2024-07-25

Information Geometric Approaches to Optimal Experimental Design and Reduced-Order Modeling in Ocean Acoustics

Jay C. Spendlove
Brigham Young University

Information Geometry Approaches to Optimal Experimental Design and Reduced-Order Modeling in Ocean Acoustics

Jay C. Spendlove

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Mark K. Transtrum, Chair
Tracianne B. Neilsen
Jared P. Whitehead

Department of Physics and Astronomy

Brigham Young University

ABSTRACT

Information Geometry Approaches to Optimal Experimental Design and Reduced-Order Modeling in Ocean Acoustics

Jay C. Spendlove
Department of Physics and Astronomy, BYU
Master of Science

This thesis demonstrates the application of information geometry to problems in underwater acoustics. Information geometry combines the fields of information theory and differential geometry by interpreting a multiparameter model as a Riemannian manifold in an ambient data space. Information geometry tools are especially powerful in context of problems of experimental design and model selection in ocean acoustics. The application area specifically considered in this research is geoacoustic inversion, where seabed parameter values are inferred from acoustical data. Chapter 2 contains a paper submitted to the Journal of Theoretical and Computational Acoustics, which introduces information geometry tools such as the model manifold and Fisher information in context of a review of work in underwater acoustics doing parameter sensitivity analysis. An example constructing model manifolds for a sound propagation model is given in the second half of the paper. Chapter 3 contains a paper submitted to the Journal of the Acoustical Society of America, Express Letters, which constructs model manifolds for a sound propagation model and compares the information content of absolute and relative transmission loss in regards to seabed parameters, demonstrating how information geometry can be use to inform experimental design. This thesis contains the initial application of information geometry to ocean acoustics, with many more advances that can be pursued in future work.

Keywords: information geometry; underwater acoustics; model manifold; sound propagation; Fisher information; sensitivity analysis; Cramér-Rao bounds; geoacoustic inversion; reduced-order modeling; optimal experimental design; transmission loss

ACKNOWLEDGMENTS

I'd like to thank my advisors, Dr. Transtrum and Dr. Neilsen for all their help and support in my Masters. They have always been there to answer my questions and help me figure out what to try next in my research, and I have learned so much from them! I'd also like to thank my family for their support, especially my wife, Mauri, who has been a constant strength to me.

I'd also like to acknowledge the Office of Naval Research for their support of this work under contract number N00014-21-S-B001.

Contents

List of Figures

List of Tables

Chapter 1

Overview

Sound propagation in the ocean is incredibly complex, but important to understand for countless applications, including naval, ecological, and oceanographic applications. One complexity associated with underwater sound propagation is that the sound speed of the water, which acts like the index of refraction in optics, is highly variable and dependent on factors including pressure, temperature, and salinity. Additional complexity is introduced in a shallow ocean environment ($<$200 m depth), where ocean floor composition can have a large impact on sound propagation through bottom interactions. For example, seafloor compositions that are more sandy tend to be more reflective, while muddy seafloor compositions tend to be more absorptive. Having an accurate understanding of the seafloor composition is therefore important in guaranteeing accuracy in the modeling of underwater sound propagation in a shallow ocean environment.

Due to the expense and difficulty of experimentally determining seafloor composition via, for example, core sampling, a research field of great interest is the inverse problem of geoacoustic inversion, whereby seabed properties are inferred from acoustical data. Seabed properties of interest could include sediment layer height, density, sound speed, and attenuation. Acoustical data can come from active sonar sources (e.g., pings) or passive sonar sources (e.g. ship noise, wind noise). A review article by Chapman and Shang [4] provides an in depth review of developments in the area of geoacoustic inversion.

However, the accuracy of geoacoustic inversion depends significantly on model selection. Many methods for geoacoustic inversion require that you select a model for the ocean environment, such as how many sediment layers to include in the model. However, if the model selected cannot accurately represent the acoustics of the experimental setup being considered, geoacoustic inversion can lead to the inference of inaccurate parameter values; this phenomenon is called model mismatch. One specific challenge is when the model includes superfluous parameters that cannot be effectively learned from the acoustical data. These parameters are referred to as “sloppy" or unidentifiable parameters because the data and model do not contain adequate information about them. In contrast, “stiff" or identifiable parameters are those about which the data and model does contain information. For example, high frequency sound typically interacts with only the top sediment layer. Thus, if the experimental setup utilizes a high frequency sound source and the model chosen has multiple (say, three) sediment layers, acoustical data may contain sufficient information to infer sediment parameters for the top layer but not for the bottom two sediment layers. These lower sediment layer parameters are sloppy. Therefore, developing tools to enable model selection by determining what proposed model parameters are sloppy is beneficial. These tools enable answering questions such as “How many layers need to be included in this model in order to accurately represent the acoustics?" The field of information geometry provides tools for answering these questions.

Information geometry combines the fields of information theory and differential geometry, interpreting a model as a manifold, or surface. The geometry of this “model manifold" gives insight into which model parameters, given some experimental setup, are sloppy and which are stiff, and can thus be learned via geoacoustic inversion. The model manifold can inform the choice of a model to use for geoacoustic inversion by identifying progressively simpler models, associated with the finite boundaries of the model manifold, which may better represent the data. Additionally, because the model output, and thus the model manifold, depends on the experimental setup, the model manifold can be used as a tool to identify the optimal experimental setup that will provide data with the most information about parameters of interest. A model manifold containing more information about parameters manifests as a wider or more “spread out" model manifold. Experimental design parameters that could be considered are hydrophone (underwater microphone) depth, source range (how far away the hydrophone is from the sound source?), and source frequency.

This Master’s research has focused on applying information geometry methods to different problems in ocean acoustics. The scope of my contributions to the field have been to:

1. write a tutorial review paper introducing information geometry concepts such as the model manifold in context of parameter sensitivity analysis methods previously applied in ocean acoustics, and

2. provide demonstrations constructing model manifolds for underwater sound propagation models and showing how the model manifold can be used to address questions of experimental design and model selection.

In this thesis, Chapter 2 contains a review tutorial paper submitted to the Journal of Theoretical and Computational Acoustics (JTCA), introducing information geometry tools in context of a review of literature performing parameter sensitivity analyses in ocean acoustics (item #1). These information geometry tools are then demonstrated for the case of a simple sound propagation model, motivating

starting cases for use of information geometry in model selection in underwater acoustics (item #2). Chapter 3 contains a paper submitted to the Journal of the Acoustical Society of America, Express Letters (JASA-EL) which compares the information content of absolute and relative transmission loss data for seabed sediment parameters using model manifolds, demonstrating how information geometry methods can be utilized for experimental design (item #2). Chapter 4 contains a summary of contributions discussed in this thesis and future work that can be pursued to build on these efforts to apply information geometry methods to ocean acoustic modeling, including a summary of current work training neural network surrogates for the model manifold for application of powerful differential geometry tools.

Chapter 2

The outlook of information geometry for underwater acoustics applications

This chapter includes the review tutorial paper submitted to the Journal of Theoretical and Computational Acoustics (JTCA) entitled *The outlook of information geometry for underwater acoustics applications*, with authors Jay C. Spendlove, Michael C. Mortenson, Tracianne B. Neilsen, and Mark K. Transtrum. The first six sections of this paper function as the technical introduction to this thesis, introducing key concepts such as Fisher information and Cramér-Rao bounds (Sec. 2.3), and information geometry tools such as the model manifold (Sec. 2.6), as well as providing a review of relevant literature about parameter sensitivity analyses in underwater acoustics and information geometry applied to sloppy models (Sec. 2.2; Sec. 2.4; Sec. 2.5).

Section 2.7 of the paper provides a tutorial demonstrating the construction of various model manifolds to the Pekeris waveguide sound propagation model, demonstrating the application of information geometry in ocean acoustics. Introducing these information geometry concepts to the field of underwater acoustics, in context of other methods that have been utilized in the past, is the

primary purpose of this paper. Additionally, Appendices A and B of this thesis are from this paper, delineating the intricacies of the Pekeris waveguide implementation and deriving the “hard bottom" limit of the Pekeris waveguide, which is empirically observed in the model manifolds created in this paper.

Abstract

This tutorial demonstrates the use of information geometry tools in analyzing environmental parameter sensitivities in underwater acoustics. Sensitivity analyses quantify how well data can constrain model parameters, with application to inverse problems like geoacoustic inversion. A review of examples of parameter sensitivity methods and their application to problems in underwater acoustics is given, roughly grouped into ‘local’ and ‘non-local’ methods. Local methods such as Fisher information and Cramér-Rao bounds have important connections to information geometry. Information Geometry combines the fields of information theory and differential geometry by interpreting a model as a Riemannian manifold, known as the model manifold, that encodes both local and global parameter sensitivities. As an example, 2-dimensional model manifold slices are constructed for the Pekeris waveguide with sediment attenuation, for a vertical array of hydrophones. This example demonstrates how effective, reduced-order models emerge in certain parameter limits, which correspond to boundaries of the model manifold. This example also demonstrates how the global structure of the model manifold influences the local sensitivities quantified by the Fisher information matrix. This paper motivates future work to utilize information geometry methods for experimental design and model reduction applied to more complex modeling scenarios.

2.1 Introduction

Sound traveling in the ocean contains information about both the sound source and the ocean environment. Many ocean acoustics applications strive to passively infer information from recorded sounds, such as estimating the direction of arrival of the sound or determining the distance to the object producing the sound. These applications rely heavily on computational models of underwater sound propagation. A common question is how the environment should be represented in the sound propagation models. In particular, how much resolution is required in the water sound speed profile and how many layers are needed to represent the seafloor in a way that adequately captures the interaction of the sound and the sediment layers. The question of appropriate parameterization of the ocean environment has led many to study how to quantify the impact of modeling parameters on the model-data mismatch and, thus, quantify the information content and parameter sensitivities.

While common approaches for parameter sensitivity are based on perturbation theory or non-local sampling-based methods, which include Bayesian methods for obtaining posterior probability distributions, nascent methods for quantifying parameter sensitivity are found in information theory, including the Fisher information matrix (FIM) and Cramér-Rao bounds (CRB). The combination of information theory and differential geometry is the field known as information geometry. This paper serves as a tutorial about how information geometry can be used to quantify parameter sensitivity and guide the development of reduced-order identifiable models. Specific examples are provided for the application of information geometry methods to analyze the sensitivity of seabed parameters, but the method can be applied to any multi-parameter modeling problem.

First, examples are provided in Sec. 2.2 of a variety of approaches to estimating parameter sensitivity, providing background for other methods presented in this paper. Then, Sec. 2.3 introduces the Fisher information and how Cramér-Rao bounds can be obtained from it. A further review of relevant papers utilizing the FIM and CRBs in geoacoustic inversion is given in Sec. 2.4. The non-local method of rotated coordinates and its applications is discussed in greater detail in

Sec. 2.5, and its connection to the local measures of sensitivity obtained from the Fisher information matrix is explained. Sec. 2.6 then introduces the model manifold and other principles of information geometry, and then, in Sec. 2.7, an example is provided that demonstrates how model manifolds can be applied to an underwater sound propagation model. The illustrative example uses a Pekeris waveguide model with attenuation [5]. This example builds intuition about the relationship between model manifolds and CRBs and serves as an example of how information geometry can be applied to finding reduced-order models in ocean acoustics and beyond. Outlook for future applications of information geometry to problems in ocean acoustics is discussed in Sec. 2.8 which is followed by a summary (Sec. 2.9).

2.2 Background

For context, examples are now provided of other approaches to estimating parameter sensitivity such as perturbation theory, sampling-based approaches, which includes Bayesian methods to obtain posterior probability distributions, and information theoretic approaches such as Fisher information and the Cramér-Rao bound. Because uncertainty and sensitivity studies have been completed in many ocean acoustic applications over many decades, the papers cited herein should be considered as examples and not an exhaustive list. For example, the reader is referred to several other review articles that have covered the impact of the sound speed on ocean acoustic tomography [6, 7], wave propagation [8], reverberation [9], and transmission loss [10], especially in the presence of internal solitons [11] and thermocline variability [12]. We also recommend to the reader the references in Chapman and Shang [4] on the development of geoacoustic inversions to obtain estimates for seabed properties; this review article provides a detailed timeline of significant development in geoacoustic inversions including the sophisticated trans-dimensional Bayesian approach to geoacoustic inversion.

Perturbation theory approaches have been used to estimate parameter sensitivity of seabed parameters. First-order perturbation theory has been applied to analytical expressions for frequency-dependent sound pressure to obtain derivatives of pressure with respect to seabed sound speed [13], density, shear speed, and layer thickness [14, 15]. A perturbation theory approach has also been applied to investigate the sensitivity of sound propagation to changes in the water sound speed [16–19].

Sampling methods have been used for parameter sensitivity efforts in a variety of ocean acoustics applications. Many have explored how matched-field source localization [20, 21] is related to the ocean environment, including changes in water depth [22], sound speed and seabed properties [23, 24]. Matched-field geoacoustic inversion efforts have also conducted parameter sensitivity studies for different receiver configurations, including single sensors [25], vertical line arrays [9, 15, 26–30], horizontal towed arrays [31], and vector sensors [32]. Sampling-based sensitivity analyses have also been considered for different types of sound sources, including ambient noise [23, 33] and surface ships [30]. A Gibbs sampling approach for quantifying parameter sensitivity and how that leads to uncertainty in estimates from geoacoustic inversions has been thoroughly described in a pair of papers by Dosso [34, 35]. The difference between linear and nonlinear measures of parameter sensitivities was clearly delineated in a Letter to the Editor by Dosso *et al.* [36]. Many additional efforts have used sampling methods to examine how modal properties are influenced by parameter uncertainty [13, 14, 23, 25, 37–39]. Spatial shifts in sound fields have also been studied to obtain a stochastic sensitivity or global measure of sensitivity [40].

Bayesian methodology also utilizes advanced sampling methods in order to calculate posterior probability distributions (PPDs). From these PPDs, marginal and joint probability distributions can be found to evaluate parameter sensitivity and parameter coupling, respectively. Different sampling methods have been applied to obtain PPDs, such as the freeze bath method [41, 42]; Gibbs sampling [34, 35]; grid sampling [43]; directed Monte Carlo methods, e.g., genetic algorithms [44–46];

Markov chain Monte Carlo [47,48]; interacting Markov chains [49,50]; parallel tempering [51]; and reversible-jump Markov-chain Monte Carlo [52–55]. Sophisticated sampling of both the parameter space and the model space is incorporated in the trans-dimensional Bayesian approach to geoacoustic inversion [52]; the progression of geoacoustic inversions and development of trans-dimensional Bayesian approach are reviewed in Chapman and Shang [4].

Parameter sensitivity analysis has also been used to improve automated sampling algorithms, such as genetic algorithms, simulated annealing, and Tabu [29] for geoacoustic inversions. For example, genetic algorithms have been combined with the Gauss-Newton approach [45], which estimates parameter sensitivities from an approximate Hessian, and a subspace approach [56], which only considers the most identifiable portions of the parameter space. Similarly, simulated annealing methods have been modified to include adaptive temperatures for different parameters [26], adding in downhill simplex methods [57], and finding rotated coordinates to navigate the parameter space [41, 58]. The use of rotated coordinates to aid the inference process is reviewed in Sec. 2.5, along with a description of how the parameter sensitivity and coupling obtained from rotated coordinates relate to information theory concepts of the Fisher information and Cramér-Rao bounds.

While most of these sampling-based approaches to parameter sensitivity provide estimates of an average parameter sensitivity over the search space, information theory uses local parameter sensitivity analysis, which can then be extended non-locally by differential geometry. The main tool is the Fisher information, which provides a local measure of the information content about the modeling parameters in the model-data comparison [59]. Ocean acoustics studies have sought to maximize Fisher information for problems of source localization [60–62], bearing estimation [63] by optimizing off-axis targets for broadband active sonar [64], designing nonuniform linear arrays [65], optimal active sonar waveforms [66], underwater communication [67], and for autonomous underwater vehicle positioning [68]. The diagonal elements of the inverse of the Fisher information

matrix (FIM) are the Cramér-Rao bounds (CRB): the lower limit on the variance of an unbiased estimator of modeling parameters. CRBs have been used extensively in quantifying sensitivity in ocean acoustics applications; Sec. 2.4 provides a review of many examples of the application of CRBs in ocean acoustics.

