of 2
Require: Index, c, of DRT column corresponding to range direction
Require: Half-width, w (number of DRT columns), of band-stop filter

1: **function** DENOISE($\boldsymbol{f}$)
2:     $\boldsymbol{t} \leftarrow \mathcal{T}[\boldsymbol{f}]$
3:     $\boldsymbol{r} \leftarrow \mathcal{R}[\mathcal{L}(\boldsymbol{f}-\boldsymbol{t})] = \mathcal{R}[\boldsymbol{h}_\epsilon * (\boldsymbol{f}-\boldsymbol{t})]$ ▷ circular convolution
4:     $\boldsymbol{r} \leftarrow \mathcal{M}\boldsymbol{r}$ ▷ median filter (3×3 pixels)
5:     **for** $j \leftarrow c-w$ to $c+w$ **do**
6:         **for** all rows i **do**
7:             $r_{i,j} \leftarrow 0$ ▷ directional "band-stop filter"
8:     $\boldsymbol{g} \leftarrow \mathcal{L}^{-1}[\mathcal{R}^{-1}(\boldsymbol{r})] = \mathcal{F}^{-1}[\mathcal{F}[\mathcal{R}^{-1}(\boldsymbol{r})]./\mathcal{F}[\boldsymbol{h}_\epsilon]]$ ▷ element-wise division (./)
9:     **return** $\boldsymbol{g}+\boldsymbol{t}$

The algorithm was tested on three MBES datasets from the North Sea. A classification experiment on an area with three geomorphological classes was carried out to quantify the effect of the algorithm. As descriptors (features) of geomorphology we used empirical cumulative distribution functions (ECDF) of the mean curvature, Gaussian curvature, or the Laplacian, in 16×16 or 32×32 pixels blocks. As an alternative empirical distribution-based feature set we used histograms of rotation-invariant local binary patterns (same block sizes). For the corrected (DRT-filtered) DTM the classification accuracy was good: about 94 % (curvature) and 92 % (LBP) for 32×32 pixels blocks, and about 84 % (curvature) and 85 % (LBP) for 16×16 pixels blocks. For the unprocessed DTM the results were 10–28 % worse.

Attribute images derived from the corrected DTMs compared favorably to corresponding images obtained by applying conventional low-pass filters to the uncorrected DTMs or attribute images directly. Conventional filters caused greater smoothing of physical features and less reduction of artifacts compared to the DRT-based algorithm. The algorithm given above did not and should not be expected to suppress motion-induced noise completely. However, a DRT filter designed to suppress structures nearly perpendicular to the track lines was quite effective when applied locally in small, non-overlapping blocks of, e.g., 32×32 pixels. For small blocks the matrix representation for the inverse DRT can be explicitly computed in advance. This fact and the feasibility of non-overlapping block processing ensures that the method is fast.

Chapter 5

Further work

This chapter discusses in brief some possible extensions of research presented in Paper I–III. Some preliminary results are included.

5.1 Image analysis

Synthetic aperture radar: Stripe noise or linear artifacts degrade other types of remote sensing imagery besides multibeam bathymetry. One example is synthetic aperture radar (SAR) satellite imagery of the Earth's surface. In burst mode SAR imaging, particularly ScanSAR, where the radar beam is switched between different elevation angles to achieve a wider swath, an antenna pattern artifact known as scalloping may appear [129, Sec. 5.3]. Figure 5.1b and Fig. 5.1d show two original SAR image frames with 1024×1024 pixels. Scalloping is an amplitude modulation that in Fig. 5.1 is seen as thin, diffuse, horizontal stripes. The scalloping effect in Fig. 5.1 is more subtle than the MBES track line artifacts studied in Paper III (Sec. 4.3): the noise signature (at $0°$ projection angle) after applying the forward operator $\mathcal{R} \circ \mathcal{L}$ (invertible Laplace operator plus GDB transform) is faint. Nevertheless, the denoising algorithm of Paper III, with DRT values reset to zero in a narrow angular sector about the horizontal direction, produced the visibly improved images in Fig. 5.1b and Fig. 5.1d. The procedure was the same as for the MBES data in Paper III, except that the median filter in the DRT domain was left out (the median filter was only intended to remove motion-induced noise in MBES data). The application of the invertible Laplace operator was an essential step; applying the DRT directly to the original images did not yield satisfactory results. The work on the descalloping filter is described in more detail in [127]. Although limited in scope, the SAR tests suggest that the proposed denoising method may be useful also in other remote sensing applications.

Radon transform inversion formula: In Paper III, a key problem is to solve the equation

$$g = \mathcal{R}(\Delta_\epsilon f) = \mathcal{R}(f * h) \tag{5.1}$$

with respect to f in the 2D discrete setting. Here, f represents a restored image, g is the given output of a filter applied to a noisy image in the Radon domain, and $\Delta_\epsilon \equiv \mathcal{L}$ is an approximation to the Laplace operator obtained by convolution with the PSF h. It is therefore useful to note the relationship between the Laplace operator and the Radon transform in $\mathbb{R}^n$. For $f : \mathbb{R}^n \to \mathbb{R}$,
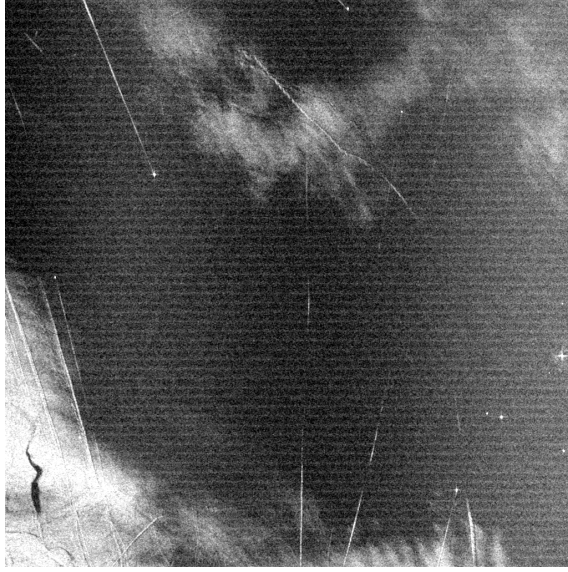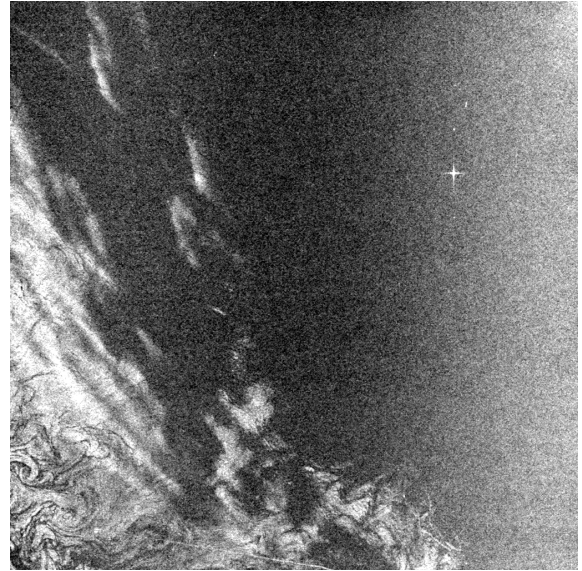
(a) Original SAR image (I)　　　　(b) DRT-filtered SAR image (I)
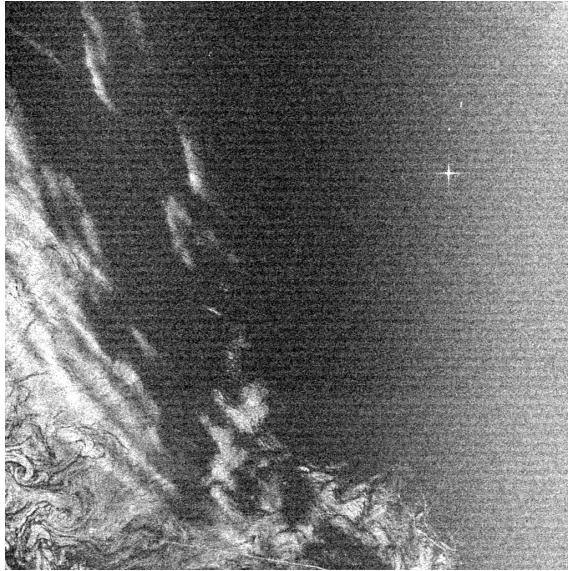
(c) Original SAR image (II)　　　　(d) DRT-filtered SAR image (II)

Fig. 5.1: Effect of the descalloping filter applied to two 1024×1024 pixels frames from an Envisat Advanced Synthetic Aperture Radar image of the southern North Sea [127] ©ESA/KSAT.

the fractional Laplace operator is defined by

$$\mathcal{F}\left[(-\Delta)^{\alpha/2} f\right](\xi) = |2\pi\xi|^{\alpha}\, \mathcal{F}\left[f\right](\xi), \tag{5.2}$$

and the Riesz potential $I^{\alpha} f$ on $\mathbb{R}^{n}$ is defined by

$$\mathcal{F}\left[I^{\alpha} f\right](\xi) = |2\pi\xi|^{-\alpha}\, \mathcal{F}\left[f\right](\xi),\ \alpha < n. \tag{5.3}$$

For $g : C^{n} \to \mathbb{R}$, where C^{n} is the unit cylinder defined in (2.99), the Fourier transform of $g(\theta, s)$ acts by definition only on the second argument, s, so that

$$\mathcal{F}\left[I^{\alpha} g\right](\theta, \sigma) = |2\pi\sigma|^{-\alpha}\, \mathcal{F}\left[g\right](\theta, \sigma),\ \alpha < 1, \tag{5.4}$$

where $\theta \in S^{n-1}$. In particular,

$$I^{-2}g = -\frac{\partial^2}{\partial s^2}g(\theta, s) \equiv -g''. \tag{5.5}$$

If w is a smooth (infinitely differentiable) and rapidly decreasing function[1], then for $\alpha < n$, inversion formulas are given by [130, Theorem 2.5]

$$w = c_n I^{-\alpha}\mathcal{R}^+ I^{\alpha-n+1}\mathcal{R}w \tag{5.6}$$

where c_n is a constant. Choosing $\alpha = n - 1$ yields

$$w = c_n I^{1-n}\mathcal{R}^+\mathcal{R}w \tag{5.7}$$

where $I^{1-n} = (-\Delta)^{(n-1)/2}$. The simplest case is $n = 3$, for which

$$w = -c_3 \, \Delta \, \mathcal{R}^+\mathcal{R}w. \tag{5.8}$$

Suppose $w = \Delta f$, where f is a function to be reconstructed. If (5.8) has a unique solution for f, it must then be obtained directly by

$$f = -c_3\mathcal{R}^+ g, \tag{5.9}$$

where $g = \mathcal{R}\,\Delta f$ is a given (known) function. For $n = 2$, equation (5.6) with $\alpha = 1$ yields

$$w = c_2(-\Delta)^{1/2}\mathcal{R}^+\mathcal{R}w. \tag{5.10}$$

Suppose again that $w = \Delta f$, with $g = \mathcal{R}\,\Delta f$ a known function. Since $(-\Delta)^{1/2}(-\Delta)^{1/2} = -\Delta$ by (5.2), it follows that a solution of

$$(-\Delta)^{1/2}f = -c_2\mathcal{R}^+ g \tag{5.11}$$

is a solution of (5.10). Formally, we may write a reconstruction formula for f as

$$f = -c_2(-\Delta)^{-1/2}\mathcal{R}^+ g = -c_2 I^1(\mathcal{R}^+ g). \tag{5.12}$$

A necessary condition here is that $\mathcal{F}[f](0) = 0$. Taking the Fourier transform of (5.12) we obtain

$$\mathcal{F}[f](\xi) = -c_2\,\mathcal{F}\left[I^1(\mathcal{R}^+ g\right] = -c_2\,|2\pi\xi|^{-1}\,\mathcal{F}\left[\mathcal{R}^+ g\right](\xi). \tag{5.13}$$

Next we recall a dual version of the projection-slice theorem [130, Theorem 2.4]: For $g \in \mathcal{S}(C^n)$ an even function,

$$\mathcal{F}\left[\mathcal{R}^+ g\right](\xi) = d_n\,|2\pi\xi|^{1-n}\,\mathcal{F}[g](\xi/\,|\xi|\,,|\xi|) \tag{5.14}$$

---

[1]Technically, for w an element of the Schwartz space $\mathcal{S}(\mathbb{R}^n)$.

where d_n is a constant. Using this result in (5.13) for $n = 2$ gives

$$\mathcal{F}[f](\xi) = -c_2 d_2 \, |2\pi\xi|^{-2} \, \mathcal{F}[g](\xi/\, |\xi| \, , |\xi|).$$ (5.15)

When (5.12) holds, equation (5.15) yields a direct and exact formula for the Fourier transform $\mathcal{F}[f]$, and for f itself via the inverse Fourier transform. This relationship may be explored further in both the continuous and discretized domains.

Local discrete Radon transform: Paper III describes the use of the GDB Radon transform in an image filter algorithm. The GDB algorithm is fast due to its recursive form. For image size $N \times N$, the number of additions in the GDB algorithm is $4N^2 \log_2 N$, which for large N is much lower than for conventional $\mathcal{O}(N^3)$ DRTs (e.g., discrete approximations to the classical Radon transform). For example, $N = 128$ gives $4N^2 \log_2 N = 458752$, while there are 8388608 non-zero elements (ones) in the matrix representation of the GDB algorithm (a ratio of $1 : 18$). For $N \leq 64$, however, matrix multiplication using an efficient low-overhead library may be faster than the recursive algorithm. Moreover, for $N \leq 64$ (at least), the inverse transform can be obtained non-iteratively because the matrix representation of the inverse can be computed explicitly and stored in memory. The solution to the normal equations $\boldsymbol{R}^T \boldsymbol{R} \boldsymbol{f} = \boldsymbol{R}^T \boldsymbol{d}$ can be obtained by multiplication with the Moore-Penrose pseudoinverse of the forward matrix. The matrix representation of the GMRES or Neumann series inverse algorithms can be computed to arbitrary accuracy (number of iterations) in a one-time operation for a given image dimension N. Our tests show that this gives a significant performance gain in block processing or sliding window processing of larger images. The cost of inversion by matrix multiplication is independent of the accuracy with which these matrix representations are obtained.

The GDB algorithm obtains the sum of pixel values along certain prescribed graphs. At step k ($k = 1, \ldots, \log_2 N$), the algorithm sums the contributions from the partial graphs in pairs of adjacent subimages, each of which has size $N \times 2^{k-1}$. It may be possible to use the same recursive scheme for computing many local (sliding window) DRTs in parallel. The partial graphs would then not be zero-padded but include pixels in neighboring windows, which implies that more bookkeeping is required. Local DRTs can be used to accentuate linear features that are oriented in a particular angular sector, i.e. by zeroing the DRT domain outside the sector before applying the inverse transform. This approach may be useful for detecting thin structures with a predominant orientation, such as seismic fault lines, fractures, or perhaps blood vessels. In such applications we would use the matrix representation of the DRT inverse as described above.

Figure 5.2 shows an application of a local DRT filter for fault line detection in a seismic image (5.2a). This seismic section has many vertically oriented faults. A DRT filter was applied in sliding windows to the image in Fig. 5.2b, which is a seismic coherence attribute derived from the original image. The window size was 32×32 pixels. The filter was implemented in the DRT domain by muting all DRT values is a $90°$ sector centered on the horizontal, and retaining only values in the highest $10\,\%$ range in the remaining $90°$ sector centered on the vertical. These high intensity parts correspond to linear structures. In this example the inverse DRT was computed using the Moore-Penrose pseudoinverse matrix. Figure 5.2c shows the result applying the local DRT filter, and Fig. 5.2d shows the filtered image merged with the original seismic section to

(a) Seismic profile

(b) Coherence attribute
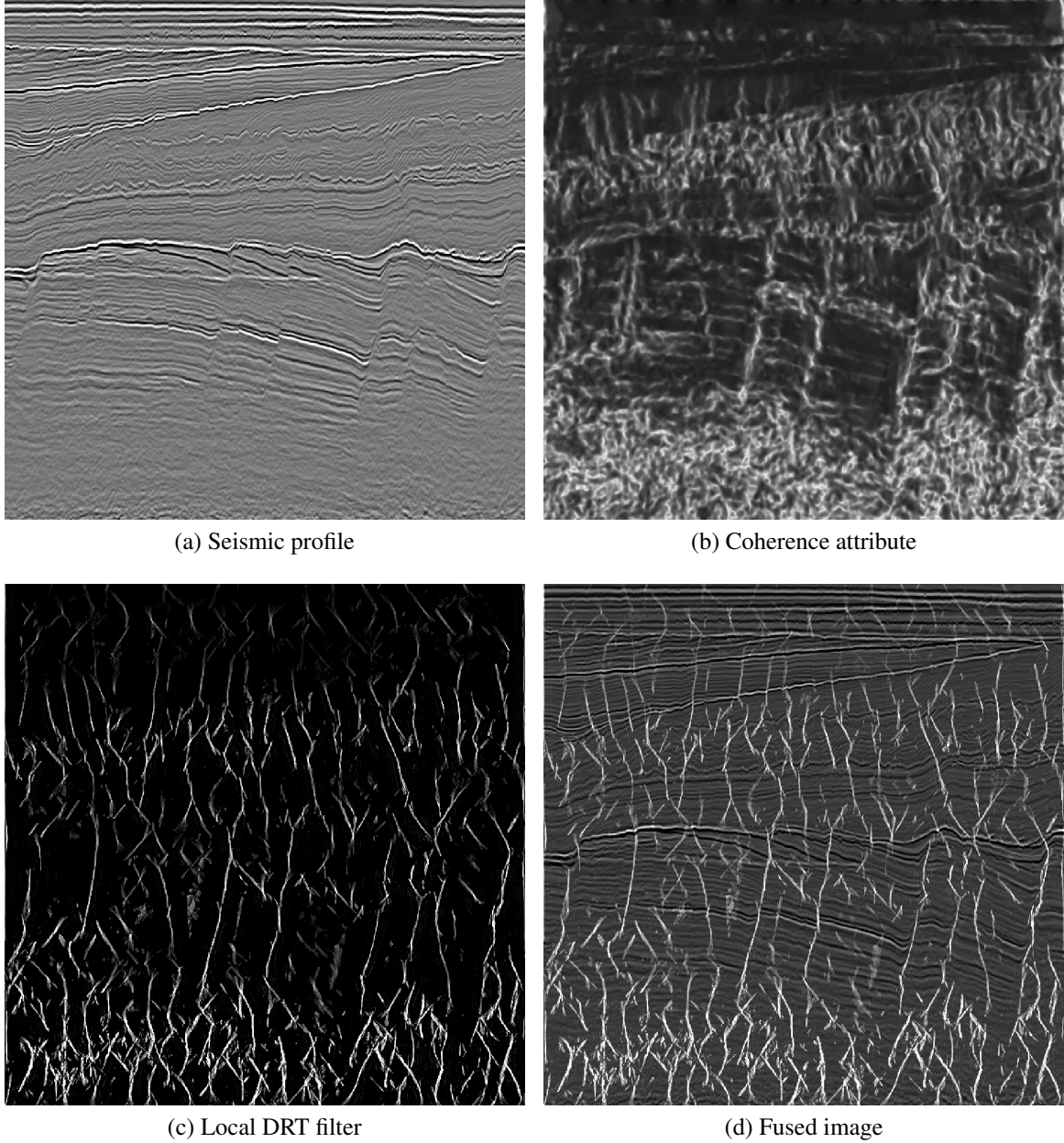
(c) Local DRT filter

(d) Fused image

Fig. 5.2: Fault line detection in a seismic image using a local DRT filter.

produce an overlay. This result is only preliminary, but suggests that the DRT filter may be a useful aid for such tasks.

Gradient operators: In Paper III (Sec. 4.3) we made use of a Laplace-like invertible edge detection operator $\mathcal{L}$. The operator $\mathcal{L}$ was defined as the circular convolution with a PSF $\boldsymbol{h}_M^0 + \boldsymbol{\eta}_M(\epsilon)$, where $\boldsymbol{h}_M^0$ denotes the standard Laplacian PSF symmetrically zero-padded to size $M \times M$, and $\boldsymbol{\eta}_M(\epsilon)$ is an $M \times M$ matrix with random elements η_{ij} such that $|\eta_{ij}| \leq \epsilon \ll 1$. As any forward-inverse operator pair may in principle be used for making image-transform-based filters, it may also be of interest to construct invertible gradient operators in a similar way. Standard gradient operators $\nabla_{\boldsymbol{v}} g$ have a larger kernel (null space) than the standard Laplacian. Nevertheless, one may again use convolution with a PSF on the form $\boldsymbol{h}_M^0 + \boldsymbol{\eta}_M(\epsilon)$, where now

(a) Test image   (b) Sobel operator   (c) Invertible gradient operator

(d) Deconvolved Sobel   (e) Inverse gradient
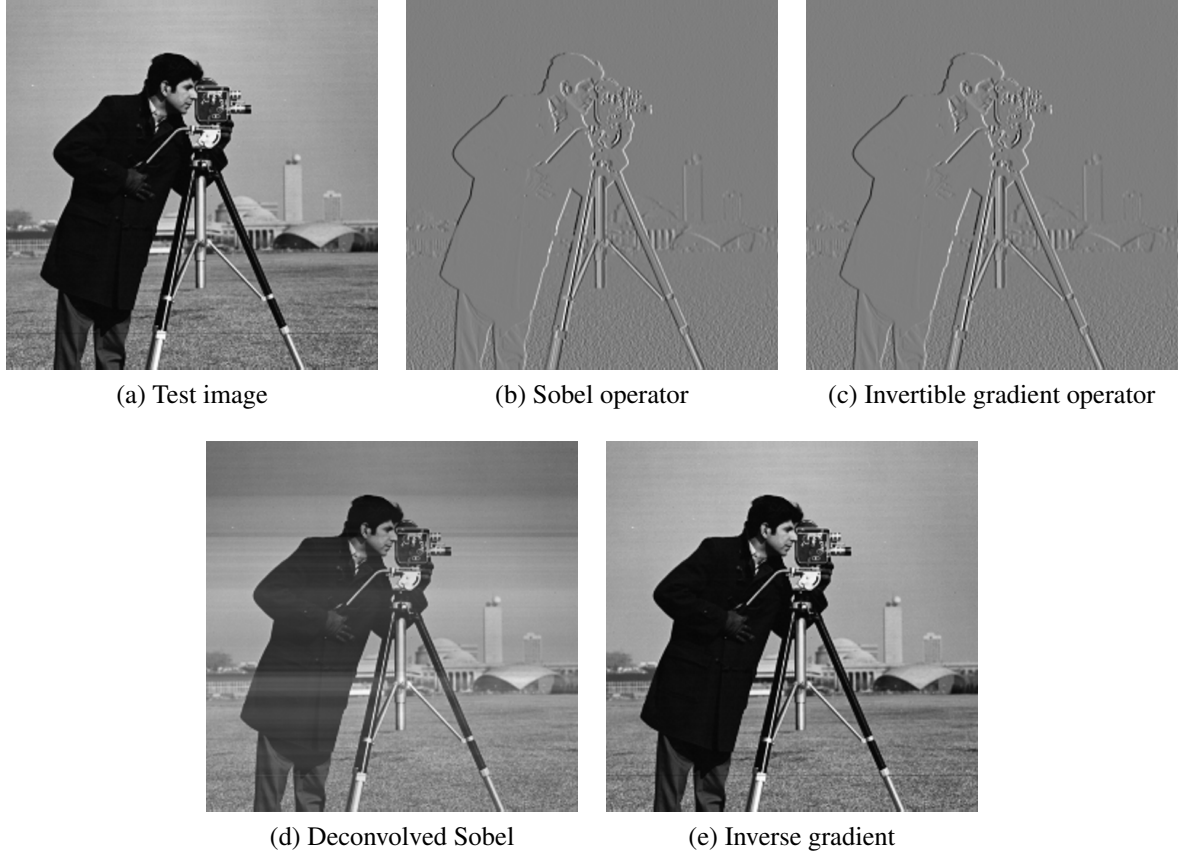
Fig. 5.3: The horizontal Sobel operator (b) and proposed (horizontal) gradient operator (c) applied to the standard cameraman test image (a) with intensity range 0–1. Applying the inverse gradient operator to (c) yields (e) (no rescaling of intensities). The RMS error with respect to the original image (a) is $1.4 \cdot 10^{-14}$, and the residual 2-norm is $6.7 \cdot 10^{-11}$. For comparison, the Sobel gradient image (b) was deconvolved using the regularized filter algorithm [112] [131]. The deconvolved image (d) has been rescaled to the intensity range 0–1, but the actual range is between -0.64 and 0.68; the RMS error is 0.48 and the residual 2-norm is 244.4.

$\boldsymbol{h}_M^0$ is formed by zero-padding the Sobel operators, i.e.,

$$\boldsymbol{h} = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \tag{5.16}$$

and its transpose. The effect is shown in Fig. 5.3, where the horizontal gradient of a test image has been computed using the Sobel operator and the modified Sobel operator. The gradient images were subsequently deconvolved; only the modified gradient allows perfect reconstruction of the original image.

5.2   MBES data and models

Scattering strength estimation: The classification method presented in Paper II (Sec. 4.1) is based on measuring the seabed scattering strength as a function of incidence angle. The MBES data employed in Paper II did not include the full beam echo time series. The data contained

only sufficiently many samples around the bottom detection point to map a contiguous set of samples to the seafloor across the whole swath, i.e., to form a side-scan image. This limitation implies extra uncertainty in the estimation of the scattering cross section. First, even if the beam pattern is known, it is not known precisely where the ensonified area lies within the beam footprint at a particular time t. Second, the volume scattering component in the tail of the echo is not taken into account. Combined, this shows that the (averaged) scattering strength should be estimated based on the total backscattered energy (full bottom echo).

Another point to consider is the spatial orientation of the ensonified area with respect to the incident wave field. In Paper II, we used the depth soundings to compute the terrain surface normal vector, $\boldsymbol{n}$, and true incidence angle, θ. The ensonification function $\Omega(\boldsymbol{r}, \boldsymbol{r}_0)$, however, depends not only on the angle between $\boldsymbol{n}$ and the incident wave vector. For MBES array geometries, the peak intensity of the incident wave field varies much in the along-track direction and only slightly in the across-track direction. For a given angle, θ, the surface normal $\boldsymbol{n}$ defines a cone of incidence vectors with incidence angle θ. If we fix a vector $\boldsymbol{t}$, representing the direction (near the seafloor) of a particular beam, and imagine rotating a terrain surface facet while keeping the angle $\theta = \angle(\boldsymbol{t}, \boldsymbol{n})$ fixed, the ensonified area will rotate with respect to the high and low intensity parts of the incident field.

Approximate formulas to obtain the the scattering strength, that take into account the effects mentioned here—beam pattern, total energy, and terrain surface orientation— were derived during the preparations for Paper II. The next step should be to implement these formulas in the data processing chain and apply them to MBES data with full bottom echo times series. The corrections may be small where surface scattering dominates volume scattering, and where the seafloor is fairly smooth and slowly varying with respect to the spatial resolution of the MBES system. However, the method should be relevant for high-frequency, high-resolution instruments used in littoral zones and shallow waters.

MBES channel data: With the assistance of Kongsberg Maritime (Horten, Norway) and the Norwegian Defence Research Establishment (FFI, Kjeller/Horten, Norway), we have obtained an MBES type EM 710 dataset that includes not only full beam times series (water column data in Kongsberg Maritime terminology), but also the channel data from each transducer element. Unfortunately work on this dataset has not progressed far due to time constraints. However, the dataset provides new opportunities for research on scattering strength estimation, seabed classification, and the effects of beamforming and bottom detection algorithms.

5.3   Classification and inversion

Bayesian inversion: Paper II (Sec. 4.1) considered a compound PDF $f_{Y|\Theta}(y \mid \theta)$ for the observed, corrected intensity y as a function of incidence angle θ. It was assumed that the two components of the measured complex pressure are independent Gaussian random variables (cf. Sec. 3.3.2), and that the seabed is characterized by some unknown angle-dependent distribution $f_{S|\Theta}(s \mid \theta)$ for the (averaged) scattering cross section, regarded as a random variable, S. Furthermore, it was assumed that the effect of propagation through the water column could be compensated by a deterministic propagation loss factor. The resulting compound distribution

for the corrected intensity is

$$f_{Y|\Theta}(y \mid \theta) = \int_0^\infty \frac{1}{s} e^{-y/s} f_{S|\Theta}(s \mid \theta) \, \mathrm{d}s. \tag{5.17}$$

By forming a scattering strength feature vector where each component is the average in a fixed angular segment, and by invoking the multidimensional CLT, error bounds were derived for the two-class classification problem. The error bounds and the feature vector PDFs depend on the angle-dependent mean, $\varsigma(\theta) = \mathrm{E}[S \mid \Theta = \theta]$, and the variance, $\nu^2(\theta) = \mathrm{Var}[S \mid \Theta = \theta]$, of the scattering strength. To make predictions about the classifier performance for different types of sediments as a function of the intrinsic variance $\nu^2(\theta)$, a six-parameter theoretical model for seabed scattering strength [79] has been implemented (see also Sec. 3.2.2).

A natural extension of this work would be to apply inverse methods to estimate the model parameters from the measured data, i.e., to determine which physical properties of the seabed are consistent with the data. One possible approach is Bayesian inference theory [4] (see also Sec. 3.1), where the unknown model parameters, $\boldsymbol{m} \in \mathbb{R}^n$, are taken as random variables, and the main objective is to compute $P(\boldsymbol{m} \mid \boldsymbol{d})$, the posterior probability density (PPD) for $\boldsymbol{m}$ given the observed data $\boldsymbol{d} \in \mathbb{R}^K$ (the feature vector). Examples in the literature include [132], on long-range, low-frequency acoustic reverberation and propagation inversion. By analogy, this work may be a starting point by replacing the reverberation-level-versus-range data of [132] with the angle-dependent scattering strength vector of Paper II. From the PPD one may obtain, e.g., the maximum a posteriori (MAP) estimate,

$$\mathrm{MAP} = \underset{\boldsymbol{m} \in \mathbb{R}^n}{\mathrm{argmax}} \, P(\boldsymbol{m} \mid \boldsymbol{d}) \tag{5.18}$$

and the mean model

$$\overline{\boldsymbol{m}} = \int_{\mathbb{R}^n} \boldsymbol{m} P(\boldsymbol{m} \mid \boldsymbol{d}) \, \mathrm{d}\boldsymbol{m}. \tag{5.19}$$

Furthermore, one may derive quantities that characterize parameter uncertainties and correlations, in particular the 1D marginal distributions

$$P(m_i \mid \boldsymbol{d}) = \int_{\mathbb{R}^n} \delta(m_i' - m_i) P(\boldsymbol{m}' \mid \boldsymbol{d}) \, \mathrm{d}\boldsymbol{m}' \tag{5.20}$$

and the correlation matrix $R_{ij} = C_{ij}/\sqrt{C_{ii}C_{jj}}$, where

$$\boldsymbol{C} = \int_{\mathbb{R}^n} (\boldsymbol{m}' - \overline{\boldsymbol{m}})(\boldsymbol{m}' - \overline{\boldsymbol{m}})^T P(\boldsymbol{m}' \mid \boldsymbol{d}) \, \mathrm{d}\boldsymbol{m}' \in \mathbb{R}^{n \times n}. \tag{5.21}$$

The integrals in (5.19)–(5.21) may be computed by Monte-Carlo methods, where

$$P(\boldsymbol{m} \mid \boldsymbol{d}) = \frac{p(\boldsymbol{d} \mid \boldsymbol{m}) P(\boldsymbol{m})}{\int_{\mathbb{R}^n} p(\boldsymbol{d} \mid \boldsymbol{m}') P(\boldsymbol{m}') \, \mathrm{d}\boldsymbol{m}'} \tag{5.22}$$

by Bayes' rule, and the form of $p(\boldsymbol{d} \mid \boldsymbol{m})$ must be postulated or determined. We consider

log-intensity data in K angular segments, and define as in Paper II

$$\varsigma_L(\theta) = \mathrm{E}[\log S|\Theta = \theta] \tag{5.22a}$$

$$\nu_L^2(\theta) = \mathrm{Var}[\log S|\Theta = \theta]. \tag{5.22b}$$

The functions $\varsigma_L(\theta)$ and $\nu_L^2(\theta)$ are unknown quantities, but the former may be replicated using a theoretical scattering model with a model vector $\boldsymbol{m} \in \mathbb{R}^n$ as input; this function will be denoted $\varsigma_L(\theta; \boldsymbol{m})$. For the kth segment (Paper II, Appendix),

$$p_{\log Y}(w; k) = \int_0^\infty \frac{e^w}{s} e^{-e^w/s} \langle p_{S|\Theta}(s|\theta)\rangle_k \, \mathrm{d}s, \ k = 1, \ldots, K \tag{5.23}$$

and

$$\mathrm{E}[\log Y; k] = \langle \varsigma_L \rangle_k - \gamma \tag{5.24}$$

where γ is the Euler-Mascheroni constant, and $\langle \cdots \rangle_k$ means averaging with respect to incidence angle θ over the kth segment, i.e.,

$$\langle A \rangle_k \equiv \frac{1}{\theta_k - \theta_{k-1}} \int_{\theta_{k-1}}^{\theta_k} A(\theta) \, \mathrm{d}\theta. \tag{5.25}$$

For the variance, we have

$$\mathrm{Var}[\log Y; k] = \langle \nu_L^2 \rangle_k + \frac{\pi^2}{6} + \mathrm{Var}_\Theta [\varsigma_L; k]. \tag{5.26}$$

The segments can be chosen such that the third term, $\mathrm{Var}_\Theta [\varsigma_L; k] = \langle \varsigma_L^2 \rangle_k - \langle \varsigma_L \rangle_k^2$, may be neglected. With the simplifying assumption that log-intensity measurements from different scattering elements are statistically independent [drawn from (5.23)], the PDF of the feature vector is asymptotically normal with diagonal covariance matrix, $\boldsymbol{\Sigma}$,

$$L(\boldsymbol{m}) = p(\boldsymbol{d} \mid \boldsymbol{m}) = \frac{1}{(2\pi)^{K/2}\sqrt{\det \boldsymbol{\Sigma}}} \exp \left\{ -\frac{1}{2} [\boldsymbol{d} - \boldsymbol{d}(\boldsymbol{m})]^T \boldsymbol{\Sigma}^{-1} [\boldsymbol{d} - \boldsymbol{d}(\boldsymbol{m})] \right\}, \tag{5.27}$$

where

$$\boldsymbol{d}(\boldsymbol{m}) = \begin{bmatrix} \langle \varsigma_L(\theta; \boldsymbol{m})\rangle_1 - \gamma & \ldots & \langle \varsigma_L(\theta; \boldsymbol{m})\rangle_K - \gamma \end{bmatrix}^T \tag{5.28}$$

and

$$M\boldsymbol{\Sigma} \approx \mathrm{diag}\left(\langle \nu_L^2 \rangle_1, \ldots, \langle \nu_L^2 \rangle_K\right) + \frac{\pi^2}{6}\boldsymbol{I}_K. \tag{5.29}$$

The factor M is a constant that depends on the experiment, namely the number of (independent) observation vectors $\begin{bmatrix} Y_1 & \ldots Y_K \end{bmatrix}$ that are averaged to produce a feature vector. Averaging is thus a way to control the variance. However, it cannot in general be assumed that the theoretical model or the measurements are accurate. In practice, the model may have too few degrees of freedom to replicate observations, based on simplifying (perhaps erroneous) assumptions. The variance should therefore also account for theory errors. For fixed, given data, the PDF (5.29) is the likelihood function, $L(\boldsymbol{m})$, that measures the model-data misfit. The likelihood function

(5.27) is on the form $L(\boldsymbol{m}) \propto \exp\left[-\Phi(\boldsymbol{m})/2\right]$, where $\Phi(\boldsymbol{m})$ is the misfit function

$$\Phi(\boldsymbol{m}) = \sum_{k=1}^{K} \log\left[\langle \nu_L^2 \rangle_k + \frac{\pi^2}{6}\right] + M \sum_{k=1}^{K} \frac{(d_k - \langle \varsigma_L(\theta; \boldsymbol{m}) \rangle_k + \gamma)^2}{\langle \nu_L^2 \rangle_k + \pi^2/6}. \tag{5.30}$$

Setting $\langle \nu_L^2 \rangle_k = 0$ for $k = 1, \dots, K$ corresponds to observing a seafloor characterized by a single, fixed scattering cross section, when measurement and theory errors are small compared to the statistical fluctuations associated with random scattering with Gaussian statistics. In this case, the first term in (5.30) is model independent, and the misfit function may be reduced to

$$\Phi(\boldsymbol{m}) = \frac{6M}{\pi^2} \sum_{k=1}^{K} (d_k - \langle \varsigma_L(\theta; \boldsymbol{m}) \rangle_k + \gamma)^2. \tag{5.31}$$

In general, however, the unknown variances must be determined. One possible strategy for such problems is to expand the model space to include the K quantities $\langle \nu_L^2 \rangle_1, \dots, \langle \nu_L^2 \rangle_K$ as additional model (nuisance) parameters [133]. This increases the dimension of the integrals in (5.19)–(5.22), so it is probably advantageous to keep K low. Combined with the scattering strength estimation method outlined in Sec. 5.2, this development provides a starting point for geoacoustic parameter estimation using MBES data.

Hybrid acoustic features: As noted in Ch. 3, there are two main approaches to statistical seabed classification using MBES data. The first approach is to use scattering strength versus incidence angle to discriminate between seabed types [98–101]. The second approach is to extract features from sidescan images. Texture measures derived from gray level co-occurrence matrices (GLCMs) are an important class of such image features [38, 114–116]. GLCMs are obtained by quantizing images to, say, L pixel values (gray levels). For any rectangular neighborhood in the image, one may form an $L \times L$ matrix $\boldsymbol{G}(\Delta x, \Delta y)$, such that G_{ij} is the relative frequency of two pixels with gray levels i and j respectively being separated by pixel distance $(\Delta x, \Delta y)$ [134]. From $\boldsymbol{G}(\Delta x, \Delta y)$ one may derive several attributes which represent the character of the texture [113, 134]. Texture attributes characterize fluctuations and spatial correlations in seabed scattering strength, and may be combined with first-order statistics derived from the local distributions of pixel values.

Sidescan image classification invariably requires that the angular dependence of scattering strength be removed by some normalization and correction procedure. Conversely, statistical fluctuations are ignored when deriving features only from mean scattering strength versus angle. In Paper II, a hybrid method is outlined, but not developed, to retain both types of features. The method applies to beam data that have been corrected for propagation loss and beampattern variations, and mapped to a gray level range. The features of Paper II are based on a subdivision of the incidence angle range, $0 < \theta_1 < \dots < \theta_{K-1} < \theta_{\max}$, such that for a fixed K the variation with angle is effectively minimized within each segment. Within each angular segment, for a sequence of N_p consecutive pings, one may form an analogous co-occurrence matrix $\boldsymbol{G}(\Delta x, \Delta \theta)$, where Δx is along-track separation between pairs of samples, and $\Delta \theta$ is the difference in estimated incidence angle. A special case is $\Delta x \approx 0$, where only separation in angle is considered. Combining texture attributes derived from $\boldsymbol{G}(\Delta x, \Delta \theta)$ with the scattering

strength vector $\boldsymbol{d}$ of Paper II may help improve classification accuracy. In particular, it may be useful when dividing the sonar swath into smaller sectors treated separately so as to improve the lateral spatial resolution of the classification procedure.

Terrain classification: Work on terrain classification was commenced in Paper III, using two sets of features: rotation-invariant LBPs and local empirical cumulative distribution functions (ECDFs) of intrinsic geometrical properties. The invariant properties of these features make them well suited for classification of geomorphology, and this should be explored on other datasets. In the classification experiments of Paper III, only terrain features were employed. The natural next step is to integrate terrain features and acoustic features to see if better classification accuracy can be attained.

Semi-supervised learning: Supervised classification methods, including the seabed mapping method of Paper II, assign each observation $\boldsymbol{d}$ to one of a predefined set of classes $\{\,\omega_1, \dots, \omega_c\,\}$. In the Bayesian approach a class ω is represented by the PDF $p(\boldsymbol{d} \mid \omega)$, which is estimated from training data. When applied to acoustic seabed data, two interconnected problems must be addressed. The first problem is that a dataset may contain observations that do not fit into any of the predefined classes (inadequate training data). The second is the problem of delimiting sediment classes which physically belong to a continuum, whether characterized by acoustic or sedimentological properties. Mathematically, delimiting classes is straightforward; it is implemented by the decision boundaries, i.e., the set of points $\boldsymbol{x}$ in feature space such that $p(\boldsymbol{x} \mid \omega_i)P(\omega_i) = p(\boldsymbol{x} \mid \omega_j)P(\omega_j)$ for any pair of class indices $i \neq j$. The decision boundary between two classes is therefore strongly influenced by the (co-)variances of the training data, but it does not inform us about the physical transition between classes.

Suppose an observation $\boldsymbol{d}$ has low likelihood $p(\boldsymbol{d} \mid \omega)$ with respect to all the predefined classes. The evidence factor

$$p(\boldsymbol{d}) = \sum_{j=1}^{c} p(\boldsymbol{d} \mid \omega_j)P(\omega_j)$$

appearing in Bayes' theorem will however also be low, and a class may still get a high posterior probability $P(\omega \mid \boldsymbol{d})$. In this situation it may be appropriate to introduce a new class or label the observation as an unclassified outlier. Here we briefly outline an algorithm for growing the number of classes; it is a natural extension of the classification method of Paper II. The aim is to exploit the spatial relationship between observations as well as the statistical similarity. Therefore we index each observation $\boldsymbol{d}(\boldsymbol{r})$ by its geographical position $\boldsymbol{r}$. Assume that there are c sediment classes with corresponding estimated PDFs $p(\boldsymbol{d}, \omega_i)$, $i = 1, \dots, c$. Specify a threshold T for the likelihood: if

$$p(\boldsymbol{d} \mid \omega_i) < T,$$

then $\boldsymbol{d}$ should be rejected from class ω_i regardless of the likelihood with respect to the other classes. Next form a binary representation of the survey area S by thresholding: Put

$$S = D \cup D^c,$$

where

$$D = \{\, \boldsymbol{r} \mid p(\boldsymbol{d}(\boldsymbol{r}) \mid \omega_i) < T \text{ for all } i = 1, \dots, c \,\},$$

comprise the geographical regions with low likelihood, and assign ones and zeros to elements in D an D^c respectively. This georeferenced set should be smoothed with a block median filter before extracting the connected components $R_1, \ldots, R_L$ of D. The distribution in feature space of observations $\boldsymbol{d}$ inside a connected region R_i is potentially multimodal; let $\boldsymbol{d}_m$ be the global maximum (most probable value). Next run a spatial filter on R_i defined as follows. Specify a geographical radius r_0, and define a window ("ball") with respect to position $\boldsymbol{q}$ as

$$B(\boldsymbol{q}; r_0) = \{\, \boldsymbol{d}(\boldsymbol{r}) \mid \| \, \boldsymbol{r} - \boldsymbol{q} \| < r_0 \,\}.$$

We then define the statistical center of R_i as the position which minimizes the distance from $\boldsymbol{d}_m$ in the sense that

$$\boldsymbol{q}_c = \operatorname*{argmin}_{\boldsymbol{q} \in R_i} \langle \rho(\boldsymbol{d}, \boldsymbol{d}_m) \mid \boldsymbol{d} \in B(\boldsymbol{q}, r_0) \rangle, \tag{5.32}$$

where ρ is a similarity measure and $\langle A \rangle$ denotes the arithmetic mean of A. Two natural choices for ρ are

$$\rho(\boldsymbol{d}, \boldsymbol{d}_m) = \| \boldsymbol{d} - \boldsymbol{d}_m \|$$
$$\rho(\boldsymbol{d}, \boldsymbol{d}_m) = (\boldsymbol{d} - \boldsymbol{d}_m)^T \, \Sigma^{-1} \, (\boldsymbol{d} - \boldsymbol{d}_m),$$

i.e. the Euclidean distance and the Mahalanobis distance, and Σ is the sample covariance matrix for data in R_i. For each R_i, observations in the set $B(\boldsymbol{q}_c; r_0)$ can be used as training data for new candidate classes; hence one obtains the new PDFs $p(\boldsymbol{d} \mid \omega_{c+1}), \ldots, p(\boldsymbol{d} \mid \omega_{c+L})$. The whole procedure can be repeated if necessary until the area of D is below some threshold. Finally, statistical tests must be applied on all the PDFs to determine if some of them are similar and should be merged into a single class. This method determines the number of classes through the adjustment of a single parameter, the threshold T. The greater the likelihood an observation is required to have, the finer the division into seabed types will be.

Across-track resolution: With the ASC method proposed in Paper II, the across-track resolution hinges in practice not on the number of independently classified observations, but on the ability to detect class transitions within the swath. As the examples shown in Paper II indicate (Paper II, Fig. 11), it is possible to do reasonably accurate classification using limited angular sectors. One way to proceed would therefore be to construct an algorithm that detects discontinuities in the scattering strength function, $S(\theta)$, and suggests a partition of the incidence angle domain based on this. Each partition (sector) would then be classified independently.

Time-domain echo classification: Paper I discusses the QTC ASC algorithm in detail. Many of the limitations of the software seem due to reasonable design choices to make the software robust and straightforward to use without extensive expert manipulation. However, Paper I mentions several possible enhancements, including

- using more than three features, e.g. including total backscattered energy;

- using ensemble properties such as measures of ping-to-ping variability (see e.g. [135]);

- including spatial information in the clustering algorithm, e.g. by adjusting the prior prob-

(a) Bathymetry and transect
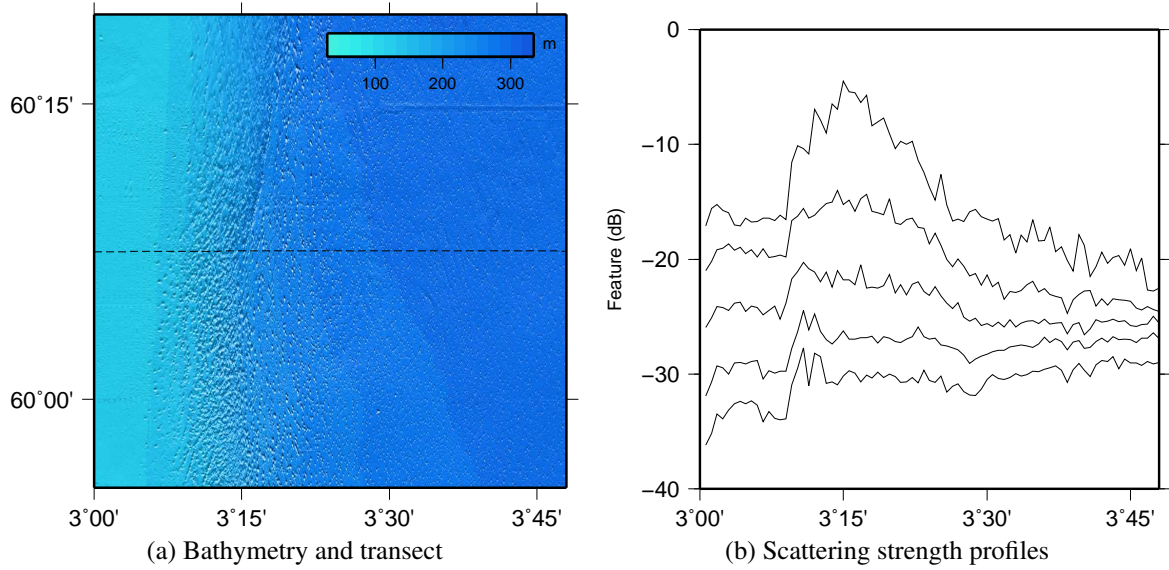
(b) Scattering strength profiles

Fig. 5.4: (a) Dashed line shows east-west transect along which backscatter feature vectors are plotted as a function of longitude in (b): The feature vectors have five components, one for each curve. Each curve shows scattering strength as a function longitude along the transect. The topmost curve corresponds to small incidence angles, the lowest curve to high angles (from [136]).

abilities using spatial relationships;

- performing feature extraction and reduction by a piecewise constant function approximation of the corrected echo time series;

- using the QTC methodology on the vertical incidence beam of multibeam echo sounders, thereby overcoming a break-down in accuracy over significantly sloping terrain (presently a fundamental limitation);

- using the EM algorithm for faster processing (compared to the stochastic cluster algorithm of QTC).

5.4 Marine geology

There are some noteworthy features of the MBES backscatter data from the sloping area between the North Sea plateau and the western side of the Norwegian Channel, as shown in Fig. 5.4. The dashed line in Fig. 5.4a represents a west-to-east transect along which the trend in backscatter data is shown in Fig. 5.4b. To make Fig. 5.4b, scattering strength vectors as defined in Paper II, with $K = 5$ components, were computed in a narrow geographical region centered on the EW transect. The vectors were sorted according to ascending eastern longitude, and averaged over samples with approximately equal latitude (averaging reduces fluctuations and produces a clearer trend). The result was the five curves in Fig. 5.4b, one for each component (angular segment), as a function of distance along the transect. The slope covers a depth range of about 100–300 m; the plateau to the west is covered with sand and the trench to the east is covered with soft marine sediments. We assumed that the grain size distribution would change

gradually down the slope and cause a corresponding gradual change in scattering strength. This is partly confirmed by Fig. 5.4b, but there is a sharp increase in backscatter levels at about 03° 10' E, particularly at small incidence angles. This discontinuity corresponds with one of the class boundaries found in Paper I [10], and with a change in seafloor morphology at the edge of the plateau. The trend in Fig. 5.4 is consistently seen in data from three different surveys covering the sloping area.

As discussed in Paper I and II, the seabed morphology on the slope is rough, with an irregular pattern of elongated depressions attributed to a combination of water current and fluid seepage erosion [137, 138]. The presence of vertically migrating shallow gas may be a partial explanation for the observed high scattering strength on the upper slope. The results from Paper I and II, obtained with different data and methods, further indicate that there is a significant sand fraction in the top sediment layer at the foot of the slope in the westernmost part of the trench, but the MBES results (Paper II) indicate more sand than the SBES results. Fluid seepage up through the shallow geological layers may cause seabed sediments to be lifted into suspension, and suspended sand particles moved by currents may accumulate at the bottom of the slope. This hypothetical erosion and transport process may help explain the acoustic sediment classification of the western trench. The few core samples from this area have a high sand fraction in the top layer. The discrepancy between MBES and SBES classification results may be due to the difference in frequency; the bottom interaction of the high-frequency MBES is more confined to the surficial layer. However, more observations, in particular more sediment samples, are needed to resolve these questions.

Chapter 6

Conclusion

The nature of marine sediments reflects past and present physical processes on the seafloor. It is also a critical factor for benthic ecology, shallow water acoustics, and oceanic geotechnical engineering. The aim of this thesis has been to develop an acoustical remote sensing technique for seabed classification across extended areas. Available tools to accomplish this include modern multibeam echo sounders and sidescan sonars for data acquisition, and machine learning and inverse theory for classification (prediction) and parameter estimation, respectively. The emphasis of this thesis has been on machine learning and echo sounders. From experience with data interpretation, it is particularly useful to merge information about seabed acoustic response and physical composition on the one hand, with detailed, full-coverage bathymetric and geomorphological data on the other. The volume of such data is increasing as sonar technology advances and more MBESs are deployed. This should motivate further development of seabed classification methods.

The main contributions of this thesis, described in three research papers, are the following:

- A new regional map of surficial sediments in the northern North Sea ($59°30'N - 61°N$) was presented. Analysis of acoustic classification results and supporting data provides information about erosional and depositional processes in the Norwegian Channel: Glacial erosion on the near-shore side, fluid seepage and water current erosion on the western slope, and deposition of fine-grained suspended sediments in the central part.

- A supervised classification method for MBES data was presented. The method uses the angle dependence of the seabed scattering strength to discriminate between sediment types. Theoretical and experimental analyses suggest that good classification accuracy and spatial resolution are possible even with limited training data. The method should also be applicable to sidescan sonar or other wide-swath systems.

- A composite image processing transform was introduced and used to reduce noise and artifacts in MBES bathymetry. The algorithm significantly improves the quality of, e.g., attribute images derived from digital terrain models. This data enhancement makes it possible to perform accurate terrain classification with invariant, intrinsic, geometric features, or, alternatively, well-established texture operators (LBPs).

Further results derived from Paper III include local DRT filters applied to fault line detection in seismic images, invertible gradient operators, and enhancement of synthetic aperture radar

images. The DRT-based filter may also be useful for reducing other forms of stripe noise occurring in remote sensing imagery. There is also more to explore concerning the use of invertible difference operators and alternative inverse DRT algorithms. In particular, the relationship between the (pseudo-differential) Laplace operator and the classical exact inversion formulas for the Radon transform should be further investigated (Sec. 2.3.2).

A natural extension of Paper II would be to implement model-based Bayesian inversion; one approach is outlined in Sec. 5.3. By numerically estimating the PPDs of the model parameters, one may obtain both point estimates of seabed geoacoustic parameters, parameter correlations, and marginal distributions that indicate how much information the data yield about individual parameters. However, the high-frequency acoustics of seabeds is complex. A realistic model for a two-phase porous medium, with sound speed and density fluctuations, anisotropic roughness spectrum, gradients and boundary layers, will require many parameters that may be difficult to resolve using MBES data. Many of these parameters are also hard to measure by other means, either in situ or in the laboratory, making it difficult to validate the forward model and inversion results. Thus, theory (model) errors and measurement (calibration) errors are two important challenges for the inverse theoretic approach. In many applications, such as marine geology, the sedimentological descriptions (e.g., grain size distribution) typically associated with training data are more relevant than acoustic parameters. Statistical classification may then be a just as useful approach.

Appendices

A   The acoustic far-field

This appendix reviews the derivation of the acoustic far-field from a plane array in a homogeneous medium. It is mainly based on [12, Ch. 7].

In a region $\Omega \subset \mathbb{R}^3$ with constant mass density and no acoustic sources except possibly on the boundary $\partial\Omega$, low-intensity sound propagation is governed by the homogeneous linear wave equation

$$\frac{\partial^2 p}{\partial t^2} - c^2(\boldsymbol{r})\nabla^2 p = 0, \tag{A.1}$$

where c is the sound speed. Using the integral representation

$$p(\boldsymbol{r}, t) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{p}(\boldsymbol{r}, \omega) e^{i\omega t} \, \mathrm{d}\omega$$

where $\hat{p}(\boldsymbol{r}, \omega)$ is the temporal Fourier transform and $\omega = 2\pi\nu$, one obtains from (A.1) the elliptic Helmholtz equation for $\hat{p}(\boldsymbol{r}, \omega)$:

$$\nabla^2 \hat{p}(\boldsymbol{r}, \omega) + k^2(\boldsymbol{r})\hat{p}(\boldsymbol{r}, \omega) = 0, \tag{A.2}$$

where $k^2(\boldsymbol{r}) = \omega^2/c^2(\boldsymbol{r})$. The solution to (A.2) may be written as the Kirchhoff integral

$$\hat{p}(\boldsymbol{r}, \omega) = \int_{\partial\Omega} \left[G_\omega(\boldsymbol{r}, \boldsymbol{r}_0) \frac{\partial}{\partial n_0} \hat{p}(\boldsymbol{r}_0, \omega) - \hat{p}(\boldsymbol{r}_0, \omega) \frac{\partial}{\partial n_0} G_\omega(\boldsymbol{r}, \boldsymbol{r}_0) \right] \mathrm{d}S_0, \tag{A.3}$$

where $\boldsymbol{r}_0$ ranges over $\partial\Omega$ and $\partial/\partial n_0$ denotes the derivative in the direction of the outward surface normal at $\boldsymbol{r}_0$. The formula (A.3) holds for any solution to the equation

$$\nabla^2 G_\omega(\boldsymbol{r}, \boldsymbol{r}_0) + k^2 G_\omega(\boldsymbol{r}, \boldsymbol{r}_0) = -\delta(\boldsymbol{r} - \boldsymbol{r}_0), \tag{A.4}$$

where $\boldsymbol{r}_0$ is a fixed parameter and ∇^2 operates with respect to $\boldsymbol{r}$. The Green's function $G_\omega(\boldsymbol{r}, \boldsymbol{r}_0)$ can be interpreted as the field at $\boldsymbol{r}$ due to a point source at $\boldsymbol{r}_0$. For simplicity, it will be assumed that the array plane is not moving except at the transducer elements. This is to say that the array plane is perfectly rigid (has infinite impedance). In a homogeneous medium (c is constant), the appropriate Green's function is

$$G_\omega(\boldsymbol{r}, \boldsymbol{r}_0) = \frac{1}{4\pi|\boldsymbol{r} - \boldsymbol{r}_0|} e^{ik|\boldsymbol{r} - \boldsymbol{r}_0|} + \frac{1}{4\pi|\boldsymbol{r} - \bar{\boldsymbol{r}}_0|} e^{ik|\boldsymbol{r} - \bar{\boldsymbol{r}}_0|}, \tag{A.5}$$

where $\bar{r}_0$ denotes the reflection of r_0 about the array plane. The field due to a point source at r_0 is therefore a spherical wave plus a perfectly reflected spherical wave which mathematically can be interpreted as being due to an image source at $\bar{r}_0$. A key property of (A.5) is that $\partial G_\omega/\partial n = \partial G_\omega/\partial n_0 = 0$ at $\partial\Omega$ due to the field symmetry about the array plane [note the manifest reciprocity $G_\omega(r, r_0) = G_\omega(r_0, r)$]. Therefore, the second term in the integrand of (A.3) vanishes.