Information geometry combines the advantages of local methods, such as the FIM and CRBs, and of non-local sampling-based methods such as rotated coordinates by quantifying both local and global model parameter sensitivities. Information geometry interprets multi-parameter models as a Riemannian manifold (referred to as the model manifold) where model parameters act as coordinates on the manifold. While previous studies in ocean acoustics have used Riemannian manfiolds [62, 69, 70], these applications primarily use principles of Riemannian geometry to obtain measures of distance between cross spectral density matrices for matched field processing and signal detection schemes, which differs from our use of a Riemannian model manifolds to explore environmental parameter identifiability in that our distance metric is motivated by information theory.

Information geometry has been shown to be particularly powerful in context of the paradigm of sloppy models. In a sloppy model, the model output is primarily affected by a few key parameter combinations, while other parameters do not appreciably change the model output. The characteristic feature of sloppy models is log-linear spaced eigenvalues of the FIM. Parameter combination to which the model output is most sensitive are known as identifiable, or stiff parameter combinations; data is most informative about stiff model parameters. Unidentifiable, or sloppy, parameter combinations are parameter combinations to which the model is not sensitive; sloppy parameters are not well constrained by data. Sloppiness has been studied in a variety of fields including systems biology [71–78] (such as biochemical reaction networks [79–87], cell state dynamics [88, 89], and tumor growth [90]), neuroscience [91–95], chemical engineering [96–100], power systems [101–104], ecology [83, 105, 106], solid state physics [107], deep learning [108–111],

economics [112–114], quantum simulation [115, 116], nuclear physics [117–119], cosmology [120], ocean acoustics [121], and more [109, 122–124]. Information geometry methods can also leverage sloppy model behavior to obtain reduced-order models which capture the relevant physics. For example, some parameter combinations will be “practically unidentifiable," meaning that their removal from the model will not significantly decrease model accuracy. Information geometry methods for model reduction, such as the manifold boundary approximation method [72, 125], have been applied to a variety of models [72, 84, 103, 104, 119, 125–127], and will be discussed in greater detail in Sec. 2.8.

2.3 Fisher information and Cramér-Rao bounds

A local measure of parameter sensitivity provides an estimate of how small changes in a parameter value impact the models output. Two powerful tools from information theory characterize the local behavior of a parameterized model in the context of inverse problems: the Fisher information matrix and the matrix inequality known as the Cramér-Rao bound. The FIM quantifies how well the model resolves parameter estimates from information in the data, and the CRB places a lower bound on the uncertainty of an unbiased estimator for the parameters of a model. This tutorial begins in Sec. 2.3.1 with a derivation of the form of the FIM used in this paper, followed by a presentation of the FIM in terms of a quadratic expansion about the best fit on a cost surface. The CRB and other measures derived from the FIM which bound the uncertainty of inferred parameters are introduced in Sec. 2.3.2. Throughout this tutorial, vectors are denoted with bold face. For example, $\mathbf{r} = [r_1, r_2, ...]$.

2.3.1 Fisher information matrix

The Fisher information matrix, denoted as $\mathscr{I}$, is defined as the expectation of the Hessian of the log likelihood, l:

$$\mathscr{I} = -\left\langle \frac{\partial^2 l}{\partial \theta_\mu \partial \theta_\nu} \right\rangle \equiv -\langle \partial^2_{\mu\nu} l \rangle. \tag{2.1}$$

For this tutorial, we consider a forward model that yields some output $y_\theta(t_i)$ based on modeling parameters $\boldsymbol{\theta} = [\theta_1, ..., \theta_N]$ at a value of the independent variable t_i. For an ocean acoustics sound propagation model, for example, the model output $y_\theta(t_i)$ may be sound pressure, transmission loss, or power reflection coefficients; and $\boldsymbol{\theta}$ may contain any combination of environmental, source, and receiver parameters. The independent variable t_i is determined by the choice of "experimental design," such as different source-receiver ranges or hydrophone depths on a vertical line array. In this section, a generic model is assumed, but application to an ocean acoustics model is provided in Sec. 2.7.

Inference problems typically compare model predictions with data. The i-th data sample d_i is assumed to have a mean equal to the model $y_\theta(t_i)$ with some Gaussian additive noise:

$$d_i \sim N(y_\theta(t_i), s_i^2) \tag{2.2}$$

where the noise has variance s_i^2. In general, s_i^2 is unknown and may include both random (aleatoric) noise and systemic (epistemic) noise that might come from a measurement apparatus. To compare the model with data, the data-model residual is defined:

$$r_i = \frac{d_i - y_\theta(t_i)}{s_i} \quad , \quad \text{with} \quad r_i \sim N(0,1). \tag{2.3}$$

The probability distribution of r_i is

$$P(r_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} r_i^2}. \tag{2.4}$$

For M measurements, corresponding to $\boldsymbol{t} = [t_1, t_2, ..., t_M]$ and $\boldsymbol{r} = [r_1, r_2, ..., r_M]$, the multivariate Gaussian probability distribution of $\boldsymbol{r}$ is

$$P(\boldsymbol{r}) = \prod_i^M \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}r_i^2}. \tag{2.5}$$

The log-likelihood l is defined as the natural logarithm of $P(\boldsymbol{r})$:

$$l = \ln(P(\boldsymbol{r})) = -\frac{M}{2}\ln(2\pi) - \frac{1}{2}\sum_i^M r_i^2. \tag{2.6}$$

The Hessian of the log-likelihood has elements

$$\partial^2_{\mu\nu} l = \sum_i^M (-\partial_\mu r_i \, \partial_\nu r_i - r_i \, \partial^2_{\mu\nu} r_i). \tag{2.7}$$

Thus, the FIM, as given in Eq. 2.1, has elements

$$\mathscr{I}_{\mu\nu} = -\langle \partial^2_{\mu\nu} l \rangle = \sum_i^M \langle \partial_\mu r_i \, \partial_\nu r_i \rangle + \sum_i^M \langle r_i \, \partial^2_{\mu\nu} r_i \rangle. \tag{2.8}$$

Note that the derivatives of the residual, $\partial_\mu r_i$, can be expressed in terms of derivatives of the model $y_\theta(t_i)$ with respect to θ_μ:

$$\partial_\mu r_i = -\frac{1}{s_i} \partial_\mu y_\theta(t_i), \tag{2.9}$$

which does not depend on noisy data $d_i \sim N(y_\theta(t_i), s_i^2)$. Therefore, the expectation operator in the first term drops out (i.e., $\langle \partial_\mu r_i \, \partial_\nu r_i \rangle = \partial_\mu r_i \, \partial_\nu r_i$). The second term can be rewritten as $\langle r_i \rangle \, \partial^2_{\mu\nu} r_i$, so for zero mean data the second term goes to zero. Thus, for these assumptions about the noise, $\mathscr{I}_{\mu\nu}$ simplifies to

$$\mathscr{I}_{\mu\nu} = \sum_i^M \partial_\mu r_i \, \partial_\nu r_i. \tag{2.10}$$

By substituting Eq. 2.9 into Eq. 2.10, the elements of the FIM can be written as

$$\mathscr{I}_{\mu\nu} = \sum_i^M \frac{1}{s_i^2} \partial_\mu y_\theta(t_i) \, \partial_\nu y_\theta(t_i). \tag{2.11}$$

Thus, the FIM can be expressed as

$$\mathscr{I} = J^T S^{-2} J, \tag{2.12}$$

where J is the Jacobian matrix of derivatives of $y_\theta(t_i)$ with elements

$$J_{i\mu} = \frac{\partial y_\theta(t_i)}{\partial \theta_\mu} = \partial_\mu y_\theta(t_i), \tag{2.13}$$

and S is a diagonal matrix with elements s_i. A different noise assumption in Eq. 2.2 will lead to a different form of the FIM.

As previously stated, s_i is unknown and experiment dependent. Therefore, for the purposes of model analysis, it is convenient to assume $s_i = s$ for all i, so that $S^{-2} = s^{-2}\,\mathbb{I}$, and

$$\mathscr{I} = s^{-2} J^T J \propto J^T J. \tag{2.14}$$

Thus in calculation of the FIM in practice, the s^{-2} term can be neglected. This convention is adopted herein.

This mathematical derivation of the FIM may be further appreciated by a graphical discussion that provides a more intuitive understanding. This discussion begins by introducing a function $\boldsymbol{Y}(\boldsymbol{\theta})$, known as the model map:

$$\boldsymbol{Y}(\boldsymbol{\theta}) = \begin{bmatrix} y_\theta(t_1) \\ y_\theta(t_2) \\ \vdots \\ y_\theta(t_M) \end{bmatrix}, \tag{2.15}$$

where M is the number of measurements. For N parameters in $\boldsymbol{\theta}$, the model map $\boldsymbol{Y}(\boldsymbol{\theta})$ forms a mapping from an N-dimensional parameter space into an M-dimensional data space:

$$\boldsymbol{Y}(\boldsymbol{\theta}) : \mathbb{R}^N \to \mathbb{R}^M. \tag{2.16}$$

In a modeling scenario, predictions of $y_\theta(t_i)$ are compared to experimental data d_i in a cost function, sometimes called an error or loss function, which measures how well model prediction values match observed data. One option is the least-squares cost, which can be motivated in a maximum likelihood sense. It can be shown that the ordinary least squares estimator for a linear

regression model maximizes the likelihood under the assumption of Gaussian additive noise, which is used here. The least-squares cost is:

$$C(\boldsymbol{\theta}) = \frac{1}{2}\sum_{i}^{M}(y_{\theta}(t_i) - d_i)^2. \tag{2.17}$$

The set of modeling parameters that minimizes the cost given some data is referred to as the best fit parameterization $\hat{\boldsymbol{\theta}}$.

The graph of the cost $C(\boldsymbol{\theta})$ defines an N-dimensional surface, called a cost surface, whose peaks correspond to parameter values that poorly fit the data and whose valleys correspond to parameter values that fit the data well. N-dimensional surfaces are unwieldy to visualize, so two-dimensional slices of cost surfaces are often used to gain a "topographical" intuition of the peaks and valleys of the cost surface. One such two-dimensional slice in the neighborhood of a best fit $\hat{\boldsymbol{\theta}}$ is shown in Fig. 2.1 for a cartoon cost surface; this cost surface is used for illustration throughout the present section.

Cost surfaces for nonlinear models, such as those found in modeling ocean acoustics, often include narrow valleys, flat plateaus, and multiple local minima. However, by zooming in on a sufficiently small region around a local minimum, the cost surface is approximately quadratic with lines of constant $C(\boldsymbol{\theta})$ forming ellipse-shaped contours about the minimum. The axes of such ellipses provide insight into the sensitivity of model parameters.

The local geometry about a minimum is described by a Taylor series of the forward model expanded about the best fit:

$$\tilde{\boldsymbol{Y}}(\boldsymbol{\theta}) = \boldsymbol{Y}(\hat{\boldsymbol{\theta}}) + J\boldsymbol{\delta\theta} + \mathscr{O}(\boldsymbol{\delta\theta})^2, \tag{2.18}$$

where $\boldsymbol{\delta\theta} = \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}$ and J is the same M by N Jacobian matrix defined above in Eq. 2.13. In this linear approximation, the cost function becomes

$$\tilde{C}(\boldsymbol{\theta}) = C(\hat{\boldsymbol{\theta}}) + \frac{1}{2}\boldsymbol{\delta\theta}^T J^T J\boldsymbol{\delta\theta} \tag{2.19}$$

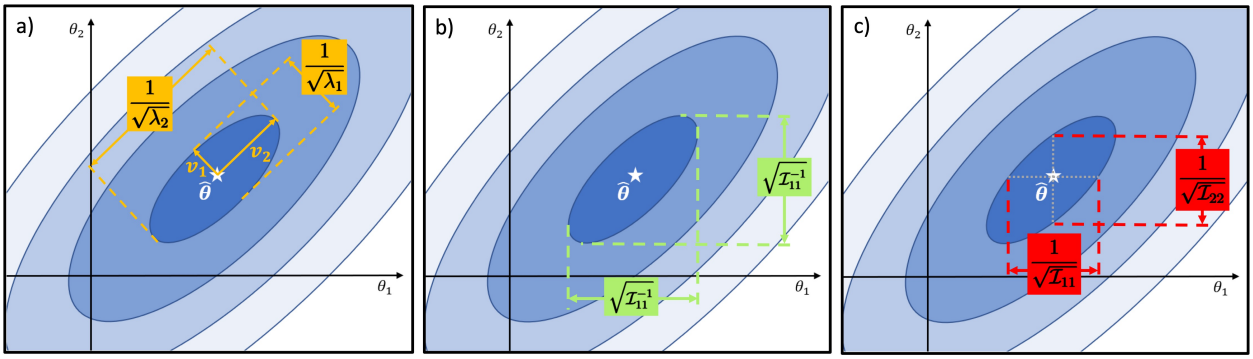

Figure 2.1 Two-dimensional cost surface cross-section in the θ_1-θ_2 parameter plane showing the neighborhood about a local minimum. Notice the elliptical contours of equal cost about the minimum at the best-fit parameters $\hat{\boldsymbol{\theta}}$, marked by a white star. The semi-minor and semi-major axes of the contours correspond to the FIM eigenvector directions $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, respectively. The inverse of the square root of corresponding eigenvalues of the FIM, λ_1 and λ_2, are proportional to the semi-minor and semi-major widths of the cost ellipses. The width of the ellipses projected onto the θ_1 and θ_2 axes are proportional to $\sqrt{\mathscr{I}^{-1}_{jj}}$, the square root of the corresponding diagonal elements of the inverse FIM (a covariance matrix). $\mathscr{I}^{-1}$ is calculated when finding the CRB. Finally, the widths of the ellipses through the best fit point and parallel to the parameter axes are proportional to $1/\sqrt{\mathscr{I}_{jj}}$ the inverse of the square root of the corresponding elements of the FIM, also known as the conditional standard deviation.

using the fact that the gradient is zero at the best fit. The second term contains the Fisher information matrix $\mathscr{I}$ as defined in Eq. 2.14, and thus the $\mathscr{I}$ describes the curvature of the cost surface at the best fit. Note that the FIM does not depend on the data, only the derivatives of the forward model; thus the lower bounds on the informativity of potential experimental data can be established by analyzing the model alone.

2.3.2 Bounding parametric uncertainty with the FIM

Information about model parameter sensitivities can be extracted from the FIM in several ways, and understood in context of the cost surface in Fig. 2.1. Specifically, the eigenvalues and eigenvectors of the FIM quantify a model's linearized sensitivity to changes of parameter combinations. Because the FIM is often ill conditioned, these are typically obtained by a singular value decomposition (SVD) of J,

$$J = U\Sigma V^T, \tag{2.20}$$

where U and V are unitary matrices, and Σ is a diagonal matrix containing the singular values of J. The FIM can be expressed as

$$\begin{aligned} \mathscr{I} &= J^T J = (U\Sigma V^T)^T (U\Sigma V^T) = V\Sigma U^T U\Sigma V^T \\ &= V\Sigma^2 V^T, \end{aligned}$$

which can be recognized as an eigenvalue decomposition of $\mathscr{I}$. From this form, the eigenvalues of the FIM are the squared singular values of the Jacobian, and the columns of V are the eigenvectors of the FIM. These eigenvectors $\boldsymbol{v_i}$ are the independent parameter combinations that diagonalize the covariance matrix, $\mathscr{I}^{-1}$, which is the inverse of the FIM.

The eigenvectors of the FIM are in general linear combinations of the original parameters and give a natural basis for quantifying the model sensitivities. As previously mentioned, sloppy models are characterized by roughly log-linear spacing of eigenvalues of the FIM, which manifests as ellipses with high aspect ratios on the cost surface (Fig. 2.1). The eigenvector corresponding to the largest eigenvalue of $\mathscr{I}$ indicates the most identifiable parameter combination, represented by $\boldsymbol{v_1}$ in Fig. 2.1(a), which lies along the semi-minor axis of the ellipse about the best fit. Intuitively, moving in the direction of $\boldsymbol{v_1}$ on the cost surface leads to the most rapid change in model output. In contrast, the eigenvector with the smallest eigenvalue is the sloppiest parameter combination, corresponding to $\boldsymbol{v_2}$ in Fig. 2.1(a), which lies along the semi-major axis of the ellipse. The inverse of the square root of corresponding eigenvalues of the FIM, λ_1 and λ_2, are proportional to the semi-minor and semi-major widths of the cost contour ellipses, as marked in yellow in Fig. 2.1(a).