If coordinates are chosen such that the array elements are laid out in the plane $z = 0$ and radiating into the half-plane $z > 0$, then $\partial/\partial n = -\partial/\partial z$, and the velocity of the array surface is given by a function on the form $u = \begin{bmatrix} 0 & 0 & u(x, y, t) \end{bmatrix}^T$. Moreover, $r_0 = \bar{r}_0$ at $\partial\Omega$ ($z = 0$). The relationship between fluid particle velocity and pressure is $\rho\partial u/\partial t = -\nabla p$, which is conservation of momentum to first order in p and u. From all this follows that (A.3) reduces to

$$\hat{p}(r, \omega) = -2ik\rho c \int_\mathbb{R} \int_\mathbb{R} \hat{u}(x_0, y_0, \omega) \frac{1}{4\pi|r - r_0|} e^{ik|r-r_0|} \, dx_0 \, dy_0 \qquad \text{(A.6)}$$

with $r_0 = \begin{bmatrix} x_0 & y_0 & 0 \end{bmatrix}^T$. The factor $R = |r - r_0| = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$ may be written as

$$r\left[1 - 2\frac{xx_0 + yy_0}{r^2} + \frac{x_0^2 + y_0^2}{r^2}\right]^{1/2}$$

where $r = |r|$. The far-field approximation is valid for r sufficiently large such that the third term $(x_0^2 + y_0^2)/r^2 \ll 1$ can be ignored. This is to say that the support of $\hat{u}(x_0, y_0, \omega)$ (the extent of the active array) is small compared to the range r. Expanding the square root to first order in the second term and introducing spherical coordinates r, φ and ϑ (with $z = \cos\vartheta$ etc.) yields

$$R \simeq r - (x_0 \cos\varphi + y_0 \sin\varphi)\sin\vartheta \text{ and } \frac{1}{R} \simeq \frac{1}{r}.$$

This approximation gives the result [12, Sec.7.4]

$$\hat{p}(r, \omega) \simeq -ik\rho c \frac{e^{ikr}}{2\pi r} \int_\mathbb{R} \int_\mathbb{R} \hat{u}(x_0, y_0, \omega) e^{-i(k_x x_0 + k_y y_0)} \, dx_0 \, dy_0 \qquad \text{(A.7)}$$

with $k_x = k\sin\vartheta\cos\varphi$ and $k_y = k\sin\vartheta\sin\varphi$. Equation (A.7) shows that the directivity of the far-field is given by the spatial Fourier transform of the velocity $\hat{u}(x_0, y_0, \omega)$.

B Ray theory

This appendix provides a coordinate-independent derivation of the pressure amplitude along a ray (characteristic curve) in first order ray theory.

The high-frequency, ray-theoretical approximation is obtained by inserting the Wentzel-Kramers-Brillouin-Jeffreys (WKBJ) trial solution

$$\hat{p}(r, \omega) = e^{i\omega\tau(r)} \sum_{j=0}^{\infty} \frac{A_j(r)}{(i\omega)^j} \qquad \text{(B.1)}$$

into the Helmholtz equation (A.2) and dropping all $\mathcal{O}(1/\omega^k)$ terms for $k \geq 1$. This yields the

eikonal equation

$$|\nabla\tau|^2 - \frac{1}{c^2(\boldsymbol{r})} = 0, \tag{B.2}$$

where c is the speed of sound, and the first-order transport equation

$$2\nabla\tau \cdot \nabla A_0 + \left(\nabla^2\tau\right) A_0 = 0. \tag{B.3}$$

The eikonal equation (B.2) has the form $F(\boldsymbol{r}, \boldsymbol{q}) = 0$, with $\boldsymbol{q} \equiv \nabla\tau$ and $F : \mathbb{R}^6 \to \mathbb{R}$ defined by $F(\boldsymbol{r}, \boldsymbol{q}) = |\boldsymbol{q}(\boldsymbol{r})|^2 - 1/c^2(\boldsymbol{r})$. We write a point in $\mathbb{R}^6$ as $\boldsymbol{v} = \begin{bmatrix} \boldsymbol{r} & \boldsymbol{q} \end{bmatrix}$, with $\delta\boldsymbol{v} = \begin{bmatrix} \delta\boldsymbol{r} & \delta\boldsymbol{q} \end{bmatrix}$ a small perturbation. Assuming F is differentiable at $\boldsymbol{v}$, then

$$F(\boldsymbol{v} + \delta\boldsymbol{v}) - F(\boldsymbol{v}) = DF(\boldsymbol{v}) \cdot \delta\boldsymbol{v} + \mathcal{O}(\|\delta\boldsymbol{v}\|) \tag{B.4}$$

by definition of the differential map $DF(\boldsymbol{v}) : \mathbb{R}^6 \to \mathbb{R}^6$. Similarly, $\delta\boldsymbol{q} = D\boldsymbol{q}(\boldsymbol{r})\cdot\delta\boldsymbol{r}+\mathcal{O}(\|\delta\boldsymbol{r}\|)$, and in Euclidean space the differential map is given by the Jacobian matrix, so

$$DF(\boldsymbol{v}) = \begin{bmatrix} \nabla_{\boldsymbol{r}}F(\boldsymbol{v}) & \nabla_{\boldsymbol{q}}F(\boldsymbol{v}) \end{bmatrix}.$$

Both terms on the left hand side of (B.4) must be zero on the solution surface of (B.2). In the limit $\|\delta\boldsymbol{v}\| \to 0$, this allows us to write (B.4) as

$$\frac{\delta\boldsymbol{r}}{\lambda(\boldsymbol{r})} \cdot \left[\lambda(\boldsymbol{r})\nabla_{\boldsymbol{r}}F(\boldsymbol{v}) + D\boldsymbol{q}(\boldsymbol{r}) \cdot \lambda(\boldsymbol{r})\nabla_{\boldsymbol{q}}F(\boldsymbol{v})\right] = 0 \tag{B.5}$$

where $\lambda : \mathbb{R}^3 \to \mathbb{R}$ is an arbitrary function. Next, the vector $\boldsymbol{w} \equiv \lambda(\boldsymbol{r})\nabla_{\boldsymbol{q}}F(\boldsymbol{v})$ may be represented by the equivalence class of curves $\gamma : \mathbb{R} \to \mathbb{R}^3$ such that $\gamma(0) = \boldsymbol{r}$ and $\gamma'(0) = \boldsymbol{w}$. The second term in (B.5) may therefore be written

$$D\boldsymbol{q}(\boldsymbol{r}) \cdot \boldsymbol{w} = \lim_{s\to 0} \frac{\boldsymbol{q}(\boldsymbol{r} + s\boldsymbol{w}) - \boldsymbol{q}(\boldsymbol{r})}{s} = \lim_{s\to 0} \frac{\boldsymbol{q}(\gamma(s) - \mathcal{O}(s^2)) - \boldsymbol{q}(\gamma(0))}{s}$$
$$= = (\boldsymbol{q} \circ \gamma)'(0) = \frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}s}(0).$$

The curve, or ray, $\gamma : s \mapsto \boldsymbol{r}(s)$ is defined by

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s} = \lambda(\boldsymbol{r})\nabla_{\boldsymbol{q}}F(\boldsymbol{v}), \tag{B.6}$$

while equation (B.5) may be reduced to

$$\lambda(\boldsymbol{r})\nabla_{\boldsymbol{r}}F(\boldsymbol{v}) + \frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}s} = 0. \tag{B.7}$$

Finally, for τ we have

$$\frac{\mathrm{d}\tau}{\mathrm{d}s} = D\tau \cdot \boldsymbol{w} = \nabla\tau \cdot \lambda\nabla_{\boldsymbol{q}}F. \tag{B.8}$$

For the eikonal equation, we have $\nabla_{\boldsymbol{q}}F = 2\nabla\tau$ and $\nabla_{\boldsymbol{r}}F = (-2/c^3)\nabla c$. By choosing $\lambda(\boldsymbol{r}) = c(\boldsymbol{r})/2$, the general set of equations (B.6)–(B.8) become in this instance the familiar ray

equations

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s} = c(\boldsymbol{r})\boldsymbol{q} \tag{B.9a}$$

$$\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}s} = -\frac{1}{c^2(\boldsymbol{r})}\nabla c(\boldsymbol{r}) \tag{B.9b}$$

$$\frac{\mathrm{d}\tau}{\mathrm{d}s} = \frac{1}{c(\boldsymbol{r})}. \tag{B.9c}$$

From (B.9a) and (B.2) it is seen that the ray tangent vectors $\mathrm{d}r/\mathrm{d}s$ have unit length, and s is the arc length along the rays. Since $\boldsymbol{q} = \nabla\tau$, the tangent vectors are normal to the surfaces of constant travel time τ, i.e., the wave fronts. In MBES data processing, the initial directions of the ray tangent vectors are given by the main response axes of the beams (Sec. 2.1.3).

In the stratified model, using cylindrical coordinates (ρ, φ, z) where ρ is horizontal range, we have $\mathrm{d}q_\rho/\mathrm{d}s = 0$ and $\mathrm{d}\rho/\mathrm{d}s = c(z(s))q_\rho$. Consequently $q_\rho(s) = \eta$, the ray constant. Since $\mathrm{d}\rho/\mathrm{d}s = \cos\theta$, where θ is the angle between the ray vector and the horizontal plane, the ray constant is $\eta = \cos\theta/c(z)$ for any z. The inverse function theorem now implies that $(\mathrm{d}s/\mathrm{d}\rho)(\rho(s)) = [(\mathrm{d}\rho/\mathrm{d}s)(s)]^{-1} = 1/\eta c(z)$, so that

$$\frac{\mathrm{d}z}{\mathrm{d}\rho} = \lim_{\Delta\rho\to 0}\frac{1}{\Delta\rho}\sqrt{(\mathrm{d}s/\mathrm{d}\rho)^2(\Delta\rho)^2 + \mathcal{O}((\Delta\rho)^3) - (\Delta\rho)^2} = \frac{1}{\eta c}\sqrt{1 - \eta^2 c^2}.$$

Moreover, $\mathrm{d}s/\mathrm{d}\rho = (c\,\mathrm{d}\tau/\mathrm{d}z)(\mathrm{d}z/\mathrm{d}\rho)$, and the forms used for numerical integration with respect to z become

$$\frac{\mathrm{d}\rho}{\mathrm{d}z} = \frac{\eta c}{\sqrt{1 - \eta^2 c^2}} \tag{B.10a}$$

$$\frac{\mathrm{d}\tau}{\mathrm{d}z} = \frac{1}{c\sqrt{1 - \eta^2 c^2}}. \tag{B.10b}$$

If we choose $\lambda = c^2(\boldsymbol{r})/2$ in (B.6)–(B.8), then $\mathrm{d}\tau/\mathrm{d}s = 1$, so $s = \tau$ up to a constant which may be taken as zero. The initial value of a ray tangent vector may be specified as

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\tau}(\tau = 0) = c(z_0)\boldsymbol{u}$$

where $\boldsymbol{u} \in S^2$, the unit sphere, and $\boldsymbol{r}(\tau = 0)$ is the source point. Let $\gamma_{\boldsymbol{u}}$ denote the solution to the initial value problem (B.6)–(B.8), and let Γ denote the map

$$\Gamma : S^2 \times \mathbb{R}^+ \to \mathbb{R}^3, \ (\boldsymbol{u}, \tau) \mapsto \gamma_{\boldsymbol{u}}(\tau). \tag{B.11}$$

For $V \subset S^2$, $\Gamma(V \times \{\tau\}) \subset \mathbb{R}^3$ is an isochrone surface, and a travel time interval $(0 < \tau_1 < \tau_2 < \infty)$ defines a "ray tube" $\Gamma(V, [\tau_1, \tau_2])$. We introduce coordinates on V via any smooth bijective map $\chi : V \to \mathbb{R}^2$ and write $\chi(\boldsymbol{u}) = \begin{bmatrix}\varphi(\boldsymbol{u}) & \vartheta(\boldsymbol{u})\end{bmatrix}$. The transport equation (B.3) may be integrated using Green's first identity for twice differentiable scalar functions ψ, τ on $\Omega \subset \mathbb{R}^3$, i.e.,

$$\int_\Omega (\nabla\tau \cdot \nabla\psi + \psi\nabla^2\tau)\,\mathrm{d}V = \int_{\partial\Omega}\psi(\nabla\tau \cdot \boldsymbol{n})\,\mathrm{d}S \tag{B.12}$$

where $\partial\Omega$ is the boundary of Ω and $\boldsymbol{n}$ is the normal vector on $\partial\Omega$. After multiplying (B.3) by A_0, the transport equation may be written

$$\nabla\tau \cdot \nabla(A_0^2) + A_0^2 \nabla^2\tau = 0. \tag{B.13}$$

Hence from (B.12), with $\psi = A_0^2$, it follows that

$$\int_{\partial\Omega} A_0^2 (\nabla\tau \cdot \boldsymbol{n}) \, \mathrm{d}S = 0. \tag{B.14}$$

In particular, for a ray tube $\Omega = \Gamma(V, [\tau_1, \tau_2])$, the integrand in (B.14) is non-zero only on the isochrones, where $\nabla\tau \cdot \boldsymbol{n} = \pm 1/c$. Hence,

$$\int_{\Gamma(V,\tau_1)} \frac{A_0^2}{c} \, \mathrm{d}S = \int_{\Gamma(V,\tau_2)} \frac{A_0^2}{c} \, \mathrm{d}S. \tag{B.15}$$

Now we set $\tau_2 = \tau_1 + \Delta\tau$ and integrate the function $f = A_0^2/c^2$ over the ray tube. Applying the change of variables formula [139, Sec. 9.3] with the map $\Phi = \Gamma \circ (\chi^{-1} \times \mathrm{id}) : V \times \mathbb{R}^+ \to \mathbb{R}^3$ gives

$$\int_{\Gamma(V,[\tau_1,\tau_2])} f(\boldsymbol{r}) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = \int_{\chi(V) \times [\tau_1,\tau_2]} f(\Phi(\varphi, \vartheta, \tau)) \, |\det D\Phi(\varphi, \vartheta, \tau)| \, \mathrm{d}\varphi \, \mathrm{d}\vartheta \, \mathrm{d}\tau. \tag{B.16}$$

In the limit $\Delta\tau \to 0$, the RHS divided by the LHS in (B.16) gives

$$\lim_{\Delta\tau \to 0} \frac{\Delta\tau}{\Delta\tau} \frac{\int_{\chi(V) \times \tau_1} f(\Phi) \, |\det D\Phi| \, \mathrm{d}\varphi \, \mathrm{d}\vartheta}{\int_{\Gamma(V,\tau_1)} c \frac{A_0^2}{c^2} \, \mathrm{d}S} = \frac{\int_{\chi(V) \times \tau_1} f(\Phi) \, |\det D\Phi| \, \mathrm{d}\varphi \, \mathrm{d}\vartheta}{\int_{\Gamma(V,\tau_1)} \frac{A_0^2}{c} \, \mathrm{d}S} = 1. \tag{B.17}$$

By (B.15), the denominator is independent of τ. Accordingly, the numerator must also be constant for all time slices. Since this holds for arbitrary $V \subset S^2$, the integrand must be constant along rays, i.e.,

$$\frac{A_0^2}{c^2} \, |\det D\Phi(\varphi, \vartheta, \tau)| = \frac{A_0^2(\tau = 0)}{c^2(z_0)} \, |\det D\Phi(\varphi, \vartheta, 0)| \, . \tag{B.18}$$

This equation determines the evolution of the ray amplitude. The matrix of $D\Phi$ is the Jacobian $\boldsymbol{J}_\tau = \partial\boldsymbol{r}/\partial(\varphi, \vartheta, \tau)$. With respect to arc length s, $\mathrm{d}\boldsymbol{r}/\mathrm{d}\tau = c \, \mathrm{d}\boldsymbol{r}/\mathrm{d}s$, so that

$$\det \boldsymbol{J}_s \equiv \det \partial\boldsymbol{r}/\partial(\varphi, \vartheta, s) = \frac{\det \boldsymbol{J}_\tau}{c}.$$

An alternative version of (B.18) is therefore

$$A_0(s) = A_0(0) \sqrt{\frac{c}{c(z_0)} \frac{\det \boldsymbol{J}_s(0)}{\det \boldsymbol{J}_s}}. \tag{B.19}$$

With $\boldsymbol{r}$ given in cylindrical coordinates and φ, ϑ being the spherical angles, the Jacobian deter-

minant for the stratified model, $c = c(z)$, reduces to

$$\det \boldsymbol{J}_s = \frac{\partial \rho}{\partial \vartheta} \frac{\partial z}{\partial s} - \frac{\partial \rho}{\partial s} \frac{\partial z}{\partial \vartheta}$$
$$= \rho \sqrt{\left(\frac{\partial z}{\partial \vartheta}\right)^2 + \left(\frac{\partial \rho}{\partial \vartheta}\right)^2}. \tag{B.20}$$

Equations (B.19)–(B.20) are the formulas used for numerically computing the amplitude reduction due to geometrical spreading.

C Complex pressure

This appendix reviews the complex representation of signals from the point of view of Titchmarsh's theorem [140]. The definition used here for the Fourier transform of p, $\hat{p}(\nu) \equiv \mathcal{F}[p(t)](\nu)$, is

$$\hat{p}(\nu) = \int_{-\infty}^{\infty} p(t) e^{-2\pi i t \nu} \, \mathrm{d}t$$
$$p(t) = \int_{-\infty}^{\infty} \hat{p}(\nu) e^{2\pi i t \nu} \, \mathrm{d}\nu.$$

The Hilbert transform of a function $p : \mathbb{R} \to \mathbb{C}$ is defined by

$$\mathcal{H}[p](t) = -\frac{1}{\pi} \lim_{\epsilon \to 0} \int_{\epsilon}^{\infty} \frac{p(t + \tau) - p(t - \tau)}{\tau} \, \mathrm{d}\tau = \frac{1}{\pi} \mathcal{P}(\frac{1}{t}) * p, \tag{C.1}$$

provided the limit exists. $\mathcal{P}$ denotes the Cauchy principal value and $*$ is the convolution operator. $\mathcal{H}$ is a linear operator. Mathematically, the so-called analytic signal p_a is formed by means of the Hilbert transform:

$$p_a(t) = p(t) + i\,\mathcal{H}[p](t). \tag{C.2}$$

The required properties of the Hilbert transform are provided by the following

Theorem (Titchmarsh [140]). *If $f : \mathbb{R} \to \mathbb{C}$ is square-integrable, then the following statements are equivalent:*

1. *The Fourier transform $\hat{f}(\nu) = \mathcal{F}[f(t)](\nu)$ is zero for $\nu < 0$;*

2. *There is a holomorphic extension $f(t+iy)$ of $f(t)$ in the upper complex half-plane such that the restriction to any horizontal line is square-integrable, i.e,*

$$\int_{-\infty}^{\infty} |f(t + iy)|^2 \, \mathrm{d}x < \infty, \ y > 0. \tag{C.3}$$

3. *The real and imaginary parts of $f(t)$, $\Re f$ and $\Im f$, form a Hilbert transform pair, i.e.,*

$$\Im f(t) = \frac{1}{\pi} \mathcal{P}(\frac{1}{t}) * \Re f(t)$$
$$\Re f(t) = -\frac{1}{\pi} \mathcal{P}(\frac{1}{t}) * \Im f(t).$$

The Fourier transform of

$$g(t) = p(t) + i \, \mathcal{F}^{-1} \left[-i \operatorname{sgn}(\nu) \hat{p}(\nu) \right] \tag{C.4}$$

is

$$\hat{g}(\nu) = \begin{cases} 0 & \text{if } \nu < 0 \\ 2\hat{p}(\nu) & \text{if } \nu \geq 0. \end{cases} \tag{C.5}$$

Hence if g is square-integrable, Titchmarsh's theorem applies, and since $\mathcal{H}$ is linear, Point 3 implies that

$$\mathcal{H}[g](t) = -ig(t). \tag{C.6}$$

Hence

$$-ig(t) = \mathcal{H}[p](t) + i \, \mathcal{H}[f(t)], \tag{C.7}$$

where $\hat{f}(\nu) \equiv -i \operatorname{sgn}(\nu) \hat{p}(\nu)$. If p is real, then $\hat{p}(-\nu) = \overline{\hat{p}(\nu)}$. It follows that also $\hat{f}(-\nu) = \overline{\hat{f}(\nu)}$, and that $f(t)$ is purely real. Then by definition, $\mathcal{H}[f]$ is real, and taking the real part of (C.7) yields

$$\mathcal{H}[p](t) = \Im g(t) = \mathcal{F}^{-1} \left[-i \operatorname{sgn}(\nu) \hat{p}(\nu) \right]. \tag{C.8}$$

A discrete-time Hilbert transform may be based on this formula.

At a fixed position in space, the acoustic pressure field emitted by an echo sounder is a function of time, $p(t) = p_{\boldsymbol{x}}(t) = p(t, \boldsymbol{x})$, with Fourier transform $\hat{p}(\nu)$. For a short-duration, narrow-band pulse, e.g., a gated sine-wave with center frequency ν_c in the ultrasound range, a natural representation is

$$p(t) = a(t) \cos\left[2\pi\nu_c t - \phi(t) \right]. \tag{C.9}$$

The amplitude and phase-modulating functions, $a(t)$ and $\phi(t)$, are assumed to be band-limited, real-valued functions. By band-limited we mean that $\hat{a}(\nu)$ [$\hat{\phi}(\nu)$] is square-integrable and that there is a positive real number Ω such that $\hat{a}(\nu) = 0$ for $|\nu| > \Omega$. Hence,

$$\int_{-\infty}^{\infty} |\hat{a}(\nu)|^2 \, \mathrm{d}\nu = \int_{-\Omega}^{\Omega} |\hat{a}(\nu)|^2 \, \mathrm{d}\nu = K < \infty. \tag{C.10}$$

Furthermore, with $z = t + iy$, an extended function $a \colon \mathbb{C} \to \mathbb{C}$ that approaches $a(t)$ for $y \to 0$ is

$$a(z) = \int_{-\Omega}^{\Omega} \hat{a}(\nu) e^{2\pi i z \nu} \, \mathrm{d}\nu. \tag{C.11}$$

From complex analysis it is known that provided (C.10) holds, such functions are entire, i.e. everywhere analytic (holomorphic) [141, Sec. 19.1]. Let $\phi(z)$ be the corresponding complex analytic extension of $\phi(t)$, and consider the complex function

$$g(z) = a(z) e^{i[2\pi\nu_c z - \phi(z)]}. \tag{C.12}$$

Products, differences, and compositions of analytic functions are analytic. Since $\exp z$ and z as well as $a(z)$ and $\phi(z)$ are entire functions, it follows that $g(z)$ is entire too. The function $g(t + iy)$ is square-integrable for fixed $y > 0$, i.e., on horizontal lines in the upper complex

half-plane: First note that

$$\int_{-\infty}^{\infty} |g(t)|^2 \, \mathrm{d}t = \int_{-\infty}^{\infty} |a(t)|^2 \, \mathrm{d}t = \int_{-\infty}^{\infty} |\hat{a}(\nu)|^2 \, \mathrm{d}\nu = K < \infty \tag{C.13}$$

by the Plancherel theorem and (C.10). For fixed y, the integrand in (C.12) is a function of a real variable, which implies that the complex conjugate $\overline{a(z)}$ can be found by conjugating the integrand. Hence

$$\begin{aligned}
\int_{-\infty}^{\infty} |g(t+iy)|^2 \, \mathrm{d}t &= \int_{-\infty}^{\infty} a(z)\overline{a(z)} \, \mathrm{d}t \\
&= \int_{-\infty}^{\infty} \int_{-\Omega}^{\Omega} \int_{-\Omega}^{\Omega} \hat{a}(\nu)e^{2\pi i \nu z}\overline{\hat{a}(\nu')}e^{-2\pi i \nu' \bar{z}} \, \mathrm{d}\nu \, \mathrm{d}\nu' \, \mathrm{d}t \\
&= \int_{-\Omega}^{\Omega} \int_{-\Omega}^{\Omega} \hat{a}(\nu)\overline{\hat{a}(\nu')}e^{-2\pi y(\nu+\nu')} \left[\int_{-\infty}^{\infty} e^{2\pi i t(\nu-\nu')} \, \mathrm{d}t \right] \mathrm{d}\nu \, \mathrm{d}\nu'.
\end{aligned}$$

The factor in square brackets represents a δ-distribution. Therefore,

$$\int_{-\infty}^{\infty} |g(t+iy)|^2 \, \mathrm{d}t = \int_{-\Omega}^{\Omega} |\hat{a}(\nu)|^2 e^{-4\pi y \nu} \, \mathrm{d}\nu < K e^{4\pi \Omega y} < \infty. \tag{C.14}$$

By the Titchmarsh theorem applied to

$$g(t) = a(t)\cos\left[2\pi\nu_c t - \phi(t)\right] + ia(t)\sin\left[2\pi\nu_c t - \phi(t)\right], \tag{C.15}$$

it follows from (C.13) and (C.14) that

$$\mathcal{H}\left[a(t)\cos\left(2\pi\nu_c t - \phi(t)\right)\right] = a(t)\sin\left(2\pi\nu_c t - \phi(t)\right). \tag{C.16}$$

Consequently, $g(t)$ can be identified with the analytic signal $p_a(t)$. Moreover,

$$p(t) = \Re p_a(t) = \Re\left[\Pi(t)e^{2\pi i \nu_c t}\right], \tag{C.17}$$

with $\Pi(t)$ defined as

$$\Pi(t) = a(t)e^{-i\phi(t)} = p_a(t)e^{-2\pi i \nu_c t}. \tag{C.18}$$

$\Pi(t)$ is the complex pressure (signal) with envelope $|\Pi(t)| = |p_a(t)|$.

Bibliography

[1] D. R. Jackson and M. D. Richardson, *High-Frequency Seafloor Acoustics*, ser. Monograph series in underwater acoustics. New York, NY: Springer, 2007.

[2] D. M. Lawrence, *Upheaval from the Abyss: Ocean Floor Mapping and the Earth Science Revolution*, 1st ed. New Brunswick, NJ: Rutgers University Press, Feb. 2002.

[3] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York, NY: Wiley-Interscience, 2001.

[4] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*, 1st ed. Philadephia, PA: SIAM: Society for Industrial and Applied Mathematics, 2005.

[5] X. Lurton, *An Introduction to Underwater Acoustics: Principles and Applications*, 2nd ed. Berlin, Heidelberg: Springer, 2010.

[6] L. Otto, J. Zimmerman, G. Furnes, M. Mork, R. Sætre, and G. Becker, "Review of the physical oceanography of the North Sea," *Netherlands Journal of Sea Research*, vol. 26, no. 2-4, pp. 161–238, 1990.

[7] N. Winther and J. Johannessen, "North Sea circulation: Atlantic inflow and its destination," *Journal of Geophysical Research — Oceans*, vol. 111, C12018, doi:10.1029/2005JC003310, 2006.

[8] H. De Haas, E. Okkels, and T. Van Weering, "Recent sediment accumulation in the Norwegian Channel, North Sea," *Norges Geologiske Undersøkelse Bull.*, vol. 430, pp. 57–65, 1996.

[9] H. P. Sejrup, I. Aarseth, H. Haflidason, R. Løvlie, Å. Bratten, G. Tjøstheim, C. F. Forsberg, and K. L. Ellingsen, "Quaternary of the Norwegian Channel: glaciation history and paleooceanography," *Norsk geologisk tidsskrift*, vol. 75, no. 2-3, pp. 65–87, 1995.

[10] E. J. Eidem and K. Landmark, "Acoustic seabed classification using QTC IMPACT on single-beam echo sounder data from the Norwegian Channel, northern North Sea," *Continental Shelf Research*, vol. 68, pp. 1–14, Oct. 2013.

[11] D. H. Johnson and D. E. Dudgeon, *Array signal processing: concepts and techniques*, 1st ed., ser. Prentice Hall processing series. Englewood Cliffs, NJ: Prentice Hall, 1993.

[12] P. M. Morse and K. U. Ingard, *Theoretical acoustics*. Princeton, NJ: Princeton University Press, 1986.

[13] L. M. Brekhovskikh and Y. P. Lysanov, *Fundamentals of Ocean Acoustics*, 3rd ed., ser. Modern acoustics and signal processing.   New York, NY: Springer, 2003.

[14] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics*, 2nd ed., ser. Modern acoustics and signal processing.   New York, NY: Springer, 2011.

[15] G. V. Frisk, *Ocean and Seabed Acoustics: a Theory of Wave Propagation*, 1st ed.   Englewood Cliffs, NJ: PTR Prentice Hall, 1994.

[16] N. Bleistein, J. K. Cohen, and J. W. Stockwell, Jr., *Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion*, 1st ed., ser. Interdiciplinary applied mathematics.   New York, NY: Springer, 2001.

[17] R. E. Francois and G. R. Garrison, "Sound absorption based on ocean measurements. Part 2: Boric acid contribution and equation for total absorption," *J. Acoust. Soc. Amer.*, vol. 72, no. 6, pp. 1879–1890, 1982.

[18] Kongsberg Maritime AS, "EM 710 / Multibeam echo sounder / High resolution seabed mapping system," Fact sheet, Feb. 2011. [Online]. Available: http://www.km.kongsberg. com/ks/web/nokbg0397.nsf/AllWeb/71E3E270C5D26C47C1256FD30040DA3E/$file/ 164939ae_em710_product_specification.pdf?OpenElement

[19] M. A. Ainslie, *Principles of Sonar Performance Modeling*, 1st ed., ser. Springer-Praxis books in geophysical sciences.   Berlin Heidelberg: Springer-Verlag, 2010.

[20] D. R. Jackson and K. B. Briggs, "High-frequency bottom backscattering: Roughness versus sediment volume scattering," *The Journal of the Acoustical Society of America*, vol. 92, no. 2, pp. 962–977, 1992. [Online]. Available: http://scitation.aip.org/content/ asa/journal/jasa/92/2/10.1121/1.403966

[21] D. R. Jackson and A. N. Ivakin, "Scattering from elastic sea beds: First-order theory," *The Journal of the Acoustical Society of America*, vol. 103, no. 1, pp. 336–345, 1998. [Online]. Available: http://scitation.aip.org/content/asa/journal/jasa/103/ 1/10.1121/1.421093

[22] K. L. Williams, J. M. Grochocinski, and D. R. Jackson, "Interface scattering by poroelastic seafloors: First-order theory," *The Journal of the Acoustical Society of America*, vol. 110, no. 6, pp. 2956–2963, 2001. [Online]. Available: http://scitation.aip.org/content/asa/journal/jasa/110/6/10.1121/1.1414704

[23] S. T. McDaniel, "Fluctuations of the acoustic field Bragg-backscattered from rough surfaces," *Waves in Random Media*, vol. 6, no. 3, pp. 229–249, 1996.

[24] E. Jakeman and P. N. Pusey, "Significance of K distributions in scattering experiments," *Physical Review Letters*, vol. 40, no. 9, pp. 546–550, 1978.

[25] F. W. J. Olver, "Bessel functions of integer order," in *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, M. Abramowitz and I. A. Stegun, Eds.   National Bureau of Standards, 1964.

[26] E. Jakeman and P. N. Pusey, "A model for non-Rayleigh sea echo," *IEEE Trans. Antennas Propag.*, vol. 24, no. 6, pp. 806–814, 1976.

[27] E. Jakeman, "On the statistics of K-distributed noise," *Journal of Physics A: Mathematical and General*, vol. 13, pp. 31–48, 1980.

[28] D. A. Abraham and A. P. Lyons, "Novel physical interpretation of K-distributed reverberation," *IEEE J. Ocean. Eng.*, vol. 27, no. 4, pp. 800–813, Oct. 2002.

[29] A. P. Lyons, S. F. Johnson, D. A. Abraham, and E. Pouliquen, "High-frequency scattered envelope statistics of patchy seafloors," *IEEE J. Ocean. Eng.*, vol. 34, no. 4, pp. 451–458, Oct. 2009.

[30] K. D. Ward, "Compound representation of high resolution sea clutter," *Electron. Lett.*, vol. 17, no. 16, pp. 561–563, Aug. 1981.

[31] C. J. Oliver, "A model for non-Rayleigh scattering statistics," *Optica Acta*, vol. 31, no. 6, pp. 701–722, 1984.

[32] C. Oliver and S. Quegan, *Understanding synthetic aperture radar images*.   Raleigh, NC: SciTech Publishing Inc., 2004.

[33] C. J. Oliver, "Correlated K-distributed clutter models," *Optica Acta*, vol. 32, no. 12, pp. 1515–1547, 1985.

[34] L. Hellequin, J.-M. Boucher, and X. Lurton, "Processing of high-frequency multibeam echo sounder data for seafloor characterization," *IEEE J. Ocean. Eng.*, vol. 28, no. 1, pp. 78–89, Jan. 2003.

[35] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th ed.   Elsevier Academic Press, 2009.

[36] J. Preston and R. Kirlin, "Comment on 'Acoustic seabed classification: improved statistical method'," *Can. J. Fish. Aquat. Sci.*, vol. 60, pp. 1299–1300, 2003.

[37] J. M. Preston, A. C. Christney, L. S. Beran, and W. T. Collins, "Statistical seabed segmentation—from images and echoes to objective clustering," in *Proc. Seventh European Conf. Underwater Acoustics ECUA 2004*.   Delft, The Netherlands: Delft University of Technology, Jul. 2004, pp. 813–818.

[38] J. Preston, "Automated acoustic seabed classification of multibeam images of Stanton Banks," *Applied Acoustics*, vol. 70, no. 10, pp. 1277–1287, Oct. 2009.

[39] G. Schwarz, "Estimating the dimension of a model," *Annals of Statistics*, vol. 6, no. 2, pp. 461–464, 1978.

[40] H. Bhat and N. Kumar, "On the derivation of the Bayesian Information Criterion," 2010. [Online]. Available: http://nscs00.ucmerced.edu/~nkumar4/BhatKumarBIC.pdf

[41] I. Madsen and J. Tornehave, *From Calculus to Cohomology. De Rham cohomology and characteristic classes.* Cambridge, UK: Cambridge University Press, 1997.

[42] B. Clarke, E. Fokoué, and H. H. Zhang, *Principles and Theory for Data Mining and Machine Learning*, 1st ed., ser. Springer Series in Statistics. New York, NY: Springer, 2009.

[43] A. C. Jensen and A. S. Solberg, "Fast hyperspectral feature reduction using piecewise constant function approximations," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 4, pp. 547–551, Oct. 2007.

[44] J. S. Weszka, C. R. Dyer, and A. Rosenfeld, "A comparative study of texture measures for terrain classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, no. 4, pp. 269–285, Apr. 1976.

[45] U. C. Herzfeld and C. A. Higginson, "Automated geostatistical seafloor classification—principles, parameters, feature vectors, and discrimination criteria," *Computers & Geosciences*, vol. 22, no. 1, pp. 35–52, Feb. 1996.

[46] M. F. J. Wilson, B. O'Connel, C. Brown, J. C. Guinan, and A. J. Grehan, "Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope," *Marine Geodesy*, vol. 30, no. 1-2, pp. 3–35, May 2007.

[47] M. Einsiedler and T. Ward. (2015, Mar.) Functional analysis, spectral theory, and applications. Draft version. [Online]. Available: https://dl.dropboxusercontent.com/u/2098511/FAnotes.pdf

[48] B. Jähne, *Digital Image Processing*, 6th revised and extended ed. Berlin Heidelberg New York: Springer, 2005.

[49] W. A. Götz and H. J. Druckmüller, "A fast digital Radon transform—an efficient means for evaluating the Hough transform," *Pattern Recognition*, vol. 29, no. 4, pp. 711–718, Apr. 1996.

[50] M. L. Brady, "A fast discrete approximation algorithm for the Radon transform," *SIAM J. Comput.*, vol. 27, no. 1, pp. 107–119, 1998.

[51] W. H. Press, "Discrete Radon transform has an exact, fast inverse and generalizes to operations other than sums along lines," *Proc. Natl. Acad. Sci. USA (PNAS)*, vol. 103, no. 51, pp. 19 249–19 254, Dec. 2006.

[52] K. Landmark, A. H. S. Solberg, F. Albregtsen, A. Austeng, and R. E. Hansen, "A Radon-transform-based image noise filter - with applications to multibeam bathymetry," *IEEE Trans. Geosci. Remote Sens*, vol. 53, no. 11, pp. 6252–6273, Nov. 2015, published online 15 June 2015.

[53] Y. Saad and M. H. Schultz, "GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems," *SIAM. J. Sci. Statist. Comput.*, vol. 7, no. 3, pp. 856–869, Jul. 1986.

[54] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, ser. Frontiers in applied mathematics.   Philadephia, PA: Soc. Ind. Appl. Math., 1995, vol. 16.

[55] I. C. F. Ipsen and C. D. Meyer, "The idea behind Krylov methods," *Amer. Math. Monthly*, vol. 105, pp. 889–899, 1997.

[56] R. B. Lehoucq and D. C. Sorensen, "Deflation techniques for an implicitly restarted Arnoldi iteration," *SIAM. J. Matrix Anal. & Appl.*, vol. 17, no. 4, pp. 789–821, Oct. 1996.

[57] R. B. Lehoucq, D. C. Sorensen, and C. Yang. (1998) ARPACK Users' Guide: Solution of Large Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods. Web version available. [Online]. Available: http://www.caam.rice.edu/software/ARPACK/UG/ug.html

[58] J. T. Anderson, D. Van Holliday, R. Kloser, D. G. Reid, and Y. Simard, "Acoustic seabed classification: current practise and future directions," *ICES Journal of Marine Science*, vol. 65, no. 6, pp. 1004–1011, 2008.

[59] C. J. Brown and P. Blondel, "Developments in the application of multibeam sonar backscatter for seafloor habitat mapping," *Applied Acoustics*, vol. 70, no. 10, pp. 1242–1247, Oct. 2009.

[60] P. Blondel, *The Handbook of Sidescan Sonar*.   Chichester, UK: Springer Praxis Publishing, 2009.

[61] J. V. Barrie and K. W. Conway, "Seabed characterization for the development of marine renewable energy on the Pacific margin of Canada," *Continental Shelf Research*, vol. 83, no. SI, pp. 45–52, Jul. 2014.

[62] K. L. Ellingsen, J. Gray, and E. Bjørnholm, "Acoustic classification of seabed habitats using the QTC VIEW system," *ICES Journal of Marine Science*, vol. 59, pp. 825–835, 2002.

[63] R. Freitas, S. Silva, V. Quintino, A. Rodrigues, K. Rhynas, and W. Collins, "Acoustic seabed classification of marine habitats: studies in the western coastal-shelf area of Portugal," *ICES Journal of Marine Science*, vol. 60, pp. 599–608, 2003.

[64] K. Haris, B. Chakraborty, B. Ingole, A. Menezes, and R. Srivastava, "Seabed habitat mapping employing single and multi-beam backscatter data: A case study from the western continental shelf of India," *Continental Shelf Research*, vol. 48, pp. 40–49, 2012.

[65] R. J. Kloser, J. D. Penrose, and A. J. Butler, "Multi-beam backscatter measurements used to infer seabed habitats," *Continental Shelf Research*, vol. 30, no. 16, pp. 1772–1782, Sep. 2010.

[66] A. Rattray, D. Ierodiaconou, L. Laurenson, S. Burq, and M. Reston, "Hydro-acoustic remote sensing of benthic biological communities on the shallow South East Australian continental shelf," *Estuarine Coastal and Shelf Science*, vol. 84, no. 2, pp. 237–245, Sep. 2009.

[67] A. Micallef, T. P. Le Bas, V. A. I. Huvenne, P. Blondel, V. Huhnerbach, and A. Deidun, "A multi-method approach to benthic habitat mapping of shallow coastal areas with high-resolution multibeam data," *Continental Shelf Research*, vol. 39-40, pp. 14–26, May 2012.

[68] T. Medialdea, L. Somoza, R. León, M. Farrán, G. Ercilla, A. Maestro, D. Casas, E. Llave, F. Hernández-Molina, M. Fernández-Puga, and B. Alonso, "Multibeam backscatter as a tool for sea-floor characterization and identification of oil spills in the Galicia Bank," *Marine Geology*, vol. 249, pp. 93–107, 2008.

[69] C. Wienberg and A. Bartholomä, "Acoustic seabed classification in a coastal environment (outer Weser Estuary, German Bight)-a new approach to monitor dredging and dredge spoil disposal," *Continental Shelf Research*, vol. 25, no. 9, pp. 1143–1156, 2005.

[70] A. Bartholomä, "Acoustic bottom detection and seabed classification in the German Bight, southern North Sea," *Geo-Marine Letters*, vol. 26, pp. 177–184, 2006.

[71] E. Hammerstad, "Backscattering and seabed image reflectivity," Kongsberg Maritime, EM technical note, Jan. 2000. [Online]. Available: http://www.km.kongsberg.com/ks/web/nokbg0397.nsf/AllWeb/C2AE0703809C1FA5C1257B580044DD83/$file/EM_technical_note_web_BackscatteringSeabedImageReflectivity.pdf?OpenElement

[72] R. Eckhoff, "Back scatter corrections," by personal communication, Kongsberg Maritime, Aug. 2009, unpublished.

[73] D. Sternlicht and C. De Moustier, "Remote sensing of sediment characteristics by optimized echo-envelope matching," *Journal of the Acoustical Society of America*, vol. 114, no. 5, pp. 2727–2743, 2003.

[74] M. Snellen, K. Siemens, and D. Simons, "Model-based sediment classification using single-beam echosounder signals," *Journal of the Acoustical Society of America*, vol. 129, no. 5, pp. 2878–2888, 2011.

[75] C. De and B. Chakraborty, "Model-based acoustic remote sensing of seafloor characteristics," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3868–3877, 2011.

[76] L. Fonseca and L. Mayer, "Remote estimation of surficial seafloor properties through the application angular range analysis to multibeam sonar data," *Mar. Geophys. Res.*, vol. 28, no. 2, pp. 119–126, Jul. 2007.

[77] K. Haris, B. Chakraborty, C. De, R. G. Prabhudesai, and W. Fernandes, "Model-based seafloor characterization employing multi-beam angular backscatter data — A comparative study with dual-frequency single beam," *J. Acoust. Soc. Am.*, vol. 130, no. 6, pp. 3623–3632, Dec. 2011.

[78] B. Chakraborty, H. W. Schenke, V. Kodagali, and R. Hagen, "Seabottom characterization using multibeam echosounder angular backscatter: An application of the composite roughness theory," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2419–2422, Sep. 2000.

[79] "High-frequency ocean environmental acoustic models handbook," University of Washington, Seattle, WA, Applied Physics Laboratory technical report APL-UW TR 9407, 1994.

[80] K. L. Williams, "An effective density fluid model for acoustic propagation in sediments derived from biot theory," *J. Acoust. Soc. Am.*, vol. 110, no. 5, pp. 2276–2281, Nov. 2001.

[81] J. Yu, S. A. Henrys, C. Brown, I. Marsh, and G. Duffy, "A combined boundary integral and Lambert's Law method for modelling multibeam backscatter data from the seafloor," *Continental Shelf Research*, vol. 103, pp. 60–69, 2015.

[82] A. Amiri-Simkooei, M. Snellen, and D. Simons, "Principal component analysis of single-beam echo-sounder signal features for seafloor classification," *IEEE Journal of Oceanic Engineering*, vol. 36, no. 2, pp. 259–272, 2011.

[83] F. Madricardo, J. Tegowski, and S. Donnici, "Automated detection of sedimentary features using wavelet analysis and neural networks on single beam echosounder data: A case study from the Venice Lagoon, Italy," *Continental Shelf Research*, vol. 43, pp. 43–54, 2012.

[84] P. A. Van Walree, J. Tegowski, C. Laban, and D. G. Simons, "Acoustic seafloor discrimination with echo shape parameters: A comparison with ground truth," *Continental Shelf Research*, vol. 25, pp. 2273–2293, 2005.

[85] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning. Data Mining, Inference and Prediction*, 2nd ed., ser. Springer Series in Statistics. New York, NY: Springer, 2009.

[86] C. K. Wentworth, "A scale of grade and class terms for clastic sediments," *The Journal of Geology*, vol. 30, no. 5, pp. 377–392, 1922. [Online]. Available: http://www.jstor.org/stable/30063207?origin=JSTOR-pdf

[87] ——, "Fundamental limits to the sizes of clastic grains," *Science*, vol. 77, no. 2009, pp. 633–634, Jun. 1933.

[88] R. L. Folk, "A review of grain-size parameters," *Sedimentology*, vol. 6, no. 2, pp. 73–93, 1966.

[89] E. L. Hamilton and R. T. Bachman, "Sound velocity and related properties of marine sediments," *J. Acoust. Soc. Am.*, vol. 72, no. 6, pp. 1891–1904, Dec. 1982.

[90] R. L. Folk and W. C. Ward, "Brazos River bar: a study in the significance of grain size parameters," *Journal of Sedimentary Petrology*, vol. 27, no. 1, pp. 3–26, Mar. 1957.

[91] R. Folk, "The distinction between grain size and mineral composition in sedimentary-rock nomenclature," *Journal of Geology*, vol. 62, no. 4, pp. 344–359, 1954.

[92] The Geological Survey of Norway (NGU) / The Norwegian Mapping Authority (Kartverket). (2006) Fagområde: Løsmassegeologi. SOSI standard—generell objektkatalog versjon 4.0 (object catalogue for Quaternary geology maps). [Online]. Available: http://www.statkart.no/Documents/Standard/SOSI-standarden%20del% 201%20og%202/SOSI%20standarden/Losmasse.pdf

[93] G. Mavko, T. Mukerji, and J. Dvorkin, *The Rock Physics Handbook: Tools for Seismic Analysis in Porous Media*, 1st ed. Cambridge, UK: Cambridge University Press, 1998.

[94] K. B. Briggs, A. P. Lyons, E. Pouliquen, L. A. Mayer, and M. D. Richardson, "Seafloor roughness, sediment grain size, and temporal stability," in *Underwater Acoustic Measurements: Technologies and Results*, J. P. Papadakis and L. Bjørnø, Eds., Heraklion, Crete, Jun. 2005, pp. 337–343, also available as Naval Research Laboratory technical report no. NRL/PP/7430-05-6.

[95] L. Fonseca, C. Brown, B. Calder, L. Mayer, and Y. Rzhanov, "Angular range analysis of acoustic themes from Stanton Banks Ireland: A link between visual interpretation and multibeam echosounder angular signatures," *Applied Acoustics*, vol. 70, no. 10, pp. 1298–1304, Oct. 2009.

[96] H. Matsumoto, R. Dziak, and C. G. Fox, "Estimation of seafloor microtopographic roughness through modeling of acoustic backscatter data recorded by multibeam sonar systems," *J. Acoust. Soc. Am.*, vol. 94, no. 5, pp. 2276–2287, Nov. 1993.

[97] C. De and B. Chakraborty, "Model-based acoustic remote sensing of seafloor characteristics," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3868–3877, Oct. 2011.

[98] G. Canepa and N. G. Pace, "Seafloor segmentation from multibeam bathymetric sonar," in *Proceedings of the Fifth European Conference on Underwater Acoustics, ECUA 2000*, M. E. Zakharia, P. Chevret, and P. Dubail, Eds., vol. 1. Lyon, France: European Commision, Jul. 2000, pp. 361–367.

[99] Z.-H. Michalopoulou, D. Alexandrou, and C. de Moustier, "Application of neural and statistical classifiers to the problem of seafloor characterization," *IEEE J. Ocean. Eng.*, vol. 20, no. 3, pp. 190–197, Jul. 1995.

[100] L. J. Hamilton and I. Parnum, "Acoustic seabed segmentation from direct statistical clustering of entire multibeam sonar backscatter curves," *Continental Shelf Research*, vol. 31, no. 2, pp. 138–148, Feb. 2011.

[101] B. Chakraborty, V. Kodagali, and J. Baracho, "Sea-floor classification using multibeam echo-sounding angular backscatter data: A real-time approach employing hybrid neural network architecture," *IEEE J. Ocean. Eng.*, vol. 28, no. 1, pp. 121–128, Jan. 2003.

[102] R. Che Hazan, D. Ierodiaconou, and J. Monk, "Evaluation of four supervised learning methods for benthic habitat mapping using backscatter from multi-beam sonar," *Remote Sensing*, vol. 4, no. 11, pp. 3427–3443, Nov. 2012.

[103] L. Hellequin, X. Lurton, and J.-M. Augustin, "Postprocessing and signal corrections for multibeam echosounder images," in *OCEANS '97. MTS/IEEE Conference Proceedings*, vol. 1, Halifax, NS, Oct. 1997, pp. 23–26.

[104] G. Lamarche, X. Lurton, A.-L. Verdier, and J.-M. Augustin, "Quantitative character-ization of seafloor substrate and bedforms using advanced processing of multibeam backscatter—Application to Cook Strait, New Zealand," *Continental Shelf Research*, vol. 31, pp. S93–S109, 2011.

[105] J. H. Clarke, "Towards remote seafloor classification using the angular response of acous-tic backscattering: A case study from multiple overlapping GLORIA data," *IEEE J. Ocean. Eng.*, vol. 19, no. 1, pp. 112–127, Jan. 1994.

[106] Y. Le Gonidec, G. Lamarche, and I. C. Wright, "Inhomogeneous substrate analysis using EM3000 backscatter imagery," *Marine Geophysical Researches*, vol. 24, pp. 311–327, 2005.

[107] G. Le Chenadec, J. Boucher, and X. Lurton, "Angular dependence of K-distributed sonar data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1224–1235, May 2007.

[108] G. Le Chenadec, "Analyse de descripteurs énergétiques et statistiques de signeaux sonar pour la charactérisation des fonds marins (An analysis of energetical and statistical features of sonar signals for the characterization of seabeds)," Thèse de doctorat (Ph.D. thesis), L'Université de Bretagne Occidentale, Jul. 2004, in French. [Online]. Available: http://archimer.ifremer.fr/doc/00000/242/

[109] D. G. Simons and M. Snellen, "A Bayesian approach to seafloor classification using multi-beam echo-sounder backscatter data," *Applied Acoustics*, vol. 70, no. 10, pp. 1258–1268, Oct. 2009.

[110] R. T. Bachman, "Acoustic and physical property relationships in marine sediment," *J. Acoust. Soc. Am.*, vol. 78, no. 2, pp. 616–621, Aug. 1985.

[111] P. A. Van Walree, M. Ainslie, and D. Simons, "Mean grain size mapping with single-beam echo sounders," *Journal of the Acoustical Society of America*, vol. 120, no. 5, pp. 2555–2566, 2006.

[112] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Upper Saddle River, NJ: Pearson/Prentice Hall, 2008.

[113] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 4, pp. 610–621, Nov. 1973.

[114] T. B. Reed and D. Hussong, "Digital image processing techniques for enhancement and classification of SeaMARC II side scan sonar imagery," *J. Geophys. Res.*, vol. 94, no. B6, pp. 7469–7490, Jun. 1989.

[115] P. Blondel and O. Gómez Sichi, "Textural analyses of multibeam sonar imagery from Stanton Banks, Northern Ireland continental shelf," *Applied Acoustics*, vol. 70, no. 10, pp. 1288–1297, Oct. 2009.

[116] V. A. I. Huvenne, P. Blondel, and J.-P. Henriet, "Textural analyses of sidescan sonar imagery from two mound provinces in the Porcupine Seabight," *Marine Geology*, vol. 189, no. 3-4, pp. 323–341, Sep. 2002.

[117] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, Jan. 1996.

[118] T. Ojala and M. Pietikäinen, "Unsupervised texture segmentation using feature distributions," *Pattern Recognition*, vol. 32, no. 3, pp. 477–486, 1999.

[119] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.

[120] M. Pietikäinen, T. Ojala, and Z. Xu, "Rotation-invariant texture classification using feature distributions," *Pattern Recognition*, vol. 33, no. 1, pp. 43–52, 2000.

[121] T. Kohonen, "The self-organizing map," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464–1480, 1990.

[122] R. Bø, L. Rise, and D. Ottesen, "Elongate depressions on the southern slope of the Norwegian Trench (Skagerrak): morphology and evolution," *Marine Geology*, vol. 146, pp. 191–203, 1998.

[123] M. Hovland, "Elongated depressions associated with pockmarks in the western slope of the Norwegian Trench," *Marine Geology*, vol. 51, pp. 35–46, 1983.

[124] L. Breiman, *Probability*, ser. Classics in applied mathematics, vol. 7.   Philadelphia, PA: SIAM, 1992.

[125] M. P. do Carmo, *Differential geometry of curves and surfaces*.   Englewood Cliffs, NJ: Prentice-Hall, 1976.

[126] T. F. Chan and J. Shen, *Image processing and analysis: variational, PDE, wavelet, and stochastic methods*.   Philadelphia, PA: SIAM, 2005.

[127] K. Landmark and A. H. S. Solberg, "Reducing scalloping in synthetic aperture radar images using a composite image transform," in *Proc. SPIE 9643, Image and Signal Processing for Remote Sensing XXI, 96431B*, Oct. 2015. [Online]. Available: http://dx.doi.org/10.1117/12.2194952

[128] C. C. Paige and M. A. Saunders, "LSQR: An algorithm for sparse linear equations and sparse least squares," *ACM. Trans. Math. Softw.*, vol. 8, no. 1, pp. 43–71, Jun. 1982.

[129] G. Francheschetti and R. Lanari, *Synthetic Aperture Radar Processing.* Boca Raton, FL: CRC Press, 1999.

[130] F. Natterer and F. Wübbeling, *Mathematical Methods in Image Reconstruction.* Philadephia, PA: Soc. Ind. Appl. Math., 2001.

[131] DECONVREG. MathWorks Matlab R2014a Documentation. [Online]. Available: http://www.mathworks.com/help/images/ref/deconvreg.html

[132] S. E. Dosso, P. L. Nielsen, and C. H. Harrison, "Bayesian inversion of reverberation and propagation data for geoacoustic and scattering parameters," *J. Acoust. Soc. Am.*, vol. 125, no. 5, pp. 2867–2880, May 2009.

[133] S. E. Dosso and M. J. Wilmut, "Data uncertainty estimation in matched-field geoacoustic inversion," vol. 31, no. 2, pp. 470–479, Apr. 2006.

[134] F. Albregtsen. (2008, Nov.) Statistical Texture Measures Computed from Gray Level Cooccurence Matrices. Dept. Informatics, Univ. Oslo. [Online]. Available: http://www.uio.no/studier/emner/matnat/ifi/INF4300/h08/undervisningsmateriale/glcm.pdf

[135] L. Hamilton, P. Mulhearn, and R. Poeckert, "Comparison of RoxAnn and QTC-View acoustic bottom classification system performance for the Cairns area, Great Barrier Reef, Australia," *Continental Shelf Research*, vol. 19, no. 12, pp. 1577–1597, 1999.

[136] K. Landmark, A. H. S. Solberg, C.-I. C. Nilsen, A. Austeng, and R. E. Hansen, "Bayesian seabed classification using angle-dependent backscatter data from multibeam sonar," in *11th European Conference on Underwater Acoustics 2012 (ECUA 2012)*, ser. Proceedings of the Institute of Acoustics Volume 34 Pt.3. Institute of Acoustics (IOA), Oct. 2012, pp. 417–424.

[137] M. Hovland, "Elongated depressions associated with pockmarks in the western slope of the Norwegian Trench," *Marine Geology*, vol. 51, no. 1-2, pp. 35–46, Feb. 1983.

[138] R. Bøe, L. Rise, and D. Ottesen, "Elongate depressions on the southern slope of the Norwegian Trench (Skagerrak): morphology and evolution," *Marine Geology*, vol. 146, no. 1-4, pp. 191–203, Apr. 1998.

[139] J. E. Marsden and M. J. Hoffman, *Elementary Classical Analysis*, 2nd ed. New York, NY: Freeman, 1993.

[140] E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals*, 2nd ed. Oxford: Oxford at the Clarendon Press, 1948, ch. 5.4-5.5.

[141] W. Rudin, *Real and Complex Analysis*, third (international) ed. Singapore: McGraw-Hill, 1987.

Paper I:

Acoustic seabed classification using QTC IMPACT on single-beam echo sounder data from the Norwegian Channel, northern North Sea

Acoustic seabed classification using QTC IMPACT on single-beam echo sounder data from the Norwegian Channel, northern North Sea

Ellen Johanne Eidem and Knut Landmark

Abstract

Sediment mapping is important for understanding the physical processes, the impact of human activity, and the conditions for marine life on the seabed. For this purpose, the seabed classification tool QTC IMPACT analyses statistical variations in single-beam echo sounder data. QTC was applied in a large and physically diverse area of the Norwegian Channel, between 59°30' N and 61°N, to produce a new sediment map and to verify the QTC algorithm. The results were interpreted using ground truth (grain size analyses of 40 gravity cores and five grab samples), multi-beam echo sounder bathymetry (MBES), and seismo-acoustic profiles. Surficial sediments were divided into five classes: (1) mud and silt, (2) a variety of clay, silt and sand, (3) sandy mud with gravel, (4) sand with gravel, and (5) clay and sandy clay. Along the Norwegian coast, where MBES imagery shows evidence of glacial erosion, the surficial sediment distribution is variable. The echo shape analysis of QTC did not produce a natural partition of the data, and statistical assumptions did not always hold. Sediment classification was therefore sensitive to the choice of cluster algorithm. However, QTC produced the most physically plausible results on a large scale compared to other cluster algorithms. Class boundaries were consistent with supporting data. One exception is a transition from muddy to sandy sediments not visible in seismo-acoustic data. A possible explanation is that seabed fluid seepage and water current erosion cause sand particle transport into the western part of the channel. The study confirms the capability of QTC in a complex environment, but there are some possible improvements.

Index Terms

Acoustic seabed classification, marine geology, North Sea, Norwegian Channel, sediments, single-beam echo sounder.

I. INTRODUCTION

Acoustic seabed classification (ASC) is a technology for mapping surficial seabed properties and sediment distribution with echo sounders [1]. The idea is that a change in sediment composition often implies a change in acoustic properties. This results in a systematic difference in the recorded echoes, provided the data have been corrected for extraneous influences, such as variable water depth or instrument settings. ASC may resolve fine variations in seabed properties with a high spatial resolution, which makes it a valuable complement to sediment sampling. Information about seabed composition is useful in a range of problems, including sonar performance prediction, the original purpose of this work, benthic habitat mapping and marine resource management [2]–[5], environmental monitoring [6], [7], and geotechnical engineering [8]. A scientific understanding of the physical processes that form the seabed is also a goal in itself, and new acoustic techniques for seabed mapping are instrumental to achieve this.