The FIM also provides insight into the statistical properties of inferred parameters (i.e., the original parameters of the model). The Cramér-Rao bound inequality states that the inverse of the FIM, $\mathscr{I}^{-1}$, is a covariance matrix that forms a lower bound on the uncertainty of parameter estimations. For an unbiased estimator, the Cramér-Rao bound inequality is

$$\mathrm{Cov}(\hat{\boldsymbol{\theta}}) \geq \mathscr{I}^{-1} = V\Sigma^{-2}V^T \tag{2.21}$$

where $\mathrm{Cov}(\hat{\boldsymbol{\theta}})$ is the covariance of the estimator of $\hat{\boldsymbol{\theta}}$, and the matrix inequality $A \geq B$ is understood to mean that the matrix $A - B$ is positive semidefinite. An estimator that achieves this bound is said to be fully efficient. Large-valued entries of $\mathscr{I}^{-1}$ indicate high parameter uncertainty.

Often of interest are the diagonal entries of this covariance matrix, as they correspond to the variances in the estimate of the individual parameters. The square root of the diagonal elements of $\mathscr{I}^{-1}$ are referred to as the Cramér-Rao bounds (CRB), and indicated as $\sigma_{\mathrm{CRB},j}$ here:

$$\sigma_{\mathrm{CRB},j} = \sqrt{(\mathscr{I}^{-1})_{jj}}. \tag{2.22}$$

The CRB gives the lower bounds on the uncertainty in inferring individual parameters from data; in other words, the CRB gives the smallest standard deviations in inferred values possible given a model and data. The CRB for θ_1 and θ_2 are shown in green in Fig. 2.1(b), corresponding to the widest part of the constant cost ellipse in the two parameter directions.

Large CRBs have two causes. First, a model might be insensitive to changes in an individual parameter, such as is the case for parameters deep beneath the seafloor in a transmission loss model at sufficiently high frequencies. Second, a large CRB can result from strong correlation among parameters. In physical systems, correlation is typically due to physical mechanisms that manifest similar effects in model predictions. In such cases the effect of one model parameter may be cancelled by tuning another. The model is often sensitive to many of the parameters individually but is insensitive to coordinated changes in combinations of parameters.

Alternatively, the uncorrelated parameter uncertainty could be ascertained if all other model parameters were fixed to their true values. The theoretical uncorrelated parameter sensitivity, called the conditional standard deviation (conditional STD) and indicated by $\sigma_{\text{cond},j}$, can be calculated from the FIM as

$$\sigma_{\text{cond},j} = \frac{1}{\sqrt{\mathscr{I}_{jj}}}. \tag{2.23}$$

The conditional STD for a parameter is always less than or equal to $\sigma_{\text{CRB},j}$, meaning that $\sigma_{\text{cond},j}$ underestimates the uncertainty in inferring parameters from data. Disparities between $\sigma_{\text{CRB},j}$ and $\sigma_{\text{cond},j}$ are due to correlations among parameters. In terms of the cost surface, the conditional STD corresponds to the width of the constant-cost ellipse measured through the best fit and parallel to a parameter axis, denoted by red in Fig. 2.1(c). In the case of a model with uncorrelated parameters, the axes of the ellipses of constant cost align with the parameter axes, meaning that $\sigma_{\text{CRB},j}$ and $\sigma_{\text{cond},j}$ are equal.

Because the FIM is independent of the observed data $\boldsymbol{d}$ and depends only on derivatives of the forward model, the lower bounds on the informativity of potential experimental data can be established by analyzing the model alone. Through the CRB, the FIM establishes the best-case variance for each model parameter, quantifying the information content, and determining whether it is feasible to infer individual parameters to the desired resolution from given data.

2.4 Local measures of parameter sensitivity

The local measure of parameter uncertainty in the FIM and CRB (Eq. 2.22) have been applied to a variety of applications in ocean acoustics. In the literature, the context of either the experimental setup or the inversion approach typically determines how the FIM and CRB are obtained. Where circumstances permit, analytical expressions for the FIM and CRB are employed. When analytical models are not available, calculation of the FIM and CRB relies upon numerical gradients of computational models taken with respect to parameters. In this section, examples are provided to illustrate the breadth of applications of the CRB to problems in ocean acoustics, with a focus on work related to quantifying parameter sensitivities for the ocean environment.

CRBs have been calculated analytically for acoustic reflection measurements [128, 129], source receiver geometry [130], array shape calibration [131, 132], and source localization [133–136], including in deep water [137, 138]. For source localization specifically, CRBs have been used for direction-of-arrival estimation using noise covariance matrices [139], range and bearing localization with two hydrophones [140], underdetermined systems using sparse linear arrays [141, 142], virtual sensor arrays to handle unsynchronized sensors [143], and semi-blind source localizations [144]. CRBs obtained analytically have also been used to quantify seabed parameter sensitivies using ambient ocean sound [145, 146].

Variations of the CRB have also been used, including hyrbrid CRBs for water sound speed fluctuations [147] and direction of arrival estimates [148], a stochastic CRB for direction-of-arrival estimates using vector sensors [134], and Bayesian CRBs for source ranging [149] and seabed parameter estimation [150]. The citations in Baggeroer [134] provide additional references on the development and use of stochastic formulations of the CRB.

When analytical models are not available, calculations of the FIM and CRB rely upon numerical gradients of computational models taken with respect to parameters. However, gradients can be difficult to obtain. Some success has been found with analytical gradients [14, 45] and implicit adjoint methods [151] although both require a problem-specific gradient derivation. One interesting example comes from Hawkes and Nehorai [152] who develop an efficient CRB based measure of estimation accuracy for 3-dimensional source localization on a distributed system involving a two stage CRB calculation, one which can be performed analytically and the other requiring numerical derivatives. While some success has been found with calculating the FIM and CRBs using finite-differencing, as in Gebbie and Siderius [150] for quantifying seabed parameter sensitivities, finite-differencing has typically been difficult due to the need for appropriate step sizes for different parameters; too large or too small of step sizes makes the method inaccurate [46]. Mortenson *et al.* [121], recently presented a strategy for parameter preconditioning followed by Richardson extrapolation that renders feasible a general finite-difference methodology for finding numerical gradients in sound propagation models, applied to quantify seabed parameter sensitivity.

2.5 Non-local measures of parameter sensitivity

The eigenvectors of the FIM (and more completely the corresponding covariance matrix) provide a set of basis vectors that are aligned with the curvature of the cost surface and show how the modeling parameters are coupled around a best fit point $\hat{\boldsymbol{\theta}}$. While these basis vectors describe only the local cost surface, there have also been efforts to non-locally estimate a covariance matrix. The eigenvalues of this covariance matrix provide a spatially averaged view of parameter coupling which provides a coordinate transformation to better align model parameters with global features of the cost surface, improving optimization efficiency.

Such a coordinate transformation was introduced to the underwater acoustics community in 1995 by Collins and Fishman [58]. The coordinate transformation was obtained from the eigenvectors of a covariance matrix K, calculated by integrating the gradient of the cost function E over the parameter search space Ω:

$$K = \int_{\Omega} \nabla E (\nabla E)^T d\Omega \tag{2.24}$$

An SVD of K yields eigenvectors $\{\boldsymbol{v_j}\}$ and eigenvalues λ_j. The $\{\boldsymbol{v_j}\}$ are referred to as rotated coordinates, and the associated eigenvalues λ_j indicate the relative importance of $\boldsymbol{v_j}$ in describing the curvature of the cost surface. In other words, the rotated coordinates correspond to the orthogonal transformation that diagonalizes the covariance matrix of the cost function gradient.

In practice, the integral in Eq. 2.24 can be difficult to estimate. Nevertheless, even approximations of K, for example by using an initial sparse sampling of parameter space, have been shown to provide $\{\boldsymbol{v_j}\}$ that greatly increase the efficiency of optimization algorithms. In addition, a reduced-order model may be obtained by only using $\boldsymbol{v_j}$ with λ_j greater than a threshold to navigate the space. The advantages of using rotated coordinates in a sampling algorithm increases as the number of modeling parameters increases [58].

Rotated coordinates were first used to improve self-starters for range-dependent modeling [153–155] and in simulated annealing algorithms [58,156–161]. Early work with rotated coordinates in simulated annealing included source localization [159, 162]; geoacoustic inversions to find characteristics of sediment layers in the seabed from acoustic data [155–157], including one study of a Biot-Stoll model of the sediment [161]; and focalization studies that included both source and environmental parameters [158, 160].

A method for obtaining rotated coordinates using the model covariance matrix was employed in Jaschke and Chapman [41] to do matched-field inversions using the freeze bath method. This approach has been used in Bayesian optimizations as well [163]. Other studies have also used a non-local derivative based sampling to inform parameter step sizes, such as in [164].

The calculation of rotated coordinates from a non-local estimate of a covariance matrix can be considered an analogous generalization of the local information obtained from the Fisher information matrix. For a nonlinear model, the orientation of contours and valleys in the cost surface changes over parameter space. Thus while the rotated coordinates roughly align with the curvature of the cost surface, they are an “averaged" quantity that does not perfectly align with the cost surface near a specified point $\boldsymbol{\theta}$. Thus rotated coordinates are useful for efficiently navigating a large parameter space during an optimization, but the Fisher information is needed for information geometry.

2.6 Information Geometry

Information geometry combines the fields of information theory and differential geometry by interpreting a multi-dimensional model as a high-dimensional manifold embedded in the space of potential predictions, known as the model manifold. Each point on the model manifold can be interpreted as a probability distribution, and distance along the manifold relates to statistical distinguishability. Information geometry is especially powerful for parameter sensitivity analysis

when interpreted in context of the paradigm of sloppy models, as introduced in Sec. 2.1. The model manifold is introduced in Sec. 2.6.1; it quantifies parameter sensitivities on local, intermediate, and global scales, and gives a path for obtaining reduced-order models by identifying relevant parameter combination limits. Connections between the model manifold and the Fisher information matrix are discussed in Sec. 2.6.2, followed by an example calculating model manifolds for a sound propagation model, the Pekeris waveguide with sediment attenuation, in Sec. 2.7.

2.6.1 The model manifold

The model manifold is the embedding of a model into data space, or prediction space. The model manifold is the set of all possible predictions of the model for different values of the parameters and forms a hyper-surface, or manifold, in data space. Using the convention introduced earlier of a model $y_{\boldsymbol{\theta}}(t_i)$ with parameters $\boldsymbol{\theta}$ and independent variable t_i, we construct the model manifold for a toy model to illustrate the process. The toy model is the sum of two decaying exponentials:

$$y_{\boldsymbol{\theta}}(t_i) = e^{-\theta_1 t_i} + e^{-\theta_2 t_i} \quad , \quad \theta_1, \theta_2 \geq 0. \tag{2.25}$$

This model has $N = 2$ modeling parameters: $\boldsymbol{\theta} = [\theta_1, \theta_2]$.

Often models $y_{\boldsymbol{\theta}}(t_i)$ are visualized as functions of the independent variable, e.g. over t, for a fixed values of the parameter $\boldsymbol{\theta}$; this view displays the model's input-output relations. For the sum of exponentials model, the input-output relations are shown in Fig. 2.2(a) for three different parameter choices $\{\boldsymbol{\theta}^A, \boldsymbol{\theta}^B, \boldsymbol{\theta}^C\}$.

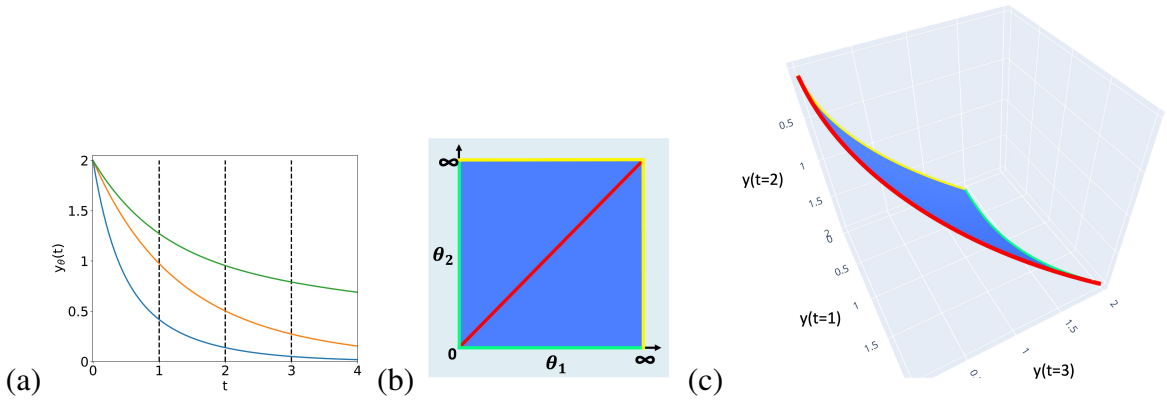

Figure 2.2 (a) Input-output relations of the sum of exponentials model, with different model realizations using different values of $\boldsymbol{\theta}$. (b) Parameter space for the sum of exponentials model. θ_1 and θ_2 can vary from 0 to ∞. (c) Model manifold for the sum of exponentials model. The model has compacted and folded in half the original parameter space, as seen in the model manifold. Colored boundaries of the model manifold correspond to the colored lines on the parameter space plot.

While input-output relations show how the model changes continuously in the independent variable, data are often obtained at M discrete measurements of the independent variable t_i, e.g., measurements at hydrophones of different depths on a vertical array. This choice makes use of the model map in Eq. 2.15 natural; and Eq. 2.25 can be expressed in terms of the model map also as

$$\boldsymbol{Y}(\boldsymbol{\theta}) = e^{\theta_1 \boldsymbol{t}} + e^{\theta_2 \boldsymbol{t}}. \tag{2.26}$$

Therefore, the model manifold can be constructed by varying model parameters $\boldsymbol{\theta}$ and plotting the model map $\boldsymbol{Y}(\boldsymbol{\theta})$ in an M-dimensional data space.

For the sum of exponentials model, θ_1 and θ_2 are varied as shown in the parameter space Fig. 2.2(b). A specific set of model parameters $\boldsymbol{\theta}^l$ corresponds to a point on the model manifold. The model manifold is shown in Fig. 2.2(c) for the case of three experimental design values of $\boldsymbol{t} = [1,2,3]$. Each of the three axes in data space corresponds to an element of $\boldsymbol{Y}(\boldsymbol{\theta})$ for each value of the independent variable t_i, and the model manifold is a 2-dimensional surface because there are $N = 2$ model parameters being varied.

Interestingly, the infinite, unbounded parameter space in Fig. 2.2(b) has been compressed to a finite, bounded model manifold; these boundaries of the model manifold, colored red, green, and yellow in Fig. 2.2(c), correspond to “reduced models" where one parameter “drops out" of the model. For example, the yellow boundary corresponds to taking either θ_1 or θ_2 to infinity, which makes the corresponding term vanish, making the reduced model $\boldsymbol{Y}(\boldsymbol{\theta}) = e^{-\theta_1 \boldsymbol{t}}$. Similarly, the green boundary corresponds to taking θ_1 or θ_2 to zero which makes the corresponding term equal to one, making the reduced model at that boundary $\boldsymbol{Y}(\boldsymbol{\theta}) = 1 + e^{-\theta_1 \boldsymbol{t}}$. Note that the model is symmetric across the $\theta_1 = \theta_2$ line in parameter space (e.g., the model gives the same output for $\boldsymbol{\theta} = (1,3)$ and $\boldsymbol{\theta} = (3,1)$). This symmetry is reflected in the model manifold, where the parameter space has been folded in half over the line $\theta_1 = \theta_2$ in the model manifold. The red boundary is where $\theta_1 = \theta_2$ and the model reduces to be $\boldsymbol{Y}(\boldsymbol{\theta}) = 2e^{-\theta_1 \boldsymbol{t}}$. These reduced-order models associated with the boundaries of the model manifold become applicable in the parameter regimes that map to regions near the boundaries.