Physics-based ASC methods estimate seabed acoustical parameters by fitting simulated data to observations [9]–[11]. Such methods are sensitive to instrument calibration errors and theory errors. Statistical methods search the data for systematic differences between distinct seabed types [5], [12]–[14]. Such methods use pattern classification algorithms [15], [16] and work best where observations divide naturally into a set of discrete classes, the nature of which may not be fully known in advance. Empirical relations have been used to estimate mean grain size from e.g. measured echo energy [17], but it is not straightforward to link acoustical response with lithology. Sound scattering strength is sensitive to the seabed surface roughness [18], which may be influenced by coverage with benthic organisms and bioturbation, or sediment transport and attendant bedform development due to currents or ocean waves. As highlighted by e.g. Wienberg and Bartholomä [7], an ASC system is best employed in conjunction with other data to avoid interpretational mistakes, particularly side scan sonar images (in their case) or multi-beam echo sounder (MBES) bathymetry.

The objectives of this paper are twofold. First, to present a new map of the surficial seabed composition in a large and physically diverse area of the Norwegian Channel (NC), northern North Sea, with water depths from 100 to 700 m. The primary tool has been the statistical ASC software QTC IMPACT [19], applied to single-beam echo sounder (SBES) data. QTC software is a widely used ASC tool which has proven successful in other studies [3], [7], [8], [20]–[22]. The second objective of the paper is to evaluate the QTC approach and result, including underlying assumptions, using an extensive supporting data set. Analysis of supporting data, especially ground truth (grain size analyses), is key to the interpretation and verification of the ASC results.

The paper first introduces the physical setting of the study area (Section II). The established Quaternary stratigraphy, on which our interpretation of seismo-acoustic data is based, is briefly explained. Section III-A summarizes the data acquisition, in particular the pre-processing of the SBES data, and the preparation and analysis of the sediment samples. Section III-B examines the principles of the QTC method, which include a method for depth compensation

of echoes, a data reduction step, and a clustering algorithm derived from Bayesian classification theory. Section III-B3 explains our approach to validate the QTC clustering algorithm. Section IV presents the raw classification results, which required merging two datasets with different pulse lengths, and the results of the cluster validity analysis. The ASC method is an unsupervised one, so interpretation of the classes is subsequently done by comparison with the distribution of the sediment samples. The ASC classes are finally compared with parametric sonar profiles and geomorphology. Section V discusses the significance of the ASC results, the performance of QTC, and compares with previous work. The paper ends with the main conclusions from this work (Section VI).

II. STUDY AREA AND SEAFLOOR GEOLOGY

The NC is a depression in the continental shelf along the coast of southern Norway, with depths up to 700 m in the Skagerrak area (Fig. 1). It is a major topographic feature of the North Sea, which is a shallow continental sea bounded by the British Isles, Norway, and the northern European continent [23]. The western slope of the NC is one of the major pathways for inflowing North Atlantic water. Mixed oceanic and fresh water flows out of the North Sea as the northbound Norwegian Coastal Current in the eastern and central NC [24]. Thus the circulation pattern of the North Sea is influenced by the topography of the NC. Conversely, the NC is an important trap for water-borne sediments in the North Sea [25].

The study area spans about 23 000 km^2 from the western Norwegian shoreline to the eastern margin of the North Sea plateau, between 59°30'N and 61°N latitude and to the east of 02°30'E (Fig. 1). The Quaternary sediments in this area consist of alternating layers of till and marine or glaciomarine sediments. The established Quaternary stratigraphy [28] is based in part on a drilled core sample (0-219 m below seafloor) from the Troll Field (Fig. 1). Of interest in this paper are the B1 unit and the A unit. Unit B1 is connected with seabed moraines in the eastern NC. Unit A is the top sediment layer in the central NC, truncated by the eastern moraines and, about 5-7 km from the shore, the crystalline basement (Fig. 2). Unit B1 is a till dating to the maximum of the Weichselian glaciation, corresponding to a homogenous diamicton in the Troll core with about 30 % coarse material (grain size > 63 μm) and high shear strength [28], [29]. Unit A is a horizontally bedded formation with sediments deposited from the onset of the last deglaciation to the present. The largest accumulations of unit A sediments are found outside the Korsfjorden inlet and at the foot of the North Sea Plateau (up to 60 m, compared to 17 m in the Troll core) (Fig. 3). The topmost layer in unit A, Holocene, has a variable thickness of typically 3-7 m as observed in present parametric sonar data (Fig. 3). In the Troll core, the topmost layer of unit A is a 4.5 m thick marine deposit consisting of silty clay with high water content [28], [29].

Rise and Rokoengen [30] mapped the surficial sediments of the adjoining (partly overlapping) area between 60°30'N and 62°N, using seafloor photographs and sediment sample analyses. They divided the seabed sediments into four groups, from sand with coarse material to sandy clay (Fig. 1), and found increasingly more fine-grained sediments towards the central axis of the channel. Large local variations on the Måløy Plateau were interpreted as an effect of glacial activity.
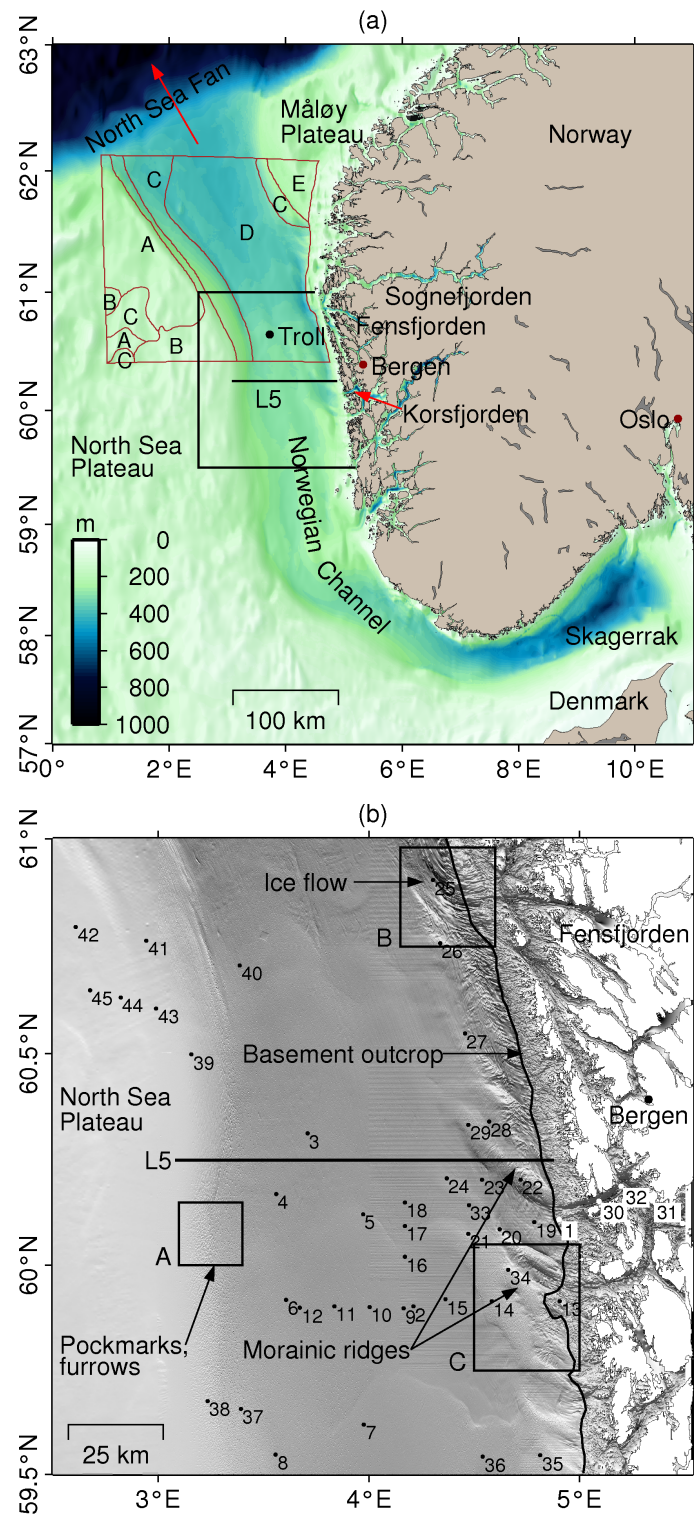
Fig. 1. (1a) Overview and bathymetry of the Norwegian Channel. The study area to the west of Bergen is outlined. The Troll core is located at 60.64°N, 3.72°E. The brown polygons represent the sediment classes of Rise et al. (1984) [26]: sand with coarse material (A); very well-sorted sand (B); silty-clayey sand (C); silty, sandy clay (D); and variable (E). Map based on bathymetric data from [27, last accessed Dec. 2012]. The line L5 is the transect of the seismic profile shown in Fig. 2. (1b) Bathymetric relief of the study area based on MBES data, with principal morphological features indicated. The location of the 45 grain size analysed sediment samples (sample numbers given in Table 1) are also shown. The rectangular areas (A)-(C) are enlarged in Figs. 12 and 14.
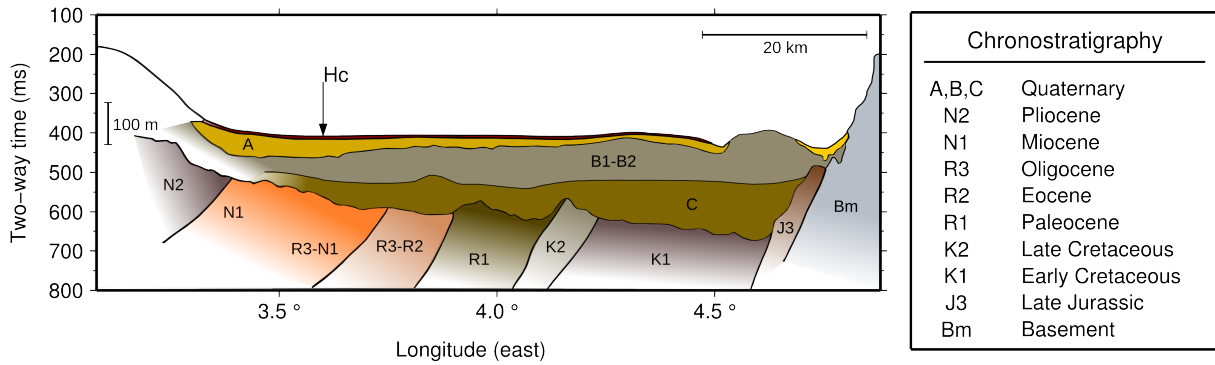
Fig. 2. Shallow stratigraphy (schematic) along seismic line L5 (Fig. 1). The interpretation is based on the stratigraphic work of Sejrup et al. [28], Rise and Rokoengen [30], and Sigmond [31]. The principal Quaternary units, all unconsolidated, are A (pelite), B1 (diamicton), B2 (sandy pelite over diamicton), and C (pelite). The topmost layer of unit A is a Holocene marine deposit (Hc). Crystalline basement (Bm) is exposed in the east, closest to shore. The total thickness of Quaternary sediments is about 150-200 m assuming a fixed average sound speed of 1800 m/s. The average sound speed has been estimated at 39 locations using seismic refraction and reflection data from sonobuoys (unpublished work by FFI). The vertical length scale has been calculated using 1800 m/s and is not accurate for the water or bedrock layers.

III. MATERIAL AND METHODS

A. Data acquisition and preparation

All the data were acquired by the Norwegian Defence Research Establishment (FFI), on FFI's research vessel M/S H. U. Sverdrup II, during the years 2003-2008. The MBES was a 95 kHz Kongsberg Maritime EM 1002. The EM 1002 has 111 beams and $2° \times 2°$ 3 dB beamwidth. The SBES was a 38 kHz Kongsberg Maritime EA 600 with a 3 dB beamwidth of $7.1°$.

The SBES data were acquired using — arbitrarily — two pulse lengths (0.512 ms and 1.024 ms), and pre-processed with EchoView 4.40 (Myriax, previously SonarData, Tasmania). EchoView splits the raw data files into separate sonar (Sv—volume backscattering strength) and navigation files that can be read by QTC IMPACT. The transformation from raw data to backscattering strength consisted of automatically calibrating the data with respect to absorption factor (9.2-10 dB/km), transmit power (800-2000 kW), and sound speed (1500 m/s or as measured); the values used for each data file were as in the first ping. The absorption factor was computed according to the Francois and Garrison model [32]. Further calibration factors were transducer gain, two-way beam angle, and 3 dB beamwidth (all user set). It was not possible to compensate for different pulse lengths. The SBES and MBES data were collected using a dense survey pattern with track line separation 500-600 m or less (full-coverage MBES surveys).

High-resolution seismic equipment (two 40 in^3 airguns with single-channel streamer) and a parametric sonar were used to map the Quaternary layer sequence. The parametric sonar was a Kongsberg Maritime Topas PS 018 transmitting a 20 ms 2-4 kHz chirp pulse. These data were acquired on a coarse survey grid with a total track length of about 2500 km. In total 39 sonobuoys were dropped during the seismic surveys to acquire wide-angle seismic data.

In total 40 single gravity cores with lengths between 0.4 and 2.8 m and a diameter of 63 mm were acquired at water depths of 170-670 m. Five grab samples were collected with a

(a) Unit A thickness

(b) Holocene thickness
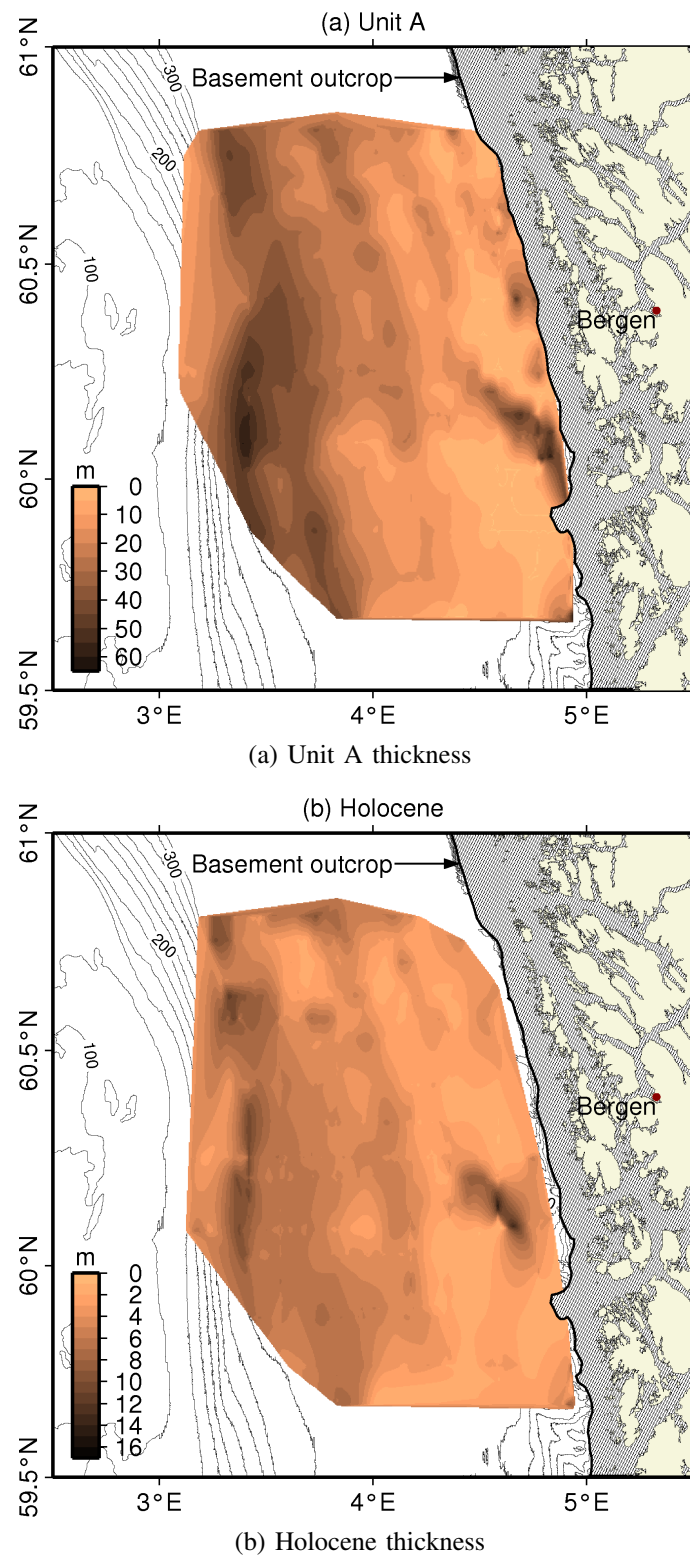
Fig. 3. Thickness (isopach maps) of unit A sediments in the Norwegian Channel based on interpretation of parametric sonar data (unpublished work by FFI). (3a) Total thickness of unit A. Two-way travel time (seabed to top Weichselian moraine) was converted to meters using an average sound speed of 1550 m/s. (3b) Thickness of surficial Holocene sediments, assuming an average sound speed of 1500 m/s. The interpretation of base Holocene was difficult in the western part of the channel. Seismic refraction analysis of sonobuoy data constrains the sound speed in unit A, but the uncertainty in the thickness estimates is at least 3-4 %. For reference, bathymetric contours are also shown.

Van Veen grab sampler at water depths of 110-180 m. The sediment samples were sent to the University of Bergen (UiB) in four batches (2006, 2007 and two in 2008) for quantitative analyses. The gravity cores were analysed with respect to grain size distribution, sound speed, gamma density, shear strength and partly magnetic susceptibility. Only grain size distributions were obtained for the five grab samples. An additional 32 grab samples collected with the Van Veen grab sampler in the early phase of the study were analysed rudimentarily and given a qualitative description by FFI and the Geological Survey of Norway (NGU). The analysis was limited because the study was initially conceived mainly as an MBES survey program.

The UiB grain size laboratory analysis consisted of two steps. The cores were subsampled at ca. every 20-40 cm (2006-2007) and 4-85 cm (2008), and approximately 10-20 g samples were collected for each analysis. The samples were submerged in distilled water and agitated for 48-72 hours, before sieving through progressively smaller meshes: 1 mm, 0.25 mm (2008 only), 0.150 mm (2007 only), 0.125 mm, and 0.063 mm. The sediments collected on each sieve were dried and reweighed. The distribution of sediments smaller than 0.063 mm was analysed down to 2 m (2006, 2008) and 1 m (2007 only) using a Micrometrics Sedigraph 5100 (2006 only) and Micrometrics Sedigraph III 5120 (2007-2008). Nine of the ten core samples containing sediments coarser than 1 mm according to the first analysis, were re-sieved for classification purposes in 2008 using in addition a 2 mm mesh to establish the amount of gravel.

The grain size laboratory analysis of the five grab samples (2008) consisted of collecting and sieving approximately 100-400 g material of each sample. The meshes had sizes 4, 2, 1, 0.5, 0.25, 0.125 and 0.063 mm. One of the samples was analysed down to 2 μm using the Micrometrics Sedigraph III 5120.

Core and grab samples with less than 2 % gravel were classified according to the Folk classification system [33] based on the grain size distribution. Sediment samples with at least 2 % gravel were classified according to a modified version of the Folk classification system, where no division between 2 and 30 % gravel exist, and the class gravelly, sandy mud (gsM) is introduced; consequently the original Folk class gravelly mud (gM) is reduced to cover samples containing maximum 10 % sand and 30 % gravel. The classification was performed on the top subsample of each core, with mean depth ranging from 1 to 16 cm. In average the mean depth was 5.8 ± 3.1 cm (the wavelength of the acoustic signals for a 38 kHz echo sounder is 4 cm). The classification is based on the mass percentage of gravel (grain size > 2.0 mm), sand (grain size 63 μm-2 mm), silt (grain size 2-63 μm), and clay (grain size < 2 μm). The sand and gravel content of all the subsamples at each site was calculated to find layers in the gravity cores; of special interest has been the composition of the surface layer.

B. Acoustic classification

QTC IMPACT analyses the shape of seabed echoes [19]. The digitally sampled echoes are first aligned and summed over consecutive pings (stacking), a common noise reduction step. QTC has an automatic bottom detection algorithm for aligning echoes before summation. The shape analysis produces 166 parameters (features) which are stored in a feature vector ($\boldsymbol{x}$). A representative subset of feature vectors (the "catalogue data") is used to compute a principal component transform. By discarding all but three principal components, each observed feature vector is subsequently reduced to a data (observation) vector $\boldsymbol{d}(\boldsymbol{x}) \in \mathbb{R}^3$.

QTC IMPACT version 3.40 and 3.50 was used in this study. The stack length was five pings. Bad traces were manually removed using a built-in data editor. The reference depth

was set to 300 m; the observed depth range was from 100 to 600 m. The data with pulse lengths $\tau = 0.512$ ms and $\tau = 1.024$ ms were processed separately. The several depressions (pockmarks) observed were not removed. The classified observations were decimated with a 200 m $\times$ 200 m block mode filter using Generic Mapping Tools [34], and plotted along with ground truth data on shaded relief maps of MBES bathymetry for evaluation. Only blocks with a unique class were accepted. The classification results for the two pulse lengths were merged into one map in a similar process. Finally, the classes were given a sedimentological description based on classification of the 40 core and five grab samples.

1) Echo compensation: An echo time series depends on intrinsic properties of the seabed (roughness and volume scattering strength, reflection coefficient, sound attenuation), but also extraneous factors (source strength, receiver sensitivity, pulse length, frequency, beamwidth, propagation loss, water depth, and seabed slope). The QTC algorithm normalizes the stacked echoes to unit peak amplitude, so only echo shape is taken into account [19]. Besides seabed slope, this leaves the effect of water depth on the echo length as the principal extraneous factor. The procedure to eliminate this factor is described in [35]. Briefly, the echo length is parameterized by water depth D, beamwidth Ψ, pulse length τ, and sediment penetration depth. For a given frequency f, QTC computes a standard penetration depth p_{ref}. This is used to compute a reference echo length $T_{ref} = T_{ref}(D, \Psi, \tau; p_{ref})$, where D is the only parameter that varies from ping to ping. Each echo is resampled at a rate $N/T_{ref}(D)$, where the fixed number of samples N is set by the operator. With respect to a reference depth D_{ref}, this resampling is equivalent to a depth-dependent time dilation $t \mapsto [T_{ref}(D_{ref})/T_{ref}(D)]\, t$, whereby shallow echoes are stretched out in time and vice versa. This makes it possible to compare echo shapes at varying depths.

2) Unsupervised classification and statistical assumptions: The QTC algorithm is based on a normal mixture model [19], [36]. By assumption, the probability density function (PDF) for an observation $\boldsymbol{d}$ is

$$p(\boldsymbol{d} \mid \boldsymbol{q}, \boldsymbol{w}; c) = \sum_{j=1}^{c} p(\boldsymbol{d} \mid \omega_j, \boldsymbol{q}_j) P(\omega_j), \tag{1}$$

where $p(\boldsymbol{d} \mid \omega_j; \boldsymbol{q}_j)$ is a normal PDF associated with the class denoted ω_j. $P(\omega_j)$ is the unknown prior probability or relative class frequency, and c is the number of classes. Here we let $\boldsymbol{q}_j$ denote the nine unknown parameters of $p(\boldsymbol{d} \mid \omega_j, \boldsymbol{q}_j)$ in three dimensions, and write $\boldsymbol{q} = [\boldsymbol{q}_1, \ldots \boldsymbol{q}_c]$ and $\boldsymbol{w} = [P(\omega_1), \ldots P(\omega_{c-1})]$. The minimum-error Bayesian classification rule implies that $\boldsymbol{d}$ should be assigned to the class ω_i if the discriminant function

$$g_i(\boldsymbol{d}) = \log p(\boldsymbol{d} \mid \omega_i, \boldsymbol{q}_i) + \log P(\omega_i) \tag{2}$$

is larger than $g_j(\boldsymbol{d})$ for all $j \neq i$ [16]. For a catalogue dataset of n observations, $\boldsymbol{D} = \{\boldsymbol{d}_1, \ldots, \boldsymbol{d}_n\}$, QTC runs a stochastic optimization process, minimizing a cost function, to determine the unknown parameters $\boldsymbol{q}$ and $\boldsymbol{w}$. The cost function of QTC [19], [37] will here be written in terms of Eq. (2), thus

$$E = -2 \sum_{k=1}^{c} \sum_{i=1}^{n_k} g_k(\boldsymbol{d}_{ki}). \tag{3}$$

Here $\boldsymbol{d}_{ki}$ is the ith member of the kth class, and n_k is the number of observations currently assigned to the kth class (so $n = \sum_{k=1}^{c} n_k$). At every stage in the process, maximum likelihood

(ML) estimates $\hat{\boldsymbol{q}}_j^{ML}$ and $\hat{\boldsymbol{w}}_j^{ML}$ are computed using only the samples currently assigned to class ω_j. These estimates are accurate when the classes are well separated.

The number of classes is determined by repeating this process for different values of c to find the value that minimizes the quantity $E + M \log n$, where $M = \dim \boldsymbol{q} + \dim \boldsymbol{w} = 10c - 1$ is the number of free parameters. This maximizes the likelihood of c, given the data $\boldsymbol{D}$ and Eq. (1), but only under certain conditions that we will here make clear. From [38, Ch.] and Eq. 8 in [39] it follows that

$$-2 \log P(\boldsymbol{D} \mid c) \approx \min E + M \log n \qquad (4)$$

The conditions are the following: (1) observations are independent; (2) $\log p(\boldsymbol{D} \mid \boldsymbol{q}, \boldsymbol{w}; c)$ has a global maximum and decays quickly; (3) there is no prior information about $\boldsymbol{q}$ and $\boldsymbol{w}$; (4) n is large; (5) no value of c is a priori more likely than others in the range we consider; and (6) $\min E = E(\boldsymbol{q}^{ML}, \boldsymbol{w}^{ML})$, i.e. the true ML estimates are found. If $\min E = E(\boldsymbol{q}^{ML}, \boldsymbol{w}^{ML})$, the right hand side of Eq. (4) is the Bayesian information criterion (BIC), originally derived for models in the exponential distribution family [40]. With all $g_i(\boldsymbol{d})$ finally determined from the catalogue data, the minimum-error rule is applied to classify all observations.

3) Cluster validation: Cluster analysis can be sensitive to the choice of algorithm (and input parameters), especially if the data do not form compact, well-separated clusters. We have applied three alternative clustering algorithms and used three cluster validity indices [16] to measure the cluster compactness and separation as a function of c (the number of classes). The clustering algorithms were expectation-maximization (EM) with a normal mixture distribution, k-means, and self-organizing maps (SOM) [16], [41]. The EM algorithm was chosen because it is related to the QTC algorithm. The validity indices were the global Silhouette index S_c [16], [42], the Davies-Bouldin index DB_c [43], and the Dunn index D_c [44]. Each index is an overall cluster similarity (DB_c) or dissimilarity measure (S_c, D_c) for a given partition into c clusters. All three indices depend on a distance measure $r(\boldsymbol{x}, \boldsymbol{y})$ between pairs of observations $\boldsymbol{x}$ and $\boldsymbol{y}$. Here we used the Euclidean distance and a covariance-weighted distance. The latter was given as

$$r(\boldsymbol{x}, \boldsymbol{y})^2 = (\boldsymbol{x} - \boldsymbol{y})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{y}) \qquad (5)$$

where $\boldsymbol{\Sigma}$ is the sample covariance matrix of all observations, with each observation given relative to it's own cluster mean. Ideally DB_c should be small ($DB_c \geq 0$), S_c should be close to 1 ($-1 \leq S_c \leq 1$), and D_c should be larger than 1. This signifies that the clusters are compact and well separated.

To check if the sample distribution of each class is compatible with the normal assumption, we used the fact that a normal PDF is on the form $p(\boldsymbol{d} \mid \omega_i) \propto \exp(-r^2/2)$, where r^2 is the squared Mahalanobis distance,

$$r^2 = (\boldsymbol{d} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(\boldsymbol{d} - \boldsymbol{\mu}_i). \qquad (6)$$

Here $\boldsymbol{\mu}_i$ is the mean vector and $\boldsymbol{\Sigma}_i$ is the sample covariance matrix of the cluster to which $\boldsymbol{d}$ is assigned. For a multinormal distribution in K dimensions, r^2 follows a $\chi^2(K)$-distribution with K degrees of freedom. A relevant way to assess the QTC assumption is therefore to consider, for each class, a quantile-quantile plot of the empirical distribution of r^2 versus the $\chi^2(K)$-distribution [45]. For a multinormal distribution, the observations are expected to lie on the line with unit slope.

IV. RESULTS

A. Sediment samples

Table I lists the location, water depth, core length (recovery), percentage of clay, silt, sand and gravel, median grain size, and the Folk class of the 45 top layer grain size analysed sediment samples (one sample per site). Most of the samples consisted mainly of fine sediments; 11 out of 44 samples contained more than 2 % gravel, including one sample which contained more than 30 % gravel. The sand content of the 11 samples with coarse sediments was higher than 40 %. Of the remaining 33 samples, 22 contained less than 10 % sand. One reason why the cores contained mostly fine sediments is that gravity core samples are difficult to obtain from hard, coarse seabeds.

The sediment samples from the central NC were classified mainly as silt (Z) or mud (M); mud is a mixture of silt and clay (Fig. 4). Two samples from the eastern part of the NC were classified as clay (C) and three samples retrieved at the foot of the North Sea Plateau as sandy silt (sZ). Grab samples from the North Sea Plateau contained 26-98 % sand and 0-33 % gravel; one grab sample was classified as pure sand (S). Sand and gravel was also found on the two southern morainic ridges to the east. Here the amount of gravel was 3-26 %, which is higher than the amount of coarse material in the B1 unit of the Troll core (2-3 % above 1mm) [28]. The amount of sand was 36-52 %. Samples retrieved from the east-trending Korsfjorden were classified as clay (C), sandy clay (sC) and mud (M). Shell and shell fragments were observed in sediment samples from the North Sea Plateau. Dividing the 45 grain size analysed sediment samples into four groups based on the Folk class, and the 32 grab samples analysed rudimentarily into four equivalent groups based on the qualitative descriptions, demonstrates the diversity of the study area (Fig. 4). Although the rudimentary analysis is not precise enough to distinguish between all the Folk classes, this information both support and supplement the pattern observed from the grain size analyses. The median grain size, sometimes used to describe sediments (Freitas et al., 2003), ranges from 1ϕ to 11ϕ in the study area (Fig. 5a). Layers are present in the sediment cores that have top subsamples containing sand and gravel; typically the surficial sediments are coarser than beneath. Some cores show intermediate layers with increased sand and gravel content (Fig. 5b). The 22 sediment cores that have top subsamples with less than 10 % sand show with two exceptions no dominant layers.

B. Acoustic classes and seabed types

The need to process two pulse lengths separately made the processing and analysis more complicated; two sets of statistics and classification rules had to be computed. The full dataset for each pulse length was used as catalogue data. The statistical clustering process (Section III-B) was repeated for c in the range 4-8 (the number of classes). For a fixed value of c, the process was repeated 10 times with different random starting configurations. The number of classes (five) in the final ASC results was selected as the lowest number of classes where an increment of one did not change the pattern visually on a large scale.

The classification results for the $\tau = 1.024$ ms dataset (Fig. 6b) show that one class covers most of the central channel. The slope on the western side is divided into two parallel classes which follow the bathymetric contours. The North Sea Plateau came out as one class. To the east all three classes observed in the west are mixed together, and the picture is more complex. There is also a fifth class close to the coast and in the fiords. The classification

TABLE I

THE LOCATION (DEG. LATITUDE/LONGITUDE), WATER DEPTH, CORE LENGTH (RECOVERY), CONTENT OF CLAY ($< 2\,\mu$M), SILT (2–63 μM), SAND (63 μM–2 MM), AND GRAVEL (> 2 MM), MEDIAN GRAIN SIZE, FOLK CLASS, AND QTC CLASS (TURQUOISE, YELLOW, RED, BROWN AND PURPLE) OF THE GRAIN SIZE ANALYSED SEDIMENT SAMPLES. SAMPLES 1–40 WERE OBTAINED USING A GRAVITY CORER; 41–45 USING A GRAB.

No.	Lat	Lon	D (m)	L (cm)	Clay(%)	Silt(%)	Sand (%)	Gravel (%)	ϕ_{50}	Folk[a]	QTC
1	60.11	4.90	240	123	21.6	30.2	43.5	4.7	4.7	gsM	Outside
2	59.90	4.21	285	155	35.1	56.4	8.5	0.0	8.0	M	T
3	60.31	3.71	298	193	33.8	61.0	5.2	0.0	7.5	M	T
4	60.17	3.56	292	104	22.1	69.9	8.0	0.0	6.8	Z	T
5	60.12	3.97	279	126	31.5	66.6	2.0	0.0	7.9	Z	T
6	59.92	3.61	280	190	21.1	74.6	4.4	0.0	6.4	Z	T
7	59.62	3.98	274	276	51.3	47.4	1.3	0.0	9.1	M	T
8	59.55	3.56	247	190	28.0	59.6	12.3	0.0	6.2	sZ	T
9	59.90	4.17	277	182	18.1	78.3	3.6	0.0	6.8	Z	T
10	59.90	4.00	277	117	27.5	68.9	3.6	0.0	7.4	Z	T
11	59.90	3.84	283	114	17.0	81.7	1.2	0.0	7.1	Z	T
12	59.90	3.67	280	117	14.3	82.1	3.6	0.0	6.5	Z	T
13	59.92	4.91	199	42	23.4	21.7	49.2	5.7	3.3	gmS	Outside
14	59.91	4.58	243	198	15.0	30.4	52.0	2.5	3.4	gmS	R
15	59.92	4.36	283	204	33.0	45.1	13.2[b]	8.7[b]	7.7	sM?	T
16	60.02	4.17	283	157	34.0	63.2	2.8	0.0	8.1	M	T
17	60.09	4.17	291	167	41.5	55.8	2.7	0.0	8.5	M	T
18	60.15	4.17	294	138	33.7	63.9	2.4	0.0	8.1	M	T
19	60.10	4.78	321	208	36.0	50.1	14.0	0.0	8.1	sM	T
20	60.09	4.62	312	124	38.9	57.3	3.8	0.0	8.4	M	T(R)
21	60.08	4.47	274	193	28.3	35.6	36.0	0.0	6.8	sM	R
22	60.20	4.72	285	53	38.7	27.0	33.8	0.5	7.7	sM	R
23	60.20	4.54	315	149	28.8	68.1	3.1	0.0	7.6	Z	T
24	60.21	4.37	284	158	34.9	60.1	5.0	0.0	8.1	M	T
25	60.90	4.30	434	232	66.9	31.8	1.3	0.0	11.0	C	Y
26	60.76	4.34	326	119	20.6	13.1	62.0	4.2	2.7	gmS	Y
27	60.55	4.46	317	142	57.7	35.2	7.1	0.0	9.7	M	T
28	60.34	4.57	295	214	53.4	28.8	17.8	0.0	9.3	sM	R
29	60.33	4.47	275	148	35.7	19.0	42.3	3.0	5.9	gsM	Y
30	60.15	5.09	615	230	52.6	21.7	25.7	0.0	9.2	sC	P
31	60.10	5.49	590	163	65.1	34.4	0.6	0.0	10.3	M	P(R)
32	60.19	5.20	672	172	70.3	27.7	2.0	0.0	11.0	C	P(R,T)
33	60.14	4.47	289	149	64.0	31.8	4.2	0.0	10.3	C	P
34	59.99	4.66	236	51	29.0	11.7	50.7	8.5	2.6	gmS	R(Y)
35	59.55	4.81	211	44	28.9	16.6	40.0	14.5	3.5	gsM	R
36	59.54	4.54	260	56	19.4	12.0	43.0	25.6	2.1	gmS	R(Y)
37	59.66	3.39	220	90	13.2	58.7	28.0	0.0	4.6	sZ	T(Y)
38	59.68	3.23	168	237	45.5	50.4	4.1	0.0	8.4	M	T
39	60.50	3.16	168	143	35.0	39.3	25.7	0.0	6.6	sM	R
40	60.71	3.39	324	131	28.4	59.0	12.6	0.0	5.5	sZ	Y
41	60.76	2.94	184		3.3	9.3	87.0	0.5	3.4	zS	Y(R)
42	60.80	2.61	124			0.5[c]	66.1	33.4	0.9	sG	B(R)
43	60.61	2.99	139			25.9[c]	71.7	2.4	3.2	gmS	R
44	60.63	2.82	109			0.3[c]	92.0	7.8	2.3	gS	B
45	60.65	2.68	106			1.5[c]	98.5	0.1	2.4	S	B

[a] The modified version of the Folk classification system is used for sediment samples with more than 2 % gravel.

[b] The sample was sieved using 1 mm mesh as maximum. For the class sM, less than 2 % gravel is assumed.

[c] Four grab samples were sieved using a 63 μm mesh as minimum; this is the combined silt and clay percentage.
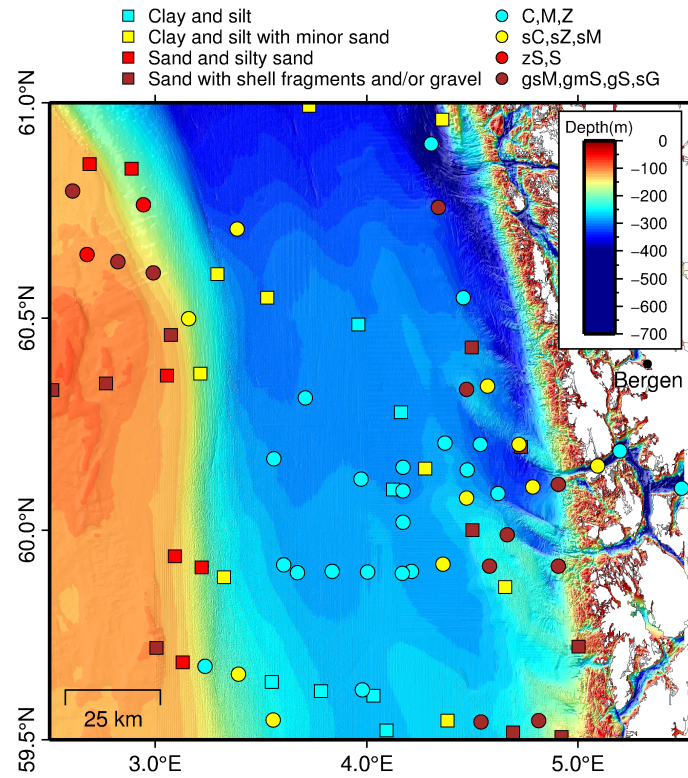
Fig. 4. The sedimentological and oceanographical diversity of the study area. The 45 grain size analysed sediment samples of the top layer (circles) and the 32 grab samples analysed rudimentarily (squares) are divided into four groups. Water depths greater than 400 m are plotted using a single color.
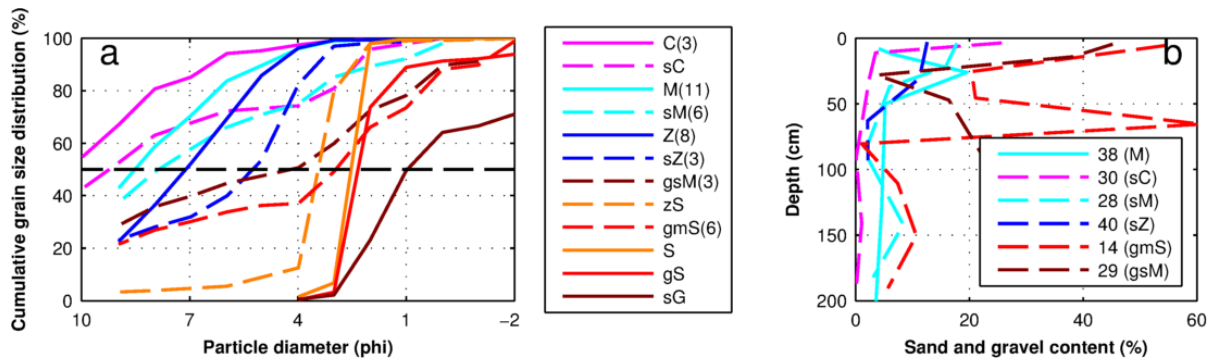

Fig. 5. (a) The average grain size distribution of top layer sediment samples with equal Folk class (the number of samples averaged in parenthesis). The grain size is expressed in units of $\phi = -2\log_2 D$, where D is the particle diameter in mm. Abbreviations: clay (C), mud (M), silt (Z), sandy clay (sC), sandy mud (sM), sandy silt (sZ), silty sand (zS), sand (S), gravelly sandy mud (gsM), gravelly muddy sand (gmS), gravelly sand (gS), and sandy gravel (sG). (b) Sand and gravel content versus depth in selected core samples.

results for the $\tau = 0.512$ ms dataset (Fig. 6a) show that the central NC is one class, just as for the $\tau = 1.024$ ms dataset. The moraines in the east stand out as one class, but with traces of other classes. There are three classes on the western slope and North Sea Plateau, just as for the $\tau = 1.024$ ms dataset.

The final classification results with an interpretation of the five acoustic classes are presented in Fig. 6c, and show that there is good agreement between crossing legs, although some discrepancies occur. The joint geographical distribution of the five acoustic classes and the grain size analysed surficial sediment samples indicates a close correspondence. The turquoise
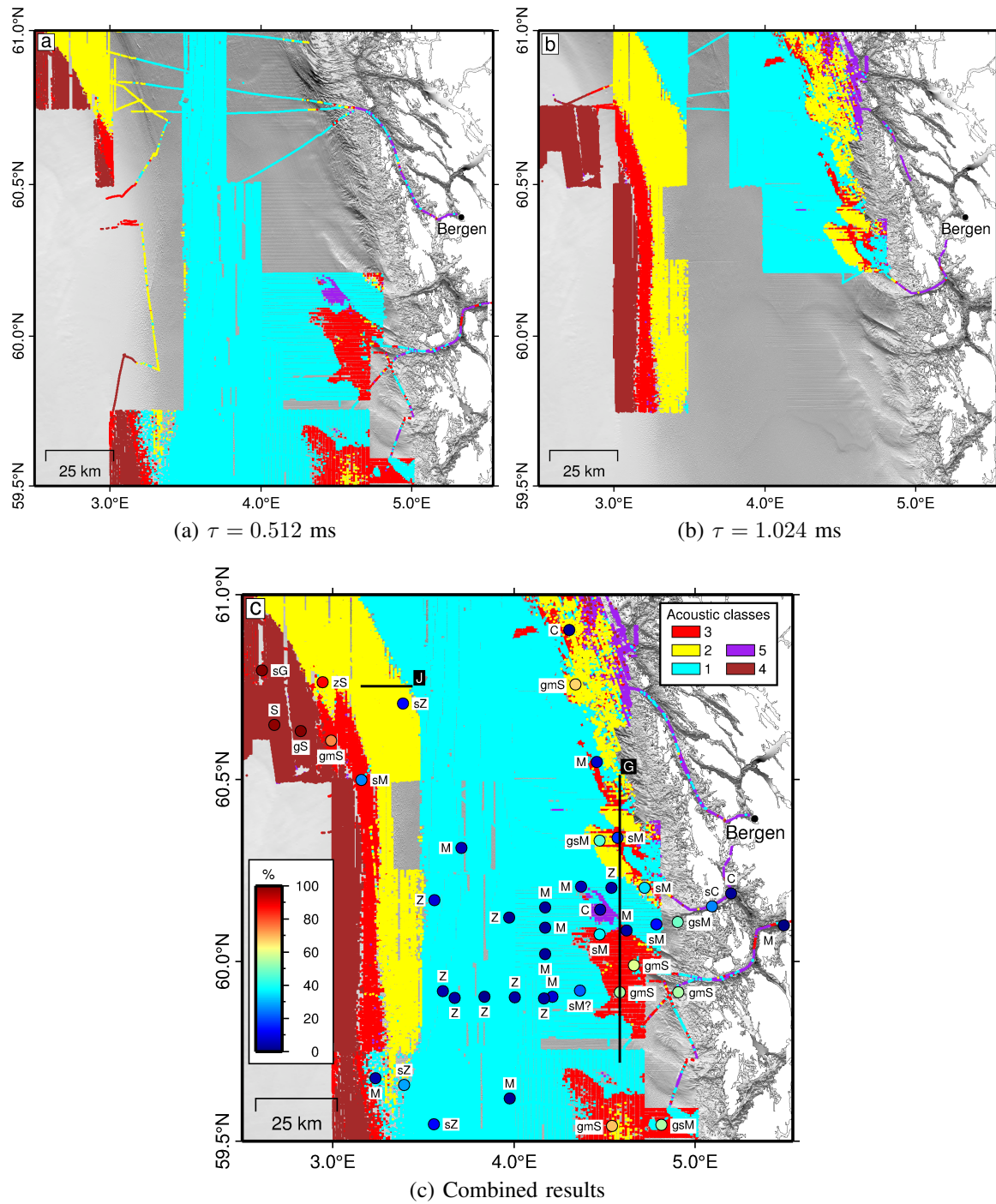
Fig. 6. The ASC results for pulse length $\tau = 0.512$ ms (a) and $\tau = 1.024$ ms (b). The colors have been deliberately chosen to match. The final classification result (c) includes a qualitative description of the five classes: (1) mud and silt, (2) a variety of clay, silt and sand, (3) sandy mud with gravel, (4) sand with gravel, and (5) clay and sandy clay. The Folk class and content of sand and gravel of cores (top subsamples) and grab samples are shown for comparison; the color scale is in percentage. The parametric sonar survey transects G and J are indicated in black.

class dominates the central channel with sediment samples classified as mud (M) and silt (Z). Four samples from this area are classified as sandy mud (sM) and sandy silt (sZ). The two sM samples are near the moraines, which may explain the content of coarser sediments (14 and 22 % sand and gravel). The two sZ samples associated with the turquoise class are in the

south-western portion of the study area; one is on the border with the area of yellow shading in Fig. 6, containing 28 % sand and gravel, and the other is from the channel, containing 12 % sand and gravel. Excluding these four core samples (#8, #15, #19, and #37, see Fig. 1b), the average content of clay, silt, sand and gravel is 33, 63, 4 and 0 %, respectively. The average median grain size is 7.8ϕ. We note that a muddy (M) core sample (#38, see Fig. 1b) matches the turquoise class, which here extends up into the elsewhere sandy western slope. This core has a sandy layer revealed in the subsample at 26 cm depth (Fig. 5), with a sand content of 19 % (less than 1 % gravel). The sandy surface layer of the two sZ samples extends down to about 50 cm depth (the top subsamples are from 3 to 16 cm depth). The thickness of the sandy surface layer of the two sM samples (#15 and #19) has not been determined due to under-sampling. Far into the channel north-west of the northern moraine, shell fragments were observed in two of the grab samples analysed rudimentarily only.

The red class covers the moraines, some small areas in the north-east, and a part of the western slope (shown as area of red shading in Fig. 6). The area contains a mixture of samples classified as sandy mud (sM), gravelly sandy mud (gsM), or gravelly muddy sand (gmS). Some rudimentarily analysed grab samples contain shells and shell fragments. The red class sediments are thus heterogeneous. The average median grain size of the five core samples from the two southern moraines (#14, #21, and #34–36) is 3.7ϕ; the average content of clay, silt, sand, and gravel is 24 %, 21 %, 44 % and 10 % respectively. We note that the four sM samples associated with the red class have median grain size values in the range of 6.6–9.3ϕ, and the sand content is 18–36 % (less than 0.5 % gravel). The fact that these four samples correspond with the red class suggests that presence of sand has a larger influence on the QTC results than the median grain size indicates. The cores associated with the red class typically have intermediate layers of coarser sediments (Fig. 5).

The yellow class is present both in the west and in the east. In the west the class extends from the slope into the deep channel, and the sediment samples associated with this class are classified as sandy silt (sZ) and perhaps silty sand (zS). In the east the yellow class occurs in an area where the nature of the sediment samples is variable. The few samples from the homogeneous parts of the yellow class make the analysis of sediment content uncertain. The most dominant tendency is only 0–4 % gravel. The percentage of clay, silt, and sand varies a lot. There is also a large variability in depth, sometimes no layering and sometimes several layers in the upper 75 cm. The sandy surface layer of the sZ core sample from the north-western channel (#40, see Fig. 1b) is about 50 cm thick (Fig. 5b).

The brown class dominates the North Sea Plateau, and based on the classification of the three sediment samples associated with this class (#42, #44 and #45, see Fig. 1b), the seabed consists of sand and coarser sediments (including shells and shell fragments). The average median grain size is 1.9ϕ, and the average contents of sand and gravel are 86 and 14 %, respectively. The mud content is less than 1 %.

Both pulse lengths were used when surveying two of the fiords, and the purple class consistently dominates these. A small area in the channel at about 60.15°N, 4.5°E and a part of the glacially eroded zone in the NE were also assigned to the purple class. The four sediment samples from the area of purple shading in Fig. 6 contains fine sediments and were classified as clay (C), sandy clay (sC), and mud (M). The average content of clay, silt, sand, and gravel is 63, 29, 8 and 0 %, respectively. The average median grain size is 10.2ϕ. Even though the clay (C) core sample from the channel (#33, see Fig. 1b) also matches the purple class, the surficial clay layer is thin. The clay content of the second to top subsample (5 cm

(a) Silhouette index

(b) Davies-Bouldin index
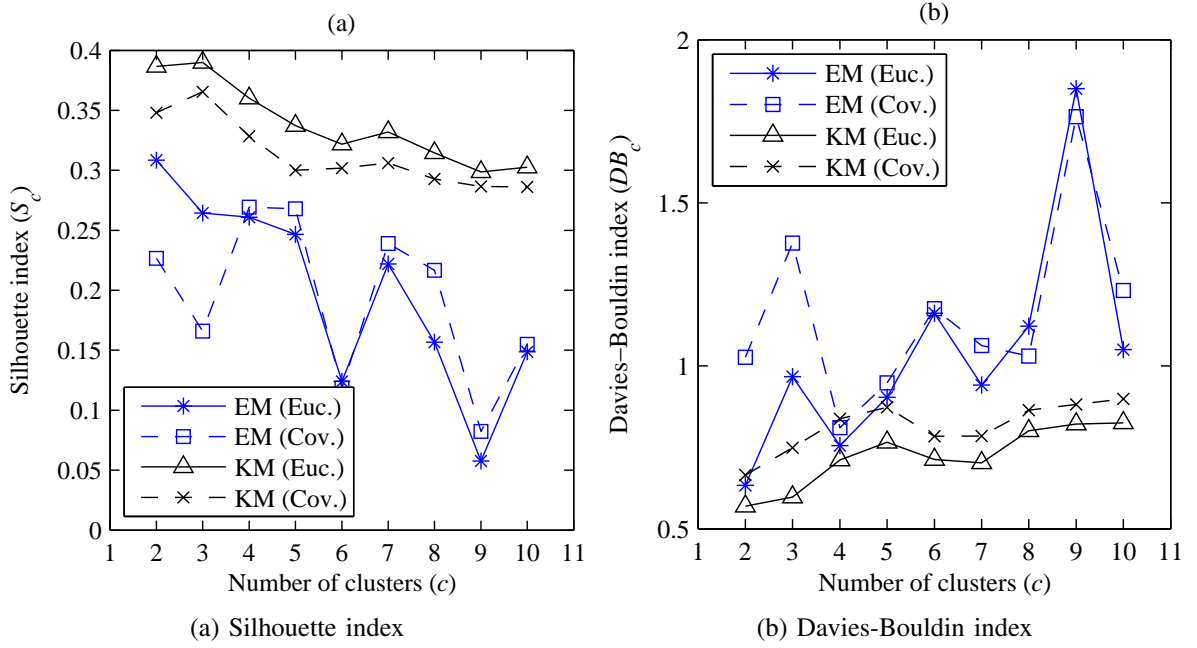
Fig. 7. Silhouette index (a) and Davies-Bouldin index (b) as a function of the imposed number of clusters for two different cluster algorithms, expectation maximization (EM) and k-means (KM). The indices were computed using both the Euclidean (Euc.) and covariance-weighted (Cov.) distance measures. A higher silhouette index S_c and a lower Davies-Bouldin index DB_c signify better separated and/or more compact clusters.

deeper) is lower (55 % compared to 64 %); this subsample is classified as mud (M). The sandy surficial layer is also thin in the sC core sample from the fiord (#30, see Fig. 1b); the second to top subsample (7.5 cm deeper) contains only 3.5 % sand (compared to 25.7 % sand in the top subsample).

C. Cluster validity

To compute the cluster validity indices (Fig. 7), the EM algorithm (with normal components), k-means, and the SOM method were applied to a subset of 10 000 randomly selected observations. Here we summarize the results for the $\tau = 0.512$ ms dataset. The validity indices for the $\tau = 1.024$ ms dataset lead to similar conclusions. The sediment class map obtained with the SOM algorithm did not show good consistency with supporting data compared to the QTC, EM, and k-means algorithms. There are several parameters in the SOM algorithm that can be tuned to improve the results, but we have instead omitted the validity indices for this algorithm (although they also point to the same conclusions).

The Dunn index was close to zero in all the experiments, which indicates that at least one pair of clusters are adjacent and poorly separated, although is sensitive to outliers. The lack of a knee or clearly optimum c-value in the graphs of Fig. 7, and the relatively good score for $c = 2$, also suggest that there is no clear cluster structure in the data. If we exclude $c = 2$ and $c = 3$ on physical grounds, as these values cannot account for the variation in the sample dataset, then four or five classes appear to be the overall best choices. The validity indices of the final QTC clustering with five classes are close to those of the EM method: $S_5 = 0.25$, $DB_5 = 0.88$ with Euclidean distances, and $S_5 = 0.25$, $DB_5 = 0.86$ with covariance-weighted distances.

Scatter plots of the two principal components for three different clusterings also suggest a lack of a clear cluster structure (Fig. 8). The overlap between the classes (particularly for QTC and EM) shows the importance of the third principal component. There are marked differences between the three clusterings (the SOM result was even more different). We have used the same colors in all three plots to ease comparison, but the physical interpretation of each color (class) will of course not be the same. This is evident in the class maps (Fig. 9), which have been processed in the same way as above using a block mode filter. The lack of cluster structure partly explains why the result is sensitive to the choice of algorithm. However, there are some features that are preserved, particularly the western boundary between the brown and yellow classes. The EM algorithm reproduces many of the features of the QTC result. For the $\tau = 1.024$ ms dataset, the QTC and EM class maps are almost indistinguishable.

For the final QTC result with five classes, the observed Mahalanobis distances (r^2) were sorted in ascending order and plotted against $F^{-1}(p_i, K)$, where $p_i = (i-0.5)/n$, $i = 1, \ldots, n$ are n equally spaced quantiles, and F^{-1} is the inverse cumulative distribution function for $\chi^2(K)$ (K is the number of observations). The quantile-quantile plots (Fig. 10) show that observations from the brown and yellow classes are not normally distributed, whereas the normal assumption is tenable for the turquoise class. The observations from the red and purple classes also deviate from the unit slope line, but less than for the brown and yellow classes. Note that removing suspected outliers did not improve the fit to the unit slope line.

D. Comparison with parametric sonar data and geomorphology

The ASC results were compared with parametric sonar profiles along several transects. There are clear changes in acoustic signature as the terrain shifts from horizontally bedded marine or glaciomarine sediments to morainic ridges. Comparing Figs. 3, 6 and 11, it is clear that the turquoise class in the central and eastern areas corresponds well with top unit A. The purple class matches a thinner layer (about 2–8 m thick) overlying the Weichselian moraine (Fig. 11, profile labelled G in Fig. 6). The top sediment sample from this layer was classified as clay.

The moraines on the eastern side of the channel have visible iceberg plough marks, as well as north-south-oriented parallel lineaments probably caused by a moving glacier (Fig. 12). The preservation, up through the Holocene, of traces of glacial erosion suggests two things: the marine sedimentation rate is low, and the seabed sediments have high stiffness. This is consistent with the interpretation of the red class (Fig. 6), and the boundary of the red class follows the outline of the moraines closely in the south-east (Fig. 11). Further north both the red and yellow classes occur and both correspond with the Weichselian till (unit B1). It is difficult to distinguish between these two classes based on parametric sonar data alone.