In summary, the model manifold is an N-dimensional surface embedded in an M-dimensional data space and yields information about global parameter sensitivities. Specifically, the boundaries of the model manifold correspond to reduced-parameter models. Distances on the model manifold can be quantified using Fisher information.

2.6.2 Fisher information matrix as a distance metric

In this section, the connection between the FIM and the model manifold are presented. The first step is to discuss how the Jacobian matrix relates to the model manifold. For a nonlinear model, the model manifold is a curved N-dimensional surface. However, a sufficiently small portion of the manifold at $\hat{\boldsymbol{\theta}}$ can be approximated by a linear surface tangent to the manifold. This tangent surface is characterized by the Jacobian matrix. The Jacobian matrix can be expressed in terms of derivatives of the model map (Eq. 2.15) with respect to the model parameters θ_μ, or as derivatives

of $y_\theta(t_i)$:

$$J = \left[\frac{\partial \boldsymbol{Y}(\boldsymbol{\theta})}{\partial \theta_1} \quad \frac{\partial \boldsymbol{Y}(\boldsymbol{\theta})}{\partial \theta_2} \quad \cdots \quad \frac{\partial \boldsymbol{Y}(\boldsymbol{\theta})}{\partial \theta_N}\right] = \begin{bmatrix} \frac{\partial y_\theta(t_1)}{\partial \theta_1} & \frac{\partial y_\theta(t_1)}{\partial \theta_2} & \cdots & \frac{\partial y_\theta(t_1)}{\partial \theta_N} \\ \frac{\partial y_\theta(t_2)}{\partial \theta_1} & \frac{\partial y_\theta(t_2)}{\partial \theta_2} & \cdots & \frac{\partial y_\theta(t_2)}{\partial \theta_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_\theta(t_M)}{\partial \theta_1} & \frac{\partial y_\theta(t_M)}{\partial \theta_2} & \cdots & \frac{\partial y_\theta(t_M)}{\partial \theta_N} \end{bmatrix}, \tag{2.27}$$

where N is the number of model parameters and M is the number of measurements. The columns of the Jacobian span the linear surface tangent to the manifold at $\hat{\boldsymbol{\theta}}$.

As given in Eq. 2.20, the Jacobian can be expressed as $J = U\Sigma V^T$, the composition of three linear functions. As an example of how these three linear functions perform transformations, consider how the Jacobian acts on a circle in a 2-dimensional parameter space centered at $\hat{\boldsymbol{\theta}}$. First, V^T is a rotation matrix that acts on the circle in parameter space and produces a circle in parameter space, but with the coordinate axes rotated. The new coordinate axes $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ are the columns of V. Second, Σ scales $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ so that the circle is now an ellipse. Finally, U is a rotation matrix that rotates and lifts the ellipse into the M-dimensional data space. This ellipse in data space has orthogonal principal axes $\sigma_1 \boldsymbol{u_1}$ and $\sigma_2 \boldsymbol{u_2}$, where $\boldsymbol{u_1}$ and $\boldsymbol{u_2}$ are the columns of U and σ_1 and σ_2 are the diagonal elements of Σ. This lifted ellipse is tangent to the model manifold at $\hat{\boldsymbol{\theta}}$. Thus, the Jacobian transforms a circle in parameter space to an ellipse in data space.

This tangent surface, which linearly approximates the model manifold at $\hat{\boldsymbol{\theta}}$, can be analyzed in terms of the four fundamental subspaces of linear algebra, helping to provide an intuitive understanding of the model manifold and connect to parameter sensitivities. The four fundamental subspaces of a linear map are the row space, the null space, the column space, and the left null space. The row space and null space make up the domain, or parameter space, of the model. The row space contains parameter directions that are identifiable at $\hat{\boldsymbol{\theta}}$. The null space contains the parameter combinations that are unidentifiable at $\hat{\boldsymbol{\theta}}$. The column space and left null space make up the codomain, or data space. The column space is the tangent space approximating the model

manifold at $\hat{\boldsymbol{\theta}}$, which is the space spanned by the columns of the Jacobian. The left null space can be understood as noise, i.e., the part of the data that cannot be captured by the model. Data are noisy and, in general, lie off of the model manifold. The noisy data can be projected onto the model manifold. Every $\boldsymbol{\theta}$ value has its own Jacobian characterizing the tangent space at that point with its accompanying four fundamental subspaces.

The Fisher information matrix is the expectation of the Hessian of the log likelihood, which was derived in Sec. 2.3 to be $\mathscr{I} \propto J^T J$ (Eq. 2.14). The FIM does not depend on data and quantifies the curvature of the cost surface. The FIM can also be thought of as quantifying the tangent of the model manifold at a given $\boldsymbol{\theta}$, in a linear approximation.

Additionally, while the FIM is often presented as an abstract statistical device, in fact it can be understood as a distance metric in data space and, therefore, on the model manifold. Understanding the FIM as a distance metric makes it easier to see its utility in parameter sensitivity analysis. To understand how the FIM is a distance metric, consider a vector $\boldsymbol{w}$ in parameter space. Through the Jacobian, this vector $\boldsymbol{w}$ corresponds to another vector in data space, $\boldsymbol{W}$,

$$J\boldsymbol{w} = \boldsymbol{W}.$$

The (Euclidean) length of $\boldsymbol{W}$ can be calculated by $||\boldsymbol{W}||_2 = \sqrt{\boldsymbol{W}'\boldsymbol{W}}$. Using the relation above,

$$\begin{aligned} ||\boldsymbol{W}||_2 &= \sqrt{\boldsymbol{W}'\boldsymbol{W}} && (2.28)\\ &= \sqrt{\boldsymbol{w}' J' J \boldsymbol{w}} && (2.29)\\ &= \sqrt{\boldsymbol{w}' \mathscr{I} \boldsymbol{w}}. && (2.30) \end{aligned}$$

This derivation shows that the FIM is the induced norm for parameter space vectors that corresponds to Euclidean distance on the model manifold. Intuitively, the FIM weights the norm between parameter space vectors by the relative stiffness or sloppiness of the parameters. For example, models are insensitive to changes in sloppy parameter combinations, meaning that a step in a sloppy direction in parameter space corresponds to a relatively small step in data space ($\sqrt{\boldsymbol{w}'\boldsymbol{w}} \gg \sqrt{\boldsymbol{w}'\mathscr{I}\boldsymbol{w}}$). The converse is true for stiff parameters.

For nonlinear models, parameter sensitivities vary across parameter space. This fact shows the utility of non-local approaches such as rotated coordinates and is the primary limitation of local sensitivity methods such as the FIM—the FIM at some $\boldsymbol{\theta}$ does not adequately describe the entire parameter space. The model manifold, however, provides a bridge between FIM local sensitivities and global parameter sensitivities. The model manifold importantly encodes the connection between local parameter sensitivities as parameters are varied, moving along the surface of the model manifold. Geodesics can be calculated along the model manifold giving a “geodesic distance": the shortest distance along the manifold between two points. This geodesic distance makes a good choice for representing statistical distinguishability [108].

In summary, the columns of Jacobian matrix span a linear space tangent to the model manifold, and the FIM is used in Eq. 2.30 to obtain a measure of distance on the model manifold and, therefore, gives information about parameter sensitivities for every point in parameter space. Geodesic distance between two points on the model manifold quantifies statistical distinguishability and shows how local parameter sensitivities change in parameter space.

2.7 Pekeris waveguide example

To demonstrate the utility of the model manifold and its connection to the FIM and CRBs, this section gives a an example of a simple sound propagation model. The sound propagation model used in this work, the Pekeris waveguide, is described in Sec. 2.7.1. Three seabed parameters are considered in this work: sediment sound speed, sediment density, and sediment attenuation. The 2-dimensional transmission loss (TL) model manifold are visualized in Sec. 2.7.2, corresponding to the traditional Pekeris waveguide at zero sediment attenuation. These manifold slices illustrate the emergent “hard bottom" reduced model corresponding to the infinite density parameter limit, and show how the manifold structure, and therefore parameter sensitivities, change for certain sediment types. For these sediment types, the local sensitivity properties for all three seabed parameters are evaluated, in Sec. 2.7.3, by calculating the FIM and quantifying parameter uncertainties with the CRB and the conditional STD. Discussion includes analysis of parameter correlations, relative parameter sloppiness and how global structure influences local parameter sensitivity. In Sec. 2.7.4, portions of the TL model manifold are constructed in the 3-dimensional parameter space and shown to illustrate the same parameter sensitivity hierarchy seen with the FIM.

2.7.1 The Pekeris waveguide model

A sound propagation model depends on the experimental design and the ocean environment. Selection of discrete samples of the independent variable (analagous to $\boldsymbol{t}$ in Sec. 2.6.1) are referred to as the experimental design and include the source depth z' (m), the receiver depth z (m), the source frequency f (Hz), and the source range r (m). The parameterization of the ocean environment is contained in $\boldsymbol{\theta}$.

For this introductory work, a Pekeris waveguide is assumed. The Pekeris model contains an upper half-space (air), a waveguide (water), and a lower half-space (seafloor). Each section is assumed to be uniform, i.e. no variation in physical properties with depth or range. The original Pekeris waveguide model paramaterizes the seafloor half-space with a uniform sound speed c_2 and density ρ_2 [165, 166]. We use an adaptation of the Pekeris model from Buckingham and Giddens [5] in which seafloor attenuation α_2 is also a parameter. For the examples given in this paper, $\boldsymbol{\theta}$ contains the seafloor parameters: c_2 (m/s), ρ_2 (g/cm^3), and α_2 (dB/λ). The water properties [water sound speed c_1 (m/s), water density ρ_1 (g/cm^3)] could be included in $\boldsymbol{\theta}$ but are held fixed in this example. (Note the convention used in this work that the subscript "1" refers to water parameters, and the subscript "2" refers to seafloor parameters.)

For the choice of $\boldsymbol{\theta} = [c_2, \rho_2, \alpha_2]$, only the cases in which the half-space parameters are greater than or equal to the water parameters are considered: $c_2 \geq c_1$, $\rho_2 \geq \rho_1$, and $\alpha_2 \geq 0$. Specific choices of values for the additional ocean parameters and the experimental design are given in Table 2.1. The ocean depth $h = 100$ m is chosen to simulate a shallow ocean environment. For the experimental design, single values of source frequency, source depth, and source range are selected, and three different receiver depths $\{z_j\} = [20, 50, 80]$ m are used. Note that there are many other possible choices for the experimental design variables that could have been used here. For this work, the modeled values are transmission loss.

The Pekeris waveguide is a range-independent normal mode sound propagation model that uses a finite summation of "proper modes" plus a branch line integral contribution. In [5], the Green's function is found in the Fourier domain and given in terms of the velocity potential ϕ:

$$\phi_\omega(r, z, \boldsymbol{\theta}) = \phi_{\text{modes}}(r, z, \boldsymbol{\theta}) + \phi_{\text{EJP}}(r, z, \boldsymbol{\theta}). \tag{2.31}$$

Table 2.1 Experimental design and ocean parameter choices used throughout all simulations.

Parameter	Value
Ocean sound speed (c_1, m/s)	1500
Ocean density (ρ_1, g/cm^3)	1.0
Ocean depth (h, m)	100
Frequency (f, Hz)	100
Source depth (z', m)	10
Source range (r, m)	3000
Receiver depths (z, m)	[20, 50, 80]

A proper mode is defined as a normal mode solution to the Pekeris dispersion relation that satisfies the radiation condition of converging to zero at infinite range. For the traditional Pekeris waveguide with no sediment attenuation, a finite number of proper modes exist, which are called "trapped modes". The number of proper modes M_0 is given in Eq. 2.7.1 of [5]

$$M_0 = \frac{k_1 h}{\pi}\sin(\alpha_c) + \frac{1}{2}, \tag{2.32}$$

where α_c is the critical angle $\alpha_c = \cos^{-1}(c_1/c_2)$ and k_1 is the ocean wave number $k_1 = \frac{2\pi f}{c_1}$. The sediment wave number for the case of zero sediment attenuation is

$$k_2 = \frac{2\pi f}{c_2}. \tag{2.33}$$

Attenuation is introduced into the Pekeris waveguide by adding a term to the sediment wave number:

$$k_2 = \frac{2\pi f}{c_2}(1 - i\gamma) \tag{2.34}$$

where γ is the "loss tangent", defined in [5] as the tangent of the phase angle of the complex number. Equation 2.204 of [166] gives the conversion from loss tangent to dB/λ as $\alpha_2 = 40\pi\,\gamma \log_{10}(e)$.

Sediment attenuation introduces additional proper modes, referred to as “dissipation modes". In this case, the total number of proper modes M, as found in Eq. 45 of [5], is

$$M \approx \left[1 - \frac{\gamma^2 \cot^4(\alpha_c)}{R^4}\right] \frac{R k_1 h \sin(\alpha_c)}{\pi} + \frac{1}{2}, \tag{2.35}$$

where

$$R = \sqrt{1 + \gamma^2 \cot^2(\alpha_c) \left\{1 + \left(\frac{k_1 h \cos(\alpha_c)}{\tanh^{-1}(b_{12})}\right)^2\right\}}$$

and b_{12} is the density ratio ρ_1/ρ_2. The vertical components of the wave numbers k_1 and k_2 are η_1 and η_2, respectively. The horizontal component of the wave numbers, p, is shared between k_1 and k_2 due to a requirement of continuity across the water-sediment interface. Each of the m proper modes has unique η and p values.

$$\eta_{1m} = \sqrt{k_1^2 - p_m^2} \quad , \quad \eta_{2m} = \sqrt{k_2^2 - p_m^2} \quad \text{for } m = 1, ..., M \tag{2.36}$$

To get the normal modes solution, η_1 has to be found for each proper mode. As outlined in [5], the Newton-Raphson method is used to iteratively solve for η_{1m} which is generally complex. (See [5] Eqs. 37-42 for more details.) After doing so, p_m can be calculated as $p_m = -\sqrt{k_1^2 - \eta_1^2}$. The p_m are then used to calculate η_{2m}, and the sum of proper modes can be calculated as given in the first part of Eq. 30 in [5]:

$$\phi_{\text{modes}}(r, z, \boldsymbol{\theta}) = \frac{Q}{2\pi} \left[2\pi i \sum_{m=1}^{M} \text{Res}[F_1(\eta_{1m}, \eta_{2m}) e^{i p_m r}]\right] \tag{2.37}$$

Equations 19 and 33 in [5] give F_1 and $\text{Res}[F_1]$ respectively:

$$F_1(\eta_1, \eta_2) = \frac{\sin(\eta_1 z_<)}{\eta_1} \frac{\eta_1 \cos[\eta_1(h - z_>)] + i b_{12} \eta_2 \sin[\eta_1(h - z_>)]}{[\eta_1 \cos(\eta_1 h) + i b_{12} \eta_2 \sin(\eta_1 h)]} \tag{2.38}$$

and

$$\text{Res}[F_1(\eta_{1m}, \eta_{2m}) e^{i p_m r}] = \frac{\eta_{1m} \sin(\eta_{1m} z) \sin(\eta_{1m} z')}{p_m [\eta_{1m} h - \sin(\eta_{1m} h)\cos(\eta_{1m} h) - b_{12}^2 \sin(\eta_{1m} h)\tan(\eta_{1m} h)]} e^{i p_m r}. \tag{2.39}$$

Q is the source strength (see Appendix A for more details), $z_< = \min(z, z')$, and $z_> = \max(z, z')$.

Reference [5] notes that the EJP branch line integral can be converted into a definite integral by change of integration variable to η_2 because "by definition, η_2 is real everywhere along the EJP branch line, with phases of $-\pi$ and zero, respectively, on the upper and lower edges of the cut". The EJP integral, given in Eq. 35 of [5], is

$$\begin{aligned}\phi_{\mathrm{EJP}}(r,z,\boldsymbol{\theta}) &= \frac{Q}{2\pi}\oint_{\mathrm{EJP}} F_1(\eta_1,\eta_2)e^{-ipr}dp \\ &= \frac{Q}{2\pi}\int_{-\infty}^{\infty}\frac{\eta_2}{p}F_1(\eta_1,\eta_2)e^{-ipr}d\eta_2 \end{aligned} \tag{2.40}$$

where p and η_1 are functions of η_2, $p=\pm\sqrt{k_2^2-\eta_2^2}$ and $\eta_1=\sqrt{k_1^2-p^2}$. The choice of the positive or negative square root for p is based on what gives a decaying exponential when p is complex, as discussed further in the Appendix A. This choice satisfies the radiation condition that the solution goes to zero at infinite range. The integral in Eq. 2.40 can be calculated numerically (see Appendix A for more details).

The acoustic pressure P can be calculated from the velocity potential ϕ_ω by the transformation,

$$P(r,z,\boldsymbol{\theta}) = i\omega\rho_1\phi_\omega(r,z,\boldsymbol{\theta}), \tag{2.41}$$

from which the transmission loss (TL) can be calculated as

$$\mathrm{TL}(r,z,\boldsymbol{\theta}) = 20\log_{10}\left(|P(r,z,\boldsymbol{\theta})|\right) \tag{2.42}$$

with units of dB re 1 m. TL is used in this work as the Pekeris waveguide model output.

In practice, geoacoustic inversions use other sound propagation models that allow for more complex ocean configurations (multiple layers, environmental parameters that vary with depth, bathymetry, etc). To illustrate how these complex models can be evaluated via the model manifold, the simpler ocean acoustics model of a Pekeris waveguide is used here to allow us to carefully and intuitively step through the information geometry approach to parameter sensitivity analysis for experimental design and model reduction in Sec. 2.7.2.

Table 2.2 Four sediment types: silt, sand, gravel, and basalt

Sediment	Sound Speed (c_2, m/s)	Density (ρ_2, g/cm^3)	Attenuation (α_2, dB/λ)
Silt	1575	1.7	1.0
Sand	1650	1.9	0.8
Gravel	1800	2.0	0.6
Basalt	5250	2.7	0.1

2.7.2 Global structure of the Pekeris waveguide model manifold

The global structure of the TL model manifold for the Pekeris waveguide is explored in this section. Specifically, 2-dimensional slices of the manifold corresponding to the traditional Pekeris model ($\alpha_2 = 0$) are visualized. The parameter space is illustrated in Fig. 2.3(a). Both c_2 and ρ_2 are constrained to be greater than or equal to the corresponding ocean parameter. To more easily visualize the parameter space, a change of axes is introduced using $\theta_1 = c_1/c_2$ and $\theta_2 = \rho_1/\rho_2$, as shown in Fig. 2.3(b). This inverts the parameter space axes such that the full range of θ_1 and θ_2 are between zero and one. For example, when $c_2 = c_1$ the quantity $\theta_1 = c_1/c_2 = 1$, and when $c_2 = \infty$ the quantity $\theta_1 = c_1/c_2 = 0$. The same is true for θ_2. This change of axes is only for ease in visualizing parameter spaces and does not change the predictions of the model. Trivial, bijective, re-parameterizations, such as this, do not change the model manifold because the model makes all the same predictions. The location of four sediment types, silt, sand, gravel, and basalt, are shown in Fig. 2.3(b), and the exact parameter values are given in Table C.2. Note that for this illustration, the sediment types are projected into the $\alpha_2 = 0$ plane.

The first portion of the TL model manifold to be visualized corresponds to the portion of parameter space where ρ_2 gets very large. This portion of parameter space is indicated on the parameter space inset in Fig. 2.4. Parameter ranges are c_2=[1501, 35000] m/s and ρ_2=[30, 10000] g/cm^3, where the upper bounds of these ranges are effectively infinite. The red line on the left of the

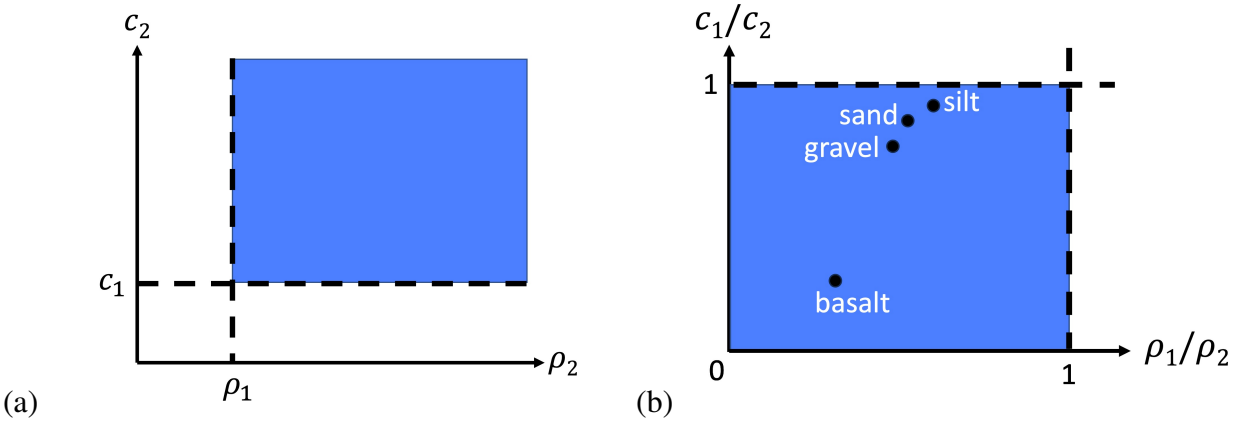

Figure 2.3 Two views of parameter space for the two-parameter Pekeris model ($\alpha_2 = 0$). (a) Parameter space for c_2 and ρ_2, where $c_2 \geq c_1$ and $\rho_2 \geq \rho_1$. (b) To more easily visualize the parameter space, a change of axes is introduced using c_1/c_2 and ρ_1/ρ_2. This inverts the parameter space axes such that the full range of c_2 and ρ_2 are between zero and one. For example, when $c_2 = c_1$ the quantity $c_1/c_2 = 1$, and when $c_2 = \infty$ the quantity $c_1/c_2 = 0$. The same is true for ρ_2. This change of axes is just for visualization of the parameter space; the parameterization of the model is not changed. Dots indicate approximate locations in parameter space for several types of sediment listed in Table 2 but projected into the $\alpha_2 = 0$ plane.

parameter space inset corresponds to $\rho_2 \to \infty$. This portion of the parameter space is mapped by the Pekeris model to the smooth model manifold shown in Fig. 2.4. The light blue line corresponds to changing c_2 for a lower ρ_2 and exhibits more variation than the yellow and light purple boundaries that correspond to changing ρ_2. In addition, more variation is seen in ρ_2 when c_2 is small (yellow line) than when c_2 is large (purple line). In the limit that $\rho_2 \to \infty$, the red line in parameter space maps to a point or corner on the model manifold.

This case illustrates how reduced-order models can be found by understanding the boundary structure of the model manifold. The red corner of the model manifold is a 0-dimensional boundary of the 2-dimensional manifold. Accordingly, this corner of the model manifold is a 0-dimensional reduced model, meaning that both c_2 and ρ_2 have dropped out of the model. This 0-dimensional reduced model is the physically interpretable case where the seafloor has become perfectly rigid and reflective, referred to here as the “hard bottom" model.

To understand how both c_2 and ρ_2 drop out of the model in this limit, first consider that, in general, at boundaries of the model manifold at least one parameter (or parameter combination) becomes totally sloppy and drops out of the model. In this situation, as ρ_2 approaches the infinite density limit, further increasing the density does not change the model prediction; ρ_2 has become sloppy and drops out of the model. Interestingly, the sound speed c_2 also becomes sloppy in this limit. As the infinite density limit is approached, changing the sound speed changes the model output less and less (seen as the manifold thinning in that dimension), until c_2 drops out of the model. Appendix **??** contains the derivation of the reduced parameter “hard bottom" model, and shows how both ρ_2 and c_2 drop out of the Pekeris waveguide equations in the infinite density limit.

The next portion of the TL model manifold visualized corresponds to the corner of parameter space with large values of c_2 and ρ_2, as shown in the parameter space inset of Fig. 2.5. Parameter ranges are $c_2 = [4500, 35000]$ m/s and $\rho_2 = [2.5, 10000]$ g/cm^3. This portion of the parameter space includes the c_2 and ρ_2 parameters corresponding to basalt (although α_2 is still zero). The

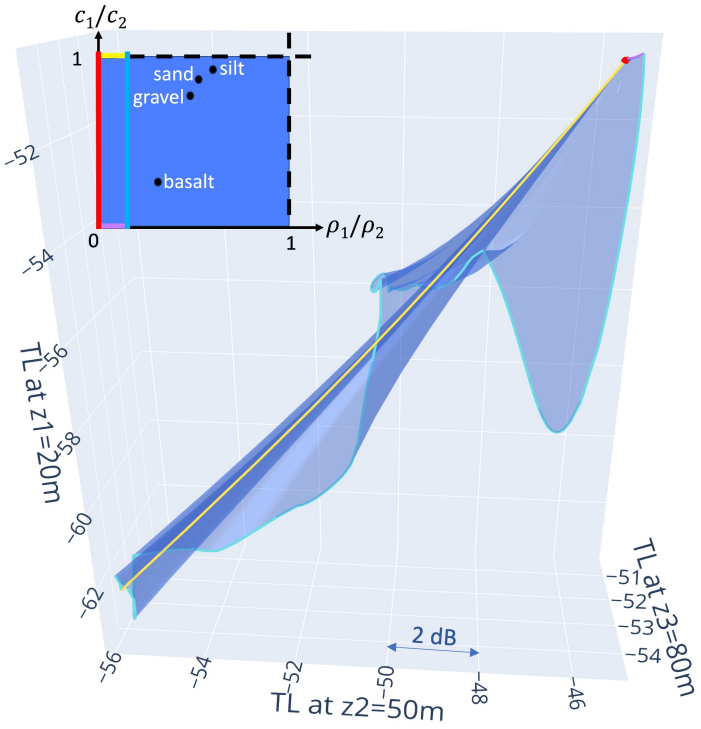

Figure 2.4 The portion of the TL model manifold corresponding to $\alpha_2 = 0$ and $\rho_2 \to \infty$, with parameter space inset. Parameter ranges are c_2=[1501,35000] m/s and ρ_2=[30,10000] g/cm^3, where the upper bounds of these ranges are effectively infinite. The colored lines in parameter space map to lines of the same colors on the model manifold. The red line in parameter space corresponding to $\rho \to \infty$ collapses to the red point on the model manifold.

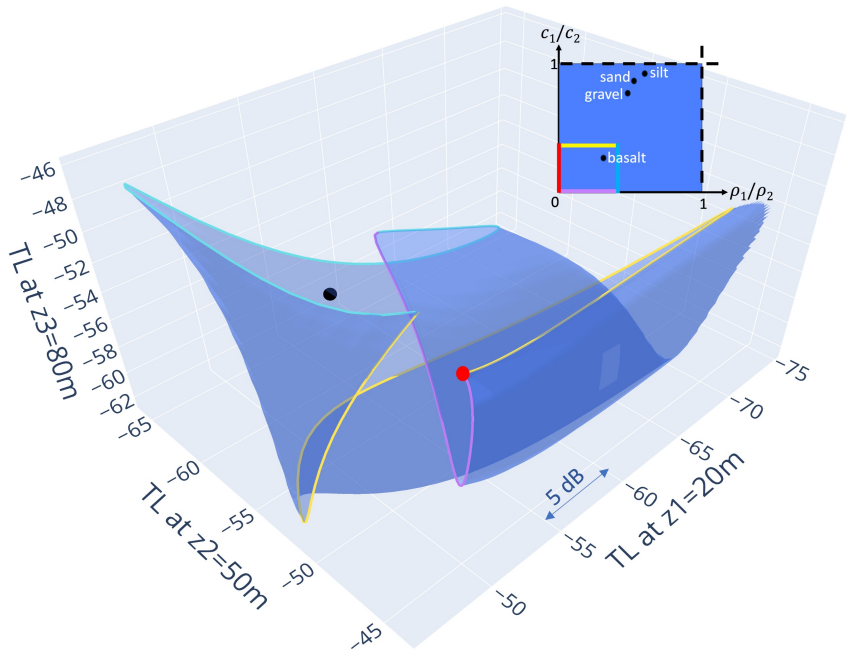

Figure 2.5 The portion of the TL model manifold corresponding to the portion of parameter space where c_2 and ρ_2 both go to infinity (for $\alpha_2 = 0$), with parameter space inset. The colored lines in parameter space map to lines of the same colors on the model manifold. Parameter ranges are c_2=[4500,35000] m/s and ρ_2=[2.5, 10000] g/cm^3. The red line in parameter space corresponding to $\rho \to \infty$ collapses to the red point on the model manifold.

model manifold shown in Fig. 2.5 is a smoothly varying 2-dimensional surface. The red dot on the manifold corresponds to the red line in parameter space, where $\rho_2 \to \infty$, the same as in Fig. 2.4. The light purple and yellow lines emanating from that point display how TL changes with variations in ρ_2 at the higher and lower values of c_2 included in this portion of the manifold. The light blue line corresponds with TL for $\rho_2 = 2.5$ g/cm^3 and varying c_2.

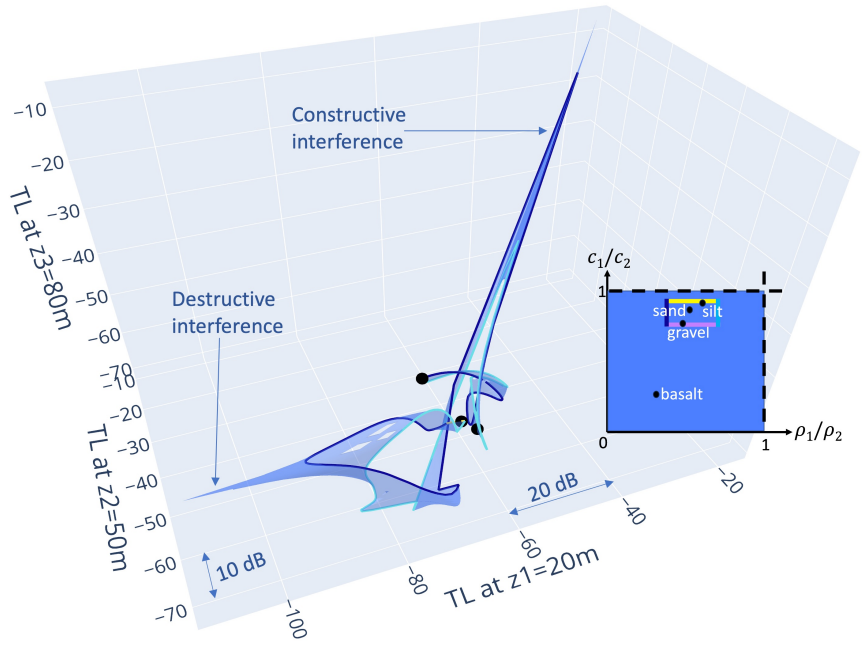

Figure 2.6 The portion of the TL model manifold corresponding to a box in parameter space around silt, sand, and gravel (for $\alpha_2 = 0$). Parameter space is inset. The colored lines in parameter space map to lines of the same colors on the model manifold. Parameter ranges are c_2=[1575, 1800] m/s and ρ_2=[1.6, 2.1] g/cm^3.

Moving away from the $c_2 \to \infty$ and $\rho_2 \to \infty$ corner, one additional section of the TL model manifold is examined in the vicinity of gravel, sand and silt (all projected into the $\alpha_2 = 0$ plane). The model manifold is shown in Fig. 2.6, with the parameter space as the inset. The parameter bounds of this portion of the manifold are c_2=[1575,1800] m/s and ρ_2=[1.6,2.1] g/cm^3. The black dots correspond to the parameter values for gravel, sand, and silt listed in Table 2 projected onto the $\alpha_2 = 0$ plane. The model manifold looks like a long, relatively narrow ribbon with two peaked extensions.