The westernmost part of unit A corresponds with the yellow class. A comparison of the ASC results with a parametric sonar profile (Fig. 13, profile labelled J in Fig. 6) shows that the transition from dominantly muddy to dominantly sandy surficial bottom sediments occurs well into the unit A area. The top subsample of the nearby gravity core from 60.71°N, 3.39°E was classified as sandy silt (sZ) with a substantial sand fraction (13 %), which confirms the presence of sand in the western NC (Fig. 6)). The morphology of the western slope of the channel is a fluted pattern of elongated depressions (Fig. 14).
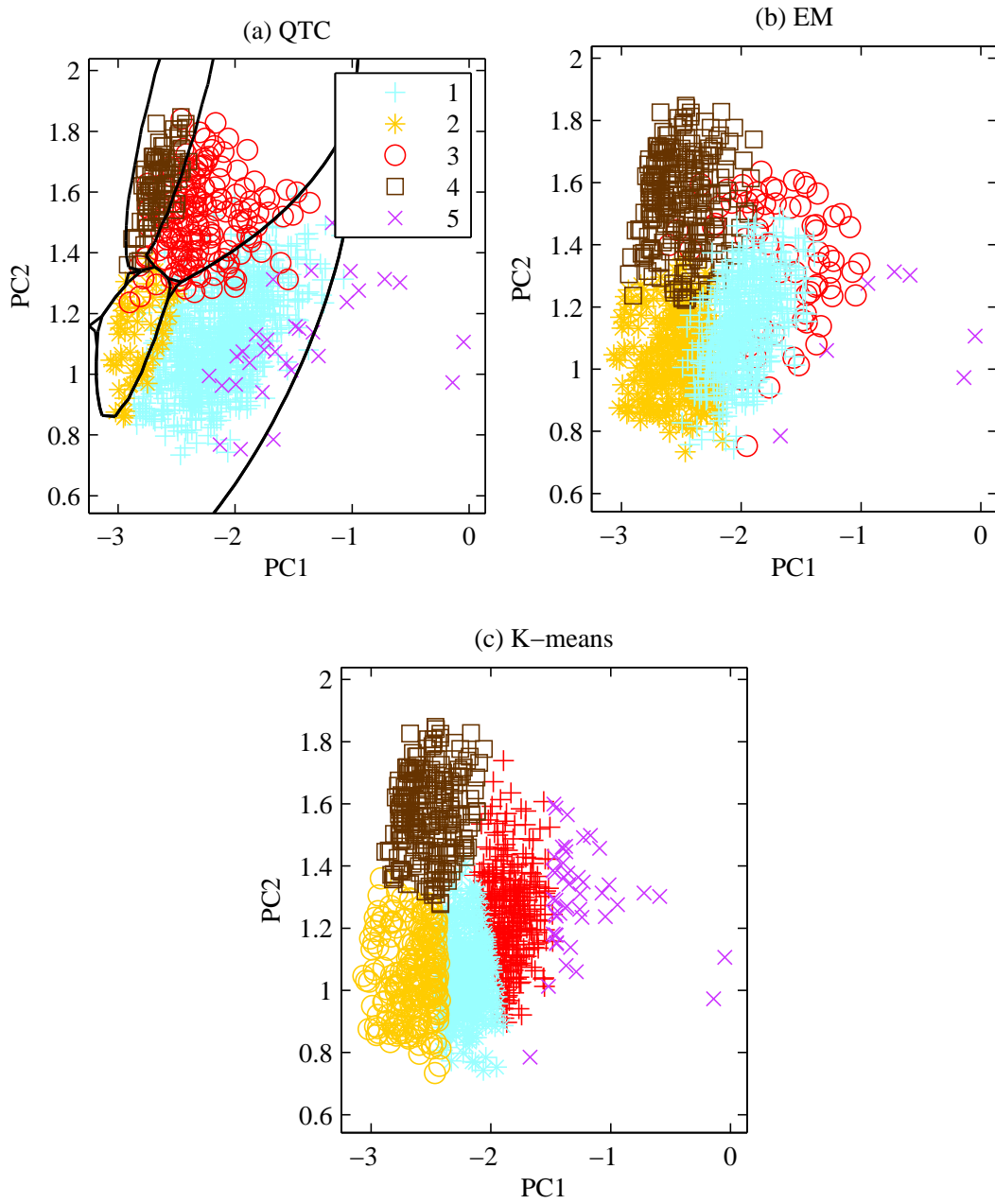
Fig. 8. Scatter plots of the two principal components (PC1 and PC2) for 1200 randomly selected observations after partition into five classes with QTC clustering (a), EM clustering (b), and k-means (c). The plot of the QTC clustering also shows the decision boundaries (points of equal posterior probability) obtained by considering only the two principal components.

V. DISCUSSION

Overall, the merging of the ASC results from the two data sets ($\tau = 0.512$ ms and $\tau = 1.024$ ms) shows a good match between the two. However, south and north on the western slope the correspondence between the red classes from the two datasets is reduced. There is also some disagreement between the red and brown classes in the north-western portion of the study area, on the North Sea Plateau. In order to improve the matching, a manual clustering may be beneficial; this has not been investigated in the present study. The
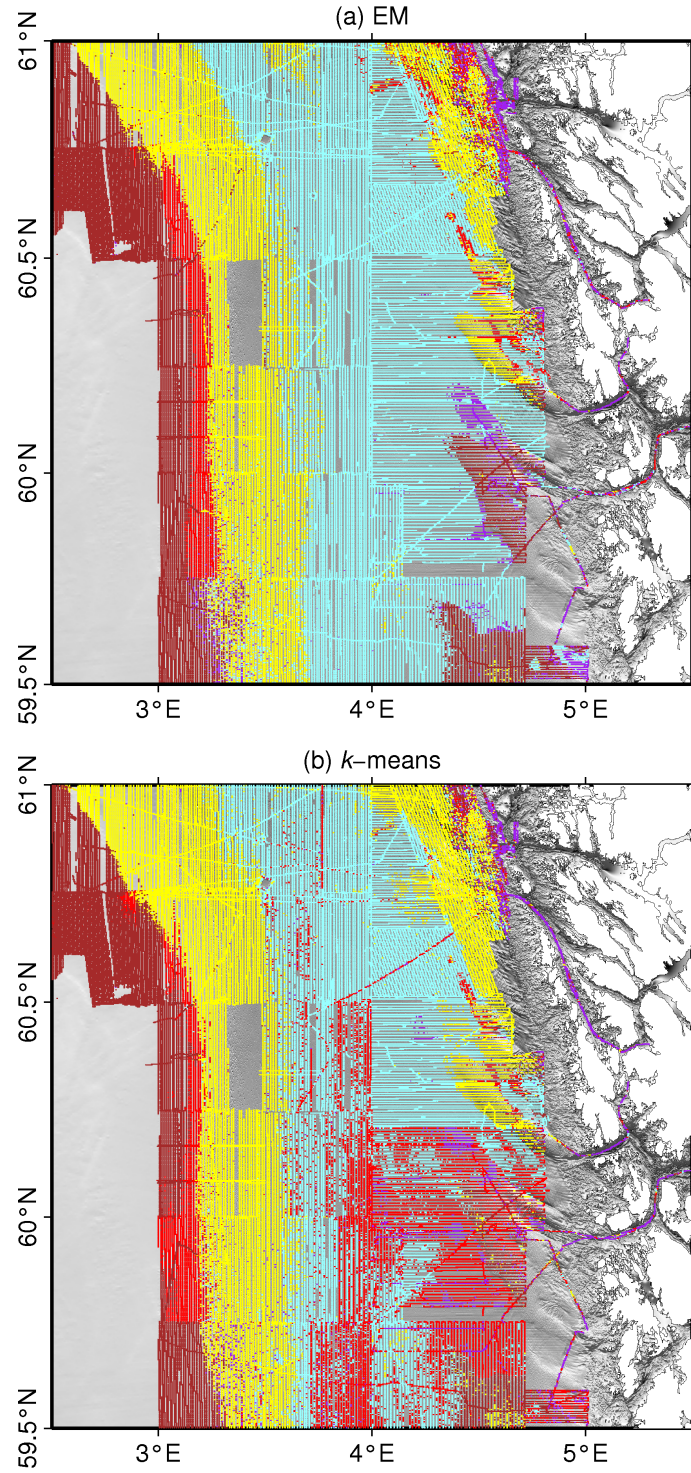
Fig. 9. Classification result (both pulse lengths combined) using EM clustering (a) and k-means clustering (b). These maps should be compared with the QTC results in Fig. 6.

final ASC map based on SBES data and interpretation of supporting data (Fig. 6c) gives a unifying and generally coherent picture of our observations, consistent with what we know about depositional and erosional processes in the region. Geographically and lithologically, the brown class corresponds well with class B of Rise and Rokoengen [30], the yellow class corresponds with their class C, and the turquoise class corresponds with their class D (Section II and Fig. 1b). In the north-east, the signs of massive glacial erosion, and probable

108
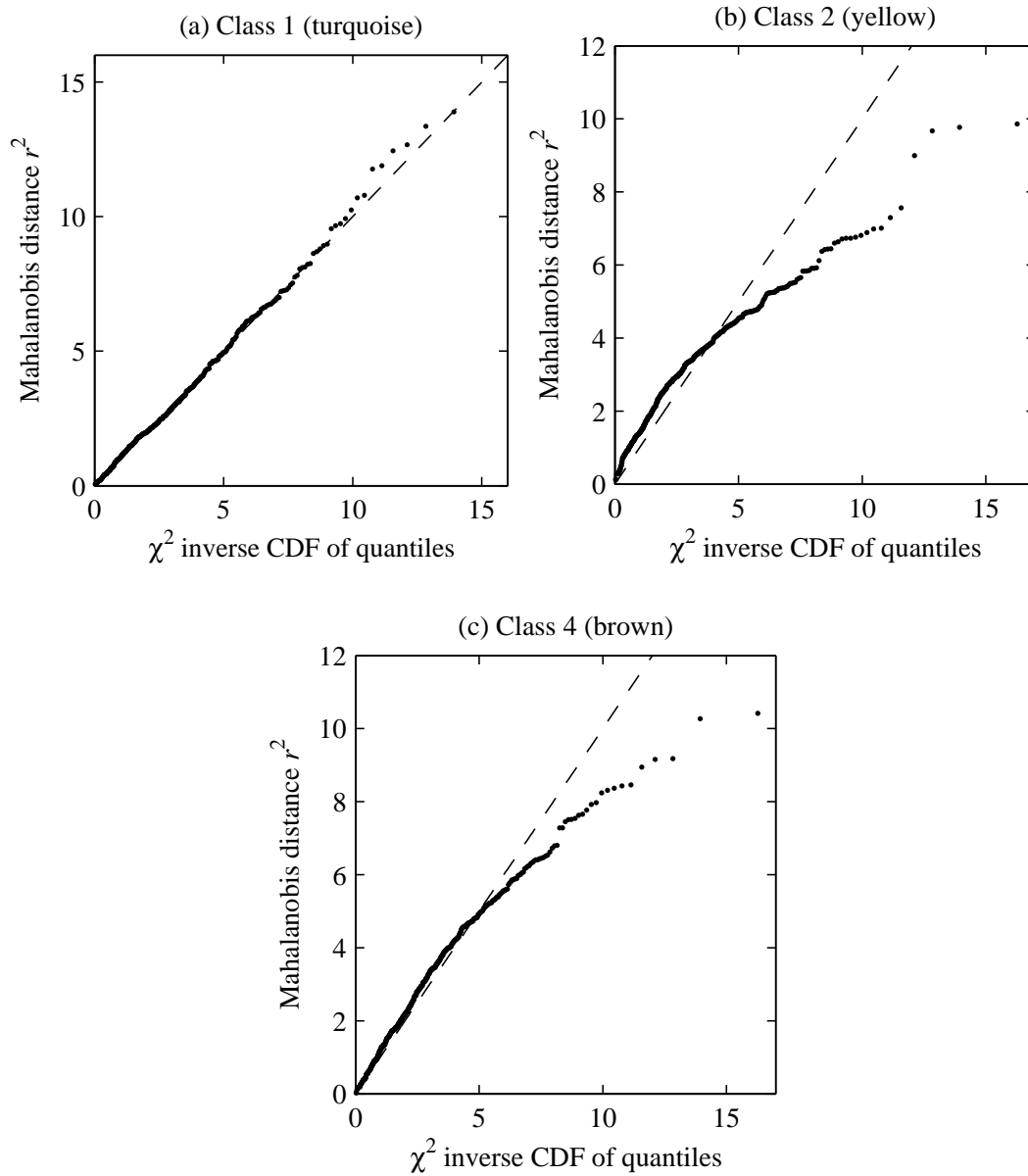
Fig. 10. Quantile-quantile plots of the observed Mahalanobis distance r^2 versus the χ^2 distribution with three degrees of freedom for 500 randomly selected observations per class. (a) turquoise class, (b) yellow class, and (c) brown class. For a multinormal distribution, the observations are expected to lie on the line with unit slope (dashed line).

presence of glacigenic sediments, explain why the ASC results are variable, with the red and yellow classes as predominant.

The ASC results were decimated with a 200×200 m block mode filter which removed fine-scale variations. The lack of consistency between the grain size analysed sediment samples associated with the yellow class may be explained by local variations in the seabed types in the north-east. The variable geomorphology also suggests that analyses of smaller regions are needed to understand the correspondence between the classification results and point sampling of the seabed. On the other hand, the QTC results may reflect more than simply the distribution patterns in sediment grain size, as stated by Wienberg and Bartholomä [7].
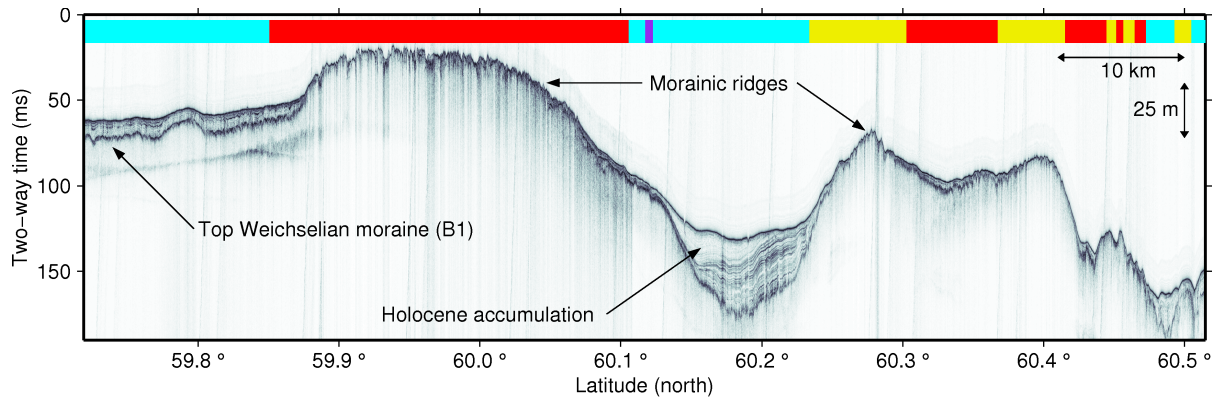
Fig. 11. Parametric sonar profile labelled G in Fig. 6 (instantaneous amplitude after deconvolution). The QTC classes are indicated with color bars. The vertical length scale is based on a sound speed of 1550 m/s, hence most accurate for unit A.
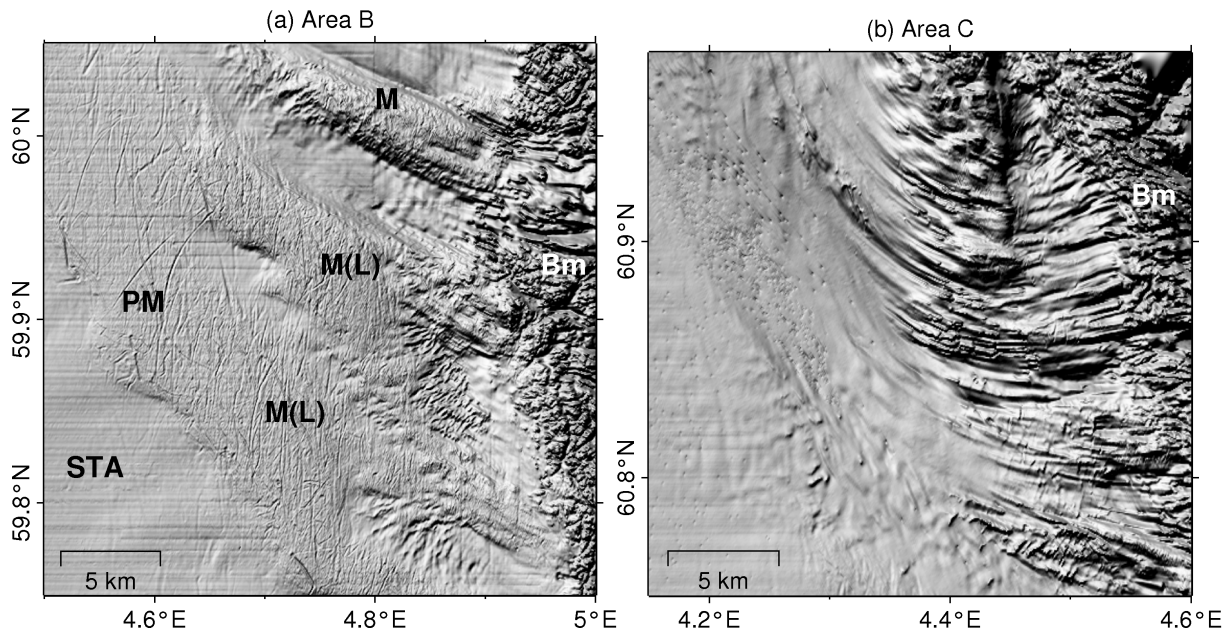
Fig. 12. Traces of glacial erosion on the eastern side of the study area (bathymetric relief based on MBES data). (a) Morainic ridges with glacial lineaments and plough marks extend north-west from the coastal zone, where basement is exposed at the seafloor. (b) The concentric patterns outside the Sognefjorden and Fensfjorden inlets suggest a movement of ice from the mainland and into the northbound Norwegian Channel Ice Stream. The figures are enlarged versions of area B and area C shown in Fig. 1b. Annotation: M: moraines; PM: plough marks; L: lineaments; Bm: basement; STA: ship track artifacts (parallel east-west-oriented lines)

The flow-like pattern seen outside the Sognefjorden and Fensfjorden inlets (Fig. 12b) suggests that this was a drainage area for the mainland ice sheet, where ice moved into the channel and was deflected northwards by the NC ice stream. Rise et al. [46] observed a pattern of semi-parallel lineations on a deeper Quaternary horizon, immediately to the north of 61°N, which they believed to be caused by a mid-Pleistocene glaciation. The observations of Rise et al. too suggest a movement of ice from the mainland north-eastwards into the channel. On the western slope, the orientation of the elogated depressions (predominantly north-south) matches the direction of inflowing Atlantic water (Fig. 14). Bøe et al. [47] interpreted similar
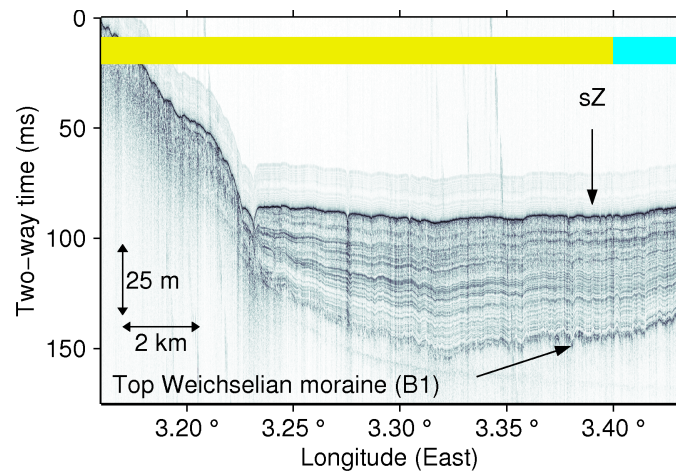
Fig. 13. Parametric sonar profile from the north-west side of the study area, labelled J in Fig. 6 (instantaneous amplitude after deconvolution). The QTC classes are indicated with color bars. The topmost layer of unit A (Holocene) appears homogeneous with little internal diffraction. A nearby gravity core shows a sandy surficial layer with the top-subsample classified as sandy silt (sZ). The vertical length scale is based on a sound speed of 1550 m/s, hence most accurate for unit A.
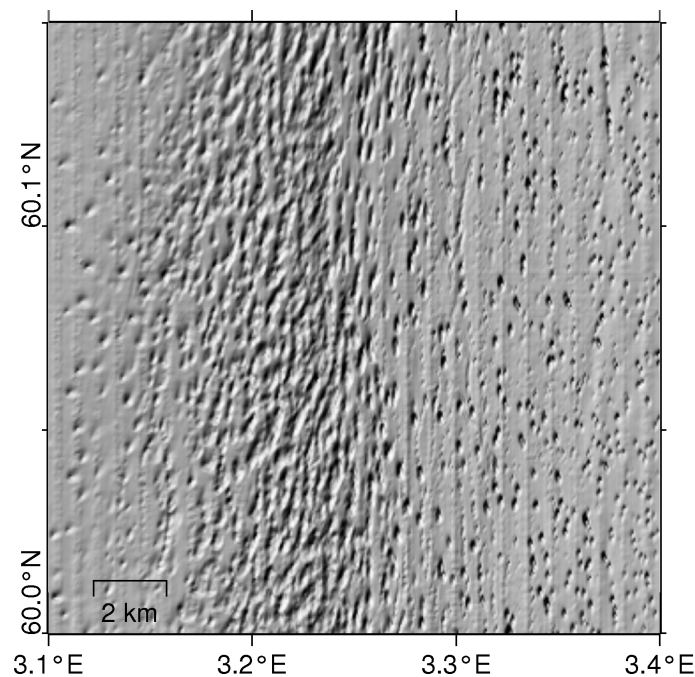
Fig. 14. Seabed morphology of area A (see Fig. 1b). The western slope of the NC has an irregular, fluted pattern of elongated depressions, a possible sign of erosion by water currents and fluid seepage. The seabed of the deep part of the NC is flat but pockmarked. Parallel north-south-oriented ship track artefacts are also visible.

patterns on the southern slope of the Skagerrak as current-modified pockmarks. Hovland [48] observed very shallow gas blankets in boomer seismic data from a part of the western slope. There were apparent vertical escape routes matching the occurrence of elongated depressions in the seabed. He concluded that vertically migrating gas is lifting sediments into suspension above the escape routes; the suspended sediments are subsequently transported away by water currents. This mechanism may explain why the yellow class extends into the deep channel

111

where we expected the turquoise class to dominate. Sand particles may accumulate on the seabed at the bottom of slope, and in decreasing amounts towards the center of the channel. The pre-Holocene part of unit A would be unaffected by this process, as it is a marine transport phenomenon. As observed in Sections IV-A and IV-B, there is a sandy surface layer (about 50 cm thick in core #38), overlying finer sediments, in cores from this area. We therefore think that ASC gives the correct indication, even though this is not evident from the parametric sonar (or seismic) data. The sharp class boundary of the ASC map is however misleading if the transition is continuous. This demonstrates a basic limitation of the method. In general there may be observations that have low likelihood with respect to all classes.

The purple class present to the north of the moraines at 60.15°N, 4.5°E stands out (Fig. 6). There are no particular bathymetric features that might explain why this area is assigned to another class than the surroundings. The one sediment sample deliberately collected here, based on initial ASC results, contains finer surficial sediments than neighbouring sediment samples. In a study by Van Vossen et al. [49] this area also stood out from its surroundings. They applied a physics-based method to low-frequency sonar data acquired at 1–2 kHz. It was found that the area generally had lower reflection coefficient (at 7°grazing angle), as expected, but higher backscattering strength than the adjacent area of turquoise shading (Fig. 6).

The lack of a clear cluster structure in the echo shape data may partly be explained by a continuous change in seabed composition along the transverse axis of the channel. However, we have found no significant differences in the class distributions when comparing data from the eastern and western sides of the study area. The question is if there are other echo features that might give better class discrimination. Several aspects of the QTC method appear to be motivated by a need to make the software robust and straightforward to use. This makes sense given the amount of tuning and level of expertise often required to successfully run scientific codes, but it also introduces some limitations. For example, normalizing echo amplitudes eliminates the need to compensate for sound attenuation in the water column and to calibrate the acoustic signals. However, the total backscattered energy is then lost, a potentially useful feature [14], [17]. Likewise, restricting the analysis to three principal components avoids all the problems of classification in high dimensions [15], and limits the amount of catalogue data needed to make robust PDF estimates, but this restriction too sheds potentially useful information. An alternative is to use aligned, digitally sampled echoes as feature vectors directly [50], [51] (each sample a feature). For such an approach we propose to use a piecewise constant function approximation [52] to reduce the number of features before running the classification algorithm.

Even if one uses all information inherent in a (stacked) echo, there are other potentially useful features. In their evaluation of QTC, Hamilton et al. [22] suggested that some measure of variability within a set of pings might be useful for identifying rock or coral bottoms, where ping-to-ping fluctuations are large. More generally, our understanding is that QTC (the single-beam variant) does not take into account the statistics of a set of contiguous observations, and the classification process proceeds without information about spatial relationships, i.e. spatial location and neighboring observations do not affect the prior probabilities. We have not explored these possibilities in this work, but think that better classification accuracy may be achieved by using spatial information.

A fundamental limitation concerns the mapping of sloping areas. The scattering strength changes significantly with incidence angle in the near-nadir regime (except for very rough surfaces). If the seabed is sloping, the symmetry of vertical ensonification is destroyed, and

this affects the echo shape. Using QTC, Von Szalay and McConnaughey [53] found that slopes exceeding 5–8°caused the classification accuracy to break down. In the study area, slopes of such magnitude only occur in the narrow basement zone and along the sides of the fiords, at least on a spatial scale of 50 m or more (the resolution of the terrain model that we have used). Except for a small area to the north-east, we do not have ASC results from the basement zone. The only class that could potentially be affected by slope errors is the purple class. Pockmarks were not removed in the pre-processing step, and it should not be necessary. The use of a block-mode filter, which picks out the most frequently occurring class within a rectangular window (block), also helps to reduce such errors. The best solution we can think of to overcome the problem of steep slopes would be to transfer the echo shape methodology to an MBES system, and at any time apply it to the beam for which the true incidence angle is closest to zero.

If there is no obvious cluster structure, as in the present data, practical considerations may limit the number of classes. When the data distribution is not unimodal, the approximation $\text{BIC} \approx -2\log p(\boldsymbol{D} \mid c)$ may be a poor one. We do not know if condition 2 [following Eq. 4] holds. Moreover, the fact that classes are not well separated induces errors in the ML estimates during the cluster process. If two seabed types have overlapping PDFs, and the two types are unequally represented in the data set, the larger class may steal members from the smaller because of unequal priors. This is another argument for including spatial information in the classification process to adjust the priors depending on location. The non-normal distribution of some classes favours increasing c to fit the data. This may explain why the BIC decreased when c was increased beyond that which supporting data indicated. Nevertheless, the parsimony term $M \log n$ counters the tendency to overfit the data in any circumstances, and there are empirical reasons to use the BIC for normal mixtures [54].

While QTC results were very similar to those of the simpler iterative EM algorithm, a qualitative comparison with supporting data has shown that QTC performed better than the k-means and SOM algorithms. Despite the above reservations, this provides support for the QTC algorithm as applied to a difficult data set.

VI. Conclusions

We have used a statistical classification method (QTC IMPACT) for single-beam echo sounders to map the surficial sediments in the Norwegian Channel. The unsupervised classification results were interpreted in view of 40 gravity cores and 37 grab samples distributed across the whole survey area, as well as MBES bathymetry and parametric sonar data. The ASC analysis has resulted in a high-resolution map where the seabed is divided into five classes: (1) silt and mud (deep, central channel), (2) a variety of clay, silt and sand (mainly north-east and lower western slope), (3) sandy mud with gravel (mainly moraines and upper western slope), (4) sand with varying amounts of gravel (predominant on North Sea Plateau), and (5) clay and sandy clay (inside fiords and one small area in the channel). The acoustic classification is highly sensitive to the sand and gravel content of the top layer. The five classes reflect the diverse erosional and depositional processes of the area, past and present: Glacial erosion on the eastern side, fluid seepage and water current erosion on the western side, and deposition of fine-grained suspended sediments in the central part.

The consistency of the results, across a large and diverse area, with respect to supporting data, different pulse lengths, and intersecting ship tracks, shows that QTC IMPACT is a reliable tool for sediment mapping. However, the lack of cluster structure in the set of

observation vectors implies a degree of arbitrariness in the final classification result. This is partly due to the nature of sediment deposition in the study area, but also suggests that additional features should be introduced for better discrimination between Folk type classes. The study also highlights the need for careful verification and cluster validity analysis of results obtained by statistical ASC methods in general.

REFERENCES

[1] J. T. Anderson, D. Van Holliday, R. Kloser, D. G. Reid, and Y. Simard, "Acoustic seabed classification: current practise and future directions," *ICES Journal of Marine Science*, vol. 65, no. 6, pp. 1004–1011, 2008.

[2] C. Brown and P. Blondel, "Developments in the application of multibeam sonar backscatter for seafloor habitat mapping," *Applied Acoustics*, vol. 70, no. 10, pp. 1242–1247, 2009.

[3] K. L. Ellingsen, J. Gray, and E. Bjørnholm, "Acoustic classification of seabed habitats using the QTC VIEW system," *ICES Journal of Marine Science*, vol. 59, pp. 825–835, 2002.

[4] R. Freitas, S. Silva, V. Quintino, A. Rodrigues, K. Rhynas, and W. Collins, "Acoustic seabed classification of marine habitats: studies in the western coastal-shelf area of Portugal," *ICES Journal of Marine Science*, vol. 60, pp. 599–608, 2003.

[5] K. Haris, B. Chakraborty, B. Ingole, A. Menezes, and R. Srivastava, "Seabed habitat mapping employing single and multi-beam backscatter data: A case study from the western continental shelf of India," *Continental Shelf Research*, vol. 48, pp. 40–49, 2012.

[6] T. Medialdea, L. Somoza, R. León, M. Farrán, G. Ercilla, A. Maestro, D. Casas, E. Llave, F. Hernández-Molina, M. Fernández-Puga, and B. Alonso, "Multibeam backscatter as a tool for sea-floor characterization and identification of oil spills in the Galicia Bank," *Marine Geology*, vol. 249, pp. 93–107, 2008.

[7] C. Wienberg and A. Bartholomä, "Acoustic seabed classification in a coastal environment (outer Weser Estuary, German Bight)-a new approach to monitor dredging and dredge spoil disposal," *Continental Shelf Research*, vol. 25, no. 9, pp. 1143–1156, 2005.

[8] A. Bartholomä, "Acoustic bottom detection and seabed classification in the German Bight, southern North Sea," *Geo-Marine Letters*, vol. 26, pp. 177–184, 2006.

[9] C. De and B. Chakraborty, "Model-based acoustic remote sensing of seafloor characteristics," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3868–3877, 2011.

[10] M. Snellen, K. Siemens, and D. Simons, "Model-based sediment classification using single-beam echosounder signals," *Journal of the Acoustical Society of America*, vol. 129, no. 5, pp. 2878–2888, 2011.

[11] D. Sternlicht and C. De Moustier, "Remote sensing of sediment characteristics by optimized echo-envelope matching," *Journal of the Acoustical Society of America*, vol. 114, no. 5, pp. 2727–2743, 2003.

[12] A. Amiri-Simkooei, M. Snellen, and D. Simons, "Principal component analysis of single-beam echosounder signal features for seafloor classification," *IEEE Journal of Oceanic Engineering*, vol. 36, no. 2, pp. 259–272, 2011.

[13] F. Madricardo, J. Tegowski, and S. Donnici, "Automated detection of sedimentary features using wavelet analysis and neural networks on single beam echosounder data: A case study from the Venice Lagoon, Italy," *Continental Shelf Research*, vol. 43, pp. 43–54, 2012.

[14] P. A. Van Walree, J. Tegowski, C. Laban, and D. G. Simons, "Acoustic seafloor discrimination with echo shape parameters: A comparison with ground truth," *Continental Shelf Research*, vol. 25, pp. 2273–2293, 2005.

[15] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*, 2nd ed. Springer, 2009.

[16] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th ed. Academic Press, 2009.

[17] P. A. Van Walree, M. Ainslie, and D. Simons, "Mean grain size mapping with single-beam echo sounders," *Journal of the Acoustical Society of America*, vol. 120, no. 5, pp. 2555–2566, 2006.

[18] D. Jackson and M. Richardson, *High-Frequency Seafloor Acoustics*. New York: Springer, 2007.

[19] J. Preston, A. Christney, L. Beran, and W. Collins, "Statistical seabed segmentation - from images and echoes to objective clustering," *Proceedings of the Seventh European Conference on Underwater Acoustics (ECUA 2004)*, pp. 813–818, 2004, delft, The Netherlands.

[20] C. Brown, B. Todd, V. Kostylev, and R. Pickrill, "Image-based classification of multibeam sonar backscatter data for objective surficial sediment mapping of Georges Bank, Canada," *Continental Shelf Research*, vol. 31, pp. 110–119, 2011.

[21] R. Freitas, F. Ricardo, F. Pereira, L. Sampaio, S. Carvalho, M. Gaspar, V. Quintino, and A. M. Rodrigues, "Benthic habitat mapping: Concerns using a combined approach (acoustic, sediment and biological data)," *Estuarine, Coastal and Shelf Science*, vol. 92, pp. 598–606, 2011.

[22] L. Hamilton, P. Mulhearn, and R. Poeckert, "Comparison of RoxAnn and QTC-View acoustic bottom classification system performance for the Cairns area, Great Barrier Reef, Australia," *Continental Shelf Research*, vol. 19, no. 12, pp. 1577–1597, 1999.

[23] L. Otto, J. Zimmerman, G. Furnes, M. Mork, R. Sætre, and G. Becker, "Review of the physical oceanography of the North Sea," *Netherlands Journal of Sea Research*, vol. 26, no. 2-4, pp. 161–238, 1990.

[24] N. Winther and J. Johannessen, "North Sea circulation: Atlantic inflow and its destination," *Journal of Geophysical Research—Oceans*, vol. 111, C12018, doi:10.1029/2005JC003310, 2006.

[25] H. De Haas, E. Okkels, and T. Van Weering, "Recent sediment accumulation in the Norwegian Channel, North Sea," *Norges Geologiske Undersøkelse Bull.*, vol. 430, pp. 57–65, 1996.

[26] L. Rise, K. Rokoengen, A. Skinner, and D. Long, "Nordlige Nordsjø. Kvartærgeologisk kart mellom 60°30' og 62°N, og øst for 1°Ø. (Northern North Sea. Quaternary geology map between 60°30' and 62°N, and east of 1°E) M 1:500.000," 1984.

[27] EMODnet, "European Marine Observation and Data Network,The European Commision," Dec. 2012.

[28] H. P. Sejrup, I. Aarseth, H. Haflidason, R. Løvlie, Å. Bratten, G. Tjøstheim, C. F. Forsberg, and K. L. Ellingsen, "Quaternary of the Norwegian Channel: glaciation history and paleooceanography," *Norsk geologisk tidsskrift*, vol. 75, no. 2-3, pp. 65–87, 1995.

[29] E. S. Andersen, S. R. Østmo, C. F. Forsberg, and S. J. Lehman, "Late- and post-glacial depositional environment in the Norwegian Trench, northern North Sea," *Boreas*, vol. 24, pp. 47–64, 1995.

[30] L. Rise and K. Rokoengen, "Surficial sediments in the Norwegian sector of the North Sea between 60°30' and 62°N," *Marine Geology*, vol. 58, pp. 287–317, 1984.

[31] E. M. Sigmond, "Geologisk kart over land-og havområder i Nord-Europa, målestokk 1:4.000.000," *Norges geologiske undersøkelse*, 2002.

[32] R. Francois and G. Garrison, "Sound absorbtion based on ocean measurements: Part II: Boric acid contributions and equation for total absorbtion," *Journal of the Acoustical Society of America*, vol. 72, no. 6, pp. 1879–1890, 1982.

[33] R. Folk, "The distinction between grain size and mineral composition in sedimentary-rock nomenclature," *Journal of Geology*, vol. 62, no. 4, pp. 344–359, 1954.

[34] P. Wessel and W. Smith, "Free software helps map and display data," *Eos Transactions American Geophysical Union*, vol. 72, no. 441, 1991.

[35] J. Preston, "Resampling sonar echo time series primarily for sediment classification, US Patent 6801474," 2005.

[36] J. Preston and R. Kirlin, "Comment on 'Acoustic seabed classification: improved statistical method'," *Can. J. Fish. Aquat. Sci.*, vol. 60, pp. 1299–1300, 2003.

[37] J. M. Preston, "Automated acoustic seabed classification of multibeam images of Stanton Banks," *Applied Acoustics*, vol. 70, pp. 1277–1287, 2009.

[38] R. Duda, P. Hart, and D. Stork, *Pattern classification*, 2nd ed. Wiley-Interscience, 2001.

[39] H. Bhat and N. Kumar, "On the derivation of the Bayesian Information Criterion," 2010. [Online]. Available: http://nscs00.ucmerced.edu/~nkumar4/BhatKumarBIC.pdf

[40] G. Schwarz, "Estimating the dimension of a model," *Annals of Statistics*, vol. 6, no. 2, pp. 461–464, 1978.

[41] T. Kohonen, "The self-organizing map," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464–1480, 1990.

[42] P. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53–65, 1987.

[43] D. Davies and D. Bouldin, "A cluster separation measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 1, no. 2, pp. 224–227, 1979.

[44] J. Dunn, "Well separated clusters and optimal fuzzy partitions," *Journal of Cybernetics,*, vol. 4, pp. 95–104, 1974.

[45] R. Johnson and D. Wichern, *Applied Multivariate Statistical Analysis*, 2nd ed.   Englewood Cliffs, NJ: Prentice Hall, 1988.

[46] L. Rise, O. Olesen, K. Rokoengen, D. Ottesen, and F. Riis, "Mid-Pleistocene ice drainage pattern in the Norwegian Channel imaged by 3D seismic," *Quaternary Science Reviews*, vol. 23, pp. 2323–2335, 2004.

[47] R. Bø, L. Rise, and D. Ottesen, "Elongate depressions on the southern slope of the Norwegian Trench (Skagerrak): morphology and evolution," *Marine Geology*, vol. 146, pp. 191–203, 1998.

[48] M. Hovland, "Elongated depressions associated with pockmarks in the western slope of the Norwegian Trench," *Marine Geology*, vol. 51, pp. 35–46, 1983.

[49] R. Van Vossen, E. J. Eidem, S. Ivansson, B. Chalindar, J. Dybedal, M. E. Colin, F. P. Benders, B. L. Andersson, B. Juhel, X. Cristol, G.-K. Olsen, J. N. Pihl, E. Tveit, S. Jespers, and M. A. Ainslie, "Improved active sonar tactical support by through-the-sensor estimation of acoustic seabed properties," *IEEE Journal of Oceanic Engineering*, vol. 39, no. 4, pp. 755–768, Oct. 2014.

[50] L. Hamilton, "Acoustic seabed segmentation for echosounders through direct statistical clustering of seabed echoes," *Continental Shelf Research*, vol. 31, no. 19, pp. 2000–2011, 2011.

[51] L. Hamilton and I. Parnum, "Seabed segmentation from unsupervised statistical clustering of entire sonar backscatter curves," *Continental Shelf Research*, vol. 31, no. 2, pp. 138–148, 2011.

[52] A. C. Jensen and A. Solberg, "Fast hyperspectral feature reduction using piecewise constant function approximations," *IEEE Geoscience and Remote Sensisng Letters*, vol. 4, no. 4, pp. 547–551, 2007.

[53] P. Von Szalay and R. McConnaughey, "The effect of slope and vessel speed on the performance of a single beam acoustic seabed classification system," *Fisheries Research*, vol. 54, no. 2, pp. 181–194, 2002.

[54] K. Roeder and L. Wasserman, "Practical Bayesian density estimation using mixtures of normals," *Journal of the American Statistical Association*, vol. 92, no. 439, pp. 894–902, 1997.

Paper II:

Bayesian seabed classification using angle-dependent backscatter data from multibeam echo sounders

Bayesian Seabed Classification Using Angle-Dependent Backscatter Data From Multibeam Echo Sounders

Knut Landmark, Anne H. Schistad Solberg, Andreas Austeng
and Roy Edgar Hansen

Abstract

Acoustical seabed classification is a technology for mapping seabed sediments. Processed multibeam sonar data yield the variation of the seabed scattering strength with incidence angle, and this paper examines the effect of this on classification. A simple Gaussian statistical model is developed for the observed scattering strength, whereby an observation is represented by a piecewise constant function of incidence angle. Provided some data for which the sediment types are known (training data), the statistics for each type can be robustly estimated. Subsequently, standard Bayesian theory is applied to classify new observations. The model was used to compute limits on classification accuracy in terms of the intrinsic scattering strength statistics of the seabed, and to predict whether a logarithmic or linear scale for the data is preferable. Systematic experiments on a North Sea data set with four sediment classes tested how the classification accuracy depends on the piecewise function approximation, incidence angle range, amount of training data, and spatial averaging (combining consecutive pings into one observation). The classifier based on Gaussian statistics performed at least as well as sophisticated algorithms with no assumptions about the data statistics. The best accuracy (95 %) was attained for logarithmic data. The amount of training data needed to achieve this was about 500 pings per class; spatial averaging could be limited to 10–20 pings. Comparable across-track spatial resolution was possible by dividing the full swath into separate independent sectors, but only at reduced accuracy (87 % or less). However, comparable accuracy may be possible by taking into account the spatial relationships of observations.

Index Terms

Bayesian methods, sonar, classification algorithms, seafloor, sediments, remote sensing.

I. INTRODUCTION

The purpose of acoustical seabed classification (ASC) and characterization is to map surficial seabed composition with sonars [1]. Multibeam echo sounders (MBES) are particularly useful, as they are now widely used for high-resolution bathymetry. MBES technology has had a significant impact on the marine sciences. Images of seabed morphology produced from multibeam bathymetry tell a detailed story of the physical processes that influence the seabed. ASC adds an extra dimension of information by analyzing and mapping out variations in the seabed acoustical response. This has proven useful, for example, in the field of marine ecology, where ASC is applied to the study of seabed (benthic) habitats [2]–[4]. Seabed ecosystems can be vulnerable to the effects of human activity, from disruptions caused by bottom trawling, to anthropogenic climate change. Between the conflicting interests of resource development and nature conservation, seabed mapping and monitoring will likely only increase in importance.

The solution of a generic classification problem has two main steps [5]: First, to preprocess the data and extract the signal characteristics that are the most effective for discriminating between different classes of signals. Second, to compute a decision rule which assigns a signal to one of several possible classes. The decision rule is often computed based on the statistical distribution of a set of signals for which the class is known, the training data (supervised learning). There are many powerful methods available for computing the decision rule [5], [6], but their relative performance is data dependent, and it is difficult to draw general conclusions about classifier performance based on just a few test data sets. Nevertheless, one should try and answer questions concerning how accurate is the decision rule, how well will the method work on new data sets (generalizability), and when might it not work.

The strategy adopted here is to develop a classification method and consider these questions experimentally as well as with the help of a simple statistical model for the observations. We focus on a property of the seabed acoustical response which can be fairly well modeled and measured. An MBES forms many narrow beams in a wide fan perpendicular to the ship's track, and therefore measures how the seabed scattering strength S varies with the incidence angle θ [7, Ch. 8]. The shape of $S(\theta)$ depends on the seabed reflection coefficient, the roughness spectrum, and the strength of sediment volume scattering [8, Ch. 13–14]. While the received signal statistics can be sensitive to changes in the ensonification pattern on the seabed [9]–[11] (an extraneous factor), properly measured scattering strength is not.

Although seabed characterization based on $S(\theta)$ can be ambiguous due to the complex interplay of scattering mechanisms, several promising results suggest that $S(\theta)$ is a useful feature for seabed classification and characterization. The angular range analysis of [12] is an inverse method for estimating acoustical parameters. It works by fitting simulated and observed features of $S(\theta)$ in several separate angular subsectors. The inversion algorithm is based on the fact that some features correlate with different scattering mechanisms; this fact may also alleviate the problem with ambiguities. Inversion results based on model-data comparison of the $S(\theta)$ function are also reported in [13], [14]. The parameter estimates of [13] are consistent with independent estimates made by inversion of a time-domain (echo shape) model applied to single beam echo sounder data in [15]. The papers [16]–[19] describe statistical or pattern recognition methods applied to the measured $S(\theta)$ function, including Bayesian classifiers, neural networks, and cluster algorithms. Simulation results in [17] showed best performance with a Bayesian classifier, albeit on data with known statistical distributions. A Markov random field based segmentation algorithm for sonar image mosaics, which also

takes into account the angular variation, was presented in [20]. Pattern recognition methods are less vulnerable to systematic errors in the data, but are best suited to situations where observations divide naturally into a set of discrete classes, which may not be fully known in advance. This problem is addressed in [21], which represents the backscatter histogram at a single angle with a normal mixture probability density function (PDF), where the number of components is determined from the acoustical data.

This paper describes a statistical supervised classification method, building on preliminary results in [22]. We consider several questions of practical importance concerning ASC applied to $S(\theta)$ measurements: accuracy (discriminating power), spatial resolution, training data requirements, statistics, and PDF estimation. In the statistical approach, an observation of $S(\theta)$ ($\theta \in [\theta_{\min}, \theta_{\max}]$) is identified not with physical seabed parameters, but by its posterior probability $P(\omega|S)$ and likelihood $p(S|\omega)$ with respect to certain classes ω. The classes are defined by ground truth data or other independent knowledge. In theory, the optimum (minimum error) classifier is the Bayes classifier [5], but this presupposes that the PDFs $p(S|\omega)$ are known. To get robust estimates of the PDFs from possibly scarce data, we propose to replace $S(\theta)$ by a vector $\boldsymbol{d} = \begin{bmatrix} d_1 & \dots & d_K \end{bmatrix}^T$, where d_k is the average scattering strength in a segment $\theta_{k-1} \leq \theta < \theta_k$. The number of components in $\boldsymbol{d}$ represents a trade-off between accurate approximation and robustness. If there are too few segments, the classification error rate increases because important information is lost by excessive smoothing. If there are too many segments, good estimation of the PDFs for $\boldsymbol{d}$ requires an unrealistic amount of training data. To achieve a good balance, the segments are determined such that the approximation error is minimized for the data set as a whole (for a fixed number of components K) [23], [24]. The observation vector $\boldsymbol{d}$ retains a simple physical interpretation, and is insensitive to all but very large errors in the estimated incidence angles. Section II-B and Section II-C explain the processing steps for obtaining the observation vectors.

The classification accuracy depends on the amount of averaging, the intrinsic scattering strength S, and the variance of S. Observation vectors result from averaging over time samples, angular segments, and consecutive pings. In Section II-D and appendix A, we compute approximate PDFs for the observation vectors, using a central limit theorem, for the simplest case of uncorrelated components. This is used to compare error bounds for classification with intensity and log-intensity data respectively. It may also be used to generate PDFs from modeled $S(\theta)$ functions.

The method was applied to a North Sea MBES data set. The survey area and supporting data are described in Section II-A. The experimental procedure is explained in Section III-A. The purpose of the experiments was to test how the classification accuracy depends on

- the dimension of the scattering strength vector $\boldsymbol{d}$;
- the amount of training data;
- the amount of spatial averaging (combining data from several pings);
- the extent of the angular sector.

Spatial averaging limits the along-track resolution of ASC mapping. Across-track resolution was also considered by dividing the full sonar swath into separate subsectors. We believe the data set is a useful testing ground for ASC methods. The survey area covers distinct sedimentary environments: a sandy plateau with low present day sedimentation; a deeper trench with thick soft marine depositions; as well as a transition zone between them. A set of sediment samples (gravity cores and grab samples) with laboratory analyses is available, and

121

TABLE I
DEFINITION OF SEDIMENT TYPES

Type	Grain size (mm)	ϕ [a]
Clay	< 0.002	> 9
Silt	0.002-0.063	4-9
Sand	0.063-2.0	(-1)-4
Gravel	> 2.0	< -1

[a] $\phi = -\log_2(D/1\,\mathrm{mm})$, where D is particle diameter.

a previous survey using an alternative ASC method and single beam echo sounder (SBES) provides a reference point [25].

Test results show that good classification accuracy and spatial resolution is possible even with limited training data. Section III-B presents the experimental results, which are discussed in Section IV.

II. METHODS AND DATA

A. Survey area

The survey area stretches from the eastern margin of the North Sea plateau and into the Norwegian Trench (Fig. 1). The North Sea is a shallow continental sea bounded by the British Isles, Norway, and the northern European continent. The Norwegian Trench (or Channel) is a depression in the continental shelf along the coast of southern Norway, with depths of about 300 m in the survey area; its recent geology is influenced by alternating glaciations and inundations during the Quaternary age [26]. The Norwegian Trench is an important trap for fine-grained water-borne sediments in the North Sea [27], and the top seabed layer is soft, silty and rich in organic material [28]. The eastern North Sea plateau is about 100 m deep, and the seabed is covered with sand and gravel [25], [29]. There is almost no present-day accumulation of water-borne, fine-grained sediments in this area. A gradual change in grain size distribution is expected down the western side of the trench. However, this slope is marked by flutes, ditches or elongated pockmarks, which suggests strong action by ocean currents or fluid seepage from below [30], [31].

The sediment sample data set (Fig. 1) confirms this picture. The mass fraction of sand in the samples increases westwards, up to more than 90 % on the plateau. The samples are labeled according to a modified version of Folk's classification system (TABLE I,II) [25], [32]. A previous study covering the area employed single beam echo sounder and the QTC Impact software [33] for seabed classification. Based on the ASC results and the same set of samples, it was concluded that four classes of seabed sediments could be discerned [25]:

1) silt and mud (deep central part of the trench);
2) sand with varying amounts of gravel (shallow plateau);
3) a variety of clay, silt, and sand (lower western slope of the trench);
4) sandy mud with gravel (upper western slope).

There are seven sediment samples inside the studied survey area, classified as either sand (S), sandy silt (sZ) or silt (Z). A fourth class designated U (unknown) was taken to represent a subarea devoid of samples. One of the sand samples (described as fine sand with shell
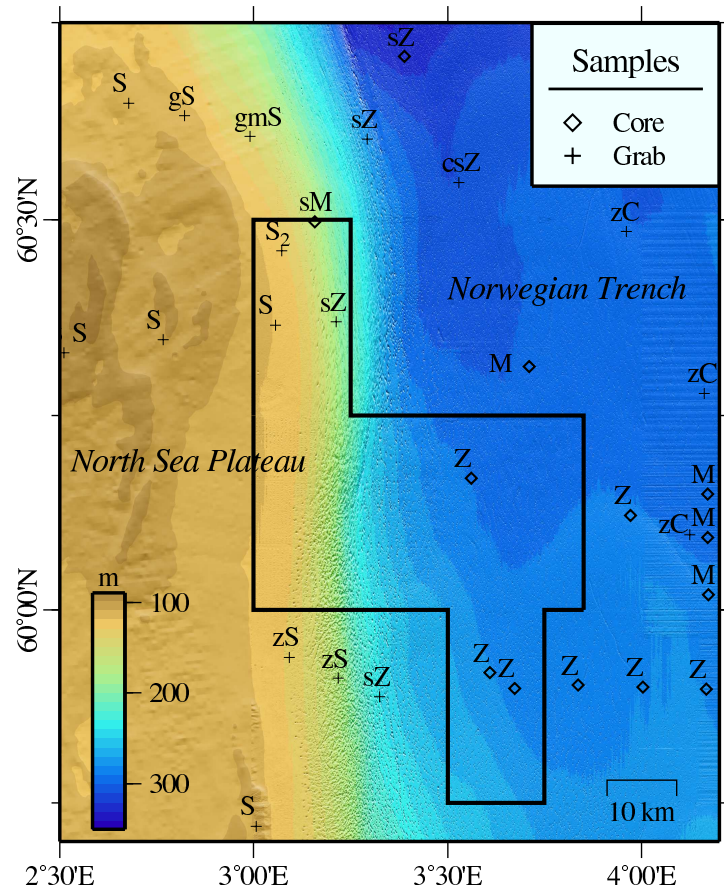
Fig. 1. Bathymetry and sediment samples (gravity cores and grabs) in and around the study area (solid lines) [25]. Letters indicate laboratory classification of samples, e.g. zS means "silty sand" (TABLE I, II).

TABLE II
COMPOSITION OF SEDIMENT SAMPLES

Abbr.	Mass fraction		Description
M	Clay+Silt > 90%	$1 : 2 <$ Clay:Silt $< 2 : 1$	Mud
C	Sand < 10 %	Clay:Silt $> 2 : 1$	Clay
Z	Sand < 10 %	Clay:Silt $< 1 : 2$	Silt
S	Sand > 90 %	Gravel < 2 %	Sand
sZ	Sand < 50 %	Clay:Silt $< 1 : 2$	Sandy silt
zS	Sand > 50 %	Clay:Silt $< 1 : 2$	Silty sand
gS	Gravel 2-30 %	Clay+Silt:Sand $< 1 : 9$	Gravelly sand
zC	Sand < 10 %	Clay:Silt $> 1 : 1$	Silty clay
sM	Sand < 50 %	$1 : 2 <$ Clay:Silt $< 2 : 1$	Sandy mud
gmS	Gravel 2-30 %	Clay+Silt:Sand $< 1 : 1$	Gravelly, muddy sand
csZ	Silt > 50 %	$1 : 2 <$ Clay:Sand $< 2 : 1$	Clayey, sandy silt

fragments) lies in a wedge-shaped area with higher backscatter levels than the surroundings. The sand class was therefore split in two: S and S_2. Hence we defined five classes in our classifier evaluation: S, S_2, Z, sZ and U.

B. Data preparation

The MBES data and sediment samples were acquired by the Norwegian Defence Research Establishment (FFI) on the research vessel M/S HU Sverdrup II. About 15 GB of data were collected during 600 hours of three separate surveys in the fall of 2005 and late winter of 2006. The three surveys together cover the contiguous area outlined in Fig. 1. The echo sounder was a Kongsberg Maritime EM 1002. The EM 1002 has a center frequency of 95 kHz; the array is semi-cylindrical and emits a fan-shaped beam with $2.0°$ along-track $3\,\mathrm{dB}$ beamwidth, and it forms 111 beams with $2.0°$ beamwidth across track. Fig. 2 illustrates the acquisition geometry of an MBES, and Fig. 3 shows a map of the backscatter data. The latter corresponds to Fig. 1, with sounding depth replaced by beam intensity, interpolated onto a regular 10 m^2 grid. The beam intensity is an approximation to the seabed scattering strength [34], [35], and not normalized with respect to angular variations. The angular coverage of the EM 1002 is up to $150°$ athwartships, but we have only used port and starboard sectors from 0-50°. The present data were acquired with a maximum beam angle (with respect to the vertical) of up to about $65°$, but many pings have a narrower angular range. This depends in particular on the water depth, refraction, and possibly also the weather conditions; outer beams may also be missing due to erroneous bottom detection. In addition, other authors [16] have observed that beampattern effects may cause unphysical backscatter levels at high incidence angles. To avoid extrapolating missing data when comparing data from different pings, observations with missing data would have to be discarded. Hence the maximum angle during data processing was set as low as $50°$ to avoid discarding many observations.

To describe the data processing, we let $y = |\Pi(t)|^2$ denote the squared envelope of the complex pressure time series. The averaged scattering cross section is defined as in [36, Ch. 16]:

$$s(\theta) = \int_\Gamma \sigma(\boldsymbol{r}; \theta)\Omega(\boldsymbol{r}; \boldsymbol{r}_0)\,\mathrm{d}^2\boldsymbol{r}, \tag{1}$$

where σ is the local scattering cross section. The integration variable $\boldsymbol{r}$ is position on the seafloor; the integral is over the ensonified area Γ of a single beam (at time t); $\Omega(\boldsymbol{r}; \boldsymbol{r}_0)$ is the ensonification function, which is the beam pattern expressed as a function of $\boldsymbol{r}$ and $\boldsymbol{r}_0$, the position of the array. It is assumed that Γ is small, so that the incidence angle θ is fixed.

For a long stationary signal, Γ is the beam footprint. The mean square envelope (in volts^2), compensated for propagation loss, is then

$$\langle y \rangle = 2s_0^2 s_r^2 s(\theta), \tag{2}$$

where s_0 is the source level and s_r is the receiver sensitivity on the beam's maximum response axis. The factor 2 occurs if s_0 is defined in terms of RMS pressure [8, App. F–G]. For a short MBES pulse of length τ, the across-track width of Γ is limited to $c\tau/2\sin\theta$, and it is not precisely known where Γ lies within the footprint (at particular time t). Nor was $\Omega(\boldsymbol{r}, \boldsymbol{r}_0)$ precisely known. Instead (2) is used with $\langle y \rangle$ replaced by the mean intensity, $\bar{y}$, in a window around the echo peak. On average, samples near the echo peak are energy scattered from resolution cells near the maximum of the ensonification function, where its value ($\Omega = 1$) is
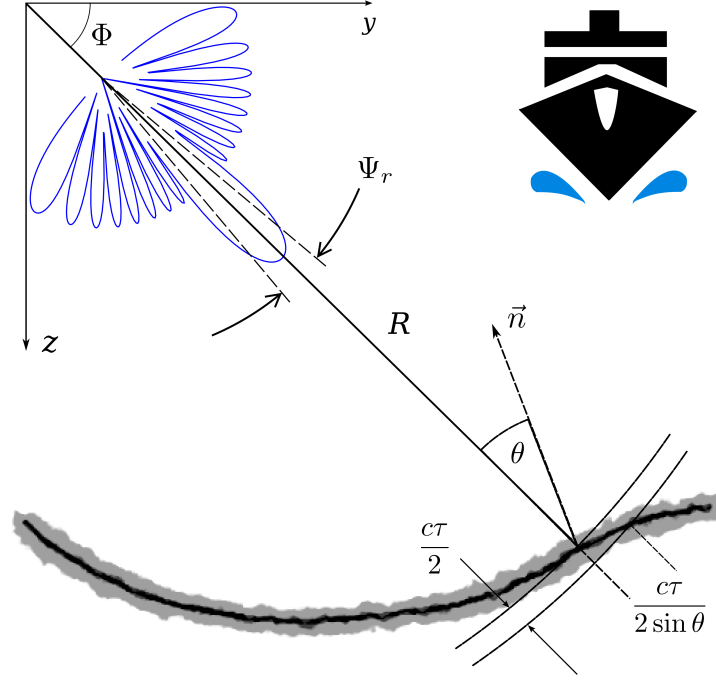
Fig. 2. MBES geometry in the across-track plane. The vessel moves along the x-axis. A single ping with pulse length τ ensonifies a strip of seafloor which is wide in y and narrow in x. The receiver array steers many beams at different depression angles Φ, yielding backscatter response at different incidence angles θ. A resolution cell has a length of $c\tau/2\sin\theta$; the beam footprint length is approximated as $R\Psi_r$. The beampattern shown here has 3 dB beamwidth $\Psi_r = 10°$ and uniform -20 dB sidelobe levels. The actual beamwidth of the EM 1002 is $\Psi_r = 2.0°$, but the precise beampattern is not known.

known by definition of Ω, s_0, and s_r. We understand that the manufacturer accounts for the factors s_0 and s_r, which have been determined experimentally [34], [35]. The corrected beam intensity is

$$y_c = \frac{\bar{y}q^2}{s_0^2 r_0^2 A},$$

(3)

where q is the one-way propagation loss. The factor A is an approximation to

$$\int_\Gamma \Omega \, d^2 \boldsymbol{r} \approx A = R\Psi_t \times \min\left[R\Psi_r, c\tau/2\sin\theta\right],$$

where R is the slant range, and Ψ_t, Ψ_r are the 3 dB transmit and receive beamwidths. The factor A ensures that $\langle y_c \rangle \approx \langle \sigma \rangle$. The bias caused by the simplification $A \approx \int_\Gamma \Omega \, d^2 \boldsymbol{r}$ is independent of seabed type. In the present work it is only the difference between seabed types which matters. The best way to determine scattering strength would be to measure the total backscattered energy from the beam footprint [37]. However, our recorded data did not include the full time series of the seabed echoes.

We used ray theory [38, Ch. 3] to compute the geometrical spreading loss and the beam direction $\boldsymbol{v}$ at the seafloor. Sound absorption was estimated using conductivity, temperature, and density (CTD) profiles and the empirical formulae of [39]. The seabed shape and surface normal $\boldsymbol{n}$ were calculated with cubic spline interpolation using depth soundings from a set of consecutive pings. (By comparison, the system software assumes that the seabed has a constant slope across the whole swath.) The incidence angle was found from the three-dimensional scalar product $\cos\theta = \boldsymbol{n} \cdot \boldsymbol{v}/\|\boldsymbol{v}\|$.
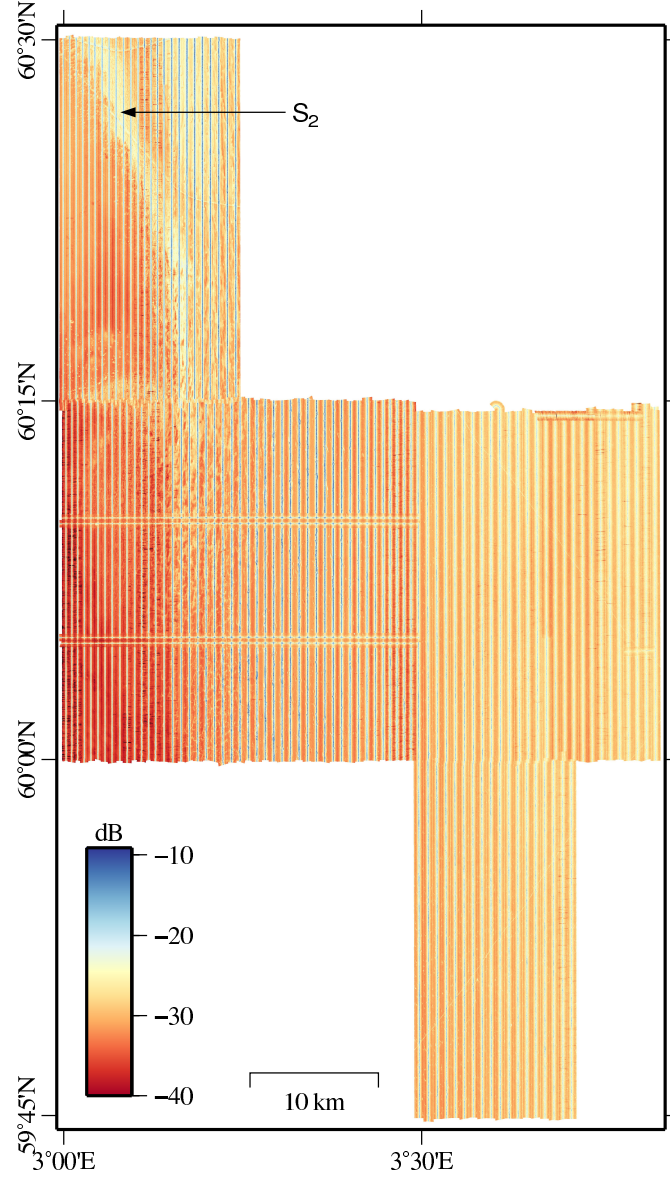
Fig. 3. Scattering strength data (beam intensity) in the survey area. An area of high backscatter levels which contains the S_2 sample (Fig. 1) is also indicated.

C. Observations

Suppose $0 < \theta_1 < \ldots < \theta_{K-1} < \theta_{\max}$ is a subdivision of the incidence angle range of the sonar swath. An observation is a K-dimensional vector $\boldsymbol{d}$, where each component d_k is the average of corrected intensity measurements (3) for incidence angles $\theta \in [\theta_{k-1}, \theta_k)$. As Fig. 4 shows, the vector $\boldsymbol{d}$ is simply a piecewise constant approximation to the full scattering strength observation. A group of measurements can be formed from N_p consecutive pings, but data from port and starboard sides of the swath are kept as separate observations.

For a seabed type ω, we shall estimate a PDF $p(\boldsymbol{d}|\omega)$ for the random vector $\boldsymbol{d}$ with respect to ω. As a rule-of-thumb [40], there should be at least ten times as many training observations for ω as there are components in $\boldsymbol{d}$. To keep K as low as possible without reducing the approximation accuracy too much, the segments $[\theta_{k-1}, \theta_k)$, $k = 1, \ldots, K$, are defined such that the total approximation error for the data set is minimized for fixed K. Hence the segments in Fig. 4 are short where the intensity variation with angle is large, and longer where the

126
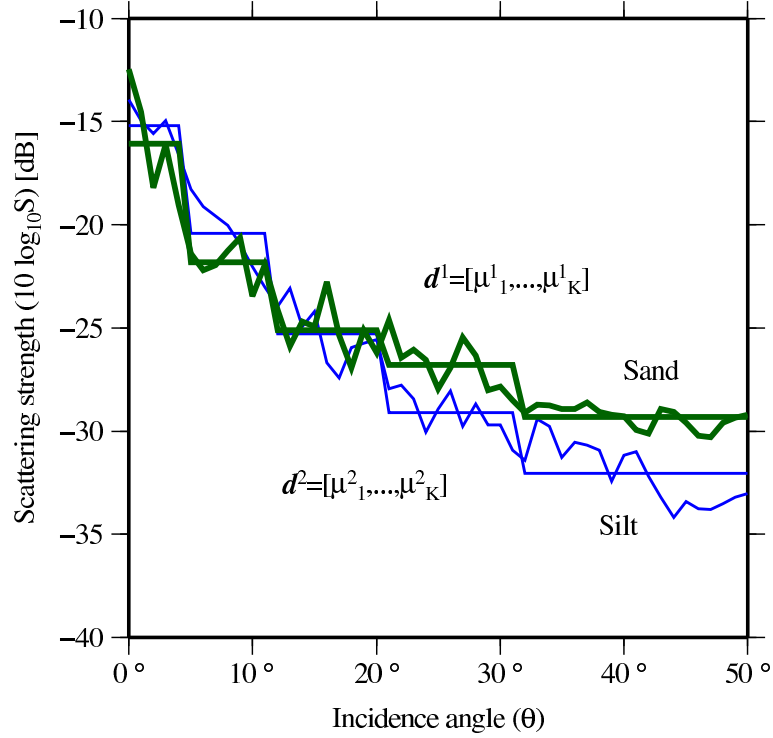
Fig. 4. Two observations of scattering strength versus incidence angle (irregular curves), along with piecewise constant approximations with $K = 5$ segments: sand (thick lines) and silt (thin lines). Each segment is an angular range corresponding to a horizontal line segment in the piecewise constant approximations. The segments are the same for both (all) classes. The K-dimensional observation vectors $\boldsymbol{d}$ comprise the mean values for the K segments. (Data from 10 consecutive pings have been used to form one vector.)

variation is smaller.

Mathematically, getting $\boldsymbol{d}$ from measurements can be viewed as a two-step binning process, defined by two non-decreasing onto functions b and c,

$$[0, \theta_{\max}] \xrightarrow{b} \{1, 2, \ldots N\} \xrightarrow{c} \{1, 2, \ldots K\}, \; K < N.$$

We let $\langle A \rangle$ denote the arithmetic mean of the elements of a discrete set A. A full observation (irregular curve in Fig. 4) is thus a vector $\boldsymbol{\varsigma} = \begin{bmatrix} \varsigma_1 & \ldots & \varsigma_N \end{bmatrix}$, where the component ς_i is the average of all measurements in the ith bin for a group of pings,

$$\varsigma_i = \langle y_c(\theta) \,|\, b(\theta) = i \rangle. \tag{4}$$

The number of bins N depends on the angular coverage and resolution. In this paper $\theta_{\max} = 50°$ and $N = 50$, so the bin length is $1°$ (about the ratio of angular coverage to number of beams). A data set of M observations $\boldsymbol{\varsigma}^1, \ldots, \boldsymbol{\varsigma}^M$ is represented by an $M \times N$ matrix

$$\boldsymbol{D} = \begin{bmatrix} \varsigma_1^1 & \cdots & \varsigma_N^1 \\ \vdots & \ddots & \vdots \\ \varsigma_1^M & \cdots & \varsigma_N^M \end{bmatrix} \overset{\xrightarrow{N \text{ bins}}}{.\big\downarrow M \text{ observations}}$$

The segments correspond to the horizontal lines in Fig. 4 and are defined by c. Hence $\boldsymbol{\varsigma}$ is replaced by the observation vector $\boldsymbol{d} = \begin{bmatrix} \mu_1 & \ldots & \mu_K \end{bmatrix}$, where $\mu_k = \langle \varsigma_i \,|\, c(i) = k \rangle$ is the mean of all components in the kth segment. For a given number of segments K, the approximation

127

error depends on c, which is determined adaptively from $\boldsymbol{D}$. We use a method from [23], [24], which minimizes the total squared approximation error for the data set $\boldsymbol{D}$. The optimization problem is

$$\min_c \sum_{k=1}^K \sum_{j=1}^M \sum_{i \in I_k} \left(\varsigma_i^j - \mu_k^j \right)^2, \tag{5}$$

where $I_k = \{\, n \mid c(n) = k \,\}$. A fast and straightforward algorithm for solving (5) can be found in [23]. The solution for c can be implemented as a linear transform in the following way. Let n_k be the number of elements in I_k, and define $N_k = \sum_{i=1}^{k-1} n_k$, $N_1 = 0$. Let $\boldsymbol{w}_k$ be the $N \times 1$ vector with $1/n_k$ in rows $N_k + 1 \ldots N_{k+1}$ and 0 elsewhere. Then

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{w}_1 & \ldots & \boldsymbol{w}_K \end{bmatrix} \tag{6}$$

is an $N \times K$ matrix such that $\boldsymbol{d} = \varsigma \boldsymbol{W}$ is the feature vector of observation ς.

The optimization (5) is performed on the whole data set (all observations and seabed types), or a representative subset, not just the training samples. The segments are the same for all seabed types. The main point is to reduce the dimension of ς (i.e. $K < N$); the PDF $p(\boldsymbol{d}|\omega)$ can be estimated more accurately when K is small. This is relevant for ASC when training observations must be picked near a few ground truth sites and therefore are limited in number. Failing to ensure a high ratio between training set size and sample space dimension may impair the classification accuracy, as is commonly seen in pattern recognition problems [5].

D. Bayesian classification

Let $p(\boldsymbol{d}|\omega_i)$ denote the PDF for feature vector $\boldsymbol{d}$ with respect to seabed class ω_i. The Bayes decision rule [5, Ch. 2] states that a new observation $\boldsymbol{x}$ should be assigned to the class with the highest posterior probability. By Bayes' theorem, this means selecting class ω_i if

$$p(\boldsymbol{x}|\omega_i)P(\omega_i) > p(\boldsymbol{x}|\omega_j)P(\omega_j) \text{ for all } j \neq i, \tag{7}$$

where $P(\omega_i)$ is the prior probability assigned to class ω_i. The PDFs may be estimated parametrically or non-parametrically. We used a multinormal model and equal prior probabilities. Thus,

$$p(\boldsymbol{d}|\omega_i) = \frac{(2\pi)^{-K/2}}{|\hat{\boldsymbol{\Sigma}}_i|^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{d} - \hat{\boldsymbol{\mu}}_i)^T \hat{\boldsymbol{\Sigma}}_i^{-1}(\boldsymbol{d} - \hat{\boldsymbol{\mu}}_i) \right], \tag{8}$$

where $\hat{\boldsymbol{\mu}}_i = (1/N) \sum_{k=1}^N \boldsymbol{d}_k^i$ is the mean vector from N training samples in class ω_i, and

$$\hat{\boldsymbol{\Sigma}}_i = \frac{1}{N} \sum_{k=1}^N (\boldsymbol{d}_k^i - \hat{\boldsymbol{\mu}}_i)(\boldsymbol{d}_k^i - \hat{\boldsymbol{\mu}}_i)^T$$

is the sample covariance matrix for observations in class ω_i.

Classification was implemented with the Matlab toolbox PRTools4 from TU Delft [41]. We have tested the three common variants of the Bayesian normal classifier: Quadratic discriminant classifier (QDC), linear discriminant (LDC), and uncorrelated (UDC). UDC is the assumption that the covariance matrices are diagonal; LDC is the assumption that the covariance matrices are equal ($\Sigma_i = \Sigma$ for all classes ω_i); QDC is general (no assumption on covariance structure). LDC or UDC may be preferred over QDC when Σ is near-singular (typically, with few and closely correlated training samples). For QDC, the decision
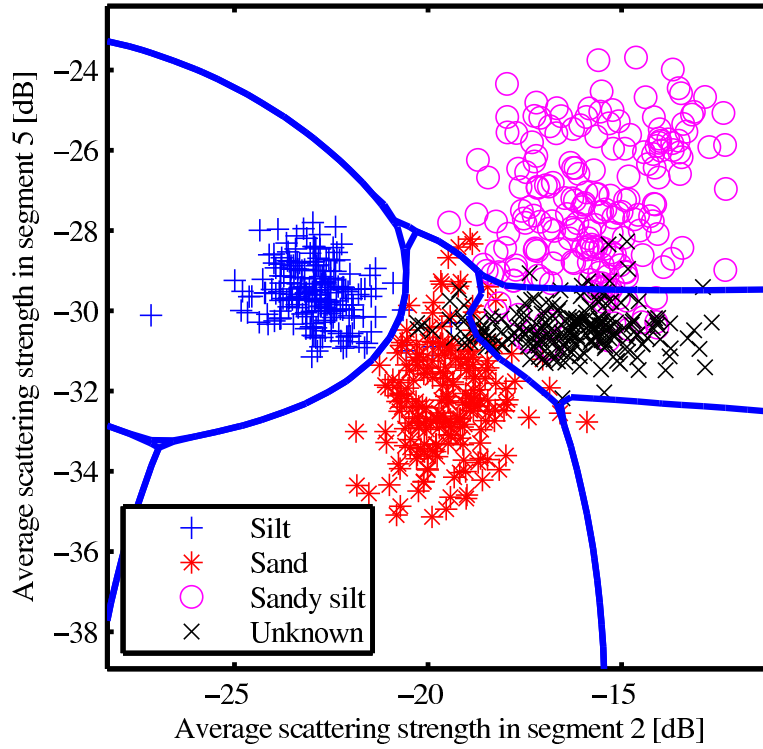
Fig. 5. Scatterplot and decision boundaries for the QDC in two dimensions. Segment 2 is the angular range $[5°, 12°\rangle$, and segment 5 is the range $[32°, 50°\rangle$. We used the forward floating search method [42] to pick the two most informative features from a five-dimensional approximation of scattering strength. The plot contains 800 observations (200 per class), which were randomly selected from large training areas as explained in Section III-A and Fig. 7.

boundaries (points of equality in (7)) are $(K - 1)$-dimensional quadric hypersurfaces, as the two-dimensional example in Fig. 5 shows.

The critical step is PDF estimation, not the assumption of normality in (8). However, using a normal model reduces the risk of overfitting the model to the data, and simplifies the PDF estimation task to computing $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}^{-1}$. By definition, a feature vector $\boldsymbol{d}$ results from a triple averaging over time samples, over pings, and over incidence angle segments. It is therefore a sum of many random vectors, each component of which is a sum of intensities from separate sonar resolution cells. With the common assumption of statistical independence, it follows below that the PDFs for the observation vectors should be asymptotically normal.

Suppose the PDF $f_{Y|\theta}(y)$ describes the distribution of beam intensity measurements at angle θ, for a particular seabed type. The seabed type is characterized by a distribution $f_{S|\Theta}(s|\theta)$ of the average scattering strength (1). To be more specific, we consider the compound model in [36, Ch. 16.3], where it is assumed that intensity measurements collected from regions with a fixed value of s are exponentially distributed (i.e. Gaussian complex pressure). The compound PDF is

$$f_{Y|\Theta}(y|\theta) = \int_0^\infty \frac{1}{s} e^{-y/s} f_{S|\Theta}(s|\theta)\, \mathrm{d}s. \tag{9}$$

Measurements in the kth segment are for random angles in $[\theta_{k-1}, \theta_k\rangle$, so we regard Θ as a

continuous random variable with uniform PDF

$$f_\Theta(\theta; k) = \begin{cases} 1/(\theta_k - \theta_{k-1}) & \text{if } \theta \in [\theta_{k-1}, \theta_k) \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

We write the expected value in segment k with respect to Θ as

$$\langle A \rangle_k = \frac{1}{\theta_k - \theta_{k-1}} \int_{\theta_{k-1}}^{\theta_k} A(\theta) \, \mathrm{d}\theta. \tag{11}$$

The marginal distribution

$$\begin{aligned} f_Y(y; k) &= \int_{\theta_{k-1}}^{\theta_k} f_{Y|\Theta}(y|\theta) f_\Theta(\theta; k) \, \mathrm{d}\theta \\ &= \int_0^\infty \frac{1}{s} e^{-y/s} \langle f_{s|\Theta}(s|\theta) \rangle_k \, \mathrm{d}s \end{aligned} \tag{12}$$

is a model for the normalized histogram of intensity measurements in the kth segment. For log-intensity measurements, (12) must be replaced with

$$\begin{aligned} f_W(u; k) &= \int_0^\infty f_{W|S}(u) \langle f_{S|\Theta}(s|\theta) \rangle_k \, \mathrm{d}s, \\ W &= \log Y, \\ f_{W|S}(u) &= \frac{e^u}{s} e^{-e^u/s}. \end{aligned} \tag{13}$$

Now form the vector $\boldsymbol{Z} = \begin{bmatrix} Y_1 & \dots & Y_K \end{bmatrix}^T$, where Y_k is a random variable with PDF as in (12) or (13). If $\boldsymbol{Z}^1, \dots, \boldsymbol{Z}^M$ are M independent random vectors distributed as $\boldsymbol{Z}$, with mean $\boldsymbol{\mu}$ and finite covariance matrix $\boldsymbol{\Gamma}$, then

$$\frac{\boldsymbol{Z}^1 + \dots + \boldsymbol{Z}^M}{\sqrt{M}} \xrightarrow{\text{D}} N(\sqrt{M}\boldsymbol{\mu}, \boldsymbol{\Gamma}), \tag{14}$$

where $\xrightarrow{\text{D}}$ means convergence in distribution [43, Ch. 11]. For sufficiently large M the PDF for the average

$$\boldsymbol{m} = \frac{\boldsymbol{Z}^1 + \dots + \boldsymbol{Z}^M}{M} \tag{15}$$

approaches the normal distribution $N(\boldsymbol{\mu}, \boldsymbol{\Gamma}/M)$. The average $\boldsymbol{m}$ (15) is a model for the observation vector $\boldsymbol{d}$, although in practice we allow for different number of measurements in different segments. The number of independent measurements increases with N_p (number of pings) and decreases with K (number of segments). With the notation

$$\varsigma(\theta) = \mathrm{E}[S|\Theta = \theta] \tag{15a}$$
$$\nu^2(\theta) = \mathrm{Var}[S|\Theta = \theta] \tag{15b}$$
$$\varsigma_L(\theta) = \mathrm{E}[\log S|\Theta = \theta] \tag{15c}$$
$$\nu_L^2(\theta) = \mathrm{Var}[\log S|\Theta = \theta], \tag{15d}$$

then from the appendix (28)–(29),

$$\mu_k \equiv \mathrm{E}[d_k] = \mathrm{E}[Y] = \langle \varsigma \rangle_k \tag{16}$$
$$M\Sigma_{kk} \approx \mathrm{Var}[Y] = 2 \langle \nu^2 \rangle_k + \langle \varsigma^2 \rangle_k + \mathrm{Var}_\Theta[\varsigma] \tag{17}$$

for linear intensity. For log-intensity, the result is (30)–(31)

$$\mu_k \equiv \mathrm{E}[d_k] = \mathrm{E}[\log Y] = \langle \varsigma_L \rangle_k - \gamma \tag{18}$$

$$M\Sigma_{kk} \approx \mathrm{Var}[\log Y] = \langle \nu_L^2 \rangle_k + \frac{\pi^2}{6} + \mathrm{Var}_\Theta\left[\varsigma_L\right]. \tag{19}$$

Here γ is the Euler-Mascheroni constant.

In the appendix we have used these formulae to calculate error bounds for discrimination between pairs of classes, parameterized by M and the standard deviation of the intrinsic scattering strength. This may in turn be used to predict whether it is better to apply the classification method to intensity or log-intensity data in general. Since we do not know the underlying distributions $f_{S|\Theta}(s|\theta)$, the maximum entropy principle [44] is invoked to make the comparisons. From (43)–(44) in the appendix,

$$\nu^2(\theta) = \left(e^{\nu_L^2} - 1\right)\varsigma^2(\theta),$$

so ν is proportional to ς, and $\nu = 0$ for $\nu_L = 0$. If we assume for simplicity that the proportionality factor (hence ν_L) is angle-independent and class-independent, then the error bound (36) can be plotted as a function of the single parameter ν_L. This gives a family of error bound curves indexed by the number M. Fig. 6 shows one example for the classes S and Z. To make this plot we have used the class means of the observed scattering strength $\varsigma_L(\theta)$ (full observations with $N = 50$ bins). These error bounds are for classification in $K = 5$ dimensions. Fig. 6 suggests that log-intensity is the better choice. The picture is similar for e.g. the pair sZ and U.

III. EXPERIMENTS AND RESULTS

A. Experimental setup

The purpose of the experiments was to see how the classification accuracy depended on

- the dimension of the scattering strength vector $\boldsymbol{d}$;
- the amount of training data;
- the amount of spatial averaging;
- the extent of the angular sector.

These factors determine the reliability, spatial resolution, and practical applicability of the method. We evaluated the classifier in two alternative ways. Both involved the following recipe:

1) Select random, independent training and test data for each known class (seabed type).
2) Estimate the PDFs (8) using the training set.
3) Count the number of misclasssified test samples in repeated experiments to obtain an average error rate, the probability of error.

The first way was with small training areas centered on the sediment sample sites. This would be the natural approach to mapping new areas where little is known about the seabed geology in advance. There are seven samples inside the survey area, representing four classes (Section II-A and Fig. 1), and each defined the center of a 3 km² block of training data. Two additional blocks of data were taken from an area far from any samples. These data represented the class U (unknown, not sampled). We defined test blocks similar in number and size in the vicinity of the sample sites; they did not overlap with the training areas, but the seabed type could nevertheless be asserted with reasonable confidence based on prior
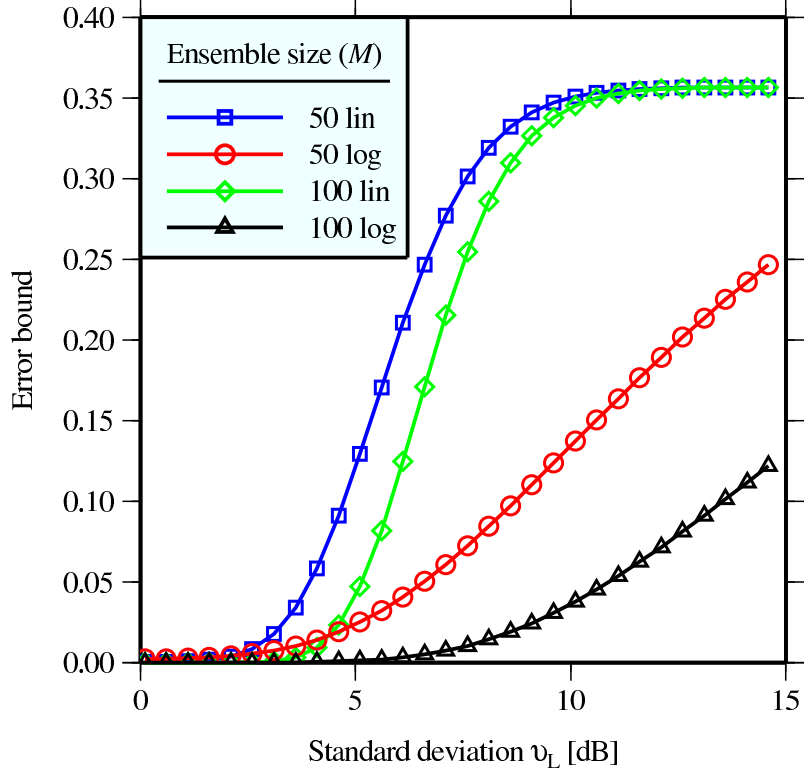
Fig. 6. Error bounds for discrimination between the S and Z classes ($K = 5$ segments). It is assumed that the variance is the same for both classes. With five segments, the values $M = 50$ and $M = 100$ correspond roughly to 10–20 pings spatial averaging. Error bounds for both intensity and log-intensity data are shown for comparison.

knowledge. Randomly drawn observations from the training areas were used to estimate the PDFs.

The second way was with more data and more careful randomization. This was an attempt to make the results and conclusions more robust. For each class we defined a large rectangular training area (rectangles with thick outlines in Fig. 7). Each training area was divided into smaller cells by a 20×20 grid subdivision (for better legibility, Fig. 7 illustrates this for 4×4 grids only). In the trials, half of the cells were assigned to training and the rest to testing. This was done in a random manner, with new random selection for each run. In each run, a specified number of training observations, ranging from 5 to 500 per class, were randomly selected from the combined set of 200 training cells and used to estimate the PDFs.

In its basic form, the present method yields two observations per group of pings, one on either side of the vessel. This limits the lateral (across-track) spatial resolution of the final class map, i.e. the number of independently classified seabed patches. To see how the method may handle environments where the seabed type varies across the swath, we extracted data from smaller, complementary angular sectors. We ran the same set of tests as above on each subsector independently of the others. The sectors were specified in two alternative ways. One approach was to define equiangular sectors, for example 0–25° and 25°–50°. The segments in each sector were computed from (5) applied to measurements in each sector separately. The number of segments K was set equal for all sectors. The other approach was to apply (5) to all data in the full angular range as before, and assign an equal number of segments to each sector. For example, with $K = 8$ (we let K be a power of two), we defined two sectors
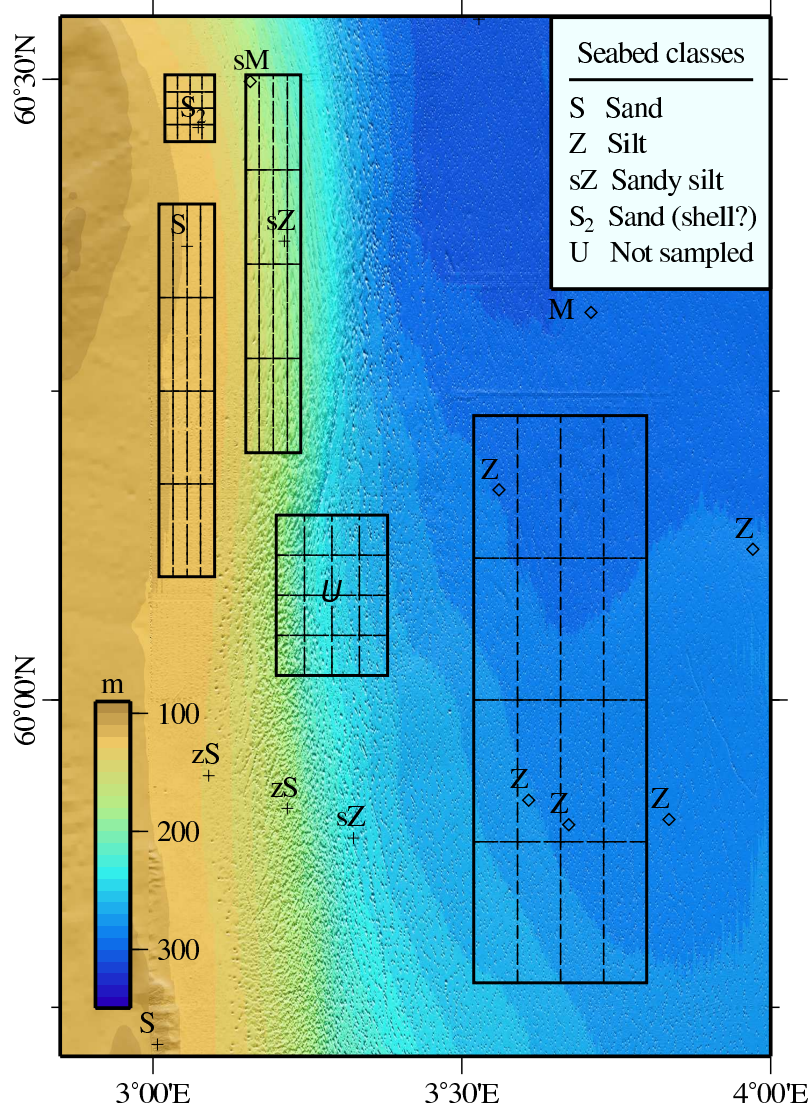
Fig. 7. Training areas (large rectangles with thick lines) with 4×4 subdivisions for five seabed classes in the survey area. In practice 20×20 subdivisions were used. For each test run, cells were randomly assigned for training and testing in a fixed proportion, and a specified number of observations were selected from the set of training cells and used to estimate the PDFs. See also TABLE I, II and Fig. 1.

with segments 1–4 and 5–8 respectively, and four sectors with segments 1–2, 3–4, 5–6, and 7–8 respectively. In the latter approach the PDF for a sector can be obtained from the full PDF by marginalizing out the segments outside the sector, which for normal distributions just means discarding the corresponding components.

B. Results

1) Accuracy: We ran the tests on $y_c(\theta)$ and $10 \log_{10} y_c(\theta)$ data, where y_c was the corrected, dimensionless beam intensity (Section II-B). The results presented here are for the classes S, Z, sZ and U. The S_2 class was left out because its small extent limits the amount of independent training and test data, and because we are less confident about the true class and nature of the S_2 data. However, when all five classes were taken into account, using small training areas around the samples, the results were again very similar to those of Fig. 8–
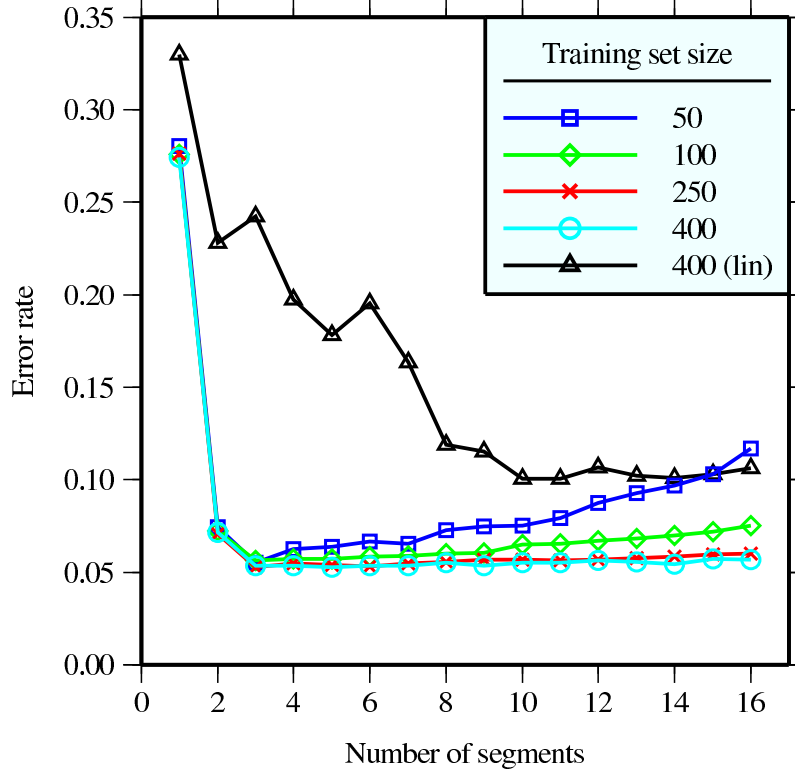
Fig. 8. Error rate vs. observation vector dimension (number of segments). The topmost curve is for intensity data, the rest are for log-intensity data (QDC, $N_p = 20$ pings averaging.)

Fig. 10. A single example (five classes, log-intensity, LDC, 50 pings averaging) can be found in [22].

As Fig. 8 shows, the best results were obtained with log-intensity (decibel) data, about 5 % less errors. Therefore, the results in Fig. 8–Fig. 11 are for log-intensity data, except for the topmost curve in Fig. 8. The estimated error rates were similar for the two approaches described in Section III-A, showing the same trends and minimum error rate. The main difference was a better error rate in very low (1–2) dimensions when using the large training areas with randomized subdivision. The results in this section have all been obtained with large training areas. The incidence angle domain was 0–50°.

The variable parameters are the number of training observations per class, the number of segments (observation vector dimension K), the number of pings (N_p) in the spatial averaging, and the classifier itself (QDC, LDC, or UDC). As the minimum error rate was attained with about five segments, we set $K = 5$ in the plots where observation vector dimension is not the independent variable. For each set of parameters, the estimated error rate is the average fraction of misclassifications from five random experiments.

Fig. 8 shows the error rate as a function of observation vector dimension. The minimum error rate is about 5 %, it is achieved with log-intensity data, and is quite stable up to 16 dimensions. However, the importance of dimensional reduction is apparent when the number of training observations is small. Even with much training data, the error rate increases in more than 16 dimensions, sometimes sharply. Fig. 9 shows the effects of training set size and spatial averaging. These results suggest that adequate estimation of the PDFs was achieved with about 50–100 observations per class, which is about 10–20 times the observation vector dimension (for $K = 5$). Increasing the number of pings in a group of measurements from 10 to
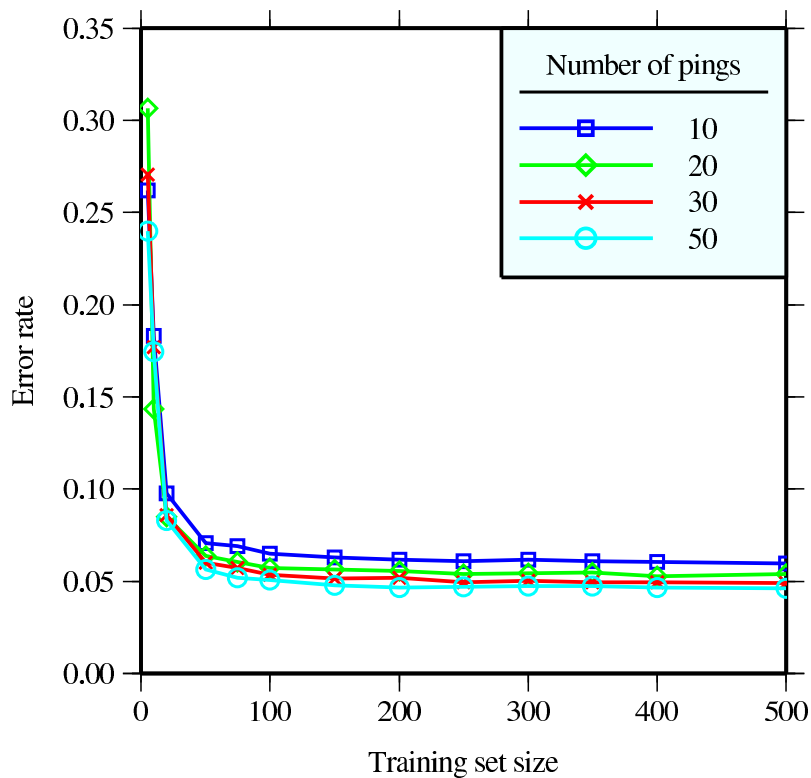
Fig. 9. Error rate vs. training set size (per class). (QDC, $K = 5$ segments.)
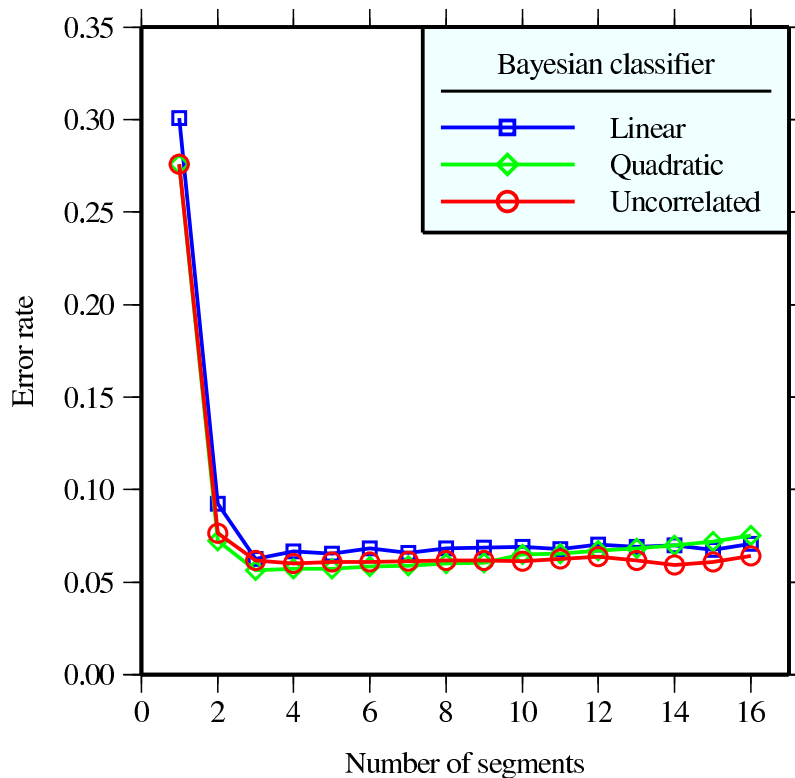

Fig. 10. Error rate vs. feature vector dimension (all classifiers). (100 training observations per class.)

50 reduces the variance, and hence the error rate, as Fig. 6 also implies. Fig. 10 compares the classifiers QDC, LDC and UDC. QDC performs best, but LDC and UDC are not much worse.
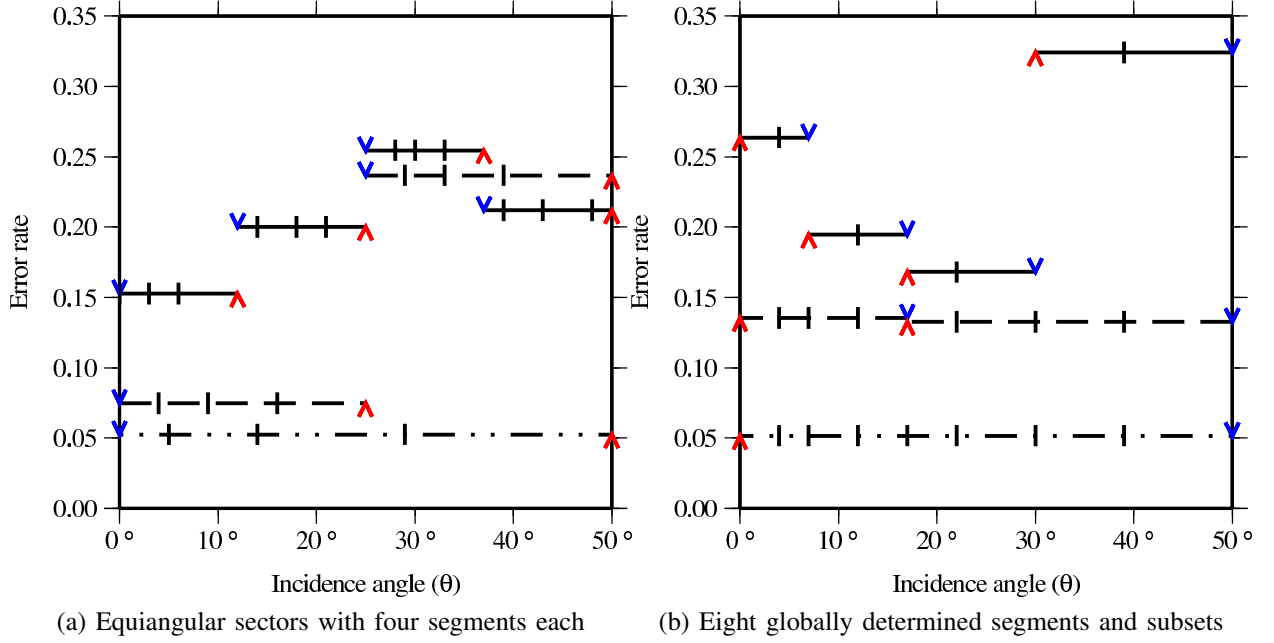
(a) Equiangular sectors with four segments each

(b) Eight globally determined segments and subsets

Fig. 11. Error rates for observations restricted to angular subsectors (QDC, 20 pings averaging, 300 training observations per class). Each horizontal line between ∧ and ∨ marks the width of a sector; the vertical bars mark the limits of segments. In (11b) there are 8 segments determined for the full angular range, with subsectors comprising two or four segments. In (11a), the sectors have been obtained by dividing the the full range into two or four equiangular parts, with four separately determined segments per sector.

In high dimensions, the reduced ratio between training set size and dimension causes more unstable estimates of the inverse covariance matrix, which affects QDC the worst. However, in 16 dimensions or less, the performance degradation of QDC is not significant.

2) Lateral resolution: For equiangular subsectors (Section III-A), the lateral (across-track) resolution is, roughly, the swath width on the seafloor divided by the number of sectors. Recall that we also tested an alternative subdivision method: determine the segments using all data (full swath), and assign an equal number of segments to each subsector. Fig. 11a and Fig. 11b show the error rates for the two approaches. The error rate for each subsector has been plotted as a horizontal line, whose extent indicates the opening angle of the sector. The vertical bars on each sector line show the breakpoints of the segments (Section II-C). In Fig. 11a, each subsector has been obtained by dividing the full swath into one, two, or four subsectors; in Fig. 11b, there are eight segments in total, and sectors with two, four, or eight segments.

For any partition into complementary sectors, the overall accuracy is limited by the accuracy of the worst sector. In these examples (Fig. 11), the best partition consist of the two sectors at 13 % error rate in Fig. 11b. This results in an overall error rate of 13 %. By comparison, the equiangular partitions in Fig. 11a do not yield acceptable error rates at moderate-to-high incidence angles. For almost every sector, the segments are longer at high angles, because the data vary less with angle in this regime. This is as predicted by scattering physics [36, Ch. 13–14]. The plots also show that using the full sector (0-50°) gives the lowest error rate.

3) Sediment map: The classifier (decision rule) was applied to the whole data set to produce a seabed class map of the survey area (Fig. 12). To make this map, the training data were restricted to 2.8 km^2 blocks containing the sediment samples shown in Fig. 1 (inside the
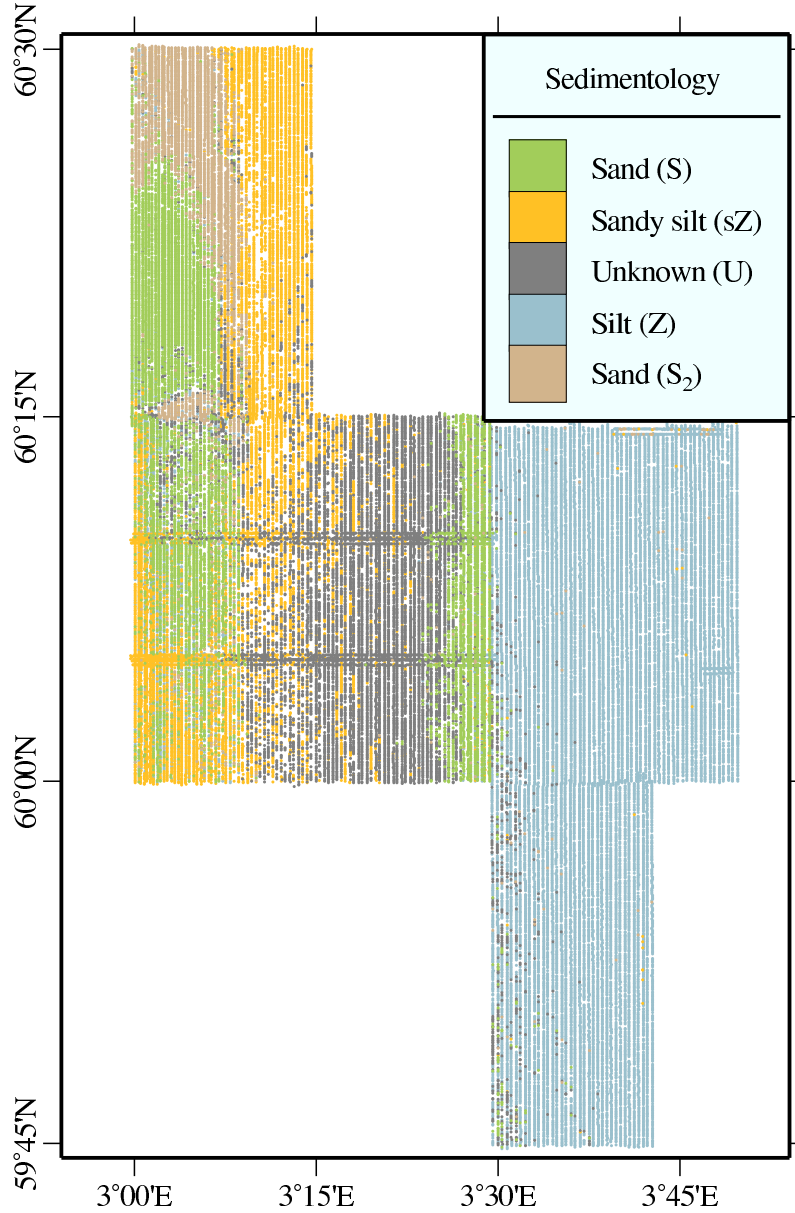
Fig. 12. Classification result for the survey area (five classes).

survey area). We included the S_2 sample, so the map has five sediment classes. The scattering strength observations were represented by a piecewise constant function approximation with $K = 5$ segments. The spatial averaging was $N_p = 20$ pings. The PDFs $p(\boldsymbol{x}|\omega_i)$ were computed using (8) with no assumptions about $\hat{\Sigma}_i$ (i.e. QDC). Each observation $\boldsymbol{d}$ was then classified according to the decision rule (7) with equal prior probabilities. Each observation was assigned the geographical coordinates of the midpoint of its angular sector (on the seafloor). The classified observations were then preprocessed with a 100 m^2 spatial block mode filter before each point was plotted with color codes to get Fig. 12. The map shows that the boundary between the sZ and U classes in particular is not crisp; sZ and U tend to blend into each other. We did anticipate that sZ and U would be similar. In fact, most of the misclassifications found during error rate estimation can be attributed to confusion between the sZ and U classes [22].

C. Statistics

In Section II-D it was argued that the PDF $p(\boldsymbol{d}|\omega_i)$ for observations within a class ω_i could be approximated by a multivariate normal model (8). This argument was based on the central limit theorem, so one should only expect it to hold if there are many statistically independent measurements for each observation vector. Moreover, the rate of convergence will be affected by the level of variance in the intrinsic scattering strength within each class. We have also assumed, based on available background information, that the training areas (Fig. 7) are homogeneous, i.e. cover a single sediment class. Since ground truth is limited, it is possible that the empirical distributions are contaminated by samples from other sediment types. Other types of outliers include those caused by acoustical shadows, and those which occur when the sonar bottom tracking algorithm fails for several consecutive pings.

The normal PDFs are on the form $p(\boldsymbol{d}|\omega_i) \propto \exp\left(-r^2/2\right)$, where

$$r^2 = (\boldsymbol{d} - \hat{\boldsymbol{\mu}}_i)^T \hat{\boldsymbol{\Sigma}}_i^{-1} (\boldsymbol{d} - \hat{\boldsymbol{\mu}}_i) \tag{20}$$

is the (squared) Mahalanobis distance. The classification accuracy is therefore directly influenced by the statistical distribution of r^2. For a true multinormal distribution, the standardized residuals $\boldsymbol{\epsilon} = \boldsymbol{\Sigma}_i^{-1/2}(\boldsymbol{d} - \boldsymbol{\mu}_i)$, where $\boldsymbol{\Sigma}_i^{-1/2}$ is the Cholesky decomposition of $\boldsymbol{\Sigma}_i^{-1}$, form an uncorrelated Gaussian random sequence with zero mean and unit standard deviation. Therefore, $r^2 = \boldsymbol{\epsilon}^T \boldsymbol{\epsilon}$ is the sum of the squares of K standard independent normal random variables, and follows a $\chi^2(K)$-distribution with K degrees of freedom (where K is the dimension of the observation vectors). Thus one way to assess the multinormal assumption is to consider a quantile-quantile plot of the empirical distribution of r^2 versus the $\chi^2(K)$-distribution [45, Ch. 4.6]. This is shown in Fig. 13 for the silt (Z) and sand (S) classes. To make these plots, $n = 500$ random observations from each class (with $K = 5$ dimensions) were drawn from the large training areas in Fig. 7. The spatial averaging was $N_p = 30$ pings. The observed values of r^2 were sorted in ascending order and plotted against $F^{-1}(p_i, K)$, where $p_i = (i-0.5)/n$, $i = 1, \ldots n$ are n equally spaced quantiles, and F^{-1} is the inverse cumulative distribution function (CDF) for $\chi^2(K)$. For a multinormal distribution, the observations are expected to lie on the line with unit slope (dashed line).

Fig. 13a and Fig. 13b reveal some large outliers. To reduce the influence of outliers on $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$, hence make the outliers stand out more, $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ were robustly estimated using the minimum covariance determinant estimator [46], [47]. Fig. 13c and Fig. 13d show the same data sets after first removing the 15 (for silt) and 16 (for sand) most extreme observations. Note that the tail of the distribution in both classes is sensitive to how many observations are considered as outliers. If too many or too few observations are excluded, the remaining extreme observations will deviate more from the unit slope curve. However, for r^2 less than about 10 or 12, the observations follow the unit slope curve very closely.

IV. DISCUSSION

The purpose of this work has been to test a method for acoustical mapping of seabed composition using measured scattering strength versus incidence angle. The questions that arose concerned how effective the $S(\theta)$ function is for seabed type discrimination; what is a suitable statistical model for this function; what spatial resolution can be achieved; and how much training data or prior knowledge about backscatter statistics is required. The main findings for the North Sea data set were:
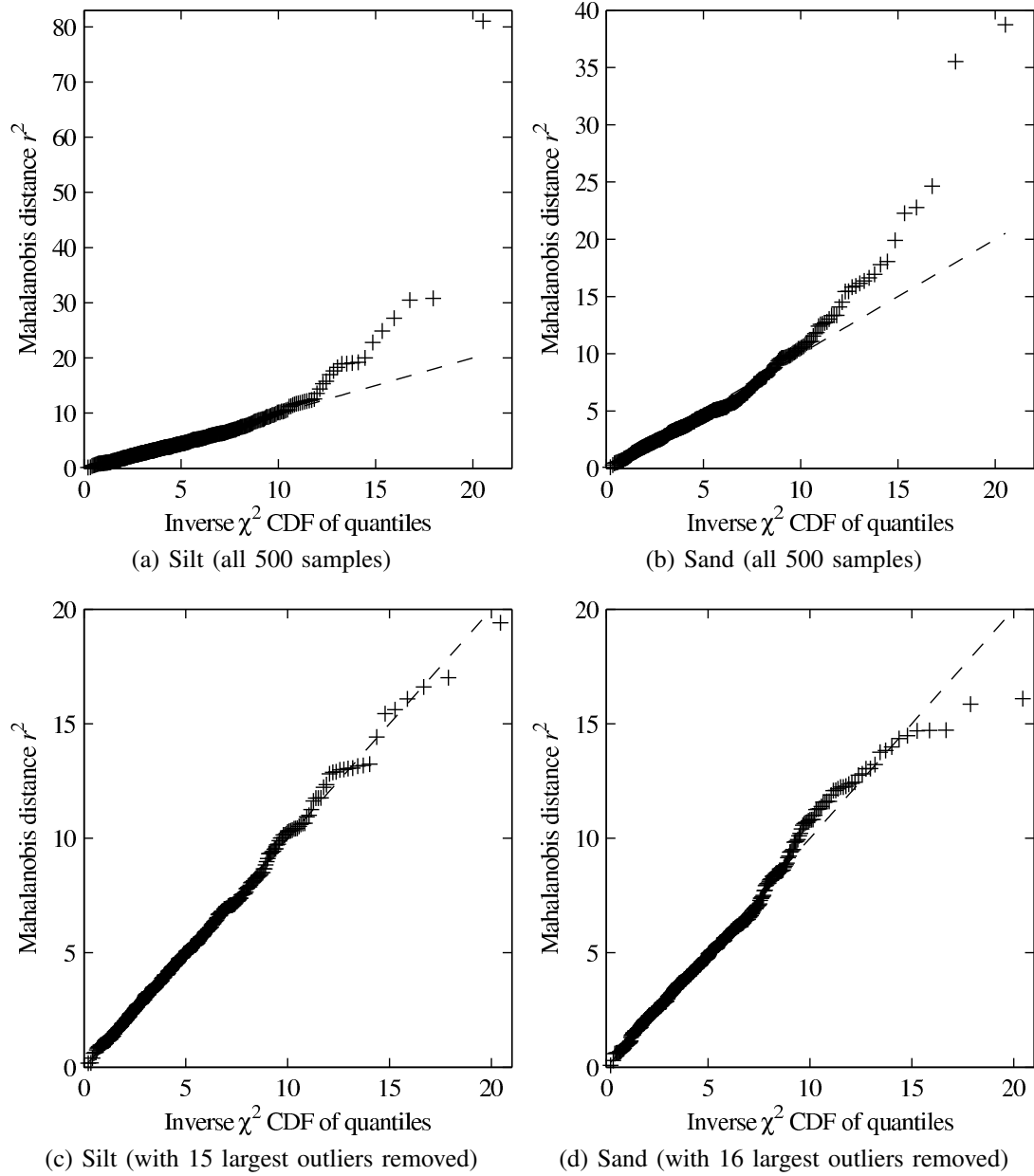
Fig. 13. Quantile-quantile plots for the distribution of the robustly determined Mahalanobis distance r^2 versus the $\chi^2(K)$-distribution with K degrees of freedom; here $K = 5$ (the dimension of the observation vectors). Fig. (a) and (b) show the result for 500 random observations from the silt (Z) and sand (S) classes respectively, drawn from the large training areas shown in Fig. 7. Fig. (c) and (d) show the result after first eliminating the 15 largest outliers for silt and the 16 largest outliers for sand. For a multinormal distribution, the observations are expected to lie on the line with unit slope (dashed line).

- The minimum error rate was about 5 %; most errors occured because we have split a large area with sparse ground truth into two classes which are physically similar.
- When the spatial averaging is decreased from 50 pings to 10 pings, the error rate increases from about 5 % to about 7 %. As this increase is not very large in absolute terms, it seems acceptable to use 10–20 pings averaging, which corresponds to a distance of about 25–50 m (5 knots, 1 Hz ping rate).
- The minimum error rate was attained with 3–5 segments in the piecewise constant approximation of $S(\theta)$. Using a single segment (overall average scattering strength) yields

a comparatively poor result, about 30 % errors. The less training data available, the worse performance in higher dimensions.

- Getting a low error rate required about 50–100 training observations per class. This agrees with the long-established rule-of-thumb that the ratio between training set size and dimension should be at least 10 (since here, $K = 3$–5) [40]. For complex classifiers, the ratio should be higher, which is a reason to choose the simple Bayesian normal classifier. With training data restricted to small areas near ground truth sites, the results were almost as good as when data were picked from large training areas in a random manner.

- The lowest error rate was achieved with QDC, the performance of LDC and UDC was marginally worse.

- For data restricted to smaller incidence angle sectors, the error rate was frequently above 15 % (Fig. 11). However, for the sectors 0–17°and 17°–50°(either side of the ship), the overall error rate was 13 % (Fig. 11b). By overall error rate we mean the highest error rate observed in any sector. If accurate classification can be done in several independent, complementary sectors, there will be more observations across the swath, hence better across-track resolution in the class map. We think that the 13 % error rate can be improved by taking into account contextual information, by weighting the prior probabilities according to the neighboring observations. Across-track resolution comparable to the along-track resolution may therefore be within reach for the present data set, but has not been examined further in this paper.

- A key question with respect to classification accuracy is how well the multinormal PDF approximates the true density. The analysis of Section III-C suggests that, excluding some outliers, the observed data are consistent with the assumptions. Each seabed type forms a cluster which is well represented by a normal PDF in the feature space, even when observations are picked at random across large areas. Because the feature vectors are formed by taking averages, the PDFs should theoretically approach normal distributions when the number of independent measurements is large. Averaging also reduces the variance of each distribution, which improves the classification accuracy.

The less training data, the more important feature reduction becomes. The present approach, piecewise constant approximation, is not the only possible solution. Alternatives include the principal components transform (PCA), polynomial approximation, and wavelet transforms. We tried to use the PCA transform on the scattering strength observations (4), and the results were similar to those of Fig. 8–Fig. 10. The piecewise constant approximation has three main advantages. First, each component of a feature vector has a physical meaning, and each component measures the same quantity, the average scattering strength. Second, taking averages over angular sectors which are typically wider than the uncertainty of the incidence angle estimates, reduces the significance of this uncertainty. Third, each feature vector component corresponds to an angular sector which ensonifies a separate patch of seafloor. Therefore there is a natural way to increase the across-track resolution of the present method, as discussed in connection with Fig. 11.

In practice, across-track resolution hinges not on the number of independently classified observations, but on the ability to detect class transitions within the swath. Therefore, one way to proceed would be to construct an algorithm which detects discontinuities in $S(\theta)$, and suggests a partition of the incidence angle domain. Then each partition (sector) would

be classified independently. The experiments of this paper (Fig. 11) merely suggest how successful classification using smaller sectors would be.