One observation about this portion of the manifold is that the mapping of sand, silt, and gravel into data space places the sediments closer to each other in data space than to other combinations of c_2 and ρ_2 that lie between them in parameter space, due to the model manifold folding over itself. This observation has consequences for parameter inference: A measured data point containing

noise will lie off the model manifold and might be equally likely to map to these different sediment types. Additionally, the relative stiffness/sloppiness of c_2 and ρ_2 can be seen on the manifold. The dark blue and light blue sides corresponding to changes in c_2 are much longer than the yellow and purple sides corresponding to changes in ρ_2 (which cannot be seen very well in the plot due to being covered by the silt and gravel dots).

The large extension, on the top right of Fig. 2.6, corresponds to a combination of c_2 and ρ_2 at which constructive interference occurs at all three receiver depths. The spike on the bottom left of Fig. 2.6 is an instance where there is a null (destructive interference) in TL for a single receiver depth at the corresponding values of c_2 and ρ_2. These extensions are interesting examples of how the model manifold portrays intermediate scale model structure. (A deeper study has indicated that the extreme nature of the larger constructive interference spike occurs due to a change in the number of proper modes used in calculating the model output; however, its presence is still illustrative of how the model manifold can capture relevant features of sound propagation such as constructive and destructive interference.)

The FIM gives information about the model at a single point and the boundaries of the manifold give information about parameters taken to their extreme limit, but neither of those gives information about the constructive and destructive interference that is an important feature of sound propagation models. Non-local measures of parameter sensitivity, such as sampling based methods or rotated coordinates discussed in Sec. 2.2 and Sec. 2.5, also fail to adequately capture such intermediate structure, as they provide somewhat of an averaged value of parameter sensitivity. Neither do these non-local methods provide information about global parameter sensitivities like boundaries of the model manifold can, such as the emergent “hard bottom" model discussed in this section. In summary, the model manifold preserves the local parameter sensitivities of a non-linear model, while also informing about larger scale model structure and global parameter sensitivities.

The TL model manifold slices shown here have all been considering the traditional Pekeris waveguide where the sediment attenuation is zero. However, this 2-dimensional manifold itself is a boundary of the TL model manifold for the three parameter Pekeris waveguide with nonzero sediment attenuation, a 2-dimensional boundary (face) of a 3-dimensional model manifold volume. This boundary corresponds to the limit where $\alpha_2 \to 0$. The next section contains calculations of local parameter sensitivities and correlations using the FIM in different parts of the 3-dimensional parameter space, and discuss how proximity to the global structures (boundaries) of the model manifold impacts local parameter sensitivities.

2.7.3 Local Pekeris parameter sensitivities using the FIM

In this section, the FIM is calculated at four different sediment types in the 3-dimensional parameter space (i.e., the Pekeris model with sediment attenuation). The parameters of the four sediment types are in Table C.2. The full 3-dimensional parameter space is shown in Fig. 2.7(a) and (b). The bottom face of the cube in Fig. 2.7(a) is at $\alpha_2 = 0$, and is the same parameter space shown in in Fig. 2.3(b). The orange, yellow, green, and blue dots represent the silt, sand, gravel, and basalt sediment types, respectively, and the empty dots below are their projection into the $\alpha_2 = 0$ plane (as used in Sec. 2.7.2). Because typical values for each θ_i vary by five orders of magnitude in SI units, a variable transformation is used to remove the impact of the different units (i.e., scales). This transformation uses the natural log of the parameters as illustrated in Fig. 2.7(b); the natural log of the parameters c_2, ρ_2, and α_2 are now on the axes, which allows the parameter uncertainties to be more easily compared. The axes start at $\log(\rho_2 = \rho_1)$ and $\log(c_2 = c_1)$ for density and sound speed, and the bottom attenuation face is an arbitrary choice of α_2 because the log range of α_2 is from $-\infty$ to ∞. The colored dots in (b) are the same as in (a).

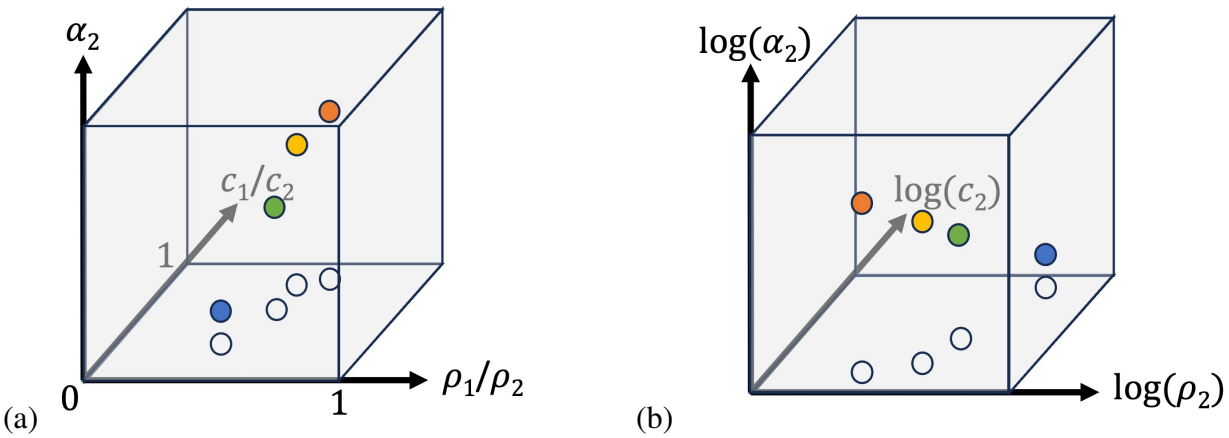

Figure 2.7 Schematics of the 3-dimensional parameter space used in the modified Pekeris waveguide. The bottom face of the cube in Fig. 2.7(a) corresponds to the parameter space at $\alpha_2 = 0$ shown in Fig. 2.3(b). The orange, yellow, green, and blue dots represent the silt, sand, gravel, and basalt sediment types, respectively, and the empty dots below are their projection into the $\alpha_2 = 0$ plane (as used in Sec. 2.7.2). To remove the dependence on units and facilitate easier comparison of uncertainties, Fig. 2.7(b) shows same parameter space but with the natural log of the parameters c_2, ρ_2, and α_2 as the axes. The axes start at $\log(\rho_2 = \rho_1)$ and $\log(c_2 = c_1)$ for density and sound speed, and the bottom attenuation face is an arbitrary choice of α_2 because the log range of α_2 is from $-\infty$ to ∞. The colored dots in (b) are the same as in (a).

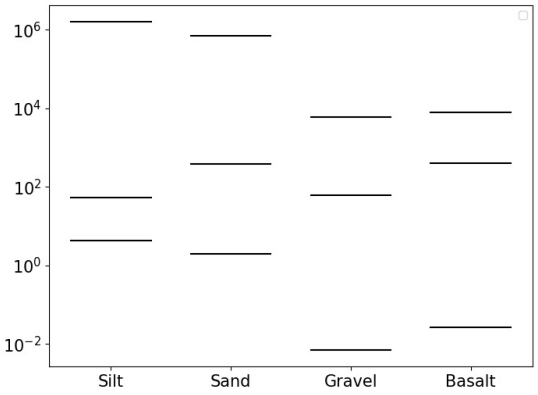

Figure 2.8 FIM eigenvalues for different ocean sediment types. Notice magnitude differences between the three eigenvalues.

As discussed in Sec. 2.3, the FIM is a local measure of parameter sensitivities. The eigenvectors of the FIM are the uncorrelated parameter combinations (at that point in parameter space), and the corresponding eigenvalues indicate the stiffness/sloppiness of that parameter combination. The FIM eigenvalues and eigenvectors are calculated at each of the four sediment types.

The eigenvalues of the FIM for each sediment type are displayed in Fig. 2.8. Larger eigenvalues indicate stiffer eigenvector parameter combinations. The three eigenvalues for each sediment span 5-6 orders of magnitude, demonstrating the phenomenon of sloppiness.

The FIM eigenvectors are visualized in Fig. 2.9 by plotting the *participation factors* for each parameter. The participation factors quantify how much each bare parameter θ_i “participates" in the eigenvectors. In this case, because the FIM is real and symmetric the participation factor is the squared components of the normalized eigenvector; see discussion and definitions in [88, 167] for more details. The eigenvectors $\boldsymbol{e_1}$, $\boldsymbol{e_2}$, and $\boldsymbol{e_3}$ are ordered from largest to smallest eigenvalues.

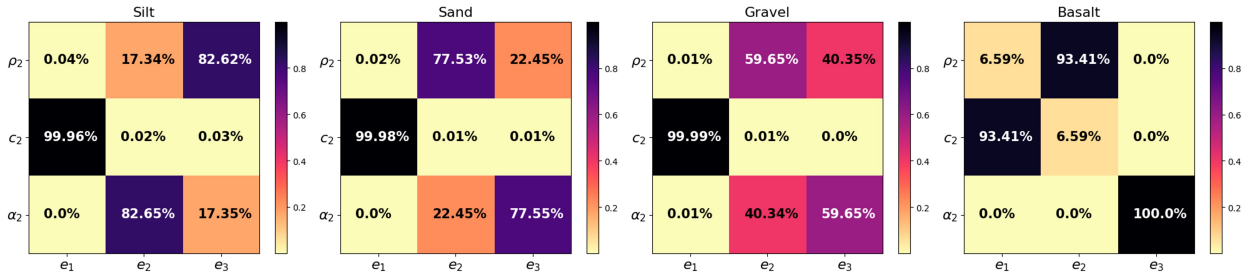

Figure 2.9 FIM eigenvector participation factors. The participation factor indicates which parameter components contribute most to the eigenvector direction (i.e., the square of the components of the normalized eigenvectors).

For silt, sand, and gravel, the first eigenvector almost completely aligns with the c_2 parameter, while a correlation exists between ρ_2 and α_2 that gets stronger as c_2 and ρ_2 increase and α_2 decreases. For basalt, which lies in a very different part of parameter space, a small correlation exists between c_2 and ρ_2, while the third eigenvector is totally aligned with α_2. Thus, the relevant parameter combinations to consider in doing parameter sensitivity analysis change depending on the region of parameter space being considered. This is a direct consequence of the nonlinearity of the model.

The Cramér-Rao bound $\sigma_{\mathrm{CRB},j}$ and conditional STD $\sigma_{\mathrm{cond},j}$ are calculated from the FIM as defined in Eq. Eq. 2.22 and 2.23, respectively. $\sigma_{\mathrm{CRB},j}$ and $\sigma_{\mathrm{cond},j}$ for c_2, ρ_2, and α_2 are presented in Table 2.3. $\sigma_{\mathrm{CRB},j}$ gives the lowest possible uncertainty of the parameter. Parameter correlations increase the uncertainty in inferring parameter values and, therefore, increase the $\sigma_{\mathrm{CRB},j}$. For example, because ρ_2 and α_2 are highly correlated in gravel (as shown in Fig. 2.9), the $\sigma_{\mathrm{CRB},j}$ for ρ_2 and α_2 in gravel are higher than they are for other sediment types. $\sigma_{\mathrm{cond},j}$ gives the estimated uncertainty if the parameters are not correlated. $\sigma_{\mathrm{cond},j}$ values roughly agree with the $\sigma_{\mathrm{CRB},j}$ values in the table but the difference between the estimates of $\sigma_{\mathrm{CRB},j}$ and $\sigma_{\mathrm{cond},j}$ for gravel due to parameter correlation is obvious.

Table 2.3 The conditional STD and CRB for each parameter from the Fisher information matrix at these four sediment types. Both the $\sigma_{\text{cond},j}$ and $\sigma_{\text{CRB},j}$ are small for stiff parameters and large for sloppy parameters. The differences between $\sigma_{\text{CRB},j}$ and $\sigma_{\text{cond},j}$ are primarily due to parameter correlations.

Sediment	CRB ($\sigma_{\text{CRB},j}$)			Conditional STD ($\sigma_{\text{cond},j}$)		
	c_2	ρ_2	α_2	c_2	ρ_2	α_2
Silt	7.98e-3	4.41e-1	2.36e-1	7.93e-4	3.96e-2	1.13e-1
Sand	5.42e-3	3.38e-1	6.24e-1	1.19e-3	4.87e-2	1.05e-1
Gravel	1.73e-2	7.57	9.21	1.29e-2	1.66e-1	2.02e-1
Basalt	1.75e-2	4.90e-2	6.14	1.16e-2	3.32e-2	5.88

For both of the metrics in Table 2.3, stiffer parameters have smaller values, while sloppier parameters have larger values. For all four sediment types, $\sigma_{\text{cond},j}$ indicates that c_2 is the stiffest parameter, followed by ρ_2, and then α_2 is the sloppiest parameter. The same ordering of parameter sensitivities are given by $\sigma_{\text{CRB},j}$ and the eigenvalues of the FIM, except in the case of silt where α_2 appears to be a stiffer parameter than ρ_2. This small difference between relative stiffness/sloppiness rankings from different metrics is most likely due to different treatments of parameter correlations; the conditional STD doesn't consider correlations, the CRB does consider correlations, and the eigenvalues of the FIM actually give the sensitivities of the relevant parameter combination (which in this case roughly align with the bare parameters, as shown in Fig. 2.9). Despite these differences, it can be concluded that for realistic sediment types, c_2 is the stiffest parameter, followed by ρ_2 and α_2.

Additionally, Table 2.3 highlights how the parameter sensitivities of individual parameters change in different parts of parameter space; these changes in parameter sensitivities can be understood in context of the global model manifold structure. The boundaries of the model manifold correspond to reduced-order models where one parameter becomes totally sloppy. Therefore, as

a parameter approaches the extreme limit corresponding to a boundary, that parameter becomes increasingly sloppy. Evidence of this is seen with α_2 for the adapted Pekeris waveguide. The attenuation α_2 for basalt is 0.1 dB/λ, about an order of magnitude smaller than the attenuation of silt, sand, and gravel (see Table C.2). Interestingly, Table 2.3 shows that the CRB and conditional STD for basalt are an order of magnitude larger than for the other sediment types. Thus as $\alpha_2 \to 0$, α_2 becomes an increasingly sloppy parameter. Therefore, proximity to the $\alpha_2 = 0$ boundary affects the locally measured sloppiness of α_2.

This same phenomenon is observed in the relative stiffness/sloppiness of c_2 for different sediment types. The first row of Table 2.3 shows that sound speed c_2 is the stiffest parameter by far for silt, evidenced by the two orders of magnitude difference between the CRB and conditional STD values for c_2 and ρ_2. Moving down the rows of Table 2.3, c_2 becomes less and less stiff until at basalt, the CRB and conditional STD for c_2 and ρ_2 are on the same order of magnitude. This increase in sloppiness of c_2 is due to the c_2 value approaching two different boundaries on the model manifold, the $c_2 \to \infty$ boundary or the $\rho_2 \to \infty$ boundary (in which both c_2 and ρ_2 become sloppy as shown in Sec. 2.7.2). This example demonstrates how local parameter sensitivities are governed by the vicinity to global structures (boundaries) on the model manifold. Additionally, this highlights the importance of considering how parameter sensitivities change over parameter space, because a single local or non-local measure of sensitivity doesn't give the complete picture.

2.7.4 Model manifold widths manifest sloppiness

The model manifold contains information about global, intermediate, and local model structure. Therefore the "hierarchy" of parameter sensitivities seen in Sec. 2.7.3 can be observed using the model manifold. This point is illustrated by focusing on the area of parameter space near sand. The FIM at sand indicates that c_2 is the stiffest parameter followed by ρ_2 and α_2.