With sufficient training data, the error rate was stable in dimensions less than $K = 16$. More sophisticated classification methods like support vector machines [6, Ch. 12] did not improve results, nor did non-parametric PDF estimation techniques. This indicates that the 5 % limit on accuracy is inherent in the data, particularly in the unclear distinction between the U and sZ classes, and also due to insufficient ground truth in these two areas. The scattering strength in the sZ area is surprisingly high, but consistently so in data from three different cruises in the area. The western boundary of the sZ region also closely follows a transition in seabed morphology (Fig. 12). On a large scale, the seabed is much more rough in the sZ and U regions, with a very irregular pattern of depressions. This area corresponds roughly to the western slope of the Norwegian Trench. The depressions may be a result of erosion by a combination of water currents and fluid seepage up through the shallow geological layers, which causes seabed sediments to be lifted into suspension. This theory was put forward by Hovland [30], who observed very shallow gas blankets in seismic boomer data from a part of the western slope, with apparent vertical escape routes matching the occurrence of elongated depressions in the seabed. A similar phenomenon has been observed further south in the Norwegian Trench (the Skagerrak) [31]. It is not possible based on two grab samples to say whether the seabed has large roughness also at the scale of ultrasonic scattering, but the presence of vertically migrating shallow gas may be a partial explanation for the observed high scattering strength.

The fact that the statistical signal model (9) is formulated in terms of the spatially averaged scattering strength s (1) (instead of σ) reflects a subtle point in ASC. Namely, the signal statistics depend on the ensonification function $\Omega(\boldsymbol{r}, \boldsymbol{r}_0)$ in a way which is not simple to compensate for. This can be a problem because Ω is not related to sediment properties, but rather to beampattern, incidence angle, pulse length and water depth. However, the present method operates only on the mean, corrected intensity $\langle y_c \rangle \approx \langle \sigma \rangle$ (3), and is therefore not sensitive to Ω.

Besides the problem of lateral resolution, the present method has two main limitations. The first is the possibility that different seabed types may have similar scattering strength functions. The second is the problem of determining how many seabed types are recorded in the data, especially when there is little available ground truth, or when there is a gradual change in sediment composition, hence no clear boundary between distinct seabed types. Our data set illustrates these dilemmas. In advance we supposed that the sediment grain size distribution would change gradually from shallow to deep water. Thus the sand region between the U and Z areas in Fig. 12 (at about $03°30'\text{E}$) was an unexpected classification result. There are three alternative explanations. There may actually be sand in the area, because sand particles that are moved by currents from the plateau possibly accumulate at the bottom of the slope. A core sample at $03°23'\text{E}, 59°40'\text{N}$ contained a large sand fraction (28 %); likewise a core sample at $03°23'\text{E}, 60,°43'\text{N}$ had a 50 cm sandy surficial layer [25]. Alternatively, there may be a silty sediment with scattering strength similar to that of sand, due to a special combination of surface roughness and scattering from suspended fluid inclusions or coarse material. The third possibility is that the likelihood $p(\boldsymbol{d}|\omega)$ of the data is small with respect to all the predefined classes $\omega \in \{\text{S}, \text{S}_2, \text{Z}, \text{sZ}, \text{U}\}$. This implies that a new class should be introduced, or else a seabed type (here: sand) with low likelihood may get a high posterior probability due to the low evidence factor $p(\boldsymbol{d})$ in Bayes' theorem. The unsupervised Bayesian approaches of [33],

[48] and [21] determine the number of classes based on a Gaussian mixture model for the complete data set. Similar solutions may extend the present method.

The amount of training data needed to achieve a 5 % error rate is about 100 observations per class using 30 pings averaging. If a 7 % error rate is acceptable, 50 observations using 10 pings averaging are sufficient. Then, in a survey with 400 m line spacing and a 1 Hz ping rate at 5 kn speed, sufficient training data could be picked from 0.25 km^2 areas surrounding the ground truth sites. Because the classification is based on a physical quantity (mean scattering strength) which is instrument independent, properly processed data from any calibrated MBES survey could in principle be used for training. Moreover, with modeled scattering strength functions $\varsigma(\theta)$ and a given level of intrinsic variance $\nu^2(\theta)$, the equations from this paper can be used to estimate theoretical PDFs, and the likelihood of observations with respect to the theoretical PDFs can be studied. From this point of view, our method is therefore related to the inverse theory approach of e.g. [12], [13].

The present results are compatible with works like [12]–[14], [16]–[19], which all obtained good results by applying classification or inverse methods to the $S(\theta)$ function. There are of course successful alternative strategies which analyze fluctuations in the backscatter levels for discriminating between sediment types. An important example is texture analysis by means of second-order statistics (gray level co-occurrence matrices) applied to corrected sidescan mosaics [48]–[51]. In this approach, the pixel values in the mosaic are quantized to, say, L values (gray levels). Then, for any rectangular neighborhood in the mosaic, one may form an $L \times L$ matrix $\boldsymbol{G}(\Delta x, \Delta y)$, such that G_{ij} is the relative frequency of two pixels with gray levels i and j respectively being separated by pixel distance $(\Delta x, \Delta y)$ [52]. From $\boldsymbol{G}(\Delta x, \Delta y)$ one may derive several attributes which represent the character of the texture [52], [53].

This kind of information is discarded when using only scattering strength versus angle. One solution, based on the angular range analysis (inversion) of [12], is to visually inspect backscatter mosaics to ensure that the inversion is applied to data from acoustically similar areas [54]. For classification too, it seems that a hybrid method is possible, where both the angular variation and the backscatter scintillations are taken into account. For example, consider again as in Section II-C a subdivision of the incidence angle range, $0 < \theta_1 < \ldots < \theta_{K-1} < \theta_{\max}$, where for a given K the subdivision is such that the variation with angle is effectively minimized within each segment. Assume that the beamformed data have been corrected for propagation loss and beampattern variations, and mapped to a gray level range. Then within each segment, for a sequence of N_p consecutive pings, one may form an analogous co-occurrence matrix $\boldsymbol{G}(\Delta x, \Delta \theta)$, where now Δx is along-track separation between pairs of samples, and $\Delta \theta$ is the difference in estimated incidence angle. A special case would be to set $\Delta x \approx 0$ and consider only separation in angle. This results in additional features for each segment besides the mean scattering strength, e.g. the entropy and homogeneity properties used in [50]. As shown in Section III-B2 and Fig. 11, it may be possible to achieve a reasonable classification accuracy also after dividing the sonar swath into smaller separate sectors, and consequently get a better lateral resolution. We hypothesize that this accuracy can be improved by augmenting the observation vector $\boldsymbol{d}$ (Section II-C) with such texture attributes.

V. Conclusion

We have presented a statistical seabed classification method for multibeam echo sounders, and discussed its practical applicability. The method uses the observed scattering strength

dependence on incidence angle $S(\theta)$. To obtain a robust statistical description from scarce training data, $S(\theta)$ is represented by a piecewise constant function approximation. Theoretical error bounds and experimental results on a North Sea data set showed that good classification accuracy and spatial resolution is possible even with limited training data.

APPENDIX A

We want to compute asymptotic error bounds for discrimination between two seabed types, ω_1 and ω_2, in terms of the mean and variance of the scattering strength. It is useful to reinterpret (9) as a multiplicative noise model similar to the product model for clutter statistics in radar images [55]. The PDF for the product of two random variables, $Y = SX$, is

$$p_Y(y) = \int_{-\infty}^{\infty} \frac{1}{|s|} p_{S,X}(s, \frac{y}{s})\, \mathrm{d}s. \tag{21}$$

If $S \geq 0$, S and X are independent, and

$$p_X(x) = e^{-x}, \tag{22}$$

then (21) reduces to (9) [with $p_S(s) = f_{s|\theta}(s)$]. X represents noise (speckle) which is exponentially distributed with unit variance.

The variance decomposition formula applied to Y and S is

$$\mathrm{Var}[Y] = \mathrm{Var}[\mathrm{E}[Y|S]] + \mathrm{E}[\mathrm{Var}[Y|S]]. \tag{23}$$

By assumption (9), $\mathrm{E}[Y|S] = S$ and $\mathrm{Var}[Y|S] = S^2$, so

$$\mathrm{Var}[Y|\Theta] = \mathrm{Var}[S|\Theta] + \mathrm{E}[S^2|\Theta]$$
$$= 2\mathrm{Var}[S|\Theta] + \mathrm{E}[S|\Theta]^2. \tag{24}$$

This result can also be found directly from the relation

$$\mathrm{E}[Y^n|\Theta] = n!\, \mathrm{E}[S^n|\Theta], \tag{25}$$

which follows from (9), but (23) is valid for other noise models as well. The PDF for $\log X$ is

$$f_{\log X}(u) = e^u e^{-e^u},$$

with $\mathrm{E}[\log X] = -\gamma$, the Euler-Mascheroni constant, and $\mathrm{Var}[\log X] = \pi^2/6$. Thus, since S and X are independent,

$$\mathrm{Var}[\log Y|\Theta] = \mathrm{Var}[\log S + \log X]$$
$$= \mathrm{Var}[\log S] + \frac{\pi^2}{6}, \tag{26}$$
$$\mathrm{E}[\log Y|\Theta] = \mathrm{E}[\log S] - \gamma. \tag{27}$$

With notation as in (11) and (16), we find using (24)-(27) that

$$\mu_k \equiv \mathrm{E}[d_k] = \mathrm{E}[Y] = \langle\varsigma\rangle_k \tag{28}$$
$$M\Sigma_{kk} = \mathrm{Var}[Y] = 2\langle\nu^2\rangle_k + \langle\varsigma^2\rangle_k + \mathrm{Var}_\Theta[\varsigma] \tag{29}$$

for linear intensity, and

$$\mu_k \equiv \mathrm{E}[d_k] = \mathrm{E}[\log Y] = \langle\varsigma_L\rangle_k - \gamma \tag{30}$$
$$M\Sigma_{kk} = \mathrm{Var}[\log Y] = \langle\nu_L^2\rangle_k + \pi^2/6 + \mathrm{Var}_\Theta[\varsigma_L] \tag{31}$$

for log-intensity. The first equality in (29) and (31) is only strictly true in the asymptotic (large M) limit. By construction (5), the breakpoints $\theta_0, \ldots, \theta_K$ are such that the variation of mean intensity with angle is minimal within a segment (for a fixed number of segments K). Therefore, the third term in (29) and (31) is small compared to the first term for most segments.

For modeled (or observed) scattering strength with associated variance at N angles, these formulae can be implemented using (6). If n_k is the number of bins in the kth segment, and $\boldsymbol{n} = \begin{bmatrix} n_1 & \ldots & n_K \end{bmatrix}^T$, then

$$\boldsymbol{\mu} = \boldsymbol{\varsigma W} \tag{32}$$

$$M\text{diag}(\boldsymbol{\Sigma}) = 2\left(\boldsymbol{\nu} \star \boldsymbol{\nu}\right)\boldsymbol{W} + \left(\boldsymbol{\varsigma} \star \boldsymbol{\varsigma}\right)\boldsymbol{W} +$$
$$+ \left(\boldsymbol{\varsigma} - \boldsymbol{m}\right) \star \left(\boldsymbol{\varsigma} - \boldsymbol{m}\right)\boldsymbol{W} \tag{33}$$

$$\boldsymbol{m} = \boldsymbol{\mu W}^T \star \boldsymbol{n}, $$

where $\boldsymbol{\varsigma} = \begin{bmatrix} \varsigma_1 & \ldots & \varsigma_N \end{bmatrix}$, $\boldsymbol{\nu} = \begin{bmatrix} \nu_1 & \ldots & \nu_N \end{bmatrix}$, and $\star$ denotes element-wise multiplication. For log-intensity, the formulae are

$$\boldsymbol{\mu} = \boldsymbol{\varsigma}_L \boldsymbol{W} \tag{34}$$

$$M\text{diag}(\boldsymbol{\Sigma}) = \left(\boldsymbol{\nu}_L \star \boldsymbol{\nu}_L\right)\boldsymbol{W} + \frac{\pi^2}{6}$$
$$+ \left(\boldsymbol{\varsigma}_L - \boldsymbol{m}\right) \star \left(\boldsymbol{\varsigma}_L - \boldsymbol{m}\right)\boldsymbol{W} \tag{35}$$

$$\boldsymbol{m} = \boldsymbol{\mu W}^T \star \boldsymbol{n}.$$

In the asymptotic limit when $p(\boldsymbol{d}|\omega_1)$ and $p(\boldsymbol{d}|\omega_2)$ are normal distributions, the Bhatacharyya bound for the probability of error is [5, Ch. 2.8]

$$P(\text{error}) \le \sqrt{P(\omega_1)P(\omega_2)}e^{-\beta}, \tag{36}$$

where

$$\beta = \frac{1}{8}\left(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1\right)^T \left[\frac{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}{2}\right]^{-1}\left(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1\right)$$
$$+ \frac{1}{2}\log\frac{\det\frac{1}{2}\left[\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2\right]}{\sqrt{\det\boldsymbol{\Sigma}_1 \det\boldsymbol{\Sigma}_2}}. \tag{37}$$

The angular segments correspond to separate patches of seabed. We therefore consider an error bound for data with independent features (the UDC assumption holds true). Consequently the covariance matrix is diagonal, with entries given by (29) or (31). Then

$$\beta = \frac{1}{2}\sum_{k=1}^{K}\left[\frac{\left(\langle\varsigma_2\rangle_k - \langle\varsigma_1\rangle_k\right)^2 M/2}{\text{Var}[Y;k]_1 + \text{Var}[Y;k]_2} + \log f\left(\sqrt{\rho_k}\right)\right], \tag{38}$$

with

$$\rho_k = \frac{\text{Var}[Y;k]_1}{\text{Var}[Y;k]_2}, \tag{39}$$

$$f(x) = \frac{1}{2}\left(x + \frac{1}{x}\right). \tag{40}$$

If $x \ge 0$, then $f(x) \ge 1$, with minimum for $x = 1$. Hence both terms inside the brackets are always greater than zero. The second term vanishes when the LDC assumption is correct.

144

When the variance becomes large, the second term dominates, and $P(\text{error})$ tends to a limit determined by the variance ratio ρ_k. The absolute upper limit (with equal prior probabilities) is $P(\text{error}) = 0.5$ (pure guessing). For log-intensity we must replace $\langle \varsigma \rangle_k$ with $\langle \varsigma_L \rangle_k$, and $\text{Var}[Y; k]$ (29) with $\text{Var}[\log Y, k]$ (31).

To compare classifiers based on intensity and log-intensity, we need to compare the means and variances for S and $\log S$ without knowledge of the underlying scattering cross section PDF $f_{S|\Theta}(s|\theta)$. For a given set of constraints, the least biased estimate for $f_{S|\Theta}(s|\theta)$ is the distribution $p(s)$ which maximizes the entropy [44], which for continuous distributions is [5, App. A.7]

$$H = -\int_{-\infty}^{\infty} p(s) \log p(s) \, \mathrm{d}s. \tag{41}$$

If we constrain the mean and variance of $\log S$,

$$\mathrm{E}[\log S|\Theta] = \int \log s f_{S|\Theta}(s|\theta) \, \mathrm{d}s = \varsigma_L(\theta),$$

$$\text{Var}[\log S|\Theta] = \int (\log s - \varsigma_L(\theta))^2 f_{S|\Theta}(s|\theta) \, \mathrm{d}s = \nu_L^2(\theta),$$

then the maximum entropy PDF for S is log-normal. Thus [56],

$$f_{S|\Theta}(s|\theta) = \frac{1}{s\nu_L\sqrt{2\pi}} \exp\left[-\frac{(\log s - \varsigma_L)^2}{2\nu_L^2}\right], \tag{42}$$

for which

$$\varsigma = e^{\varsigma_L + \nu_L^2/2}, \tag{43}$$

$$\nu^2 = \left(e^{\nu_L^2} - 1\right) e^{2\varsigma_L + \nu_L^2}. \tag{44}$$

Hence, given the scattering strength $\varsigma_L(\theta)$ from a physics model or observation, we have the necessary quantities to compare the Bhatacharyya bound (36) for Y and $\log Y$, parameterized by the variance $\nu_L^2(\theta)$ and the number M. These quantities are given by (28)–(31), (38)–(40), and (43)–(44).

REFERENCES

[1] J. T. Anderson, D. Van Holliday, R. Kloser, D. G. Reid, and Y. Simard, "Acoustic seabed classification: current practice and future directions," *ICES Journal of Marine Science*, vol. 65, no. 6, pp. 1004–1011, Sep. 2008.

[2] C. J. Brown and P. Blondel, "Developments in the application of multibeam sonar backscatter for seafloor habitat mapping," *Applied Acoustics*, vol. 70, no. 10, pp. 1242–1247, Oct. 2009.

[3] R. J. Kloser, J. D. Penrose, and A. J. Butler, "Multi-beam backscatter measurements used to infer seabed habitats," *Continental Shelf Research*, vol. 30, no. 16, pp. 1772–1782, Sep. 2010.

[4] A. Rattray, D. Ierodiaconou, L. Laurenson, S. Burq, and M. Reston, "Hydro-acoustic remote sensing of benthic biological communities on the shallow South East Australian continental shelf," *Estuarine Coastal and Shelf Science*, vol. 84, no. 2, pp. 237–245, Sep. 2009.

[5] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York, NY: Wiley-Interscience, 2001.

[6] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning. Data Mining, Inference and Prediction*, 2nd ed., ser. Springer Series in Statistics. New York, NY: Springer, 2009.

[7] X. Lurton, *An Introduction to Underwater Acoustics: Principles and Applications*, 2nd ed. Berlin/DE: Springer, 2010.

[8] D. R. Jackson and M. D. Richardson, *High-Frequency Seafloor Acoustics*. New York, NY: Springer, 2007.

[9] S. T. McDaniel, "Fluctuations of the acoustic field Bragg-backscattered from rough surfaces," *Waves in Random Media*, vol. 6, no. 3, pp. 229–249, 1996.

[10] G. Le Chenadec, J. Boucher, and X. Lurton, "Angular dependence of K-distributed sonar data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1224–1235, May 2007.

[11] A. P. Lyons, S. F. Johnson, D. A. Abraham, and E. Pouliquen, "High-frequency scattered envelope statistics of patchy seafloors," *IEEE J. Ocean. Eng.*, vol. 34, no. 4, pp. 451–458, Oct. 2009.

[12] L. Fonseca and L. Mayer, "Remote estimation of surficial seafloor properties through the application angular range analysis to multibeam sonar data," *Mar. Geophys. Res.*, vol. 28, no. 2, pp. 119–126, Jul. 2007.

[13] K. Haris, B. Chakraborty, C. De, R. G. Prabhudesai, and W. Fernandes, "Model-based seafloor characterization employing multi-beam angular backscatter data-a comparative study with dual-frequency single beam," *J. Acoust. Soc. Am.*, vol. 130, no. 6, pp. 3623–3632, Dec. 2011.

[14] B. Chakraborty, H. W. Schenke, V. Kodagali, and R. Hagen, "Seabottom characterization using multibeam echosounder angular backscatter: An application of the composite roughness theory," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2419–2422, Sep. 2000.

[15] C. De and B. Chakraborty, "Model-based acoustic remote sensing of seafloor characteristics," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3868–3877, Oct. 2011.

[16] G. Canepa and N. G. Pace, "Seafloor segmentation from multibeam bathymetric sonar," in *Proceedings of the Fifth European Conference on Underwater Acoustics, ECUA 2000*, M. E. Zakharia, P. Chevret, and P. Dubail, Eds., vol. 1. Lyon, France: European Commision, Jul. 2000, pp. 361–367.

[17] Z.-H. Michalopoulou, D. Alexandrou, and C. de Moustier, "Application of neural and statistical classifiers to the problem of seafloor characterization," *IEEE J. Ocean. Eng.*, vol. 20, no. 3, pp. 190–197, Jul. 1995.

[18] L. J. Hamilton and I. Parnum, "Acoustic seabed segmentation from direct statistical clustering of entire multibeam sonar backscatter curves," *Continental Shelf Research*, vol. 31, no. 2, pp. 138–148, Feb. 2011.

[19] B. Chakraborty, V. Kodagali, and J. Baracho, "Sea-floor classification using multibeam echo-sounding angular backscatter data: A real-time approach employing hybrid neural network architecture," *IEEE J. Ocean. Eng.*, vol. 28, no. 1, pp. 121–128, Jan. 2003.

[20] S. Dugelay, "Deep seafloor characterization with multibeam echosounders using image segmentation and angular acoustic variations," in *OCEANS '96. MTS/IEEE. Prospects for the 21st Century. Conference Proceedings*, vol. 3. Ft Lauderdale, FL: IEEE, Sep. 1996, pp. 1545–1550.

[21] D. G. Simons and M. Snellen, "A Bayesian approach to seafloor classification using multi-beam echosounder backscatter data," *Applied Acoustics*, vol. 70, no. 10, pp. 1258–1268, Oct. 2009.

[22] K. Landmark, A. H. S. Solberg, C.-I. C. Nilsen, A. Austeng, and R. E. Hansen, "Bayesian seabed classification using angle-dependent backscatter data from multibeam sonar," in *11th European Conference on Underwater Acoustics 2012 (ECUA 2012)*, ser. Proceedings of the Institute of Acoustics Volume 34 Pt.3. Institute of Acoustics (IOA), Oct. 2012, pp. 417–424.

[23] A. C. Jensen and A. S. Solberg, "Fast hyperspectral feature reduction using piecewise constant function approximations," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 4, pp. 547–551, Oct. 2007.

[24] G. Winkler and V. Liebscher, "Smoothers for discontinuous signals," *J. Nonparametric Statist.*, vol. 14, no. 1/2, pp. 203–222, 2002.

[25] E. J. Eidem and K. Landmark, "Acoustic seabed classification using QTC IMPACT on single-beam echo sounder data from the Norwegian Channel, northern North Sea," *Continental Shelf Research*, vol. 68, pp. 1–14, Oct. 2013.

[26] H. P. Seirup *et al.*, "Quaternary of the Norwegian Channel: glaciation history and paleooceanography," *Norsk geologisk tidsskrift*, vol. 75, no. 2-3, pp. 65–87, 1995.

[27] H. de Haas, E. Okkels, and T. C. E. van Weering, "Recent sedimentation and organic carbon burial in a shelf area: the North Sea," *Norges Geologiske Undersøkelse Bulletin*, vol. 430, pp. 57–65, 1996.

[28] E. S. Andersen, S. R. Østmo, C. F. Forsberg, and S. J. Lehman, "Late- and post-glacial depositional environments in the Norwegian Trench, northern North Sea," *Boreas*, vol. 24, pp. 48–63, 1995.

[29] L. Rise and K. Rokoengen, "Surficial sediments in the Norwegian sector of the North Sea between $60°30'$ and $62°N$," *Marine Geology*, vol. 58, pp. 287–317, 1984.

[30] M. Hovland, "Elongated depressions associated with pockmarks in the western slope of the Norwegian Trench," *Marine Geology*, vol. 51, no. 1-2, pp. 35–46, Feb. 1983.

[31] R. Bøe, L. Rise, and D. Ottesen, "Elongate depressions on the southern slope of the Norwegian Trench (Skagerrak): morphology and evolution," *Marine Geology*, vol. 146, no. 1-4, pp. 191–203, Apr. 1998.

[32] R. L. Folk, "The distinction between grain size and mineral composition in sedimentary-rock nomenclature," *J. Geol.*, vol. 62, no. 4, pp. 344–359, 1954.

[33] J. M. Preston, A. C. Christney, L. S. Beran, and W. T. Collins, "Statistical seabed segmentation—from images and echoes to objective clustering," in *Proc. Seventh European Conf. Underwater Acoustics ECUA 2004*. Delft, The Netherlands: Delft University of Technology, Jul. 2004, pp. 813–818.

[34] E. Hammerstad, "Backscattering and seabed image reflectivity," Kongsberg Maritime, EM technical note, Jan. 2000. [Online]. Available: http://www.km.kongsberg.com/ks/web/nokbg0397.nsf/AllWeb/C2AE0703809C1FA5C1257B580044DD83/$file/EM_technical_note_web_BackscatteringSeabedImageReflectivity.pdf?OpenElement

[35] R. Eckhoff, "Back scatter corrections," by personal communication, Kongsberg Maritime, Aug. 2009.

[36] D. R. Jackson and M. D. Richardson, *High-Frequency Seafloor Acoustics*. New York, NY: Springer, 2007.

[37] A. N. Gavrilov and I. M. Parnum, "Fluctuations of seafloor backscatter data from multibeam sonar systems," *IEEE J. Ocean. Eng.*, vol. 35, no. 2, pp. 209–219, Apr. 2010.

[38] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics*, 2nd ed., ser. Modern Acoustics & Signal Processing. New York, NY/US: Springer, 2011.

[39] R. E. Francois and G. R. Garrison, "Sound absorption based on ocean measurements. Part 2: Boric acid contribution and equation for total absorption," *J. Acoust. Soc. Amer.*, vol. 72, no. 6, pp. 1879–1890, 1982.

[40] A. K. Jain and B. Chandrasekaran, "Dimensionality and sample size considerations in pattern recognition practice," in *Handbook of statistics*, P. R. Krishnaiah and L. N. Kanal, Eds. Amsterdam: North-Holland, 1982, vol. 2, pp. 835–855.

[41] R. P. W. Duin *et al.* (2007) PRTools4 A Matlab Toolbox for Pattern Recognition. [Online]. Available: http://www.prtools.org/

[42] P. Pudil, J. Novovicova, and J. Kittler, "Floating search methods in feature selection," *Pattern Recogn. Lett.*, vol. 15, no. 11, pp. 1119–1125, Nov. 1994.

[43] L. Breiman, *Probability*, ser. Classics in applied mathematics, vol. 7. Philadelphia, PA: SIAM, 1992.

[44] E. T. Jaynes, "Information theory and statistical mechanics," *Phys. Rev.*, vol. 106, no. 4, pp. 620–630, May 1957.

[45] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, 2nd ed. Englewood Cliffs, NJ: Prentice Hall, 1988.

[46] P. J. Rousseeuw and K. V. Driessen, "A fast algorithm for the minimum covariance determinant estimator," *Technometrics*, vol. 41, no. 3, pp. 212–223, Aug. 1999.

[47] S. Verboven and M. Hubert, "LIBRA: a MATLAB library for robust analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 75, no. 2, pp. 127–136, Feb. 2005.

[48] J. Preston, "Automated acoustic seabed classification of multibeam images of Stanton Banks," *Applied Acoustics*, vol. 70, no. 10, pp. 1277–1287, Oct. 2009.

[49] T. B. Reed and D. Hussong, "Digital image processing techniques for enhancement and classification of SeaMARC II side scan sonar imagery," *J. Geophys. Res.*, vol. 94, no. B6, pp. 7469–7490, Jun. 1989.

[50] P. Blondel and O. Gómez Sichi, "Textural analyses of multibeam sonar imagery from Stanton Banks, Northern Ireland continental shelf," *Applied Acoustics*, vol. 70, no. 10, pp. 1288–1297, Oct. 2009.

[51] V. A. I. Huvenne, P. Blondel, and J.-P. Henriet, "Textural analyses of sidescan sonar imagery from two mound provinces in the Porcupine Seabight," *Marine Geology*, vol. 189, no. 3-4, pp. 323–341, Sep. 2002.

[52] F. Albregtsen. (2008, Nov.) Statistical Texture Measures Computed from Gray Level Cooccurence Matrices. Dept. Informatics, Univ. Oslo. [Online]. Available: http://www.uio.no/studier/emner/matnat/ifi/INF4300/h08/undervisningsmateriale/glcm.pdf

[53] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 4, pp. 610–621, Nov. 1973.

[54] L. Fonseca, C. Brown, B. Calder, L. Mayer, and Y. Rzhanov, "Angular range analysis of acoustic themes from Stanton Banks Ireland: A link between visual interpretation and multibeam echosounder angular signatures," *Applied Acoustics*, vol. 70, no. 10, pp. 1298–1304, Oct. 2009.

[55] K. D. Ward, "Compound representation of high resolution sea clutter," *Electron. Lett.*, vol. 17, no. 16, pp. 561–563, Aug. 1981.

[56] NIST/SEMATECH e-Handbook of Statistical Methods. [Last accessed: Oct. 15 2012]. National Institute of Standards and Technology. [Online]. Available: http://www.itl.nist.gov/div898/handbook/

[57] P. Wessel and W. H. F. Smith, "New, improved version of Generic Mapping Tools released," *EOS Trans. Amer. Geophys. U.*, vol. 79, no. 47, p. 579, 1998.

Paper III:

A Radon-transform-based image noise filter — with applications to multibeam bathymetry

A Radon-Transform-Based Image Noise Filter— with Applications to Multibeam Bathymetry

Knut Landmark, Anne H. Schistad Solberg, Fritz Albregtsen,
Andreas Austeng, and Roy Edgar Hansen

Abstract

This paper describes a linear-image-transform-based algorithm for reducing stripe noise, track line artifacts, and motion-induced errors in remote sensing data. Developed for multibeam bathymetry (MB), the method has also been used for removing scalloping in synthetic aperture radar images. The proposed image transform is the composition of an invertible edge detection operator and a fast discrete Radon transform (DRT) due to Götz, Druckmüller, and Brady. The inverse DRT is computed by using an iterative method and exploiting an approximate inverse algorithm due to Press. The edge operator is implemented by circular convolution with a Laplacian point spread function modified to render the operator invertible. In the transformed image, linear discontinuities appear as high-intensity spots which may be reset to zero. In MB data, a second noise signature is linked to motion-induced errors. A Chebyshev approximation of the original image is subtracted before applying the transform, and added back to the denoised image; this is necessary to avoid boundary effects. It is possible to process data faster and suppress motion-induced noise further by filtering images in nonoverlapping blocks using a matrix representation for the inverse DRT. Processed test images from several MB data sets had less noise and distortion compared to those obtained with standard low-pass filters. Denoising also improved the accuracy in statistical classification of geomorphological type by 10–28% for two sets of invariant terrain features.

Index Terms

Discrete transforms, image denoising, image restoration, iterative methods, remote sensing, sonar, terrain mapping.

I. INTRODUCTION

Stripe noise, track line artifacts, antenna pattern effects or motion-induced errors degrade many types of remote sensing imagery. For example, satellite images from scanning mirror systems are often affected by periodic stripes in the across-track direction [1], [2]. The primary causes of these artifacts are offsets in detector response between the forward and reverse scans, unequal response of adjacent detectors, or random fluctuations in sensor response. Detector-to-detector variations also occur in pushbroom-type systems, where a linear array is oriented perpendicular to the flight direction and records one line of pixels simultaneously. With this configuration, stripe artefacts appear in the along-track direction and are not periodic [3]. Any kind of raster data obtained from a moving platform, where the pixels are recorded line by line, geolocated and then collated together, are prone to have similar errors to some degree, particularly if the image contains data from several adjacent survey track lines.

The multibeam echo sounder (MBES) provides a good example. By electronic beamforming of acoustical arrays, an MBES measures the echo travel time and seabed reflectivity in narrow beams fanning out perpendicular to the survey vessel's track [4, Ch. 8]. Subsequent data processing produces a high-resolution digital terrain model (DTM) of the seabed and a sidescan image mosaic of seabed scattering strength. MBES technology has had a significant impact on the marine sciences; in particular, it yields much information about seabed geological processes (erosion and deposition) by mapping morphological features such as sedimentary bedforms [5]. In addition to sidescan imagery, important tools in this respect are attribute images derived from the DTM, particularly gradient, curvature or shaded relief [6]. Not only are such images powerful interpretative tools, there is also promise that terrain attributes can be used in automated statistical classification of terrain type [7] and, ultimately, sediment type [8], [9].

The origin of this work was an attempt to use terrain attributes to improve the accuracy of acoustical classification of sediments, and the realization that, unfortunately, many attributes are sensitive to slight errors or artifacts in the data. One type of noise is caused by motion sensor errors, e.g., time delays in the heave, roll, and pitch (HRP) measurements. As the vessel moves through swell, even slight errors in HRP compensation will cause artificial wavy patterns ("ribbing") in the bathymetry. Often the patterns are near-periodical and oriented perpendicular to the vessel track. A detailed investigation of this phenomenon can be found in [10]. Another source of error is the mismatch between adjacent survey lines. A vessel typically moves back and forth in a parallel "lawnmower pattern", and in open seas a single pass may take several hours, during which time the sea level changes. Even if a good tidal model is available, a residual mismatch between survey lines may persist. Sound speed profiles (SSPs) are obtained from conductivity, temperature and depth (CTD) probe casts at extended intervals (unless a towed moving vessel profiler is used). Using an inaccurate SSP for computing sound refraction leads to depth estimation errors, and such errors are most pronounced at the outer beams where adjacent swathes overlap [5]. These types of mismatch cause stripes in attribute images. In addition, many sonar systems estimate the bottom echo by peak amplitude detection in the near-nadir beams, and by interferometric methods (phase matching) at higher incidence angles. At the transition angle between the two modes, a systematic depth offset may occur, which also causes stripe noise. Finally, beam calibration errors cause stripes in sidescan data [11], a phenomenom which is analogous to the one observed in pushbroom optical systems.

Many of the methods for destriping remote sensing images have been developed for

satellite systems, and all are not directly applicable to multibeam bathymetry. The spatial filters developed in [12]–[14] are 1-D and intended for images with horizontal scan-line noise. The statistical methods of equalization [15], [16], histogram matching [17], [18], and moment matching [19] adjust each detector output to match the statistics of some reference distribution, provided each detector can be identified with a line of pixels (horizontal or vertical) in the image. To remove irregular stripes, i.e., lines where only a subset of the pixels are noise-contaminated, and nonlinear stripes where the degradation parameters change across the scene, Shen et al. [20] has proposed a method where lines are split into segments that are processed separately with a statistical destriping algorithm. It may be that these methods can be adapted to sidescan imagery (before georectification), but the artifacts to be considered here are not caused by intrinsic variations in detector response. Likewise, the imaging modes or assumptions are different from ours in the recent statistical and variational approaches reported in [1]–[3], [21], and [22].

A more general approach is to use an edge detection algorithm to find the pixels affected by stripe noise and inpainting techniques to correct them. Tsai and Chen [23] used the Canny algorithm [24] and a spline-based inpainting scheme applicable to thin horizontal or vertical stripes. Image filtering in the Fourier domain is another option when the noise is periodic or at least directive [25, Ch. 4]. This was explored in [26] for side-scan sonar, and in, for example, [27] for satellite imagery. Fourier domain filtering does not require a statistical model of the image formation process, and can be applied to both georectified and raw images, but it may cause unwanted smoothing or distortion of real features. The same is true of filtering in the wavelet domain. To minimize the smoothing effect in images with horizontal stripes, Chen et al. [27] isolated the noise frequency by averaging the 1-D power spectra from each image column, as the noise frequency tends to be constant across columns. However, the noise of primary interest in this paper has a broad spectrum.

This paper presents an alternative image-transform-based denoising filter which is motivated by the problems affecting MBES bathymetry and derived attributes. The method has also proven effective for reducing scalloping in synthetic aperture radar (SAR) images, but we concentrate here on MBES data. The forward image transform is an edge detection operator composed with a discrete Radon transform (DRT). Noise suppression takes place in the DRT domain. There are no assumptions about periodicity or orientation, and artifacts need not be parallel. A linear artifact is focused to an isolated high-intensity spot in the Radon domain. For MBES data, it is also conjectured that another noise signature is linked with motion-induced errors in the sounding data. The proposed algorithm is applied to two MBES datasets from the North Sea, and compared qualitatively with Fourier domain filters and spatial low-pass filters. To quantify the effect of the denoising algorithm, we have considered the machine learning problem of classifying different terrain morphologies, and how denoising affects the classification accuracy. For this problem we employ two sets of invariant terrain features based on the empirical distribution of local binary patterns (LBPs) and intrinsic geometric features.

A major part of this paper concerns the forward and inverse two-step image transform. One candidate for the edge operator is the Laplace operator (Laplacian), which is linear, approximately rotation invariant, edge-sensitive, and noise-sensitive. The Laplacian is also a measure of curvature, which is a recommended feature for visualizing terrain [6]. However, the Laplacian is not one-to-one, and we therefore suggest a modification to obtain an invertible edge operator. Basing the algorithm on discrete approximations of the classical forward

and inverse Radon transform formulae did not yield satisfactory results in our experiments. Filtered backprojection [28], [29], i.e., the standard fast method for inverting the linear Radon transform, causes a degree of smoothing. Discrete approximations may involve interpolation, and Beylkin [30] moreover has shown that, in a discrete setting, inversion of the Radon transform can only be approximate when it is based on a discretization of Radon's classical inversion formula. Here we consider instead the DRT of Götz and Druckmüller [31] and Brady [32] (GDB), and compute the inverse using linear algebra. In this formulation, the forward and inverse Radon transforms are discrete linear operators, where no interpolation of data is involved. Press [33] found a fast iterative inverse algorithm that can reconstruct images with arbitrary precision from the forward GDB transform data. However, a numerical analysis of the Press inverse suggests that the algorithm converges only for images of small-to-moderate size. We therefore also consider two alternative approaches to invert the GDB transform, one of which is a modification of Press' method.

Thus, this paper has four subthemes: 1) inversion of the GDB (DRT) transform; 2) construction of a linear invertible edge detection operator; 3) noise signature reduction in the GDB transform domain; and finally, 4) the effect of noise reduction on the classification of geomorphology by invariant terrain features. Section II concerns the Radon transform and its inverse, in particular the GDB algorithm (see Section II-B) and the use of iterative inverse methods (see Section II-C). Section III describes the full denoising algorithm, including the invertible Laplace operator (see Section III-A), and the suppression of noise signatures in the DRT domain (see Section III-B). Section IV describes the formation of invariant, unbiased features for terrain classification. Section V presents the MBES data (see Section V-A), examples of denoised images (see Section V-B), and the results of the classification experiment (see Section V-C). The summary and discussion (see Section VI) includes considerations on parameter choices and generalizability, elaborates further on the potential of block processing, and mentions other possible applications.

II. THE RADON TRANSFORM

A. Classical Transform and PSF

The classical Radon transform of a function $f : \mathbb{R}^2 \to \mathbb{R}$ is formed by integrating f along straight lines [29]. If $s \in \mathbb{R}$ is a signed distance from the origin, and $\boldsymbol{\theta} = \begin{bmatrix} \cos\theta & \sin\theta \end{bmatrix}$ is a unit vector, then the straight line orthogonal to $\boldsymbol{\theta}$ and parameterized by (θ, s) is $L(\theta, s) = \left\{ \boldsymbol{x} \in \mathbb{R}^2 \mid \boldsymbol{x} \cdot \boldsymbol{\theta} = s \right\}$. The Radon transform $\mathcal{R}f$ is a function on the cylinder $C = [0, 2\pi) \times \mathbb{R}$ defined by

$$(\mathcal{R}f)(\theta, s) = \int_{L(\theta,s)} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}. \tag{1}$$

The adjoint, or backprojection operator, applied to a function $g : C \to \mathbb{R}$, is defined by

$$(\mathcal{R}^+ g)(\boldsymbol{x}) = \int_0^{2\pi} g(\theta, \boldsymbol{x} \cdot \boldsymbol{\theta}) \, \mathrm{d}\theta, \tag{2}$$

which is 2π times the average of all line integrals passing through $\boldsymbol{x}$. The standard filtered backprojection algorithm for reconstructing f from $\mathcal{R}f$ is based on the fact that

$$(\mathcal{R}^+ g) * f = \mathcal{R}^+(g * \mathcal{R}f), \tag{3}$$

where $*$ denotes convolution [29]. One finds a function g such that $V \equiv \mathcal{R}^+ g \approx \delta$ is the Dirac delta function. Then, by (3), the reconstructed function is $\mathcal{R}^+(g * \mathcal{R}f)$, the backprojection
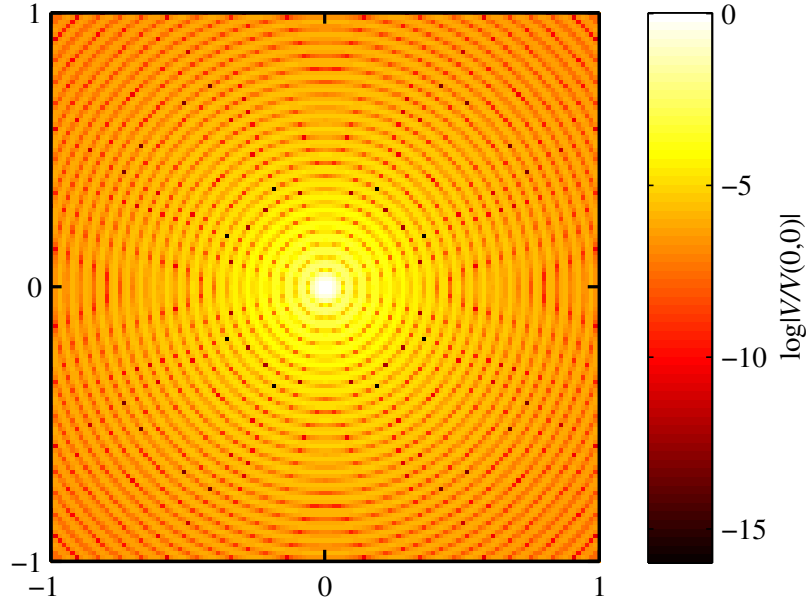
Fig. 1. PSF for filtered backprojection with a Shepp-Logan window with bandwidth $\Omega = 64$ cycles per unit length (128×128 pixels image). The number of projections is $p = 181$, which exactly meets the angular sampling criterion for this bandwidth [29].

of $\mathcal{R}f$ filtered by g; V is the point spread function (PSF), which determines how much the function f is blurred when passed through Radon transform and filtered backprojection.

For reconstructing bandlimited functions, as when f represents a digital image, g must be chosen such that the Fourier transform $\hat{V}(k) = \hat{\delta}(k) = 1/\sqrt{2\pi}$ for $|k| \leq \Omega$, i.e., the bandwidth of f. Several functions g that yield an appropriate window function for $\hat{V}(k)$ have been proposed and applied in practice [34]. One example is the Shepp-Logan window [35], for which the corresponding PSF $V(\boldsymbol{x})$ is plotted in Fig. 1 for the bandwidth $\Omega = 64$. A certain smoothing or distortion of the original function is implied by this PSF. Moreover, when applied to digital images (regular, discrete data), there will be an additional loss of accuracy from the fact that line integrals must be evaluated by interpolation between nearby pixels, or the (weighted) sum of all pixel values in a band containing the line.

B. GDB Discrete Transform

While filtered backprojection is a cornerstone of image reconstruction as in tomographic imaging, efficient evaluation of the forward Radon transform is important in its own right. For example, the classical Hough transform, used for straight line detection in image analysis [25], [36], may be implemented as a Radon transform when applied to binary images. A fast approximate method for computing line integrals was described independently in [31] and [32]. This algorithm does not involve weighted sums. Instead, each DRT element is the sum of pixel values along a certain graph which includes exactly one pixel from each column of the zero-padded image. Here $\boldsymbol{f}$ denotes the discrete set of pixel values $\boldsymbol{f} = \{ f_{i,j} \}$ indexed by row i and column j, where $f_{i,j} = 0$ for i, j outside the range $0 \leq i, j < N - 1$, and it is assumed that N is a power of 2. A graph $D_N(h, s)$ (using notation from [33]) is identified by its intercept h and rise s, meaning that it starts at $(i, j) = (h, 0)$ and ends at $(i, j) = (h + s, N - 1)$ (see Fig. 2). The construction of the graph recursive. In the first step, the image is divided into $N/2$ subimages, each two pixels wide, for which the lines $D_2(h, s)$
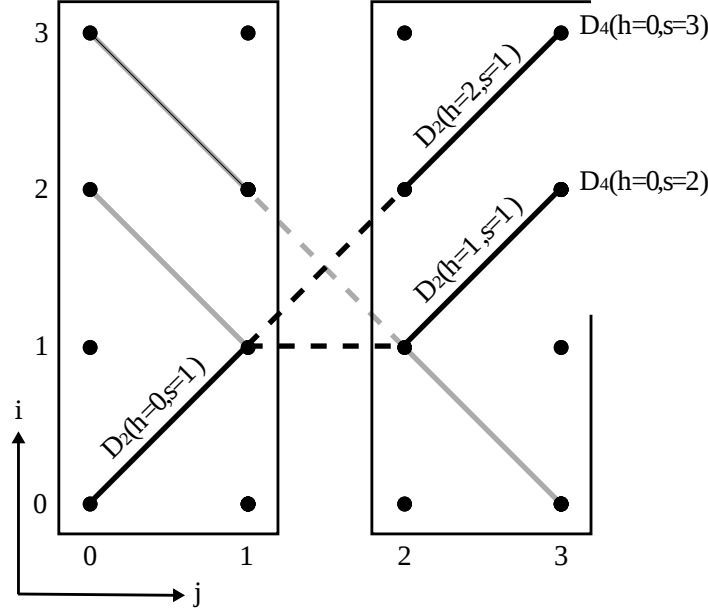
Fig. 2. Recursive construction of graphs in a 4×4 pixels image. The graphs are formed by joining graphs in the two 4×2 pixels subimages according to rule (4). The rise parameter s ranges from 0 to 1 (4×2 subimages) and 0 to 3 (4×4 image), yielding slopes from $0°$ to $45°$. When the image is flipped upside down, the graphs are mapped to lines with slopes in the range $-45°$ to $0°$ (faint lines).

are constructed. In the next step, graphs from pairs of adjacent subimages are combined, and so on, for $\log_2 N$ steps in all. Graphs from adjacent images are combined according to the following rule:

$$D_n(h, 2s) = D_{n/2}^{(L)}(h, s) \cup D_{n/2}^{(R)}(h + s, s) \tag{4a}$$

$$D_n(h, 2s + 1) = D_{n/2}^{(L)}(h, s) \cup D_{n/2}^{(R)}(h + s + 1, s) \tag{4b}$$

where it is understood that the right-hand side of a union refers to the right (R) half of the image, and correspondingly for the left-hand (L) side. A graph in a subimage therefore forms part of two graphs in the twice-as-wide subimage (see Fig. 2).

Equation (4) is applied for $0 \leq s < N$, i.e., for slopes from $0°$ to $45°$. The other line integrals are performed in the same way after first transposing or flipping the original image. The full DRT is the disjoint union of four parts (quadrants);

$$(\mathcal{R}^1 f)(h, s) = \sum_{(i,j) \in D_N(h,s)} f_{i,j} \qquad (0° \text{ to } 45°) \tag{5a}$$

$$(\mathcal{R}^2 f)(h, s) = \sum_{(i,j) \in D_N(h,s)} f_{j,i} \qquad (45° \text{ to } 90°) \tag{5b}$$

$$(\mathcal{R}^3 f)(h, s) = \sum_{(i,j) \in D_N(h,s)} f_{j,N-1-i} \qquad (-90° \text{ to } -45°) \tag{5c}$$

$$(\mathcal{R}^4 f)(h, s) = \sum_{(i,j) \in D_N(h,s)} f_{N-1-i,j} \qquad (-45° \text{ to } 0°) \tag{5d}$$

The number of non-zero elements in each quadrant is at most $N^2 + N(N - 1)/2$, and the minimum value for h is $-N + 1$, for which the associated graph intersects the image only at a corner pixel.

C. Iterative Inverses

The GDB transform (DRT) $\mathcal{R}$ is a linear operator; each output element is a sum of image pixel values. In principle $\mathcal{R}$ may therefore be represented by a $P \times Q$ matrix $\boldsymbol{R}$, with $P = 6N^2 - 2N$ and $Q = N^2$. The matrix $\boldsymbol{R}$ is only moderately sparse, and a left-inverse of $\boldsymbol{R}$ is not sparse. The inverse problem is to solve the linear system

$$\boldsymbol{R}\boldsymbol{f} = \boldsymbol{d} \tag{6}$$

with respect to the image $\boldsymbol{f} \in \mathbb{R}^Q$, when the data vector $\boldsymbol{d} \in \mathbb{R}^P$, containing the line sums, is known. (Where appropriate we will use a vector to store a two-dimensional image, by concatenating the columns into a single column of length N^2.)

1) Normal Equations: The GDB transform has a corresponding fast recursive backprojection algorithm, which, analogously to (2), sums the values of all line integrals passing through a given pixel [32]. As shown in Appendix B, the matrix representation of the GDB backprojection algorithm is precisely the transpose $\boldsymbol{R}^T$. Therefore one may consider the normal equations, i.e.,

$$\boldsymbol{R}^T \boldsymbol{R}\boldsymbol{f} = \boldsymbol{R}^T \boldsymbol{d}, \tag{7}$$

also when it is infeasible to use the explicit matrix representations. Several iterative algorithms that solve the normal equations (7) are available, e.g. the LSQR [37] and LSMR [38] algorithms.

In the proposed denoising method (see Section III-B), the transform $\boldsymbol{d}_0 = \boldsymbol{R}\boldsymbol{f}_0$ of the original image $\boldsymbol{f}_0$ is modified by a filter in the DRT domain to obtain some perturbation $\boldsymbol{d} = \boldsymbol{d}_0 + \Delta\boldsymbol{d}_0$. The vector $\boldsymbol{d}$ is generally not contained in the range $\mathcal{I}(\boldsymbol{R})$ of $\boldsymbol{R}$, which implies that (6) does not have a solution. We have the following range-null space decomposition of $\mathbb{R}^P$ (see Appendix B),

$$\mathbb{R}^P = \mathcal{I}(\boldsymbol{R}) \oplus \mathcal{N}(\boldsymbol{R}^T) \tag{8}$$

where $\oplus$ denotes the direct sum, and the range $\mathcal{I}(\boldsymbol{R})$ and the null space $\mathcal{N}(\boldsymbol{R}^T)$ are orthogonal subspaces of $\mathbb{R}^P$. A solution to (7) minimizes the residual norm $\|\boldsymbol{d} - \boldsymbol{R}\boldsymbol{f}\|$ [39, Ch. 6.9], and $\boldsymbol{d} - \boldsymbol{R}\boldsymbol{f}$ lies in the null space $\mathcal{N}(\boldsymbol{R}^T)$ of the backprojection operator.

Hence, (7) yields an optimal solution in the sense that $\|\boldsymbol{d} - \boldsymbol{R}\boldsymbol{f}\|$ is minimized. However, one should bear in mind that there is no definite best choice for $\Delta\boldsymbol{d}_0$. Rather, what matters is the quality of the restored image and the computational cost. For these reasons, we consider two alternative approaches that in experiments yielded better image reconstructions at less computational cost, and which converge faster when (6) does have a solution.

2) Press Inverse: Press [33] devised an approximate inverse for the GDB transform, based on the backprojection algorithm. The approximate inverse algorithm is a linear operator $\mathcal{B}$ that is applied recursively on downsampled versions of $\mathcal{R}^i\boldsymbol{f}$ ($i = 1, \ldots 4$), to produce a smoothed image $\boldsymbol{f}_s$. The residual $\mathcal{R}\boldsymbol{f} - \mathcal{R}\boldsymbol{f}_s$ is backprojected and high-pass filtered to obtain an image correction $\delta\boldsymbol{f}$ which is added to the smoothed image $\boldsymbol{f}_s$. Since $\mathcal{B}$ is linear, it has a matrix representation $\boldsymbol{B}$.

Suppose (6) has a solution, i.e., $\boldsymbol{d} \in \mathcal{I}(\boldsymbol{R})$. We can show that $\boldsymbol{R}$ is injective (one-to-one), therefore, the solution must be unique (see Appendix A). Define $\boldsymbol{E} = \mathbf{1}_Q - \boldsymbol{B}\boldsymbol{R}$, and let $\rho(\boldsymbol{E})$

(a) Original image   (b) Approximate inverse   (c) Three iterations (iterative inverse)

Fig. 3. Press approximate inverse (b), equation (9a), and the Press iterative inverse (c), equation (9b) with $k = 2$, applied to the GDB transform of the test image in (a) (512×512 pixels). The original image (a) has gray level intensities in the range 0–1; the approximate inverse has intensities in approximately the range 0–1.4, but has been rescaled so that white corresponds to maximum intensity in all three images. The RMS difference between (a) and (b), and (a) and (c), can be read off the graph labeled "Press (camera)" in Fig. 4 below.

denote the spectral radius of $\boldsymbol{E}$ (the largest eigenvalue magnitude). The following recurrence relation developed in [33] converges to the solution of (6)

$$\boldsymbol{f}_0 = \boldsymbol{B}\boldsymbol{d}, \tag{9a}$$

$$\boldsymbol{f}_{k+1} \equiv \left(\sum_{j=0}^{k+1} \boldsymbol{E}^j \right) \boldsymbol{B}\boldsymbol{d} = \boldsymbol{f}_k + \boldsymbol{B}\left(\boldsymbol{d} - \boldsymbol{R}\boldsymbol{f}_k \right) \tag{9b}$$

provided $\rho(\boldsymbol{E}) < 1$. This condition implies that the $N^2 \times N^2$ matrix $\boldsymbol{B}\boldsymbol{R}$ is non-singular [40, Ch. 7.10]. Fig. 3 shows the Press approximate (9a) and iterative (9b) inverses applied to the GDB transform of a standard test image. The root-mean-square (RMS) error after three iterations is about 1%.

3) GMRES: However, the analysis in Appendix C suggests that, for $N \geq 512$, a small fraction of the eigenvalues of $\boldsymbol{E}$ have magnitudes larger than unity. This implies that the convergence of (9) halts after a certain number of iterations and the residual error begins to grow as some eigenimages get amplified. Nevertheless, the initial rate of convergence is typically very good, and we would therefore like to make use of the approximate inverse $\boldsymbol{B}$.

Applying the binomial formula to the Neumann series appearing in (9b) gives

$$\boldsymbol{f}_k = \sum_{j=0}^{k} \sum_{i=0}^{j} (-1)^i \binom{j}{i} \boldsymbol{A}^i \boldsymbol{b}, \tag{10}$$

where we have defined $\boldsymbol{A} \equiv \boldsymbol{B}\boldsymbol{R}$ and $\boldsymbol{b} \equiv \boldsymbol{B}\boldsymbol{d}$. Consequently $\boldsymbol{f}_k$ lies in the Krylov subspace [41]

$$\mathcal{K}_{k+1}(\boldsymbol{A}, \boldsymbol{b}) \equiv \mathrm{span}\left\{ \boldsymbol{b}, \boldsymbol{A}\boldsymbol{b}, \ldots, \boldsymbol{A}^k \boldsymbol{b} \right\}. \tag{11}$$

The generalized minimum residual (GMRES) algorithm [42], [43] produces a sequence of iterates $\boldsymbol{x}_k$, where $\boldsymbol{x}_k$ is the solution to the least squares problem

$$\min_{\boldsymbol{x} \in \mathcal{K}_k(\boldsymbol{A}, \boldsymbol{b})} \| \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x} \|, \; k = 1, 2, \ldots \tag{12}$$

in the Euclidean norm. Hence, GMRES gives approximate solutions to the $N^2 \times N^2$ linear system obtained by left-multiplying (6) by $\boldsymbol{B}$,

$$\boldsymbol{B}\boldsymbol{R}\boldsymbol{f} = \boldsymbol{B}\boldsymbol{d}. \tag{13}$$

GMRES, like (9), involves only matrix-vector multiplications, which are implemented by the recursive algorithms $\mathcal{R}$ and $\mathcal{B}$.

The exact solution to (13), if $\boldsymbol{A}$ is nonsingular, lies in the Krylov space $\mathcal{K}_d(\boldsymbol{A}, \boldsymbol{b})$, where d is the degree of the minimal polynomial of $\boldsymbol{A}$ [41]. Note that we have only established that $\boldsymbol{A}$ is nonsingular for $N < 512$. However, $\rho(\boldsymbol{E}) < 1$ is not a necessary condition for $\boldsymbol{A}^{-1}$ to exist, it is only required that $\lambda = 1$ is not an eigenvalue of $\boldsymbol{E}$. Moreover, the necessary condition that $\boldsymbol{R}$ (and $\boldsymbol{R}^T$) is injective has been established. If $\rho(\boldsymbol{E}) < 1$, then (9) and (12) converge to the same solution, as is also evident from (32) (Appendix C).

4) Stopping Criterion: For the kth iterate x_k in (12), the relative residual norm is defined as $r_k = \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}_k\| / \|\boldsymbol{b}\|$. From (9b),

$$\frac{\|\boldsymbol{f}_{k+1} - \boldsymbol{f}_k\|}{\|\boldsymbol{B}\boldsymbol{d}\|} = \frac{\|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{f}_k\|}{\|\boldsymbol{b}\|} = r_k \tag{14}$$

For both GMRES and the Press iterative inverse the iteration is stopped when $r_k < \tau$, for some tolerance τ. In denoising applications with image size $N = 1024$, it turned out that a tolerance of $\tau \sim 0.1$ was usually enough to achieve the desired result. This corresponded to three to four iterations.

5) Comparison: To demonstrate the difference between the three methods, we applied them to the GDB transform of two test images, the cameraman image from Fig. 3 and an eigenimage of $\boldsymbol{E}$ with eigenvalue $\lambda > 1$. The eigenimage was found by Arnoldi iteration (see Appendix C). Both images have 512×512 pixels and intensity values in the range 0–1. The inverse problem has 262144 unknowns, and the dimension of $\boldsymbol{R}$ is 1571840×262144.

Fig. 4 shows the RMS value of $\boldsymbol{f} - \boldsymbol{f}_k$ [Press inverse (9)] and $\boldsymbol{f} - \boldsymbol{x}_k$ [GMRES inverse (12)] as a function of k, where $\boldsymbol{f}$ denotes the original image. The RMS error of the approximate inverse ($k = 0$) is about 10 % for both images. Fig. 4 shows that both methods apparently converge to the cameraman image when $k \leq 20$. The GMRES inverse converges to the eigenimage as well, but the Press inverse does not. Fig. 5 shows the RMS error for all three methods as a function of computing time (cameraman image only). During the initial iterations, the Press and GMRES inverses converge almost equally fast, but the amplification problem that destroys convergence for the Press inverse is now evident. The GMRES inverse approaches $\boldsymbol{f}$ to high accuracy, while the normal equation solver LSQR has slower convergence. The results are similar when plotting the residual matrix norm $\|\boldsymbol{f} - \boldsymbol{f}_k\|_2$. The preliminary conclusion from these tests is that the GMRES inverse is to be preferred because of its fast convergence when the inverse problem has a solution.

III. Denoising algorithm

The main idea is to transform the image (DTM) by an edge operator (filter) $\mathcal{L}$ followed by the DRT $\mathcal{R}$. The edge operator accentuates fine-scale intensity changes; the DRT concentrates accentuated linear features. The algorithm should cause as little distortion of real features as possible. Image reconstruction is done by applying the inverse operator $\mathcal{L}^{-1}\mathcal{R}^{-1}$. $\mathcal{R}^{-1}$ is implemented using one of the iterative methods described in Section II-C, such that $\mathcal{R}^{-1}\mathcal{R} \approx \mathbf{1}_Q$ to arbitrary desired accuracy. However, edge operators are in general not one-to-one
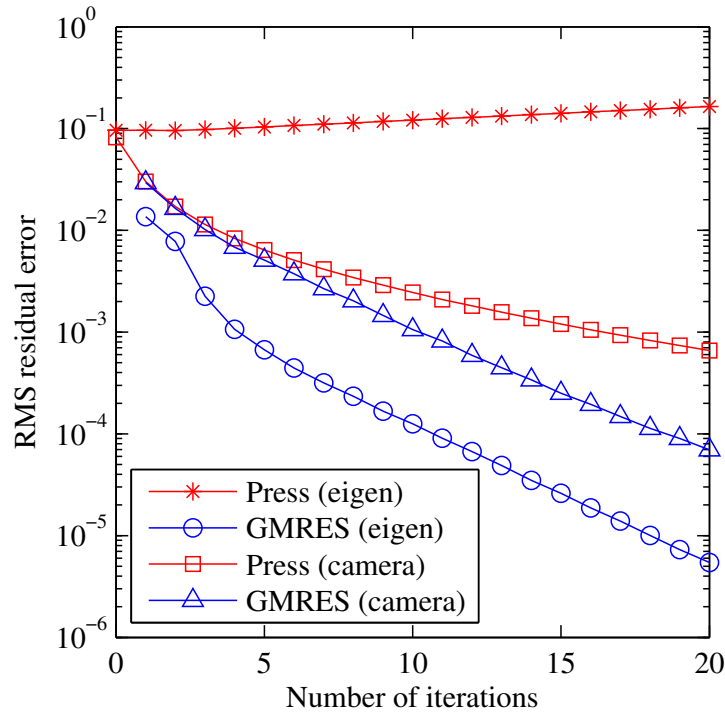
Fig. 4. RMS difference between original image and inverse iterate of the GDB transform, using (9) and (12). Results for two images are shown: the cameraman image in Fig. 3a, and an eigenimage of E with eigenvalue magnitude $|\lambda| > 1$.
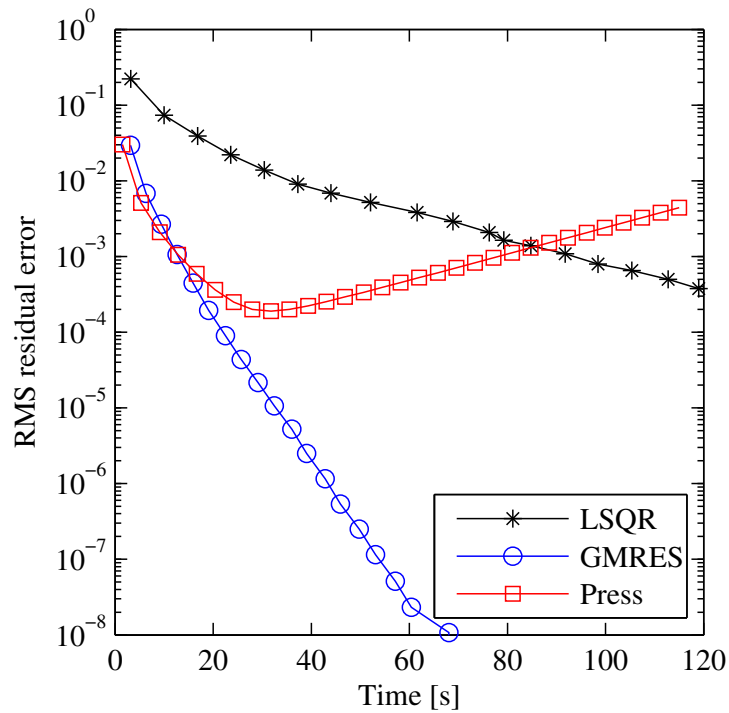

Fig. 5. RMS difference between original image (cameraman image in Fig. 3a) and the inverse iterate of the GDB transform, plotted as a function of computing time (single-core processing using Matlab on a standard desktop computer). Three inverse methods are compared: LSQR (normal equation solver), GMRES (12), and the Press inverse (9) based on the Neumann series.

mappings [44, Ch. 12.3] and hence not invertible. Therefore we first consider a suitable invertible edge operator.

A. Invertible Edge Detection Operator

For the reasons stated in Section I, our preferred operator is the Laplacian. The non-invertibility of the Laplacian is perhaps best examined in the frequency domain. A digital image $\boldsymbol{f} = [f(u,v)]_{u,v=1}^{N}$ with $N \times N$ pixels, N even, may be constructed by regularly sampling the periodic, band-limited function given by the Fourier series

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = \sum_{m=-N/2}^{N/2-1} \sum_{n=-N/2}^{N/2-1} c_{mn} e^{2\pi i(nx+my)}, \tag{15}$$

where f may be viewed as a function on the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, obtained geometrically by gluing together the two pairs of opposing edges of the unit square $[0,1] \times [0,1]$. By letting the first row in the image ($m = 1$) correspond to $y = 0$, the coefficients c_{mn} in (15) are related to the 2D discrete Fourier transform $\mathcal{F}$ by

$$c_{mn} = \frac{1}{N^2} \mathcal{F}[\boldsymbol{f}](m + N \bmod N, n + N \bmod N), \tag{16}$$

with $\mathcal{F}[\boldsymbol{f}](m,n) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} f(u,v) e^{-2\pi i(mu+nv)/N}$. A function

$$f \in L^2(\mathbb{T}^2) = \{\, g : \mathbb{T}^2 \to \mathbb{R} \mid \int_{\mathbb{T}^2} |g(\boldsymbol{x})|^2 \, \mathrm{d}\boldsymbol{x} < \infty \,\}$$

may be identified with its set of Fourier coefficients. For twice-differentiable functions in $L^2(\mathbb{T}^2)$, the continuous Laplace operator $\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2$ may be defined by

$$\Delta : c_{mn} \longmapsto -4\pi^2(m^2 + n^2)c_{mn}, \ (m,n) \in \mathbb{Z}^2. \tag{17}$$

In terms of the wave vector $\boldsymbol{k} = 2\pi(m,n)$, the transfer function is $\Delta(\boldsymbol{k}) = -\boldsymbol{k} \cdot \boldsymbol{k}$, which is a manifestation of the rotation invariance of the continuous Laplace operator. For the restricted set of functions

$$L_0^2(\mathbb{T}^2) = \left\{\, f \in L^2(\mathbb{T}^2) \ \middle| \ \int_{\mathbb{T}^2} f \, \mathrm{d}x = 0 \,\right\}, \tag{18}$$

for which $c_{00} \equiv 0$, the operator $D : L_0^2(\mathbb{T}^2) \to L_0^2(\mathbb{T}^2)$ defined by

$$D : c_{mn} \longmapsto \begin{cases} \frac{c_{mn}}{-4\pi^2(m^2+n^2)} & \text{for } (m,n) \in \mathbb{Z}^2 \setminus \{0\} \\ 0 & \text{for } m = n = 0. \end{cases} \tag{19}$$

is a right-inverse so that $\Delta D f = f$ [45, Ch. 4.4]. This suggests forming an invertible Laplace-like operator for images via (15)–(17) by perturbing the factor $(m^2 + n^2)$ in (17) away from zero, i.e.,

$$-4\pi(m^2 + n^2) \mapsto -4\pi^2(m^2 + n^2 + \epsilon^2), \ 0 < \epsilon^2 \ll 1.$$

However, while this approach works well for some images, computing the Laplacian in the frequency domain may fail because the function (15) need not be differentiable on the boundary of the unit square, whence image distortions may be incurred (one may consider, for example, the image formed by regularly sampling the function $g(x,y) = xy \exp(x^2 + y^2)$ on the unit square $[0,1] \times [0,1]$).

Instead, a more robust approach is to perturb the PSF of a discrete approximation Δ_d to the Laplacian in the spatial domain. We start with the operator implemented by circular (because of the periodicity condition) convolution with the PSF

$$\boldsymbol{h} = \frac{1}{2} \begin{bmatrix} 1/2 & 1 & 1/2 \\ 1 & -6 & 1 \\ 1/2 & 1 & 1/2 \end{bmatrix}. \tag{20}$$

The transfer function is

$$\mathcal{F}[\boldsymbol{h}] = 4\left[\cos^2\left(\hat{k}_x/2\right)\cos^2\left(\hat{k}_y/2\right) - 1\right] \tag{21}$$

$$\approx -\hat{k}^2 + \frac{3}{32}\hat{k}^4 - \frac{1}{96}\cos\left(4\phi\right)\hat{k}^4 + \mathcal{O}(\hat{k}^6), \tag{22}$$

where $\hat{\boldsymbol{k}} = 2\pi(m,n)/N$ is a scaled wave vector with polar coordinates (k, ϕ) [44, Ch. 12.5]. This operator is isotropic at long wavelengths. We have $\mathcal{F}[\Delta_d \boldsymbol{f}] = \mathcal{F}[\boldsymbol{h}]\,\mathcal{F}[\boldsymbol{f}]$, and direct deconvolution fails because of a single zero (at zero frequency) in the transfer function $\mathcal{F}[\boldsymbol{h}]$. (Here and below it is understood that the optical transfer function $\mathcal{F}[\boldsymbol{h}]$ has the same dimension as $\mathcal{F}[\boldsymbol{f}]$, obtained after zero-padding $\boldsymbol{h}$ and circularly shifting the DFT.) This problem can be amended by adding a small random component to the PSF $\boldsymbol{h}$ (not to the filtered image). We let $\boldsymbol{h}_M^0$ denote the matrix (20) symmetrically zero-padded to size $M \times M$, and let $\boldsymbol{\eta}_M(\epsilon)$ be an $M \times M$ matrix where the elements η_{ij} are random numbers such that $|\eta_{ij}| \leq \epsilon \ll 1$. The edge operator $\mathcal{L}$ is then implemented by circular convolution with the modified PSF

$$\boldsymbol{h} \rightarrow \boldsymbol{h}_M^0 + \boldsymbol{\eta}_M(\epsilon). \tag{23}$$

There is no guarantee that any randomly picked matrix on the form (23) has a zero-free transfer function, so one should check (for any dimension N only one such matrix need be found). This modified PSF does the same job as the Laplace operator as long as $\epsilon \ll 1$. There is no compromise or approximation involved in (23). It simply means replacing one operator with another which fulfills the same purpose while also being invertible. In this paper, we have used $M = 7$ and $\epsilon = 10^{-3}$.

B. Noise Suppression Filter

The GDB transform can be visualized by appropriately flipping, vertically shifting, and stitching together the four quadrants $\mathcal{R}^1 \boldsymbol{f}, \ldots, \mathcal{R}^4 \boldsymbol{f}$ into a Möbius band, as shown in Fig. 6 [33]. It is a Möbius band because the left and right edges, corresponding to vertical projections ($\pm 90°$), can be identified after flipping one of them upside down. The horizontal axis, the projection angle, maps to the rise parameter s in the GDB transform. A change in the vertical direction corresponds to a change in intercept parameter h. It is to this representation of the transformed data that we apply the noise filter.

Fig. 6a shows the GDB transform of a Laplacian image of a DTM with parallel horizontal tracks (it is actually the transposition of a DTM with vertical tracks). The track line discontinuities appear as a string of high intensity spots at the $0°$ projection angle. The other conspicuous feature is the concentration of noise at $\pm 90°$, i.e., in the direction perpendicular to the tracks. In addition, there is a more diffuse salt-and-pepper-like pattern across the whole image. We think the noise amplification at $\pm 90°$ is a manifestation of the motion-induced noise, which tends to be oriented perpendicular to the tracks.

Fig. 6b shows the effect of applying a 3×3 median filter, a standard choice for salt-and-pepper noise removal. The track line spots persist, and must be removed in a separate step. Fig. 7 shows the $0°$ column of the DRT where the track line discontinuities appear as distinct spikes in the data. When the track lines are roughly parallel as they often are in remote sensing situations, the spikes will always lie in a narrow vertical band at the angle given by the orientation of the tracks. The spikes can then be eliminated by zeroing all values in the narrow vertical band, or for example, by using a 1-D polynomial approximation for each column in the band (see Fig. 7).

Note that, when the edge operator is applied to a DTM (image) by circular convolution with (23), high pixel intensities may occur along the four edges of the resulting image. This is because the DTM may not be continuous there, in the sense of being a function on $\mathbb{T}^2$ as in (15). The edges of the image therefore get transformed to high intensity spots in the DRT domain. These spots are not artifacts in the data, and should not be affected by the DRT domain filter. To avoid boundary effects in the reconstructed DTM, the edge filter should be applied to a detrended DTM. A low-frequency approximation should first be subtracted, and subsequently added back to the final reconstructed DTM. This step removes discontinuities at the edges. As will be shown in Section V-B, standard low-pass filters tend to preserve some of the noise/artifacts that we want to remove, and this must be avoided. In their place we use a Chebyshev polynomial approximation that achieves the desired effect. The algorithm, denoted $\mathcal{T}(p, D)$, is characterized by the polynomial degree p and a downsampling factor D (see Appendix D). The complete denoising algorithm may now be summarized as follows:

> **procedure** $\boldsymbol{f} = \text{DENOISE}(\boldsymbol{f}_0)$
> $\boldsymbol{a} \leftarrow \mathcal{T}(p, D)\boldsymbol{f}_0$
> $\boldsymbol{f} \leftarrow \boldsymbol{f}_0 - \boldsymbol{a}$
> $\boldsymbol{f} \leftarrow \mathcal{L}\boldsymbol{f} = \boldsymbol{h}(M, \epsilon) * \boldsymbol{f}$ $\triangleright$ circular convolution
> $\boldsymbol{d} \leftarrow \mathcal{R}\boldsymbol{f}$
> $\boldsymbol{d} \leftarrow \text{medfilt}(\boldsymbol{d})$ $\triangleright$ 3×3 median filter
> $\boldsymbol{d} \leftarrow$ suppress peaks at track line angle
> $\boldsymbol{f} \leftarrow \mathcal{R}^{-1}\boldsymbol{d}$
> $\boldsymbol{f} \leftarrow \mathcal{L}^{-1}\boldsymbol{f} = \mathcal{F}^{-1}\left[\mathcal{F}\left[\boldsymbol{f}\right] / \mathcal{F}\left[\boldsymbol{h}\right]\right]$
>   **return** $\boldsymbol{f} + \boldsymbol{a}$
> **end procedure**

First, the trend $\mathcal{T}(p, D)\boldsymbol{f}_0$ is computed and stored in memory and subtracted from the image (DTM) f_0. The modified Laplace operator is applied to the detrended image, and the PSF is stored for later use in the inverse (deconvolution) step. Next, the forward GDB DRT is applied to the Laplacian image, and the four quadrants in (5) are merged into a single matrix as illustrated in Fig. 6. A 3×3 median filter is applied to this matrix. The median filter may not be relevant in other applications; we have used the procedure, excluding the median filter, to remove scalloping (parallel stripes) in synthetic aperture radar (SAR) images; see further remarks in Section VI. The peaks in the DRT corresponding to linear artefacts are suppressed, e.g., by zeroing the columns corresponding to the orientation of the lines. The inverse DRT is next computed using one of the iterative methods described in Section II-C. The inverse Laplacian image is computed by direct deconvolution with the PSF $\boldsymbol{h}(M, \epsilon)$ (23). Finally, the trend is added back to obtain the restored image (DTM).
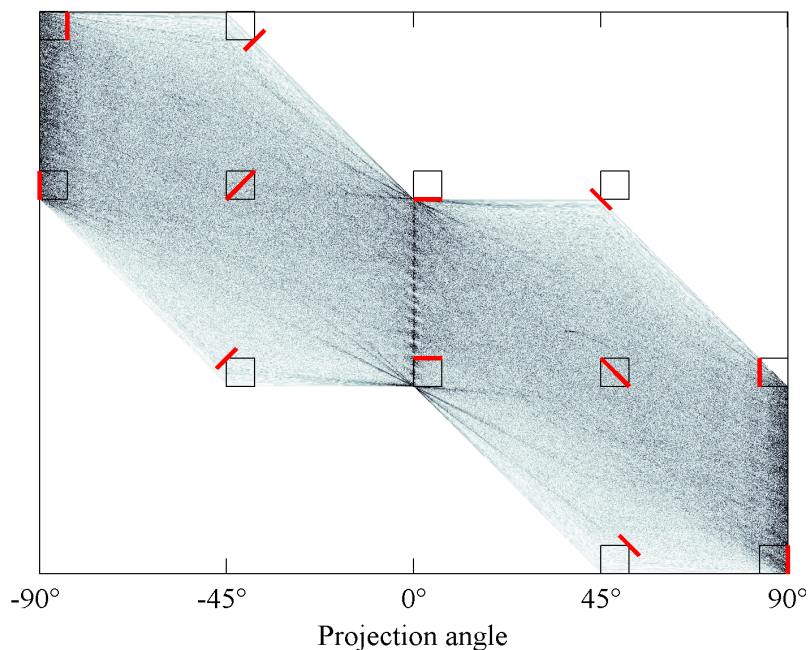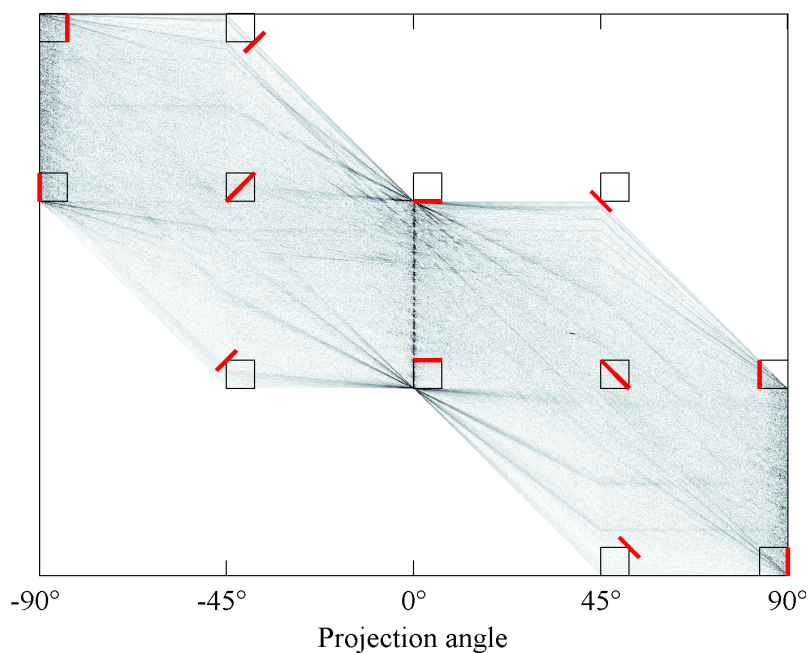
(a) DRT

(b) 3×3 median filter of DRT

Fig. 6. Typical example of the GDB transform of an edge-filtered bathymetric DTM with motion-induced noise and track line artifacts is shown in (a). (b) shows the effect of applying a 3×3 median filter in the DRT domain, i.e., to the image (a). A unit (one pixel) increase along the vertical axis corresponds to a unit change in the intercept parameter h. However, the parameter h is not constant along horizontal lines in the figure, it is only constant within each of the four quadrants. This is because the quadrant transforms $\mathcal{R}^1 f, \ldots, \mathcal{R}^4 f$ have been translated vertically so that they fit together in a Möbius band [33]. The color map is symmetric about zero, so that white corresponds to zero and black corresponds to high positive or negative values. The DTM has 1024×1024 pixels and horizontal line artifacts (tracks). The small squares symbolize the image (DTM) extent, and the red (short, thick) lines indicate the projection that corresponds to the associated pixel in the GDB domain. For example, the topmost non-zero value at $0°$ corresponds to the bottom horizontal line, while the top left value corresponds to the rightmost column in the image.

164
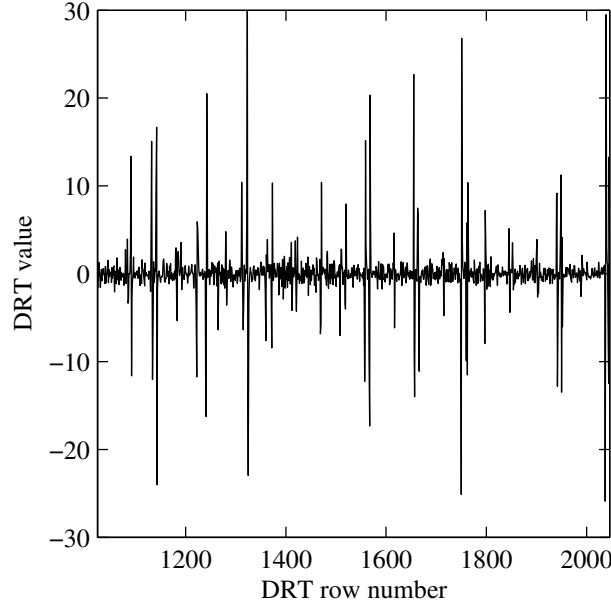
Fig. 7. GDB transform (non-zero part) at $0°$ projection angle, showing track line discontinuities as a series of spikes.

IV. UNBIASED TERRAIN FEATURES

As noted above, local geometric attributes, such as gradient or curvature, may be used to visualize landforms and geomorphology. Furthermore, images may be segmented (or classified) according to variations in morphological patterns caused by differing erosional and depositional processes. The rate at which local patterns are correctly classified will be a relative measure of image quality. Preferably, local patterns should be represented by intrinsic properties that do not depend on image orientation, intensity (or depth/elevation), or large-scale variations such as approximately constant slopes. We also require that features should not depend on externally imposed parameters, such as the direction of a derivative, the spatial relationship of pixel pairs, or the elevation angle of artificial illumination. The operator will, effectively, only determine the spatial resolution, i.e, the size of the pixel blocks used to identify local patterns.

Features with such properties may be obtained from the Laplacian Δf, which is approximately rotation invariant and insensitive to depth and constant slopes. The Laplacian is a measure of curvature, and in some situations additional information on patterns can be obtained by computing the Gaussian curvature K and the mean curvature H. K and H are intrinsic geometric properties of more general differentiable surfaces embedded in three-dimensional space [46]. For our purpose, a DTM may be considered as a regular set of samples from the graph of a function $f(x, y)$ of rectangular coordinates. For the special case of a graph surface, the formulae for K and H are

$$K = \frac{f_{xx} f_{yy} - f_{xy}^2}{\left(1 + f_x^2 + f_y^2\right)^2} \tag{24}$$

$$H = \frac{f_{xx}\left(1 + f_y^2\right) - 2 f_{xy} f_x f_y + f_{yy}\left(1 + f_x^2\right)}{2\left(1 + f_x^2 + f_y^2\right)^{3/2}}, \tag{25}$$
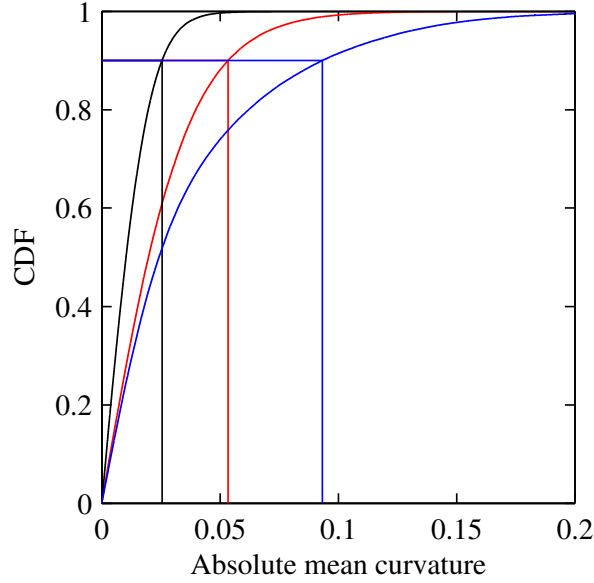
165

Fig. 8. ECDFs for the quantity $|H|$ (in a rescaled image), for three areas with different geomorphology (see Section V-C). Given the local CDF for some geometrical property, a Q-dimensional feature vector may be formed by taking the inverse CDF at Q values $0 < q_1 < \ldots < q_Q < 1$ (only $q = 0.9$ shown here).

where f_x denotes $\partial f / \partial x$ etc. [47]. The curvatures K and H have some useful local and global properties. For example, for a domain $V \subset \mathbb{R}^2$, the functional $\int_V H \sqrt{1 + |\nabla f|^2} \, \mathrm{d}x \, \mathrm{d}y$ is equivalent to the minimum curvature functional of thin-plate spline interpolation [47], and measures how much the area of the surface $f(V)$ changes in response to a unit length deformation in the surface normal direction. K and H are related to the more familiar principal curvatures, i.e., κ_1 and κ_2, by $K = \kappa_1 \kappa_2$ and $H = (\kappa_1 + \kappa_2)/2$ (at any point $p \in \Sigma$, i.e., a differentiable surface, κ_1 and κ_2 are the maximum and minimum curvatures of the family of 1D curves formed by the intersection of Σ with all planes containing the surface normal $\boldsymbol{N}_p$ at p).

An image block of size $S \times S$ pixels may be characterized by the distribution of a particular geometric property. Here we have used the empirical cumulative distribution function (ECDF), as shown in Fig. 8. Specifying Q values $0 < q_1 < \ldots < q_Q < 1$ (e.g., the 10-quantiles) yields a Q-dimensional feature vector via the inverse ECDF. One may also concatenate the feature vectors of several geometric properties, e.g. K and H or H and Δf.

For comparison and as an alternative to the ECDF, we have applied the local binary pattern (LBP) operator, a widely used texture operator that also yields distribution-based feature vectors. LBPs were introduced by Ojala, Pietikäinen and Harwood [48], building on the work on texture spectra by Wang and He [49]. The original LBP is formed by thresholding the pixel values of an 8-neighborhood with respect to the center pixel, producing a string of eight zeros and ones. The string is interpreted as a binary number in the range 0–255. The distribution (histogram) of LBPs in an image block of size $S \times S$ pixels may be used as a measure of the texture in that block [50]; the number of features equals the number of bins in the histogram. Since this gives very high-dimensional feature vectors (256 components), it is common to consider only the so-called uniform patterns corresponding to the binary strings with at most two transitions between 0 and 1 (e.g. 00111000 has two transitions and is uniform, while 01001000 has four transitions). There are 58 such patterns [51].

166

By definition, the LBP is invariant with respect to uniform (or locally uniform) changes in image intensity, but it is not rotation invariant. However, rotating an image at angles $2\pi/P$ (where $P = 8$ for an 8-neighborhood) corresponds to a circular bit shift in the patterns. Therefore, one obtains rotation-invariant distributions by identifying all patterns that are equal modulo a circular bit shift. For 8-neighborhoods, this results in 36 distinct features [52]. For the subset of uniform patterns, the corresponding number is nine.

The block size S must be large enough to ensure reliable histogram estimation; here we have used $S = 16$, the smallest size considered in [50], and $S = 32$. To compute the LBPs we have used the VLFeat library [53], which was modified to handle rotation-invariant patterns. In the VLFeat LBP algorithm, each pixel-level pattern contributes to the histograms of the four nearest blocks via bilinear interpolation.

V. TESTS AND RESULTS

A. Multibeam Bathymetry

The denoising algorithm was tested on MBES data from three areas in the northern North Sea (A, B, and C: Fig. 9). The data were acquired by the Norwegian Defence Research Establishment (FFI, Kjeller/Horten, Norway), on the research vessel M/S HU Sverdrup II, in 2005 (C), 2006 (B), and 2008 (A), respectively. Area A was surveyed with a Kongsberg Maritime (KM, Horten, Norway) EM 710 echo sounder, while areas B and C were surveyed with a KM EM 1002. The EM 1002 has a center frequency of 95 kHz; the array is semi-cylindrical and emits a fan-shaped beam with $2.0°$ along-track $3\,\mathrm{dB}$ beamwidth, and it forms 111 beams with $2.0°$ beamwidth across track. The EM 710 operates at 70–100 kHz, and has orthogonally oriented rectangular transmit and receive arrays. The EM 710 transmits in three sectors simultaneously with distinct frequencies in inner and outer sectors, and forms up to 400 $1.0° \times 0.5°$ beams on reception ($0.5°$ transmit beamwidth). The EM 710 system may generate two separate swathes per ping (at slightly different forward tilt angles), yielding a higher density of soundings than the older EM 1002 system.

The survey area stretches from the eastern margin of the North Sea plateau (water depth circa 100 m) and into the Norwegian Trench (water depth circa 300 m, see Fig. 9). The three areas A, B, and C have different sedimentary and morphological characteristics, and the data have different noise characteristics. Area A (EM 710) lies on the North Sea plateau, where the seabed is sandy. Area B covers part of the western slope of the Norwegian Trench. This area has an irregular pattern of elongated pockmarks that may be a result of fluid seepage and water current erosion [54], [55]. Area C is flat with soft, silty sediments, but there are many circular pockmarks caused by vertically migrating shallow gas. In general, noise and artifacts will be most visible in flat areas such as C, and less visible in areas with greater depth variation, such as B, which has a depth range of about 90 m.

For each area (A, B, and C), DTMs with resolutions in the range of 1.5 m to 8 m were constructed by block averaging the MBES soundings. Empty cells were filled in by taking the mean depth of the four nearest cells. All the processing results presented in the next section were obtained on grid models with 1024×1024 pixels. The EM 1002 data set had, before we received it, been put through a standard post-processing chain that included corrections for tidal variations, manual editing, and removal of outliers. The EM 1002 data have motion-induced noise which is most pronounced at the outer beams, possibly a result of imperfect sound refraction corrections; this also causes visible discontinuities between adjacent track
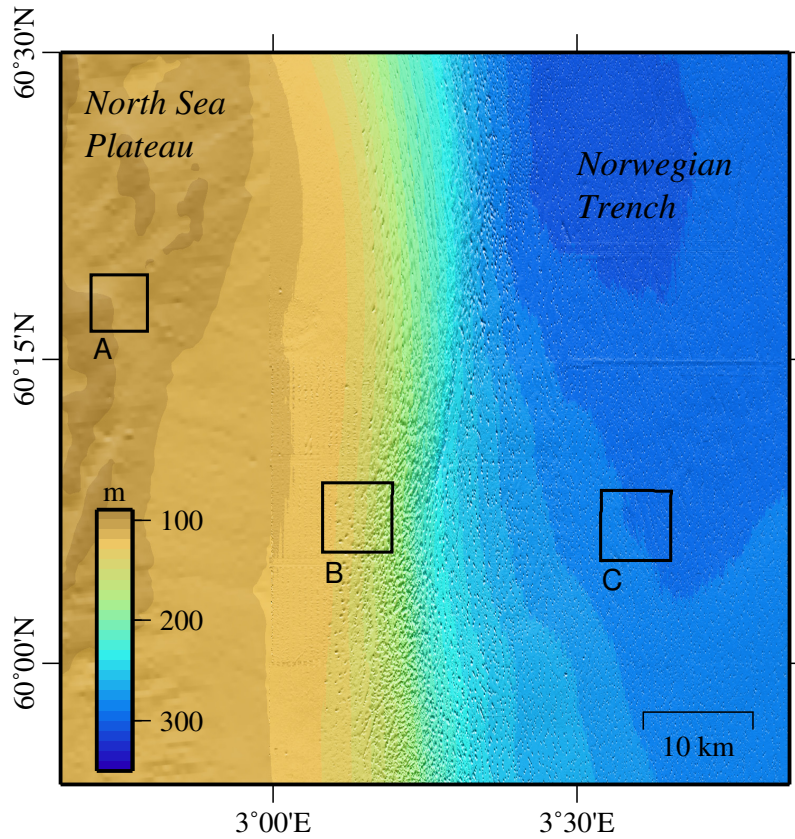
Fig. 9. Overview and bathymetry of the survey area in the northern North Sea. The denoising algorithm was tested in the areas labeled A, B, and C. Area A was surveyed with a Kongsberg Maritime (KM) EM 710 MBES, while areas B and C were surveyed with a KM EM 1002 MBES.

lines. The EM 710 data were raw, unprocessed data, where, prior to construction of DTMs we applied only a crude thresholding to exclude large outliers. The lack of tidal correction causes marked discontinuities between track lines. There is also a low-frequency swell-induced noise that appears as wavy patterns "propagating" in the direction of the ship tracks. While MBES data normally have post-processing corrections, it is useful to test a new algorithm on a challenging data set.

B. Denoising results

The DTMs for areas A, B, and C were DRT-filtered with the following settings: For areas A and B, the trend surface $\mathcal{T}f$ was computed using the Chebyshev polynomial approximation algorithm of Appendix D with the parameter values $p = 12$ and $D = 4$; for area C the parameter values were $p = 6$ and $D = 4$. Note that the approximation algorithm is based on least squares fitting, which mitigates the phenomenon of Runge (oscillations near the boundary) [56]. The modified Laplacian was computed using the PSF (23) with size $M = 7$ and perturbation level $\epsilon = 10^{-3}$. A 3×3 median filter was applied to the DRT of the Laplacian. In all three areas, the track lines are oriented approximately north-south; the track line discontinuities therefore appear as high-intensity spots at $\pm 90°$ in the DRT domain, or, equivalently, at $0°$ in the DRT of the transposed image (as in Fig. 6). To suppress linear artefacts, each DRT column was zeroed in approximately the range $\langle -1°, 1° \rangle$. (One might also use some interpolation scheme to reset the zeroed values, but this was not done in the

present work and we do not think it is important when the DRT is applied to the Laplacian of a detrended surface.) The inverse DRT was computed using the GMRES algorithm with stopping criterion $r_k = 10^{-6}$ or maximum six iterations, whichever is reached first. We found, however, that the less tight criterion of $r_k = 10^{-2}$ also yielded good results, good enough as perceived by the human eye. The rate of convergence will depend on the nature of the noise and data; we found that the algorithm converged faster when inverting the SAR images (briefly mentioned in Section VI) compared with the MBES terrain models. As discussed below (Section VI), except for the polynomial degree p the results shown here were not very sensitive to the parameters, and consequently, little parameter tuning was required. How to determine the best value for p a priori is an open question.

For comparison, various lowpass filters were applied to the DTMs or directly to attribute images, some examples of which are shown in the following. Lowpass filters were implemented as convolutional averaging filters with constant or Gaussian weights and filter diameters between 3 and 13 pixels (for Gaussian filters, the diameter was defined as 6 standard deviations). In addition, a Butterworth filter was implemented by multiplying the centered normalized spectrum of $\mathcal{F}[f - \mathcal{T}f]$ or $\mathcal{F}[\Delta f]$ (see Fig. 10b) with

$$G(\nu_1, \nu_2) = \frac{1}{(1 + \sqrt{\nu_1^2 + \nu_2^2}/D_0)^{2n}}, \tag{26}$$

where n and D_0 are adjustable parameters.

In general, we found that low-pass filters were ineffective against stripe noise caused by track line discontinuities. Depending on the size of the convolution kernel, low-pass filters can significantly reduce motion-induced noise, but only at the cost of severe smoothing of real features. The DRT filter is targeted at linear artifacts and produced comparatively good attenuation of stripe noise. The DRT filter also has a smoothing effect which reduced motion-induced noise with apparently less distortion of real features. However, severe motion-induced noise, seen in particular on the outer beams in some of the EM 1002 data, is difficult to eliminate using either low-pass filters or the DRT filter.

A partial solution to this problem was to apply a DRT filter locally by subdividing the 1024×1024 pixels DTMs into nonoverlapping blocks of 32×32, 64×64 or 128×128 pixels. This DRT filter was designed to attenuate linear artifacts in narrow sectors close to perpendicular to the track lines. Applying the filter in non-overlapping blocks ensures fast processing. Moreover, at least for block size $N \leq 64$, it is possible to use a matrix representation for the DRT inverse (see further remarks in Appendix E). It does not seem obvious in advance that disjoint block processing should work satisfactorily, i.e., without producing block artifacts in the final composite image. Indeed, local DRT filters applied directly to gradient images did produce block artifacts. In the present results, the DRT was applied to the Laplacian images, and the denoised gradient images were derived from the final, restored DTMs. These gradient images have no visible block artifacts.

Fig. 10 shows the power spectrum of the detrended original DTM $f - \mathcal{T}f$ for area A (see Fig. 10a), and the power spectrum of the Laplacian image $\Delta(f - \mathcal{T}f)$ (see Fig. 10b). These spectra reflect that the noise patterns are directed and broadbanded. For both spectra it is difficult to design a filter that eliminates the artifacts without distorting the real bathymetric features. Fig. 11 shows the frequency response of the Butterworth filter (26) with parameters $n = 2$ and $D_0 = 400$. This filter was applied to the Laplacian image of area A (cf. Fig. 10b) to produce the lowpass-filtered Laplacian image shown in Fig. 12d.
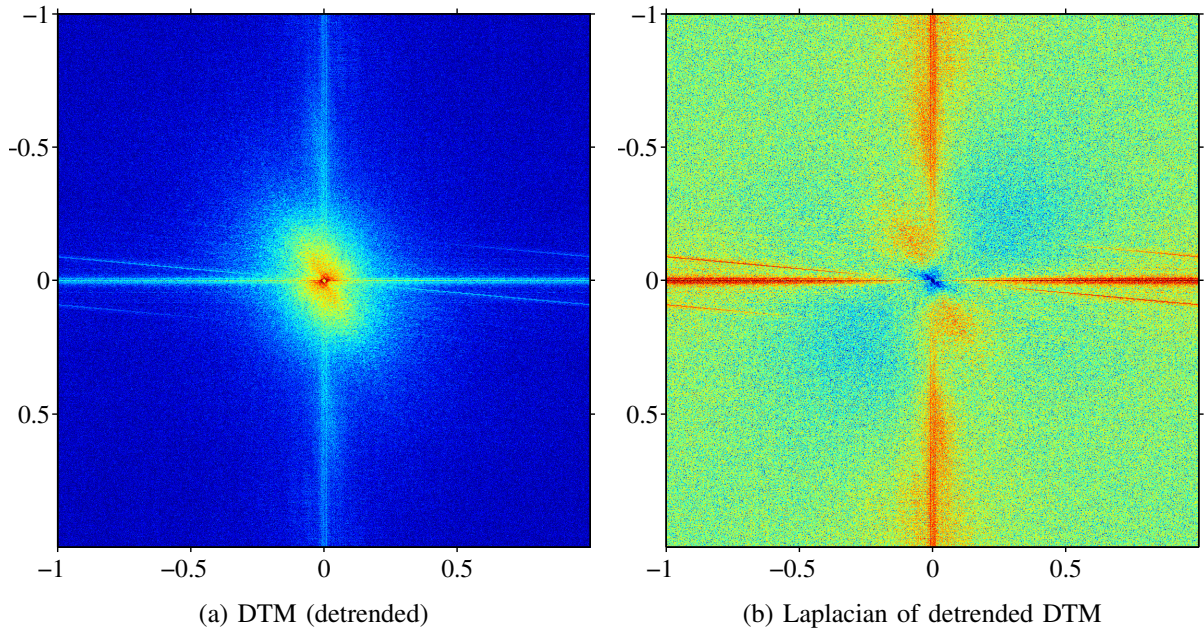
(a) DTM (detrended)

(b) Laplacian of detrended DTM

Fig. 10. Centered Fourier log-power spectra $2\log|\mathcal{F}[\cdot]|$ (normalized frequency) of detrended DTM for area A (a) and corresponding Laplacian image (b).
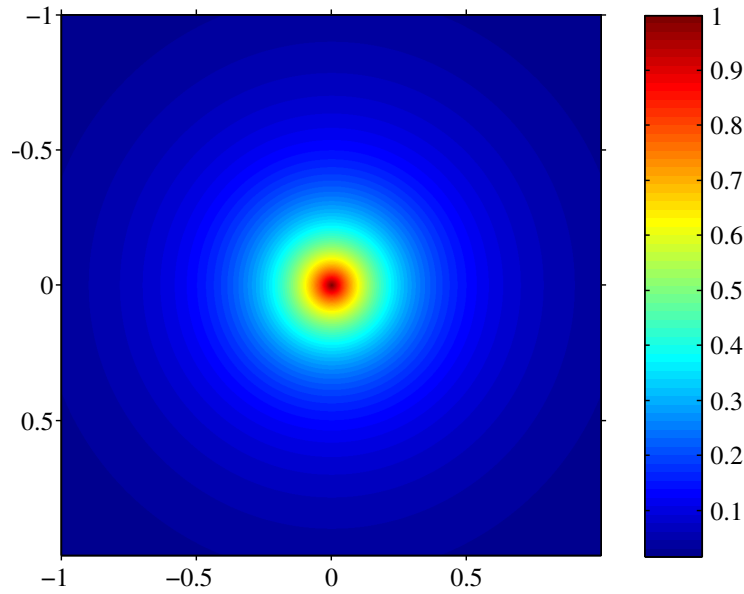

Fig. 11. Centered frequency response, in normalized frequency, of Butterworth low-pass filter (26) with $n = 2$ and $D_0 = 400$. This filter was used to produce the image in Fig. 12d below.

Fig. 12 shows the effect of the DRT filter and lowpass filters on the Laplacian $\Delta\left(f - \mathcal{T}f\right)$ for area A. In the original image (see Fig. 12a), the morphological features are poorly visible. The track line discontinuities cause parallel vertical stripes. The DRT-filtered version (see Fig. 12b) has no stripes and better defined morphology. As demonstrated in Section V-C, moreover, noise reduction is essential to achieve accurate discrimination between different geomorphological classes using second-derivative-based attributes. The low-pass-filtered images in Fig. 12c (7×7 boxcar filter applied to $f - \mathcal{T}f$) and 12d (frequency-domain Butterworth filter applied to $\Delta\left[f - \mathcal{T}f\right]$) retain the vertical stripes. In particular, it does not help to filter

the attribute image $\Delta\left(f - \mathcal{T}f\right)$ directly, nor is the effect on stripes sensitive to the filter size. It may also be seen that lowpass filters cause greater smoothing of terrain features compared to the DRT filter.

Fig. 13 shows gray level maps of bathymetry (gray-level-coded water depth) in area A before and after the DRT filter was applied. Gray level (color) maps are usually not the best way to visualize morphological details imaged by instruments like the MBES nor are they as sensitive as attribute images to the forms of noise studied here. The maps were included to show that the DRT filter did not alter the depth of real topographical features. As described in Section III-B, to ensure this, it is necessary to subtract a trend surface from the DTM before applying the Laplacian operator, to avoid having high values for the Laplacian along the boundary. If a low-frequency distortion appears in the final restored DTM, one should try and alter the degree p of the approximation operator $\mathcal{T}(p, D)$ (see Appendix D). Bear in mind that the data in this area had not been through a post-processing chain; in particular, there is no correction for tidal variations. This causes relatively large discontinuities that are not completely eliminated. The effect of the DRT filter is best seen in the attribute images.

With the given orientation of the track lines, the stripe noise is most conspicuous when taking the horizontal gradient (see Fig. 14). Again, by comparing Fig. 14b and Fig. 14c, it is seen that the DRT filter is more effective than the lowpass filter in removing stripes, whereas the latter yields a less sharp view of morphology. The straight line that runs at a steep angle across the images is a real feature, probably a pipeline. The partial attenuation of this feature is an unwanted side effect.

Fig. 15 shows an example of motion-induced noise from area C, in vertical gradient images, and the effect of block processing with the DRT filter (128×128 pixels blocks). These images show many pockmarks (craters) on an almost flat seabed. When experimenting with the standard Radon transform and filtered backprojection, we found that many of the small pockmark features were distorted by the filter, even with a high sampling rate in the angular domain. This side effect was eliminated when using the GDB transform and iterative inverse. Low-pass filters, in this example, a 5×5 pixels boxcar filter, can remove much of the motion-induced noise (see Fig. 15b), but a better result is obtained with a DRT filter (see Fig. 15c). In this example, the filter worked by zeroing the DRT in two narrow sectors symmetric about the horizontal direction.

A yet more difficult example, from area B, is shown in Fig. 16. The interference of motion-induced noise patterns, between the outer beams of adjacent track lines, creates an effect which is visible in both the horizontal (see Fig. 16a and 16b) and vertical (see Fig. 16c and 16d) gradient images. The DRT filter applied directly to area B did attenuate the vertically oriented noise patterns but not completely (see Fig. 16b). Moreover, low-pass filters were not effective in this respect. Block processing with a DRT filter, as in Fig. 15c, with block size 32×32, does markedly reduce the noise level. This result (see Fig. 16d) was achieved without any discernible smoothing or distortion of the underlying geomorphological patterns showing the ocean current-modified pockmarks of area B.

As the true terrain form and the amount of noise are not known, direct evaluation of the signal-to-noise ratio (SNR) and the level of image distortion is not possible. However, the inverse coefficient of variation (ICV) may be used as an approximate measure of SNR and hence of image quality [21], [22], [57]. The ICV for an image region W is defined as
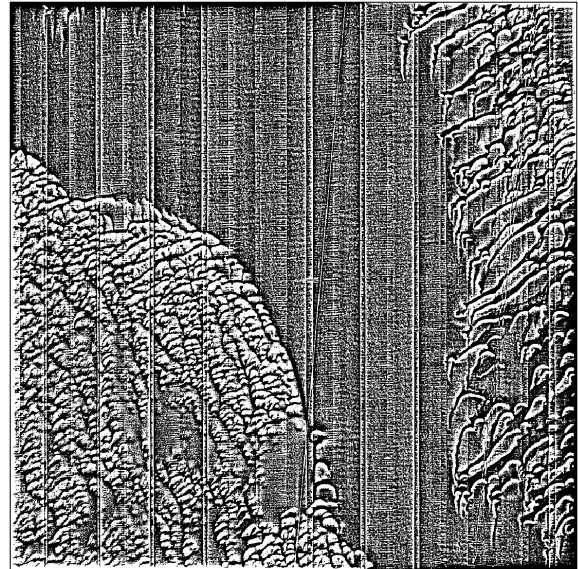
$$\text{ICV} = \frac{R_a}{R_{sd}}, \tag{27}$$

171
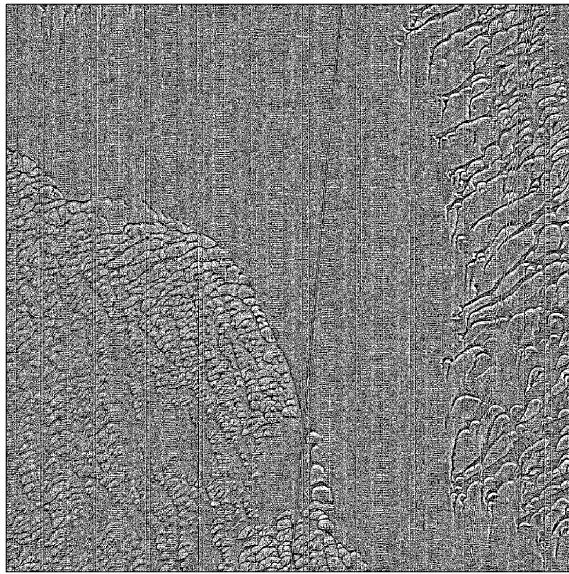
(a) Laplacian of original DTM (Δf)
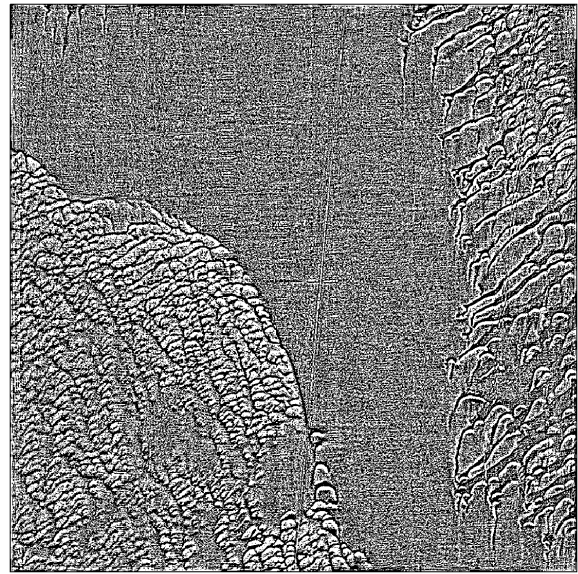
(b) Laplacian of DRT-filtered DTM
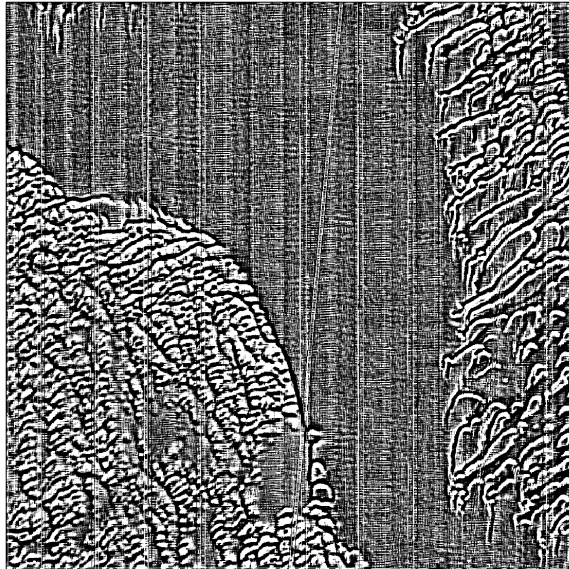
(c) Laplacian of lowpass-filtered DTM

(d) Lowpass-filtered Laplacian of original DTM

Fig. 12. The Laplacian of the DTM for area A. The color (gray) scale is identical in all four images. Low-pass filters are not effective for removing the vertical track line discontinuities; Fig. (c) was obtained after smoothing the original DTM with a 7×7 pixels boxcar averaging filter, while Fig. (d) was obtained by smoothing the Laplacian of the original DTM with a Butterworth filter.

where R_a is the mean pixel value in W and R_{sd} is the standard deviation of the pixels in W. In the works [21], [22], and [57], the ICV is computed in two or four small (10×10 pixels), homogeneous, noisy image regions, and R_{sd} is assumed to be mainly determined by noise. Here, we have computed the ICV for the whole DTM, in 8×8 pixels blocks, so that the result can be displayed as an image (see Fig. 17) showing the linear discontinuities as regions with low ICV (SNR) (to save space we only include the result for area A). Another measure employed in the above-cited works is the mean relative deviation (MRD), which is computed in noise-free regions. However, as the present images are affected by (motion-
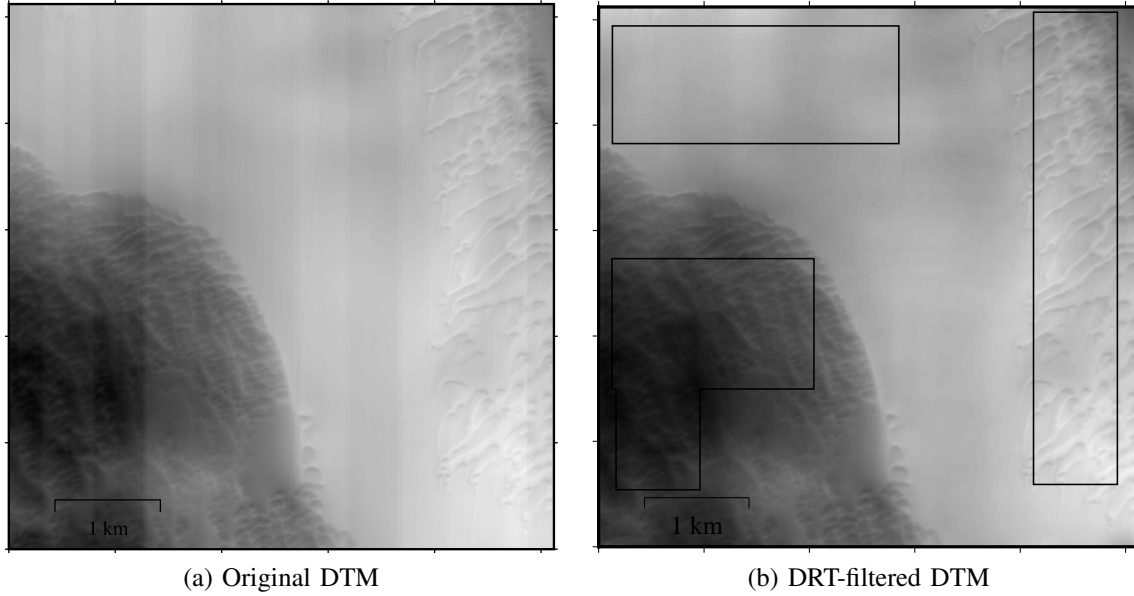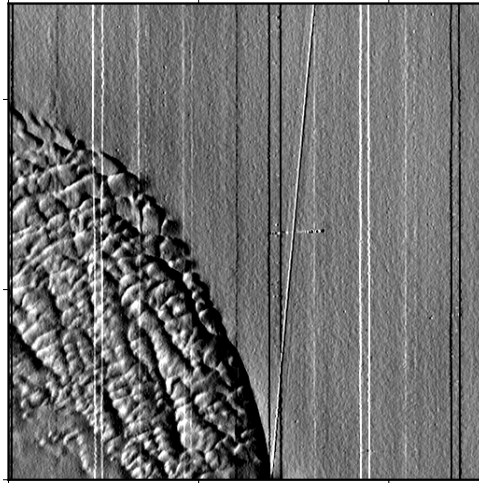
(a) Original DTM   (b) DRT-filtered DTM

Fig. 13. Gray level maps of bathymetry in area A (depth range of 94–110 m). The rectangles in Fig. 13b mark the areas representing the three classes of seafloor morphology in the classification tests (see Section V-C). Gray level (color) maps alone do normally not yield a good visual impression of morphology and noise (although moderately useful in this particular case). Here, they demonstrate that major features and absolute depth values are preserved after taking the inverse Radon transform and inverse Laplace operator.

induced) noise effectively everywhere, the MRD is not considered here. The ICV may, by definition, be high for blurred (or smoothed) images, which conflicts with the perceived sense of poor image quality. On the other hand, we emphasize that the DRT filter better preserves the image sharpness compared to low-pass filters. We have therefore also attempted to quantify the effect of denoising in terms of the practical measure of terrain classification error.

C. Effect on Classification

The classification experiment was restricted to a single area (A) to ensure that extraneous factors, such as differences in the noise characteristics or instrument resolution, did not improve the classification accuracy artificially. Three main geomorphological classes were identified, labeled, and defined by the training areas enclosed by rectangles in Fig. 13b. Each training area contains about 500 non-overlapping subcells with 16×16 pixels, for which the local ECDFs and rotation-invariant LBP histograms were computed for the quantities K (Gaussian curvature), H (mean curvature), Δf (Laplacian), and $f - \mathcal{T}f$ (depth residual). With respect to ECDFs, each cell is characterized by the set of 10-quantiles, which gives 9 features (or 18 if two geometric quantities are combined, etc.). For comparison, we have also tested the full set of uniform LBPs, which are not rotation invariant. The experiment was also run on 32×32 blocks, which yields more robust estimation of local distributions at the cost of lower spatial resolution and less training samples.

The classification results are presented here as learning curves (see Fig. 18), where the estimated classification error obtained with the Bayesian normal density classifier (quadratic discriminant function) is plotted versus the number of features D used in the classifier. A range of other classifiers were also tested, but none performed significantly better than the normal density classifier. The forward feature selection method was used to select subsets

173

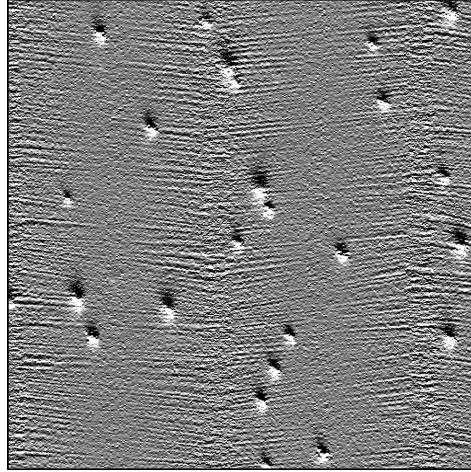(a) Horizontal gradient of original DTM ($\nabla_x f$)
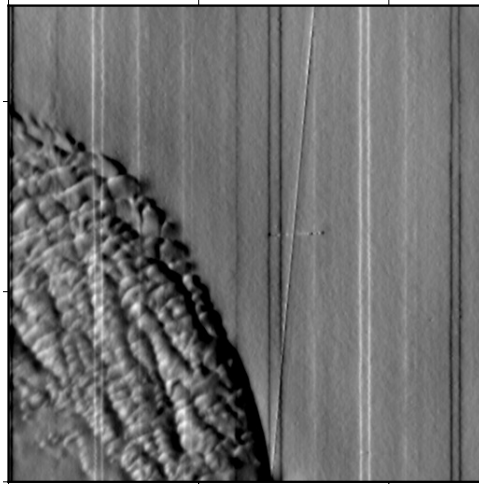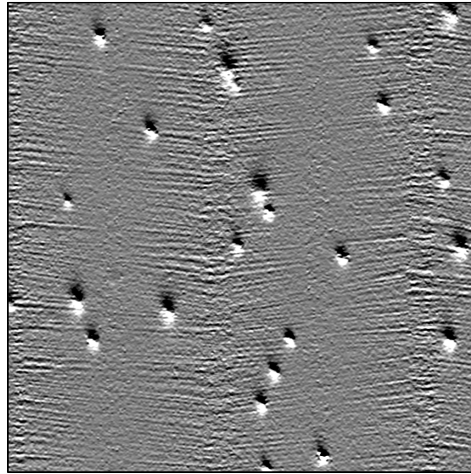

(b) Horizontal gradient of lowpass-filtered DTM
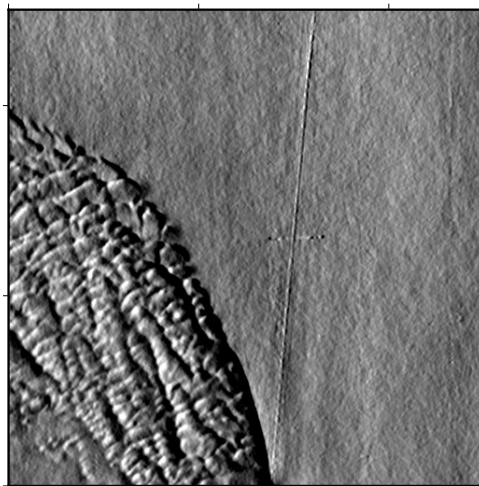
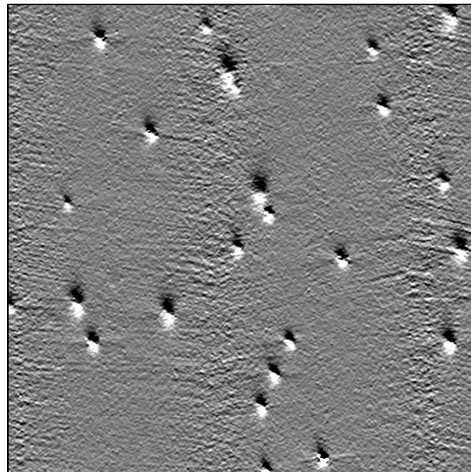(c) Horizontal gradient of DRT-filtered DTM

Fig. 14. Horizontal gradients (detail from area A) show vertical linear artifacts due to track line discontinuities and lack of tidal corrections. Fig. (b) was obtained after filtering the original 1024×1024 pixels DTM with a Butterworth lowpass filter with parameters $n = 2$ and $D_0 = 400$ [see equation (26)]. The single oblique line seen is most likely a real feature (e.g., a pipeline).