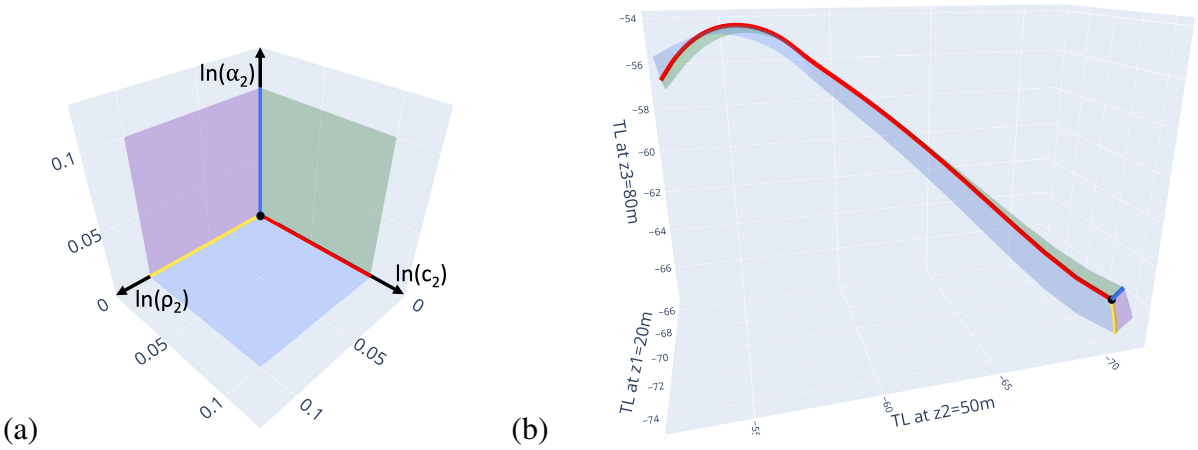

Figure 2.10 (a) A half-cube sampled in log parameter space and (b) the anisotropic widths of the TL model manifold, related to the parameter sensitivities at sand. The widths of the parameter space cube corresponds to a step of 0.1 in log space. The black dot at the corner is sand ($\rho_2 = 1.9$ g/cm^3, $c_2 = 1650$ m/s, and $\alpha_2 = 0.8$ dB/λ). The different widths of the model manifold reflect the same information contained in the FIM at sand, specifically that c_2 is the stiffest (most identifiable) parameter, followed by ρ_2 and then α_2.

Three adjacent TL model manifold sections are calculated and placed on the same axes for the faces of a half-cube in parameter space, shown in Fig. 2.10(a). Because the natural log of the parameters is used, each “step" in the three parameter directions is the same step size. A step of 0.1 in log parameters is used. Because of the different parameter sensitivities for c_2, ρ_2 and α_2, the model transforms this half-cube in parameter space into highly anisotropic manifolds, shown in Fig. 2.10(b). The red, blue, and yellow bold lines are where the three faces meet and two parameters are constant and one varies (e.g., for the red line ρ_2 and α_2 are fixed values, and c_2 varies). The lengths of these edges show the same hierarchy of stiffness as seen in Table 2.3: the red line, corresponding to varying c_2, is the longest, meaning that c_2 is the stiffest parameter because the model output responds most to a change in the sound speed. The yellow line for varying ρ_2 is next longest meaning that ρ_2 is the next stiffest parameter, followed by the shortest blue line for varying α_2, which is the sloppiest parameter.

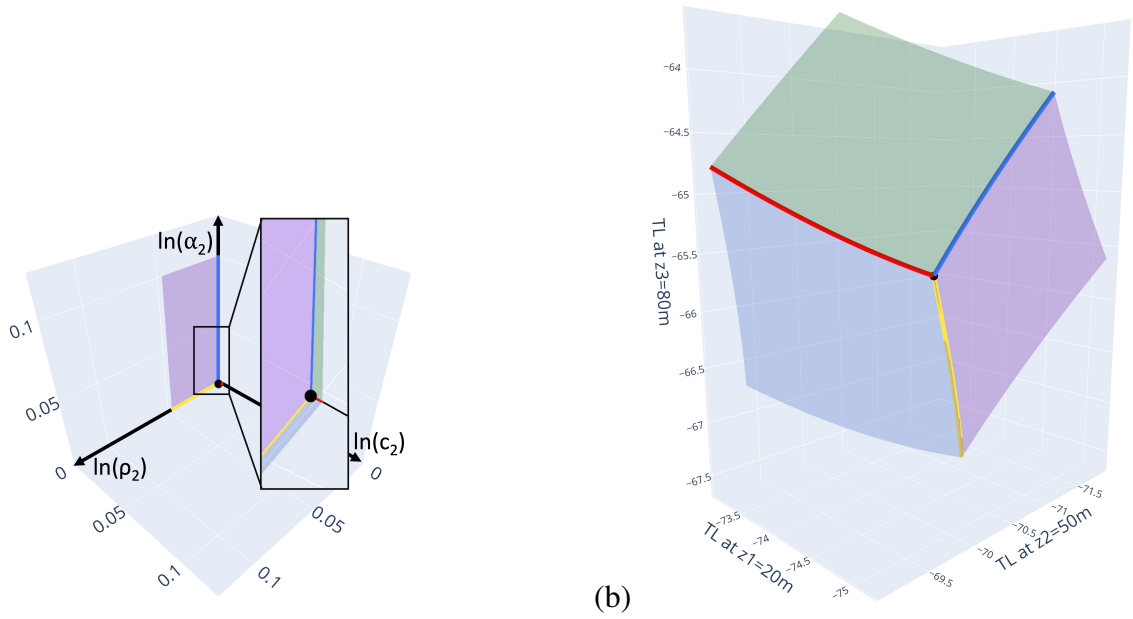

Figure 2.11 (a) Parameter space where the log parameter ranges are scaled by the conditional STD calculated from the FIM at sand. Note that the blue line here is the same length as the blue line in parameter space in Fig. 2.10(a) (0.1 log parameter step). (b) The TL model manifold is now scaled such that its widths are approximately equal.

To illustrate this parameter sensitivity hierarchy slightly differently, three model manifolds are again constructed, but the widths of the parameter ranges are scaled by the conditional STD of the parameters, such that the parameter space rectangular prism is very thin in some parameters (e.g. c_2), as shown in Fig. 2.11(a). The model manifold for these sections of parameter space, shown in Fig. 2.11(b), becomes roughly isotropic in all three parameters with this scaling. The two model manifold visualizations in Figs. 2.10(b) and 2.11(b) illustrate how the local parameter sensitivities, as found in the FIM, are similarly represented in the model manifold.

2.8 Outlook for application of information geometry

This paper has demonstrated some applications of information geometry tools to underwater acoustics, but there are many other possible applications. One application of these information geometry tools is to inform optimal experimental design, i.e., to maximize the amount of information contained in experimental data about parameters of interest. Optimal experimental design could include deciding on optimal positioning for receivers or choosing the source-receiver range to measure at, given the desired parameter to infer and, for example, the source frequency. In this scenario, visualizing slices of the model manifold may reveal that the choice of certain receiver depths spreads out the model manifold and is, therefore, more informative. Similarly, the FIM may be constructed for different experimental configurations (e.g., source-receiver geometries) and analyzed to determine parameter uncertainties. In this manner the information geometry tools provide criteria for choosing which experimental design is most informative. Other work discussing the applications of information geometry to experimental design include [168].

Additionally, more advanced future work includes using computational differential geometry methods to obtain reduced parameter models that contain only the most identifiable parameters. Using the insight that parameters become increasingly sloppy as model manifold boundaries are approached, the Manifold Boundary Approximation Method (MBAM) [125] algorithmically identifies potential reduced models corresponding to boundaries of the model manifold by calculating geodesics on the model manifold. MBAM has been successfully used for model reduction in a variety of fields including models of nuclear structure [119, 169, 170], cardiac physiology [78], power systems [101, 103, 104], biochemistry [72, 84, 125, 126], and epidemics [126]. Additionally, Paré *et al.* [171] used MBAM to find a connection between model reduction methods of balanced truncation and balanced singular perturbation approximation as limiting approximations of linear time invariant systems.

Other methods for model reduction based in information geometry have been proposed as well. Raman *et al.* [76] introduces a method for parameter reduction similar to MBAM that, instead of utilizing geodesics, simulates a particle trajectory in parameter space. The particle trajectory, starting at the best fit, is calculated via a Hamiltonian and traces out the portion of parameter space associated the most unidentifiable parameter combination, which can be removed from the model. Zhao *et al.* [172] presents an information geometry based model reduction method for models with parameter spaces of binary multivariate distributions called Confident Information First, which projects the model manifold onto a submanifold that preserves distinguishability. Holiday *et al.* [173] demonstrates a data driven, diffusion map-based method for learning an effective low-dimensional parameterization of a sloppy model, based only on the model's input-output relations; this approach is shown to be effective for a number of demonstrative and practical modeling scenarios. The present work sets the stage for the application of information geometry based model reduction methods to sound propagation models, enabling more advanced geoacoustic inversions.

Information geometry has also led to methodological advances in other areas, some of which are summarized here. Similar advances towards applications in ocean acoustics could be possible. In the area of visualization, Quinn *et al.* [111] introduced Intensive PCA, a nonlinear alternative to traditional PCA, applicable to categorical data. Teoh *et al.* [174] introduces a similar intensive embedding, known as the intensive symmetric Kullback-Leibler (isKL), for analytically embedding statistical models from an exponential family into a 2N-dimensional Minkowski space. In optimization, Transtrum *et al.* [175] improved convergence of the Levenberg-Marquardt algorithm by introducing geodesic acceleration to overcome challenges with the algorithm getting stuck in singular regimes of parameter space close to boundaries of the model manifold for sloppy models [176]. In the area of renormalization groups, Berman *et al.* [177] introduces the idea of Bayesian renormalization, interpreting the process of evolving the 'current posterior' toward the true solution as the inverse of a coarse-graining process. Under this interpretation, the Bayesian renormalization

method is introduced which coarse-grains the model according to statistical distinguishability, as defined by the Fisher information matrix. Girolami *et al.* [178] improved the efficiency of Monte Carlo sampling methods for high dimensional situations by considering the Riemannian structure of the parameter space of statistical models. Information geometry has also been used to improve Bayesian prior selection. In Mattingly *et al.* [179], the authors select an optimal prior that assumes finite amount of data, which alleviates challenges that arise when using a prior which assumes infinite data for sloppy models. LaMont *et al.* [180] proposes the learning capacity, analagous to heat capacity in statistical mechanics, to quantify model sloppiness, and additionally proposes choosing of a Bayesian prior by grouping indistinguishable models, allowing better performance when model dimension is unknown. These examples demonstrate that there are many opportunities for information geometry to influence traditional computational methods and likely have a similar effect in the area of ocean acoustics.

2.9 Summary

In summary, this tutorial review paper demonstrated the application of information geometry methods to ocean acoustics modeling. First, a review of different methods for parameter sensitivity analysis that have been applied to problems in underwater acoustics was given in Sec. 2.2. Additionally, references were given for the application of information geometry methods to sloppy models in a variety of fields, setting the stage for their demonstration in this paper. In Sec. 2.3, a tutorial explained the Fisher information matrix and its interpretation, and specifically how information regarding parameter uncertainty, such as Cramér-Rao bounds, can be extracted from the FIM. In

context of the FIM, a more thorough review of the application of Cramér-Rao bounds, a local sensitivity method, to analysis of environmental parameters in ocean acoustics was provided in Sec. 2.4. A review of rotated coordinates as a non-local method was given in Sec. 2.5, and its connections to and differences from the FIM were discussed.

The bulk of the paper introduced and demonstrated the application of information geometry methods to the problem of environmental parameter sensitivities in ocean acoustics. In Sec. 2.6, the construction of the model manifold was illustrated for a toy sum of exponentials model, and the connection between the FIM and the model manifold were discussed. A tutorial in Sec. 2.7 explained the construction of the model manifold for a simple underwater sound propagation model, the Pekeris waveguide. An adaptation of the Pekeris waveguide from [5] was used that includes an additional sediment attenuation parameter α_2. The three parameters varied were the sediment sound speed c_2, sediment density ρ_2, and sediment attenuation α_2. One parameter was held fixed while the others are varied to construct 2-dimensional TL model manifolds in the 3-dimensional data space, due to the challenges of accurately visualizing a 3-dimensional volume.

In Sec. 2.7.2 the case of the traditional Pekeris waveguide (α_2=0) was considered, and TL model manifolds were constructed in the region of parameter space near the infinite density and sound speed limits, in addition to manifolds around different sediment types (at zero attenuation). These manifolds showed how model parameters become sloppy at the model manifold boundaries, such as the infinite density boundary. Additionally the manifold section around silt, sand, and gravel showed how distinguishing some sediment types by parameter inference can be challenging. Overall, it was demonstrated how the model manifold holds information about local, intermediate, and global model structure. In Sec. 2.7.3, the FIM was calculated at four different sediment types. The eigenvectors and eigenvalues of the FIM were calculated, along with the CRB and conditional STD, giving quantitative measures of parameter correlation and sensitivities. Generally, c_2 was the stiffest parameter, followed by ρ_2 and α_2. Additionally, a discussion was provided about how

local sensitivity properties are impacted by the global structure of the model manifold. Then Sec. 2.7.4 contained TL model manifold sections that illustrate how the model manifold encodes the same local parameter sensitivity information contained in the FIM, which can be understood by considering the FIM as a distance metric on the model manifold as discussed in Sec. 2.6.2.

Finally, Sec. 2.8 provided discussion for how information geometry methods can impact problems in ocean acoustics going forward. Of particular note are the applications towards experimental design and obtaining reduced order models, to enable more accurate geoacoustic inversion. This paper introduced the information geometry method of the model manifold in context of other methods for sensitivity analysis traditionally used in problems in ocean acoustics, laying the foundation for further applications of information geometry.

In conclusion, the model manifold preserves the local parameter sensitivities of a non-linear model, while also informing about larger scale model structure and global parameter sensitivities. The model manifold builds upon local measures of parameter sensitivity, as given by the FIM, and shows how sensitivities change over parameter space. In the extreme limits of model parameters, boundaries of the model manifold correspond to physically meaningful reduced-order models, thus quantifying global parameter sensitivities. The model manifold thus demonstrates advantages over non-local methods of determining parameter sensitivity, such as sampling based methods, which provide only an averaged value of parameter sensitivity and fail to capture global sensitivities. Promising future applications of information geometry methods in ocean acoustics include experimental design and reduced order modeling for geoacoustic inversion.

Acknowledgments

The authors thank the Office of Naval Research for their support of this work under contract number N00014-21-S-B001.

Chapter 3

Information geometry analysis example for absolute and relative transmission loss in a shallow ocean

This chapter includes a paper submitted to the Journal of the Acoustical Society of America, Express Letters (JASA-EL) entitled *Information geometry analysis example for absolute and relative transmission loss in a shallow ocean*, with authors Jay C. Spendlove, Tracianne B. Neilsen, and Mark K. Transtrum. This paper demonstrates use of the model manifold to inform experimental design by comparing the differences in information content between absolute and relative transmission loss about parameters of interest. Supplementary material for this paper is included in Appendix C of this thesis, and includes sound propagation model parameters for replication and additional relative TL model manifold visualizations and related information.

Abstract

The model manifold, an information geometry tool, is a geometric representation of a model that can quantify the expected information content of modeling parameters. For a normal-mode sound propagation model in a shallow ocean environment, transmission loss (TL) is calculated for a vertical line array and model manifolds are constructed for both absolute and relative TL. For the example presented in this paper, relative TL yields more compact model manifolds with seabed environments that are less statistically distinguishable than manifolds of absolute TL. This example illustrates how model manifolds can be used to improve experimental design for inverse problems.

3.1 Introduction

A major goal of underwater acoustical modeling is to gain insight into how acoustical measurements, such as transmission loss (TL), relate to environmental parameters and, thus, estimate what environmental information is or is not encoded in acoustical data. Intuitively, if physically distinct seabeds lead to nearly identical acoustical data, those data cannot be used to distinguish the seabeds. Model parameters associated with these indistinguishable seabed properties are said to be unidentifiable. Parameter identifiability can inform both experimental design and model selection. This paper demonstrates an information geometry approach to choosing experimental design such that information about environmental parameters is maximized.

Information geometry [108, 181, 182] is a branch of mathematics that combines statistics and information theory with differential geometry. A multi-parameter model can be interpreted geometrically as a Riemannian manifold, known as the model manifold. The model manifold exists as a curved, high-dimensional hyper-surface in an ambient “data space," where each point on the manifold corresponds to a specific model prediction. The dimensionality of the model manifold is the number of parameters to be inferred, where the parameters act as coordinates on the manifold.

The dimensionality of the ambient data space corresponds to the number of model predictions. In this paper, we use a two-parameter TL model with a vertical line array (VLA) of 15 receivers, resulting in a two-dimensional (2D) model manifold embedded in a 15-dimensional (15D) ambient data space.