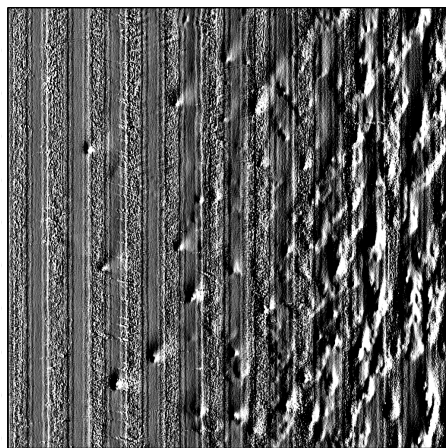(a) Vertical gradient of original DTM ($\nabla_y f$)
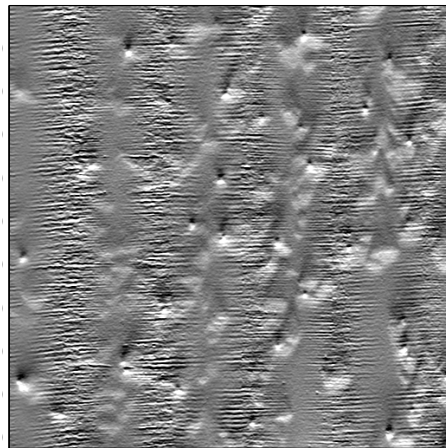

(b) Vertical gradient of lowpass-filtered DTM

(c) Vertical gradient of block-DRT-filtered DTM

Fig. 15. Vertical gradients in a flat area with pockmarks (detail from area C) show motion-induced artefacts forming a "fishbone pattern", where the overlapping areas between adjacent swathes (vertically oriented) correspond to the spine of the fish. Fig. (b) was obtained by filtering the original 1024×1024 pixels DTM with a 5×5 pixels boxcar averaging filter. Block processing the DTM with a DRT filter produced Fig. (c) (32×32 pixels block size).

(a) Horizontal gradient of original DTM

(b) Vertical gradient,original DTM (detail)
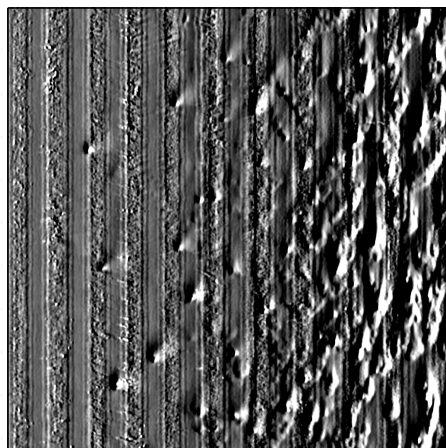
(c) Horizontal gradient, LP-filtered DTM
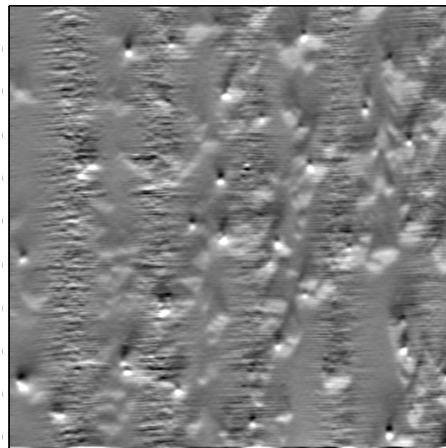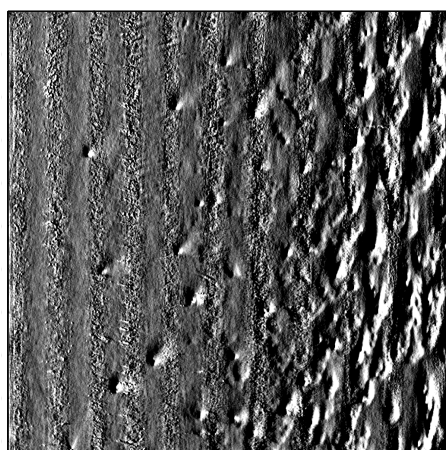
(d) Vertical gradient, LP-filtered DTM

(e) Horizontal gradient, DRT-filtered DTM
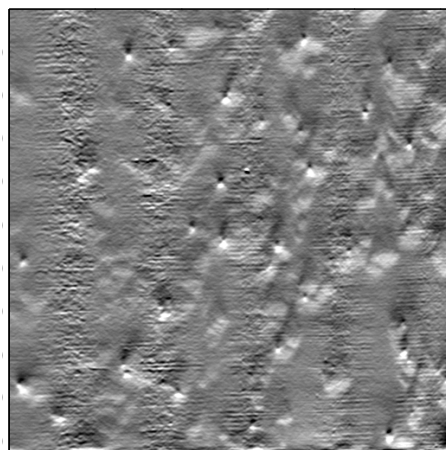
(f) Vertical gradient, DRT-filtered DTM

Fig. 16. Depth estimation errors and motion-induced noise cause vertical linear artifacts in the horizontal gradient image of area B (a). The DRT filter, applied to the 1024×1024 pixels DTM, dampens much of the stripe noise (e), whereas a lowpass filter (5×5 boxcar) is ineffective (c). However, the severe motion-induced noise, shown in Fig. (b) by taking the vertical gradient in a close-up corresponding to the lower right quadrant of of (a), is difficult to eliminate without blurring the image (d). A partial solution, which reduces noise without blurring, is to apply a DRT-filter in (nonoverlapping) blocks (here 32×32 pixels) (f); the DRT is applied to Δf (Laplacian of original DTM) and dampens nearly-horizontal linear structures. Here, one color map has been used for the left column panels and another for the right column panels, because the general scale of magnitudes is different for horizontal and vertical gradients.

(a) Original DTM
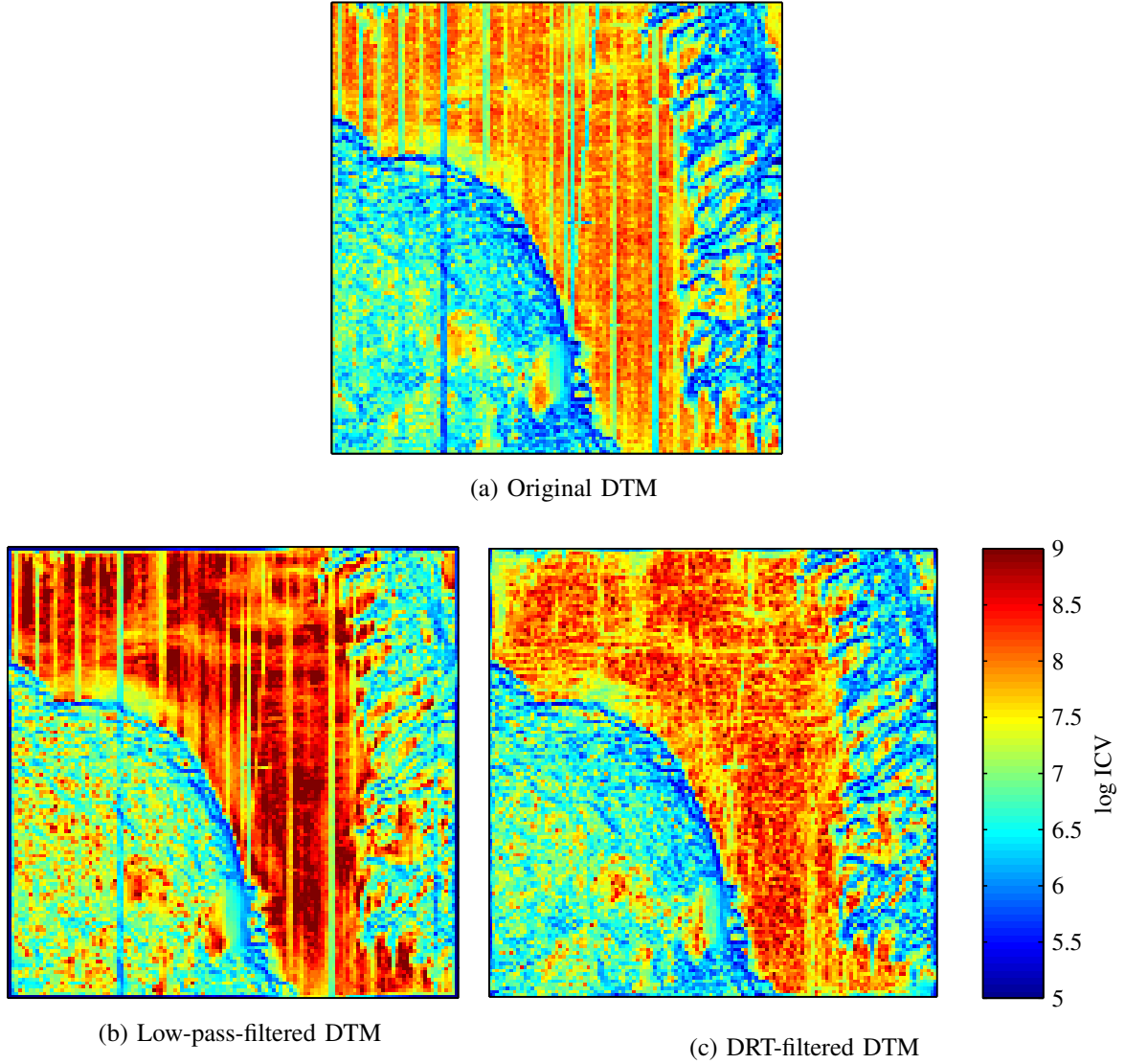

(b) Low-pass-filtered DTM

(c) DRT-filtered DTM

Fig. 17. Inverse coefficient of variation (ICV) computed for area A (cf. Fig. 12 and 13) in 8×8 pixels blocks. The ICV is a measure of the SNR, assuming that the variance of the pixels is due to noise. In (a) and (b) the linear discontinuities in the DTM are seen as regions with low ICV (SNR). In other regions, smoothing (b) produces high ICV values by definition of the ICV. However, in reality, pixel variance is due to both real terrain variations as well as noise, and smoothing affects real features.

of D features. To avoid a potential selection bias caused by performing feature selection and error estimation on the same data set (see, e.g., [58], [59]), we have used the holdout approach. The complete dataset was randomly split into a training dataset and a test dataset, with 75 % of the samples assigned for training. Feature selection was subsequently done using 10-fold cross-validation on the training set only, using the normal density classification accuracy as selection criterion. Following feature selection, the trained classifier was applied to the test data set to get the error estimate. To reduce the variance of the error estimate, the whole process was repeated 20 times (random splits) and the average error computed.

The maximum estimated classification accuracy for various types of features is listed in Table I. For DRT-filtered data, good accuracy (92–95 %) was attained for $S = 32$ using K, H, or Δf, while moderately good accuracy (83–89 %) was attained for $S = 16$. The best results were obtained using uniform LBPs applied to Δf (Fig. 18d). However, one

TABLE I
OVERALL CLASSIFICATION ACCURACY (%)

Features	Raw image		Filtered image	
	16×16	32×32	16×16	32×32
ECDF H and Δf	72.9	82.4	83.4	93.7
LBP (rot. inv.) Δf	55.3	63.3	85.4	91.5
LBP (uniform) Δf	61.6	71.4	88.6	95.2
ECDF $f - \mathcal{T} f$	74.8	75.9	74.2	73.9

should bear in mind that despite the fact that Δf is (approximately) rotation invariant, the distribution of uniform LBPs is not. Fig. 18d also demonstrates the adverse effect of having a low ratio between the number of training samples and the number of features (for $S = 32$). It is evident that noise filtering is necessary to obtain good classification results when using noise-sensitive attributes such as curvature; the improvement is as high as 30 % in some instances (see Fig. 18b). On the other hand, less reliable results may be expected when using the local distributions of the depth residual $f - \mathcal{T} f$ (Fig. 18c) primarily because the result will depend on the particulars of the trend surface $\mathcal{T} f$.

Appendix F contains a more detailed accuracy assessment, including confusion matrices, for classification with invariant geometrical properties and rotation-invariant LBPs, using the optimum number of features.

VI. SUMMARY AND DISCUSSION

The principal aim of this work has been to reduce stripe noise and motion-induced noise in remote sensing data, particularly in MBES bathymetry as a step towards more robust classification of seabed type. This paper has four subthemes: inversion of the GDB (DRT) transform, construction of a linear invertible edge detection operator, noise signature reduction in the GDB transform domain, and finally, the effect of noise reduction on the classification of geomorphology by invariant terrain features.

As an image-transform-based method, the proposed algorithm shares some of the strengths and weaknesses of Fourier domain filtering, but the noise characteristics are very different in the DRT domain. The two first steps of the algorithm, edge detection plus forward DRT, closely parallels a standard method for straight line detection in image analysis, namely to apply a gradient operator, thresholding, and the Hough transform [25] (see, e.g., [60] concerning the relationship between the Hough and Radon transforms). A straight line will correspond to a high intensity spot in the DRT domain irrespective of orientation or other image features. When lines are parallel, as in the present data and much other remote sensing imagery, the spots will align in a single vertical column in the global DRT plot (Fig. 6), but this is merely convenient, not essential. In general one could apply a computational method for finding the high intensity peaks, analogous to the `houghpeaks` function in the Matlab®Image Processing Toolbox [61]. The corresponding stripes may subsequently be suppressed by zeroing the DRT in a small window centered on each peak. The weak point in this approach is the risk of suppressing real linear features in the image, as exemplified by the supposed pipeline in Fig. 12. The proposed denoising algorithm works as a smoothing filter which, applied to the present MBES datasets, suppressed noise better than standard

(a) ECDF (H and Δf)

(b) Rotation-invariant LBPs (Δf)

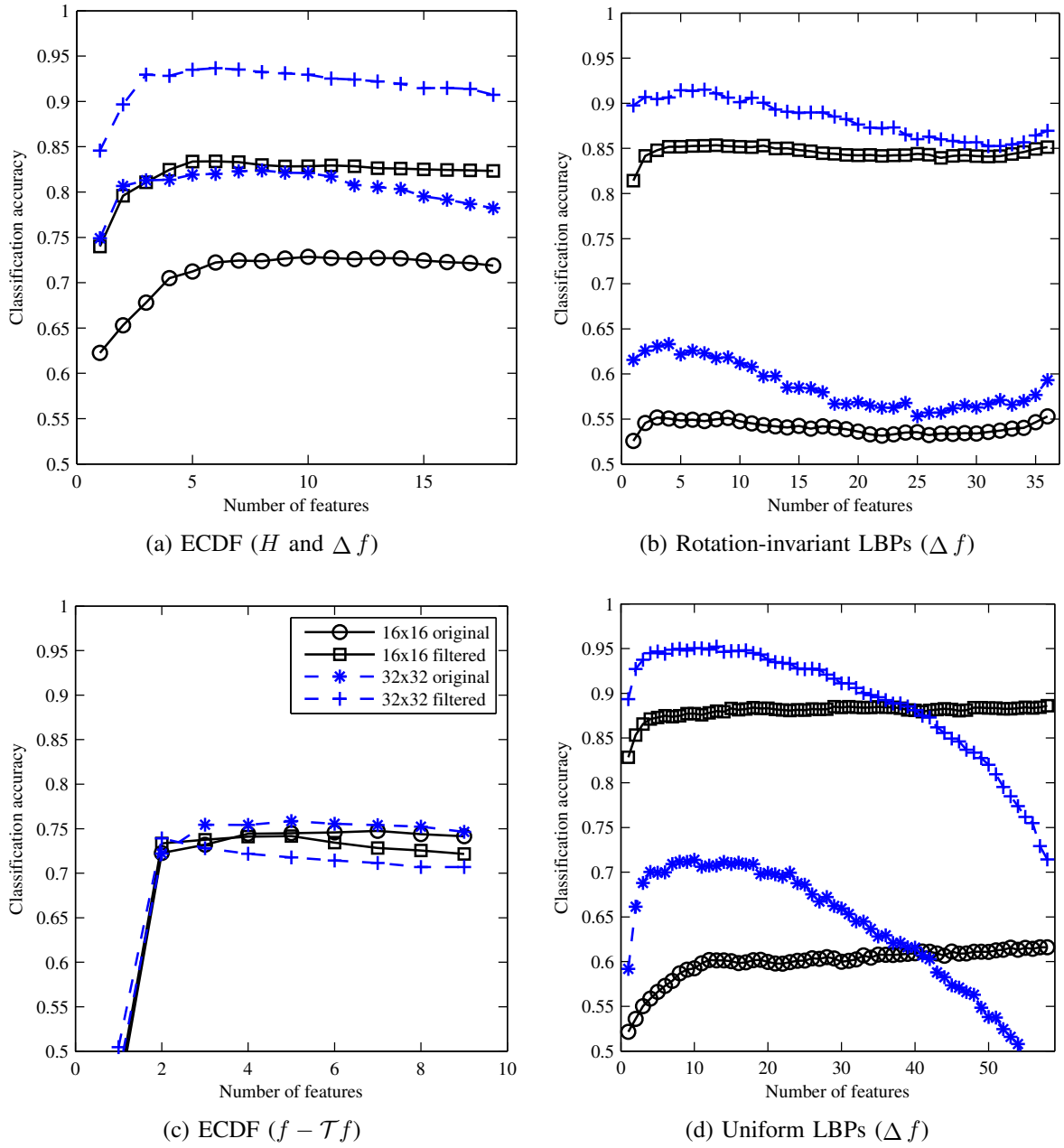(c) ECDF ($f - \mathcal{T}f$)

(d) Uniform LBPs (Δf)

Fig. 18. Learning curves for the terrain classification experiment (three geomorphological classes), before and after the DRT noise filter was applied. The legend in Fig. (c) applies to all four panels. The classification is based on local distributions (ECDFs and LBP histograms) of geometric quantities, in 16×16 and 32×32 pixels blocks.

low-pass filters and also caused less distortion of topographic features. We also found that it was simpler to design a noise filter in the DRT domain than in the Fourier domain (see Section III-B).

Compared with conventional filters in the spatial or Fourier domains, the full DRT denoising algorithm (see Section III-B) is complicated, with more adjustable parameters: the polynomial degree p and downsampling factor D of the approximation algorithm (see Appendix D), the ϵ-factor and dimension M of the modified Laplacian (23), and a stopping criterion for the iterative DRT inverse (residual error threshold or fixed number of iterations). In addition there is some freedom of choice in how to suppress the high-intensity peaks in the DRT domain.

However, little fine-tuning was needed to obtain the results presented here; only the degree p needed to be adjusted in some trials to avoid a low-frequency distortion in the restored DTM. The degree p should anyhow not be too large, so as to avoid fitting the approximation to the noise. A factor $D > 1$ both counters this tendency and reduces computer memory requirements. We think, moreover, that the fixed values of $\epsilon = 10^{-3}$ and $M = 7$ will work in other applications as well.

In addition to image size, the stopping criterion is the critical parameter with respect to computational cost. The tolerance r_k should not be smaller than necessary; $r_k \sim 10^{-2}$ or 10^{-3} produced good results for the present data. The DRT of smaller image blocks can be inverted noniteratively, i.e., by computing a pseudoinverse matrix explicitely. This may speed up execution when an image can be processed in disjoint blocks as in the two examples above (see Fig. 15c and 16d). A short discussion of these aspects can be found in Appendix E. Other iterative inverse methods that were not explored in the present work include the class of algebraic reconstruction techniques (ART), a common alternative to filtered backprojection in, e.g., medical imaging [62]. Another possibility we have not tried is to replace the approximate inverse operator $\boldsymbol{B}$ in (13) with filtered backprojection, and then apply GMRES to obtain an iterative refinement.

The classification experiment demonstrates that segmentation or classification based on local distributions of simple, invariant, second-derivative-based quantities can be an effective means for discriminating between geomorphological types. The mean and Gaussian curvatures have previously proven useful, e.g., for terrain recognition in airplane navigation, partly due to their invariant properties [63]. Curvature is, however, sensitive to noise; the Laplacian is after all used precisely as a noise detector in the DRT filter algorithm. Therefore, it is important to filter the data before classification, as Fig. 18 and Table I show. Only small differences were found when performing classification based on K, H, Δf, or a combination of the three. However, the easily interpreted quantities K and H may be useful in some situations. Indeed, it may be possible to construct features for discriminating between specific morphological types. For example, an even sandy seabed (sloping or not) has curvature $K \approx H \approx 0$, but if water currents deform the seabed by forming parallel ripples, then $K \approx 0$ but $|H| \neq 0$.

While the experimental results shown in this paper are for MBES bathymetry only, we have also used the DRT filter to remove scalloping in SAR images of the Earth's surface. Scalloping is an amplitude modulation that causes parallel stripes in burst mode SAR images [64, Ch. 5.3], particularly wide-swath ScanSAR images of ocean scenes. The procedure was the same as the DENOISE algorithm of Section III-B, minus the DRT domain median filter. (Recall that the median filter was only intended to remove motion-induced noise in MBES data.) The SAR tests suggest that the proposed method may be useful in other remote sensing applications. It is also possible to reverse the DRT domain filter so as to accentuate linear features that are oriented in a particular angular sector, i.e. by zeroing the DRT domain outside the sector before applying the inverse transform. When the DRT filter is applied locally, in sliding windows, this method can be useful for detecting thin structures with a predominant orientation, such as seismic fault lines, fractures, or perhaps blood vessels. In such applications we would use the matrix representation of the DRT inverse as described above.

Another possible development is the use of invertible gradient operators, which may be constructed in the same way as the modified Laplacian, i.e by using convolution with (23) but based instead on the PSFs of, e.g., the Sobel operators. Finally, we note that, in some applications, inversion of the GDB transform can be sped up, and memory consumption

lowered, by using the fact that each quadrant map $\mathcal{R}^i$ is injective, implying that each of the four quadrants of the GDB transform, $\mathcal{R}^1 \boldsymbol{f}, \ldots, \mathcal{R}^4 \boldsymbol{f}$, contains all the information needed to reconstruct $\boldsymbol{f}$ (see Appendix A).

APPENDIX A
INJECTIVITY OF $\mathcal{R}$

The GDB transform $\mathcal{R}$ is injective (one-to-one) if and only if $\mathcal{R} \boldsymbol{f} = 0$ implies that $\boldsymbol{f} = 0$. Suppose that $\boldsymbol{f} \in \mathbb{R}^{2n \times 2n}$, where $n = 2^k$ and $k \in \{0, 1, 2, \ldots\}$. The matrix (image) $\boldsymbol{f}$ is divided into a left half $\boldsymbol{f}^{(L)}$ and a right half $\boldsymbol{f}^{(R)}$, each of size $2n \times n$. We let x_{hs} and y_{hs} denote, in $\boldsymbol{f}^{(L)}$ and $\boldsymbol{f}^{(R)}$ respectively, the sum of pixel values on the graph $D_n(h, s)$, i.e.

$$x_{hs} = \sum_{(i,j) \in D_n(h,s)} f_{ij}^{(L)} \tag{28a}$$

$$y_{hs} = \sum_{(i,j) \in D_n(h,s)} f_{ij}^{(R)}, \tag{28b}$$

where $-2n < h < 2n$ and $0 \le s < n$. Thus $x_{00} + y_{00}$ is the sum of the bottom row, $x_{0,n-1} + y_{n,n-1}$ is the sum of the anti-diagonal, and so on. The zero-padded matrix is

$$\boldsymbol{f}_0 = \begin{bmatrix} \mathbf{0}_{(2n-1)\times n} & \mathbf{0}_{(2n-1)\times n} \\ f_{11} & \cdots & f_{1n} & f_{1,n+1} & \cdots & f_{1,2n} \\ \vdots & \nearrow^{x_{n,n-1}} & \vdots & \vdots & \nearrow^{y_{n,n-1}} & \vdots \\ f_{n1} & f_{nn} & f_{n,n+1} & f_{n,2n} \\ f_{n+1,1} & f_{n+1,n} & f_{n+1,n+1} & f_{n+1,2n} \\ \vdots & \nearrow^{x_{0,n-1}} & \vdots & \vdots & \nearrow^{y_{0,n-1}} & \vdots \\ f_{2n,1} & \xrightarrow{x_{00}} & f_{2n,n} & f_{2n,n+1} & \xrightarrow{y_{00}} & f_{2n,2n} \\ \mathbf{0}_{(2n-1)\times n} & \mathbf{0}_{(2n-1)\times n} \end{bmatrix}.$$

By construction (4), $\mathcal{R}^1 \boldsymbol{f} = 0$ implies that

$$x_{hs} + y_{h+s,s} = x_{hs} + y_{h+s+1,s} = 0. \tag{29}$$

It follows that

$$y_{-2n+1,s} = y_{-2n+2,s} = \ldots = y_{0s} = \ldots = y_{2n-1,s}, \tag{30}$$

i.e. y_{hs} must be independent of h. Since $y_{0s} = -x_{-s-1,s} = 0$, all y_{hs} are zero. Consequently all x_{hs} are also zero by the assumption $\mathcal{R}^1 \boldsymbol{f} = 0$. By (4), the complete graphs $D_{2^{k+1}}(h, s)$ are constructed by recursively combining the graphs D_{2^j} of pairs of adjacent subimages; the subimage size is $2^{k+1} \times 2^j$, for $j = 0, \ldots, k$. The above argument may now be applied to any pair of graphs $D_{2^j}^{(L)}$ and $D_{2^j}^{(R)}$. Hence if x_{hs} and y_{hs} are defined as in (28), with $n = 2^j$, and $x_{hs} + y_{hs} = 0$, then $x_{hs} = 0$ and $y_{hs} = 0$. By induction which terminates when $j = 1$, it follows that $\boldsymbol{f} = 0$.

The same conclusion is reached assuming that $\mathcal{R}^i \boldsymbol{f} = 0$ for $i = 2, 3$, or 4. Each quadrant map $\mathcal{R}^i$ is injective, and each $\mathcal{R}^i \boldsymbol{f}$ contains all information needed to reconstruct $\boldsymbol{f}$.

APPENDIX B
DECOMPOSITION OF THE DRT DOMAIN

The recursive backprojection algorithm of the GDB transform is such that, in the backprojected image, each pixel p is the sum of all DRT points $(\mathcal{R}^i)(h, s)$ whose associated graphs contain p [32]. From the linear algebra viewpoint, the GDB backprojection algorithm is the transpose $\boldsymbol{R}^T$ of the forward transform. To see this, consider that $\boldsymbol{R}^T$ maps between the vector spaces

$$\boldsymbol{R}^T : \mathbb{R}^P \to \mathbb{R}^Q,$$

where $P = 6N^2 - 2N$ and $Q = N^2$ as before. The two spaces have orthonormal basis vectors $\boldsymbol{g}_k$, $k = 1, \ldots, P$ and $\boldsymbol{e}_k$, $k = 1, \ldots, Q$ with unity in the kth component and zeros elsewhere. Let $\Pi_j \subset \{1, 2, \ldots, Q\}$ denote the set of pixels included in the jth graph ($j = 1, \ldots, P$). The components of $\boldsymbol{R}^T$ are

$$\left(\boldsymbol{R}^T\right)_{ij} = \boldsymbol{e}_i^T \boldsymbol{R}^T \boldsymbol{g}_j = \boldsymbol{g}_j^T \boldsymbol{R} \boldsymbol{e}_i$$
$$= \begin{cases} 1 & \text{if } i \in \Pi_j \\ 0 & \text{otherwise.} \end{cases}$$

Hence, for any $\boldsymbol{G} = \sum_{j=1}^{P} G_j \boldsymbol{g}_j \in \mathbb{R}^P$,

$$\left(\boldsymbol{R}^T \boldsymbol{G}\right)_i = \sum_{j=1}^{P} \left(\boldsymbol{R}^T\right)_{ij} G_j = \sum_{j:i\in\Pi_j} G_j,$$

which is the ith pixel of the backprojection of $\boldsymbol{G}$.

The vector space $\mathbb{R}^P$ can be decomposed as the direct sum of the range of $\mathcal{R}$ and its orthogonal complement, $\mathbb{R}^P = \mathcal{I}(\boldsymbol{R}) \oplus \mathcal{I}(\boldsymbol{R})^\perp$. Moreover, $\mathcal{I}(\boldsymbol{R})^\perp$ can be identified with the null space of the Hermitian adjoint $\boldsymbol{R}^+$, and $\boldsymbol{R}^+ = \boldsymbol{R}^T$ since $\boldsymbol{R}$ is real [39]. Thus, $\mathbb{R}^P$ is the orthogonal direct sum

$$\mathbb{R}^P = \mathcal{I}(\mathcal{R}) \oplus \mathcal{N}(\mathcal{R}^+), \tag{31}$$

where $\mathcal{R}^+$ as in the continuous case denotes the backprojection operator/algorithm.

APPENDIX C
ANALYSIS OF THE PRESS INVERSE

The approximate inverse algorithm of [33], $\mathcal{B}$, can be represented by a $Q \times P$ matrix $\boldsymbol{B}$, and the product $\boldsymbol{B}\boldsymbol{R}$ is square ($Q \times Q$). Provided $\boldsymbol{B}\boldsymbol{R}$ is nonsingular, and by defining the residual error matrix $\boldsymbol{E} = \boldsymbol{1}_Q - \boldsymbol{B}\boldsymbol{R}$, the identity $(\boldsymbol{B}\boldsymbol{R})^{-1}\boldsymbol{B}\boldsymbol{R} = \boldsymbol{1}_Q$ may be written

$$(\boldsymbol{1}_Q - \boldsymbol{E})^{-1} \boldsymbol{B}\boldsymbol{R} = \boldsymbol{1}_Q, \tag{32}$$

and $(\boldsymbol{1}_Q - \boldsymbol{E})^{-1} \boldsymbol{B}$ is seen to be a left-inverse of $\boldsymbol{R}$. If the spectral radius of $\boldsymbol{E}$ (largest eigenvalue magnitude) is $\rho(\boldsymbol{E}) < 1$, then [40, Ch. 7.10]

$$(\boldsymbol{1}_Q - \boldsymbol{E})^{-1} = \sum_{k=0}^{\infty} \boldsymbol{E}^k,$$

(a Neumann series) and the left-inverse becomes

$$\boldsymbol{R}_L^{-1} \equiv \left(\boldsymbol{1}_Q + \boldsymbol{E} + \boldsymbol{E}^2 + \ldots\right) \boldsymbol{B}. \tag{33}$$

TABLE II
SPECTRAL RADIUS OF $\boldsymbol{E}$

N	64	128	256	512	1024	2048
ρ	0.838	0.895	0.931	1.027	1.293	1.562
$\log_{10}\sigma$	-14.6	-14.5	-13.9	-14.4	-14.4	-14.4

From (33) follows the recurrence relation for the iteratively inverse DRT of [33], namely

$$\boldsymbol{f}_0 = \boldsymbol{B}\boldsymbol{d} \tag{34a}$$

$$\boldsymbol{f}_{k+1} \equiv \left(\sum_{j=0}^{k+1} \boldsymbol{E}^j\right)\boldsymbol{B}\boldsymbol{d} = \boldsymbol{f}_k + \boldsymbol{B}\left(\boldsymbol{d} - \boldsymbol{R}\boldsymbol{f}_k\right). \tag{34b}$$

This scheme involves only matrix-vector multiplications which are implemented by the fast, recursive algorithms $\mathcal{R}$ and $\mathcal{B}$. There is no need to compute explicitly the matrices $\boldsymbol{B}$ and $\boldsymbol{E}^j$, which is practically impossible for large images.

Although the iterative inverse (34) works very well for test images (Fig. 3) and random images [33], numerical analysis suggests that the condition $\rho(\boldsymbol{E}) < 1$ does not hold for image sizes $N \geq 512$. The spectral radius of $\boldsymbol{E}$, estimated using Arnoldi iteration [65]–[67], is shown in TABLE II. The Press approximate inverse algorithm involves a high-pass filter which we have here implemented using symmetric boundary conditions. Using other boundary conditions (zero-padding, periodicity, or replication of nearest value) results in slightly different values but does not alter the conclusion. The eigenpairs $(\lambda, \boldsymbol{g})$ estimated in this way are good approximations of true eigenpairs at the outer edge of the spectrum; for $(\lambda, \boldsymbol{g})$ corresponding to ρ in TABLE II, the residual is $\sigma \equiv \|\boldsymbol{E}\boldsymbol{g} - \lambda\boldsymbol{g}\|_2 \sim 10^{-14}$ (in operator 2-norm), which is close to machine precision. This means that the Neumann series $\sum_{j=0}^{\infty} \boldsymbol{E}^j$ does not converge, but the inverse $(\boldsymbol{1}_Q - \boldsymbol{E})^{-1} = (\boldsymbol{B}\boldsymbol{R})^{-1}$ still exists provided only that $\lambda = 1$ is not an eigenvalue of $\boldsymbol{E}$. If $\boldsymbol{E}$ has s distinct eigenvalues $\lambda_1, \ldots, \lambda_s$, Jordan decomposition gives $\boldsymbol{E} = \boldsymbol{C}\,\mathrm{diag}\,(\boldsymbol{J}_{\lambda_1,1}, \ldots, \boldsymbol{J}_{\lambda_s,t_s})\boldsymbol{C}^{-1}$, where $\boldsymbol{C}$ is a non-singular $Q \times Q$ matrix and there are $t_j = \dim_{\mathbb{C}} \mathcal{N}(\boldsymbol{E} - \lambda_j \boldsymbol{I}_Q)$ Jordan blocks for eigenvalue λ_j. The approximation error for the kth iterate in (34) may then be written

$$\begin{aligned}\boldsymbol{f} - \boldsymbol{f}_k &= \boldsymbol{f} - (\boldsymbol{1}_Q - \boldsymbol{E})^{-1}\left(\boldsymbol{1}_Q - \boldsymbol{E}^{k+1}\right)\boldsymbol{B}\boldsymbol{d} \\ &= \boldsymbol{P}^{-1}\,\mathrm{diag}\,\left(\boldsymbol{J}_{\lambda_1,1}^{k+1}, \ldots, \boldsymbol{J}_{\lambda_s,t_s}^{k+1}\right)\boldsymbol{P}\boldsymbol{f},\end{aligned} \tag{35}$$

where $\boldsymbol{P} \equiv \boldsymbol{C}^{-1}\boldsymbol{B}\boldsymbol{R} \approx \boldsymbol{C}^{-1}$. For $|\lambda| < 1$, $\lim_{k \to \infty} \boldsymbol{J}_{\lambda,i}^k = \boldsymbol{0}$ [40, Ch. 7.10], and in fact the spectrum of $\boldsymbol{E}$ is concentrated in the range $0 < |\lambda| < 1$. For $N = 512$, only four eigenvalues found by Arnoldi iteration have magnitude $|\lambda| > 1$. With respect to a basis obtained by the coordinate transformation matrix $\boldsymbol{P}^{-1}$, only the relatively few components of $\boldsymbol{f}$ corresponding to Jordan blocks with $|\lambda| \geq 1$ will not be correctly reconstructed in the limit $k \to \infty$, and the RMS error will be small. However, the amplification of the blocks $\boldsymbol{J}_{\lambda,*}^k$ for $|\lambda| > 1$ destroys convergence for sufficiently many iterations k. The reconstruction algorithm (34) will therefore not work in all instances. This is demonstrated in Fig. 4 using an eigenimage of $\boldsymbol{E}$ for $N = 512$ with $|\lambda| = \rho = 1.027$.

APPENDIX D

TREND COMPUTATION

This appendix summarizes the approximation algorithm $\mathcal{T}(p, D)$ (cf. Section III-B). The image/DTM $\boldsymbol{f} \in \mathbb{R}^{N \times N}$ is downsampled to size $N/D \times N/D$, and interpolated back to size $N \times N$ using a Chebyshev polynomial tensor product basis of total degree p. Since N is a power of 2, D is also conveniently taken to be a power of 2 ($1 \leq D < N$). We let $\{ T_k(x) \}_{k=0}^{\infty}$ denote the set of Chebyshev polynomials of the first kind on the interval $[-1, 1]$ [68, Ch. 22]. Treating $\boldsymbol{f}$ as a set of discrete samples of a function $f(x, y)$ on the square $[-1, 1] \times [-1, 1]$, interpolation is accomplished via the expansion

$$f(x, y) \approx \sum_{0 \leq j+k \leq p} w_{jk} T_j(x) T_k(y). \tag{36}$$

There are $\eta_p = \sum_{d=0}^{p} \binom{d+1}{d} = (p+2)(p+1)/2$ terms in this series. The coefficients w_{jk} are determined by a least squares fit of $N/D \times N/D$ samples. This can be implemented by multiplication with a generalized matrix inverse P of size $\eta_p \times (N/D)^2$, which may be precomputed (i.e. need only be computed once). The Chebyshev polynomials $T_0(x), \ldots, T_p(x)$ are also precomputed at N evenly spaced points on the interval $[-1, 1]$, and stored in a matrix $\boldsymbol{C} \in \mathbb{R}^{N \times (p+1)}$. The latter is accomplished by the following Matlab function given in Listing 1.

Listing 1. Compute Chebyshev polynomials (Matlab)

```
function C = chebpol(N,p)
x = -1:2/(N-1):1;
C = zeros(N,p+1);
C(:,1) = ones(N,1);
if p > 1
  C(:,2) = x';
end
if p > 2
  for k = 3:p+1
    C(:,k) = 2*x'.*C(:,k-1)-C(:,k-2);
  end
end
```

The following code fragment given in Listing 2 generates the Chebyshev polynomials C and the generalized inverse P for fixed N, p, and D.

Listing 2. Compute Chebyshev polynomials and the generalized inverse (Matlab)

```
M = N/D;
Ep = (p+1)*(p+2)/2;
C = chebpol(M,p);
L = single(zeros(M, M, Ep));
[j,k] = meshgrid(0:p,0:p);
for n = [find(j+k<=p)';1:Ep]
  L(:,:,n(2)) = ...
  C(:,j(n(1))+1)*C(:,k(n(1))+1)';
end
PI = pinv(reshape(L, M^2, Ep));
```

184

```
if D ~= 1
  C = chebpol(N,p);
end
```

If $\boldsymbol{g} \in \mathbb{R}^{N/D \times N/D}$ is the downsampled image, then the coefficients w_{jk} and trend $\boldsymbol{t} \in \mathbb{R}^{N \times N}$ may be obtained in Listing 3.

Listing 3. Compute approximate image by Chebyshev interpolation (Matlab)

```
w = PI*g(:);
t = zeros(N);
for n = [find(j+k<=p)';1:Ep]
  t = t + w(n(2)) * C(:,j(n(1))+1)...
          * C(:,k(n(1))+1)';
end
t = double(t);
```

Depending on N/D and p, this approach may require too much memory to compute and store the $(p+2)(p+1)/2 \times (N/D)^2$ matrix P. This simple solution therefore cannot be considered general, but has nevertheless worked well in our applications (both MBES and SAR) for image sizes up to $N = 2048$ (also with $D = 1$ and $p \sim 10$).

APPENDIX E
NOTE ON COMPUTING TIME

The stopping criterion influences the execution time of the DRT inverse, and the tolerance r_k (relative residual norm, Section II-C4) should not be set smaller than necessary. The code for this paper was written in C as Matlab executable (MEX) files and ran on a single 2.93 GHz CPU (core) on a Linux computer with 8 GB RAM. Less than five iterations with GMRES usually sufficed to obtain a good result, and for image size $N = 1024$ ($\sim 10^6$ unknowns and $\sim 6 \cdot 10^6$ equations) good results were obtained in about 20 s execution time, although the rate of convergence may depend on the data.

The results shown in Fig. 15c and Fig. 16d suggest that images can be processed in disjoint blocks, and program execution may then be considerably faster. The main reason is that an explicit DRT pseudoinverse matrix may be precomputed; inversion of each block is then carried out non-iteratively, by matrix multiplication, yet with arbitrarily low tolerance r_k. So for small block sizes ($N \leq 64$) we have computed the matrix representations of the three iterative algorithms of Section II-C, which is not feasible for large images due to the memory requirement. The solution to the normal equations (7) can be obtained by multiplication with the Moore-Penrose pseudoinverse. The cost of inversion by matrix multiplication is independent of the accuracy with which the matrix representations of the Press and GMRES inverses are computed (and the Press inverse is an exact left-inverse for $N < 512$). The cost of obtaining the (pseudo-)inverse matrices does depend on the accuracy, but this is a one-time-only operation.

The time for processing the $(1024/32)^2$ blocks of the DTM shown in Fig. 16d was about 9.1 s (only one quadrant of the image is shown in the figure). This includes not just the time to invert the block DRTs, but also the time to compute the forward DRTs and apply the DRT domain noise suppression filter. In this example the forward DRT was also computed using a matrix representation, i.e., of the GDB algorithm. The number of additions in the GDB

TABLE III
ACCURACY ASSESSMENT: ECDF (H AND Δf)

Class	Assignments			Total	Omiss-ions	Comm-issions	MA
	#1	#2	#2				
#1	50.1	9.5	3.4	63	20.5%	21.7%	65.4%
#2	9.3	55.1	5.6	70	21.4%	16.4%	67.7%
#3	4.4	2.0	52.6	59	10.8%	15.3%	77.5%

(a) Original image

Class	Assignments			Total	Omiss-ions	Comm-issions	MA
	#1	#2	#2				
#1	57.5	4.1	1.4	63	8.7%	4.7%	87.2%
#2	2.9	64.2	2.9	70	8.3%	7.0%	85.7%
#3	0.1	0.8	58.1	59	1.4%	7.3%	91.9%

(b) DRT-filtered image

algorithm is $4N^2 \log_2 N$, which for large N is much lower than for conventional $\mathcal{O}(N^3)$ DRTs (e.g., discrete approximations to the classical Radon transform). For $N = 128$, we have $4N^2 \log_2 N = 458752$, while there are 8388608 non-zero elements (ones) in the matrix representation of the GDB algorithm, which gives a ratio of $1 : 18$. For $N \leq 64$, however, matrix multiplication using an efficient low-overhead library may be faster than the recursive algorithm.

APPENDIX F
CLASSIFICATION ACCURACY ASSESSMENT

This appendix contains supplementary data about the classification experiments, cf. Section V-C and Table I. Table III shows the accuracy assessment for classification with ECDFs applied to invariant geometrical properties (minimum curvature and Laplacian combined) in 32×32 pixels blocks. The six best features were employed, obtained using forward feature selection, cf. Fig. 18. The experiment, including feature selection, was run 20 times, and Table III shows the average values. Hence, the values for assignments are also fractional. The mapping accuracy (MA) for a class ω is defined as

$$\text{MA} = \frac{N_{\text{cor.}}(\omega)}{N_{\text{cor.}}(\omega) + N_{\text{omi.}}(\omega) + N_{\text{com.}}(\omega)}, \tag{37}$$

where $N_{\text{cor.}}(\omega)$ is the number of correctly classified observations, and $N_{\text{omi.}}(\omega)$ and $N_{\text{com.}}(\omega)$ are the number of omissions and commissions, respectively, in class ω.

Table IV shows the corresponding results for classification with rotation-invariant LBPs in 32×32 pixels blocks using the five best features.

TABLE IV

ACCURACY ASSESSMENT: ROTATION-INVARIANT LBPs (Δf)

Class	Assignments			Total	Omiss-ions	Comm-issions	MA
	#1	#2	#2				
#1	28.4	16.4	18.2	63	54.9%	43.8%	31.4%
#2	15.7	49.5	4.8	70	29.4%	29.1%	54.8%
#3	11.9	4.0	43.1	59	26.9%	39.1%	52.6%

(a) Original image

Class	Assignments			Total	Omiss-ions	Comm-issions	MA
	#1	#2	#2				
#1	63.0	0	0	63	0%	0.1%	99.9%
#2	0.1	61.5	8.4	70	12.1%	12.7%	78.1%
#3	0.0	8.9	50.1	59	15.1%	14.3%	74.4%

(b) DRT-filtered image

ACKNOWLEDGMENT

REFERENCES

[1] M. di Bisceglie, R. Episcopo, C. Galdi, and S. Liberata, "Destriping MODIS data using overlapping field-of-view method," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 2, pp. 637–651, Feb. 2009.

[2] M. Bouali and S. Ladjal, "Toward optimal destriping of MODIS data using a unidirectional variational model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 8, pp. 2924–2935, Aug. 2011.

[3] H. Carfantan and J. Idier, "Statistical linear destriping of satellite-based pushbroom-type images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 1860–1871, Apr. 2010.

[4] X. Lurton, *An Introduction to Underwater Acoustics: Principles and Applications*, 2nd ed. Berlin, Heidelberg: Springer, 2010.

[5] J. E. Hughes Clarke, L. A. Mayer, and D. E. Wells, "Shallow-water imaging multibeam sonars: A new tool for investigating seafloor processes in the coastal zone and on the continental shelf," *Marine Geophysical Researches*, vol. 18, no. 6, pp. 607–629, Dec. 1996.

[6] M. J. Smith and C. D. Clark, "Methods for the visualization of digital elevation models for landform mapping," *Earth Surface Processes and Landforms*, vol. 30, no. 7, pp. 885–900, Jul. 2005.

[7] J. S. Weszka, C. R. Dyer, and A. Rosenfeld, "A comparative study of texture measures for terrain classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, no. 4, pp. 269–285, Apr. 1976.

[8] U. C. Herzfeld and C. A. Higginson, "Automated geostatistical seafloor classification—principles, parameters, feature vectors, and discrimination criteria," *Computers & Geosciences*, vol. 22, no. 1, pp. 35–52, Feb. 1996.

[9] M. F. J. Wilson, B. O'Connel, C. Brown, J. C. Guinan, and A. J. Grehan, "Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope," *Marine Geodesy*, vol. 30, no. 1-2, pp. 3–35, May 2007.

[10] J. E. Hughes Clarke, "Dynamic motion residuals in swath sonar data: Ironing out the creases," *Int. Hydrographic Review*, vol. 4, no. 1, pp. 6–23, 2003.

[11] L. Hellequin, J.-M. Boucher, and X. Lurton, "Processing of high-frequency multibeam echo sounder data for seafloor characterization," *IEEE J. Ocean. Eng.*, vol. 28, no. 1, pp. 78–89, Jan. 2003.

[12] R. E. Crippen, "A simple spatial filtering routine for the cosmetic removal of scan-line noise from Landsat TM P-tape imagery," *Photogramm. Eng. Remote Sens.*, vol. 55, no. 3, pp. 327–331, Mar. 1989.

[13] D. Helder, B. Quirk, and J. Hood, "A technique for the reduction of banding in Landsat thematic mapper images," *Photogramm. Eng. Remote Sens.*, vol. 58, no. 10, pp. 1425–1431, Oct. 1992.

[14] J.-J. Pan and C.-I. Chang, "Destriping of Landsat MSS images by filtering techniques," *Photogramm. Eng. Remote Sens.*, vol. 58, no. 10, pp. 1417–1423, Oct. 1992.

[15] V. R. Algazi and G. E. Ford, "Radiometric equalization of nonperiodic striping in satellite data," *Comput. Graph. Image Process.*, vol. 16, no. 3, pp. 287–295, 1981.

[16] G. Corsini, M. Diani, and T. Walzel, "Striping removal in MOS-B data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1439–1446, May 2000.

[17] B. K. P. Horn and R. J. Woodham, "Destriping Landsat MSS images by histogram modification," *Comput. Graph. Image Process.*, vol. 10, no. 1, pp. 69–83, 1979.

[18] M. Wegener, "Destriping multiple detector imagery by improved histogram matching," *Int. J. Remote Sens.*, vol. 11, no. 5, pp. 859–875, May 1990.

[19] F. L. Gadallah and F. Csillag, "Destriping multidetector imagery with moment matching," *Int. J. Remote Sens.*, vol. 21, no. 12, pp. 2505–2511, Aug. 2000.

[20] H. Shen, W. Jiang, H. Zhang, and L. Zhang, "A piece-wise approach to removing the nonlinear and irregular stripes in MODIS data," *Int. Journal of Remote Sensing*, vol. 35, no. 1, pp. 44–53, Jan. 2014. [Online]. Available: http://dx.doi.org/10.1080/2150704X.2013.860564

[21] P. Rakwatin, W. Takeuchi, and Y. Yasuoka, "Stripe noise reduction in MODIS data by combining histogram matching with facet filter," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1844–1856, Jun. 2007.

[22] H. Shen and L. Zhang, "A MAP-based algorithm for destriping and inpainting of remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1492–1502, May 2009.

[23] F. Tsai and W. W. Chen, "Striping noise detection and correction of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 12, pp. 4122–4131, Dec. 2008.

[24] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 679–698, Nov. 1986.

[25] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Upper Saddle River, NJ: Pearson/Prentice Hall, 2008.

[26] D. Wilken, P. Feldens, T. Wunderlich, and C. Heinrich, "Application of 2D Fourier filtering for elimination of stripe noise in side-scan sonar mosaics," *Geo-Marine Lett*, vol. 32, no. 4, pp. 337–347, Aug. 2012.

[27] J. Chen, Y. S. an H. Guo, W. Wang, and B. Zhu, "Destriping CMODIS data by power filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 2119–2124, Sep. 2003.

[28] R. N. Bracewell and A. C. Riddle, "Inversion of fan-beam scans in radio astronomy," *Astrophys. J.*, vol. 150, pp. 427–434, Nov. 1967.

[29] F. Natterer and F. Wübbeling, *Mathematical Methods in Image Reconstruction*. Philadephia, PA: Soc. Ind. Appl. Math., 2001.

[30] G. Beylkin, "Discrete Radon transform," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 2, pp. 162–172, Feb. 1987.

[31] W. A. Götz and H. J. Druckmüller, "A fast digital Radon transform—an efficient means for evaluating the Hough transform," *Pattern Recognition*, vol. 29, no. 4, pp. 711–718, Apr. 1996.

[32] M. L. Brady, "A fast discrete approximation algorithm for the Radon transform," *SIAM J. Comput.*, vol. 27, no. 1, pp. 107–119, 1998.

[33] W. H. Press, "Discrete Radon transform has an exact, fast inverse and generalizes to operations other than sums along lines," *Proc. Natl. Acad. Sci. USA (PNAS)*, vol. 103, no. 51, pp. 19 249–19 254, Dec. 2006.

[34] K. T. Smith and F. Keinert, "Mathematical foundations of computed tomography," *Appl. Opt.*, vol. 24, no. 23, pp. 3950–3957, Dec. 1985.

[35] L. A. Shepp and B. F. Logan, "The Fourier reconstruction of a head section," *IEEE Trans. Nucl. Sci.*, vol. 21, no. 3, pp. 21–43, Jun. 1974.

[36] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Comm. ACM*, vol. 15, no. 1, pp. 11–15, Jan. 1972.

[37] C. C. Paige and M. A. Saunders, "LSQR: An algorithm for sparse linear equations and sparse least squares," *ACM. Trans. Math. Softw.*, vol. 8, no. 1, pp. 43–71, Jun. 1982.

[38] D. C.-L. Fong and M. Saunders, "LSMR: An iterative algorithm for sparse least-squares problems," *SIAM. J. Sci. Comput.*, vol. 33, no. 5, pp. 2950–2971, 2011.

[39] D. G. Luenberger, *Optimization by Vector Space Methods*, ser. Series in decision and control. New York, NY: John Wiley & Sons, 1969.

[40] C. D. Meyer, *Matrix analysis and applied linear algebra*. Philadephia, PA: Soc. Ind. Appl. Math. (SIAM), 2000, ch. 7.10.

[41] I. C. F. Ipsen and C. D. Meyer, "The idea behind Krylov methods," *Amer. Math. Monthly*, vol. 105, no. 10, pp. 889–899, Dec. 1998.

[42] Y. Saad and M. H. Schultz, "GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems," *SIAM. J. Sci. Statist. Comput.*, vol. 7, no. 3, pp. 856–869, Jul. 1986.

[43] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, ser. Frontiers in applied mathematics. Philadephia, PA: Soc. Ind. Appl. Math., 1995, vol. 16.

[44] B. Jähne, *Digital Image Processing*, 6th revised and extended ed. Berlin Heidelberg New York: Springer, 2005.

[45] M. Einsiedler and T. Ward. (2015, Mar.) Functional analysis, spectral theory, and applications. Draft version. [Online]. Available: https://dl.dropboxusercontent.com/u/2098511/FAnotes.pdf

[46] M. P. do Carmo, *Differential geometry of curves and surfaces*. Englewood Cliffs, NJ: Prentice-Hall, 1976.

[47] T. F. Chan and J. Shen, *Image processing and analysis: variational, PDE, wavelet, and stochastic methods*. Philadelphia, PA: SIAM, 2005.

[48] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, Jan. 1996.

[49] L. Wang and D. C. He, "Texture classification using texture spectrum," *Pattern Recognition*, vol. 23, no. 8, pp. 905–910, 1990.

[50] T. Ojala and M. Pietikäinen, "Unsupervised texture segmentation using feature distributions," *Pattern Recognition*, vol. 32, no. 3, pp. 477–486, 1999.

[51] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.

[52] M. Pietikäinen, T. Ojala, and Z. Xu, "Rotation-invariant texture classification using feature distributions," *Pattern Recognition*, vol. 33, no. 1, pp. 43–52, 2000.

[53] A. Vedaldi and B. Fulkerson, "VLFeat: An open and portable library of computer vision algorithms," http://www.vlfeat.org/, 2008.

[54] M. Hovland, "Elongated depressions associated with pockmarks in the western slope of the Norwegian Trench," *Marine Geology*, vol. 51, no. 1-2, pp. 35–46, Feb. 1983.

[55] R. Bøe, L. Rise, and D. Ottesen, "Elongate depressions on the southern slope of the Norwegian Trench (Skagerrak): morphology and evolution," *Marine Geology*, vol. 146, no. 1-4, pp. 191–203, Apr. 1998.

[56] G. Dahlquist and Å. Björck, *Numerical Methods*, ser. Dover Books on Mathematics. Dover Publications, 2003.

[57] J. E. Nichol and V. Vohora, "Noise over water surfaces in Landsat TM images," *Int. Journal of Remote Sensing*, vol. 25, no. 11, pp. 2087–2093, Jun. 2004.

[58] H. Schulerud and F. Albregtsen, "Many are called, but few are chosen. feature selection and error estimation in high dimensional spaces," *Comput. Methods Programs Biomed.*, vol. 73, no. 2, pp. 91–99, Feb. 2004.

[59] C. Ambroise and G. J. McLachlan, "Selection bias in gene extraction on the basis of microarray gene-expression data," *Proc. Natl. Acad. Sci. USA (PNAS)*, vol. 99, no. 10, pp. 6562–6566, May 2002.

[60] M. van Ginkel, C. L. Luengo Hendriks, and L. J. van Vliet, "A short introduction to the Radon and Hough transforms and how they relate to each other," TU Delft, Quantitative Imaging Group Technical Report Series QI-2004-01, 2004. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.2.9419

[61] HOUGHPEAKS. MathWorks Matlab R2013b Documentation. Image processing toolbox. [Online]. Available: http://www.mathworks.se/help/images/ref/houghpeaks.html

[62] R. Gordon, R. Bender, and G. T. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography," *J. Theoret. Biol.*, vol. 29, no. 3, pp. 471–481, Dec. 1970.

[63] D. B. Goldgof, T. S. Huang, and H. Lee, "A curvature-based approach to terrain recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 11, pp. 1213–1217, Nov. 1989.

[64] G. Franchescetti and R. Lanari, *Synthetic Aperture Radar Processing*. Boca Raton, FL: CRC Press, 1999.

[65] R. B. Lehoucq and D. C. Sorensen, "Deflation techniques for an implicitly restarted Arnoldi iteration," *SIAM. J. Matrix Anal. & Appl.*, vol. 17, no. 4, pp. 789–821, Oct. 1996.

[66] R. B. Lehoucq, D. C. Sorensen, and C. Yang. (1998) ARPACK Users' Guide: Solution of Large Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods. Web version available. [Online]. Available: http://www.caam.rice.edu/software/ARPACK/UG/ug.html

[67] EIGS. MathWorks Matlab R2012b Documentation. [Online]. Available: http://www.mathworks.com/help/matlab/ref/eigs.html

[68] U. W. Hochstrasser, "Orthogonal polynomials," in *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, M. Abramowitz and I. A. Stegun, Eds. National Bureau of Standards, 1964.

[69] P. Wessel and W. H. F. Smith, "New, improved version of Generic Mapping Tools released," *EOS Trans. AGU*, vol. 79, no. 47, p. 579, Nov. 1998.