The geometric structure of the model manifold connects the information content of data to model parameters. Distance on the model manifold quantifies the statistical distinguishability of different model predictions. The Fisher information matrix is a distance metric on the model manifold, providing a connection to other information-theoretic tools for parameter identifiability analysis such as Cramér-Rao bounds. Importantly, however, the model manifold connects this local measure of identifiability with nonlocal and global properties of the parameter space. The distance between different parameter values in the ambient data space incorporates all nonlinearities of the model into a single measure of statistical distinguishability.

Additionally, the model manifold encodes global parameter sensitivities within the structure of its finite boundaries, a nonlinear effect. Model manifold boundaries correspond physically interpretable simplified models, such as completely removing sediment layers from a more complex geoacoustic profile. Algorithms such as the manifold boundary approximation method (MBAM) [72, 183] can be utilized to find these manifold boundaries, identifying reduced-order models that can retain model accuracy while being physically interpretable. The visualizations in this paper motivate starting cases for testing MBAM on ocean sound propagation models. The effectiveness of model reduction methods such as MBAM rely on the fact that model manifolds in general contain a hierarchy of widths, giving model manifolds a “ribbon-like" structure. This structure is equivalent to saying that multi-parameter models, in general, manifest parameter sensitivities spanning many orders

of magnitude, a phenomenon known as sloppiness, [184–186] which is characterized by roughly log-linear spaced eigenvalues of the Fisher information matrix. Wide dimensions of the model manifold correspond to stiff parameter combinations to which the model is most sensitive, while thin dimensions correspond to sloppy parameter combinations.

In summary, the model manifold directly connects the information content of data to model parameters in both local and global ways. Thus, the model manifold approach extends traditional, local sensitivity analyses in ways that have potential applications for model selection and experimental design. For experimental design, the model manifold depends on independent variables such as frequency and source-receiver geometry, in addition to the model parameters and, thus, quantifies the potential information content in data for different experimental setups. A more “spread out" model manifold corresponding to some experimental design choice indicates that choice of acoustical data contains more information about parameters of interest, making parameter inference more accurate. Information geometry is, therefore, a natural framework to explore questions of experimental design.

This paper demonstrates the utility of the model manifold in guiding experimental design by quantifying the impact that using relative TL has on the information content of data for the example of a 15 element vertical line array (VLA) in a shallow ocean environment. The TL model manifold is generated by a range-independent normal-mode model, and model manifold visualizations are shown of 2D slices of the data space at different receiver depths to illustrate the impact of experimental design. Principal Component Analysis (PCA) is employed to obtain representative 2D visualizations of the principal variations across a 15D model manifold. Additionally, Euclidean distances on the manifold between seabed types are calculated in the 15D data space. Visualizations of the model manifold for both absolute and relative TL for different receiver depths show that, in this example, relative TL model manifolds are more compact, with seabeds types that are less statistically distinguishable and do not follow expected similarity trends.

3.2 Method

The sound propagation model selected to compute TL is ORCA, a range-independent normal-mode model [187]. At a selected frequency f and for a specified ocean environment $\boldsymbol{\theta}$ and source-receiver configuration, ORCA calculates the depth-dependent mode functions and modal eigenvalues, which yield the Green's function:

$$p(\boldsymbol{\theta}, r, z, z_s) = \sqrt{\frac{2\pi}{r}} e^{i\pi/4} \frac{1}{\rho_s} \sum_n \frac{\bar{\phi}_n(\boldsymbol{\theta}, z)\bar{\phi}_n(\boldsymbol{\theta}, z_s) e^{ik_n(\boldsymbol{\theta})r}}{\sqrt{k_n(\boldsymbol{\theta})}}, \tag{3.1}$$

where $\phi_n(\boldsymbol{\theta}, z)$ is the n-th depth-dependent mode function for environment $\boldsymbol{\theta}$ evaluated at receiver depth z or source depth z_s, $k_n(\boldsymbol{\theta})$ is the n-th modal eigenvalue, ρ_s is the water density at the source, and r is the horizontal range between source and receiver. In addition to the explicit arguments r, z, and z_s, the Green's function implicitly depends on all of the seabed parameters, water depth and sound speed, and frequency. An incoherent sum of modes would provide a more realistic model of a broadband signal; however, in this work we follow the definition given in the original ORCA paper [187], which assumes a single frequency. Future work will explore the incoherent sum of modes formulation. The transmission loss is computed as

$$\mathrm{TL}(r, z, z_s) = -20 \log(|p(r, z, z_s)|/|p_0|), \tag{3.2}$$

in units of dB re 1 m, where $|p_0|$ is the value 1 m from the source.

This work uses a relatively simple ocean environment with an isovelocity water sound speed of 1500 m/s, water depth of 75 m, and a seabed consisting of a 35 m sediment layer over a half space. The sediment layer is parameterized using the compressional sound speed c_p, density ρ, and attenuation α_p, which is held fixed at $\alpha_p = 0.63$ dB/m/kHz in this paper. This value corresponds to the estimated value of attenuation for silt given in Jensen *et al.* [1]: for $c_p = 1575$ m/s,

$\alpha_p = 1.0$ dB/$\lambda = 0.63$ dB/m/kHz, which is larger than for most sediments. The experimental design parameters include $z_s = 6$ m, $r = 3$ km, 15 receiver depths z evenly spaced between 5 m and 75 m, and a frequency of 100 Hz. A table of the ORCA parameters used, including half space parameters, is included in Appendix C.

The model manifold is an N-dimensional surface consisting of all possible model predictions for N variable parameters, embedded in an M-dimensional data space with perpendicular axes, in this case, corresponding to TL at M choices of receiver depth. For this example, two environmental parameters, sediment sound speed c_p and density ρ, are varied (N=2), with sampled points shown in the parameter space in Fig. 3.1(a). Model manifolds are constructed by calculating the TL for each of these parameter choices at two different receiver depths (M=2). The model manifolds corresponding to TL at receiver depths of 5 m and 65 m, and 25 m and 30 m, are shown in Fig. 3.1(b)-(c), respectively. These model manifolds demonstrate the characteristic ribbon-like structure of a sloppy model due to the nonlinear model transformation. These model manifolds in Fig. 3.1(b) and (c) are qualitatively different, signaling differences in information content between channels; this observation holds for selection of any subset of receiver depths.

The model manifolds in Fig. 3.1(b)-(c) retain the same adjacency relations of the colored edges in Fig. 3.1(a), other than the apparent self-intersection of the manifold, which in general disappears in higher dimensional data spaces. (Note that any true self-intersection of the model manifold does not correspond to any continuous shorter path between parameter choices in parameter space; practically this just introduces additional ambiguities in projecting measured data onto the model manifold.)

The location for five seabed sediment types (mud, clay, silt, sand, gravel) are also indicated on the parameter space and the TL manifolds in Fig. 3.1. (The parameter values are listed in Appendix C). Distance between two points on the model manifold in data space quantifies the statistical distinguishability between those parameter values. The Euclidean distance between model outputs

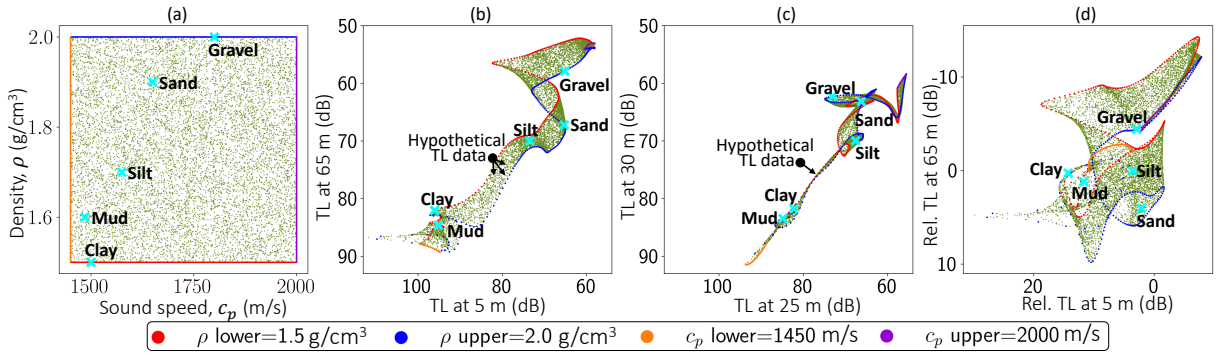

Figure 3.1 (a) Parameter space showing the sampling of seabed sound speed c_p and density ρ. The ORCA model predicts TL for each sample point (including the colored bounds) at multiple receiver depths to construct the model manifold. (b) A 2D model manifold visualization with axes of TL at receiver depths 5 m and 65 m for a 100 Hz source. Five seabed sediment types are marked on the manifold. Hypothetical observed data (black dot) contains noise and so in general does not lie on the model manifold, but it can be projected (non-uniquely) onto the manifold. (c) A 2D model manifold visualization for TL at depths 25 m and 30 m exhibits sloppiness, seen as thinness of the model manifold for TL > 72 dB. (d) Relative TL model manifold with the same axes as (a), the TL at depths 65 m and 5 m, but relative to TL at z =30 m, i.e., after subtracting TL at z =30 m. The relative TL model manifold is more compressed than the absolute TL model manifold.

for two seabed environments in data space is referred to as “seabed distance" in this paper. Seabed types with a small seabed distance, such as mud and clay in Fig. 3.1(b), exhibit a degree of similarity that can pose a challenge in geoacoustic inversion. In contrast, mud and clay have larger seabed distances from the other three seabeds—silt, sand, and gravel—making them more distinguishable from those seabed types.

Also notice that the model manifolds in Fig. 3.1(b)-(c) have a long dimension, corresponding to the more identifiable sound speed parameter, while the short dimension corresponds to the sloppier density parameter. This distinction is especially true in Fig. 3.1(c) for predictions made at nearby receiver depths of 25 m and 30 m; the very narrow portion of the manifold indicates that, for TL > 72 dB, these receiver depths contain no information about the density.

Due to measurement uncertainty and ambient noise in measured TL, data typically do not lie on the model manifold. Black dots in Fig. 3.1(b)-(c) represent hypothetical noisy TL data, and the black arrows represent possible projections of this “data" onto the model manifold. Geoacoustic inversion projects noisy data onto the model manifold by acting as a non-unique pseudo-inverse, corresponding to a specific choice of loss function in the optimization process, often the sum of squares error that minimizes the Euclidean distance between observed data and model predictions.

To reduce the noise in the data that is common to all measurements, relative TL is often used in ocean acoustic applications instead of absolute TL. One way to calculate relative TL is to subtract the TL for one channel from the other channels, reducing the effective data space dimension from M to $M-1$. Using relative TL should in practice bring noisy data closer to the model manifold. The remaining distance between the data point and the manifold could be caused by model mismatch or sources of uncertainty that vary between measurements, such as array tilt for VLAs, individual sensor noise, and sound speed variability. Thus, using relative TL for the axes of data space corresponds to a different noise model where the noise is limited to those sources that are different between measurements.

As an example, Fig. 3.1(d) shows a relative TL model manifold. The relative TL manifold is the absolute TL manifold in Fig. 3.1(b) with axes of relative TL at depths of 65 m and 5 m, with the TL at a depth of 30 m (y-axis of Fig. 3.1(c)) subtracted from each channel. Notice that the relative TL manifold appears to be more compressed than the original manifold, with smaller seabed distances than for the absolute TL manifolds. This compression of the relative TL model manifold is explored more quantitatively in Sec. 3.3. While absolute TL is often considered an exclusively positive quantity, relative TL can be positive or negative because of how it is defined here. In this application, the total dB distance in data space is most relevant, not whether the values are positive or negative.

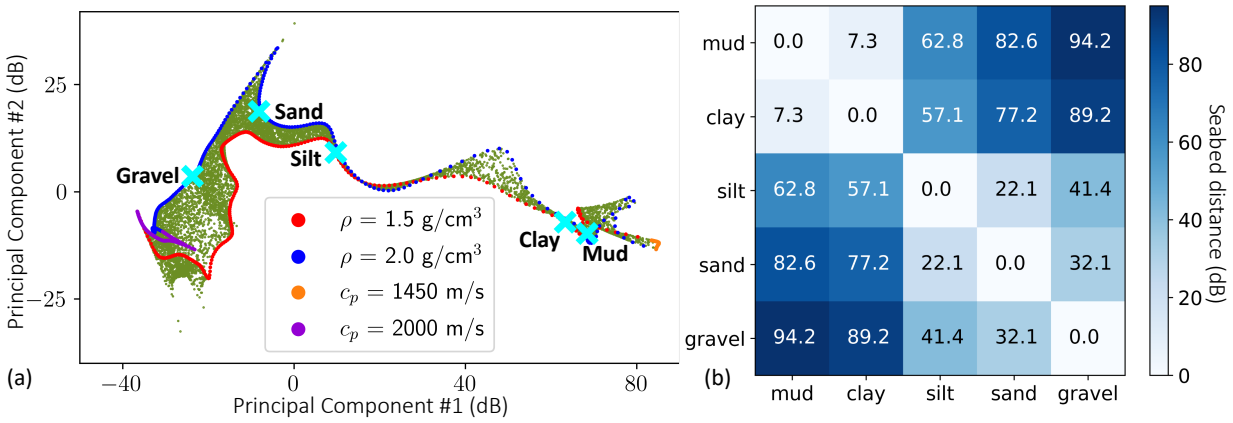

Figure 3.2 (a) PCA projection of the model manifold in 15D data space onto the two principal components with the highest variation. Units of axes are still dB. (b) Seabed distance matrix showing the Euclidean distance between seabeds on the TL manifold in the 15D data space.

To allow for more accurate discussion of seabed distances in Sec. 3.3, an alternate method for obtaining 2D visualizations of the model manifold is introduced. The 2D visualizations in Fig. 3.1(b)-(c) each correspond to observations from two receivers. Inclusions of data from additional channels (e.g., 15 receiver depths) are more informative, but difficult to visualize. Low-dimensional visualizations of high-dimensional spaces are obtained via Principal Component Analysis (PCA). PCA finds a new basis for data space aligned with the directions in which the model output varies most. Points sampled from the manifold in 15D are first translated to be centered at the origin by subtracting the mean TL of all the data points for each channel. The shifted points are collected into the columns of a matrix D, and a singular value decomposition is performed: $D = U\Sigma V^T$, where U and V are unitary matrices. The columns of V define the new basis for data space and are known as the principal components. Σ is a diagonal matrix of the singular values of D, which are the standard deviations of the sampled points projected onto each principal component. The first two principle components typically describe most of the variability in D, so the projection of the model manifold onto the first two principle components yields a low-dimensional visualization of the main features of the model manifold.

For the case of 15 receiver depths, a 2D PCA visualization of the model manifold is displayed in Fig. 3.2(a). The axes of Fig. 3.2(a) represent linear combinations of TL at all 15 depths projected onto the first two principal component directions, which captures 91.2% of the variation in the original 15D data. For Fig. 3.2(a) and the other PCA manifold visualizations in the paper, the exact values on the axes are not important, only the total distance in dB between points; the units are still dB, allowing for comparison to the widths of the manifolds shown in Fig. 3.1. The model manifold in Fig. 3.2(a) is wider and appears less folded over itself than the manifolds in Fig. 3.1(b)-(c), indicating an increase in the quantity of information relevant to distinguishing seabed parameters. The seabed distances between the five marked seabeds in the 15D data space are given in Fig. 3.2(b). These seabed distances differ slightly from the apparent distances in Fig. 3.2(a) based on the first 2D PCA model manifold because the seabed distances in Fig. 3.2(b) are calculated in the 15D data space.

3.3 Results

The impact of using relative TL instead of absolute TL on the informativity of data for a VLA is explored in this section. Relative TL model manifolds in Fig. 3.3(a)-(c) are created by subtracting the TL at reference depths of 20 m, 30 m, and 65 m, respectively, from all 15 channels. The relative TL manifold in Fig. 3.3(d) is calculated differently by subtracting from each channel the average model predictions across all 15 VLA elements, which can be thought of as subtracting the mean TL. All four relative TL manifolds are then projected into 2D using PCA. Seabed distances are calculated on the relative TL manifolds in 14D (or 15D) data space, and shown in Fig. 3.3(e)-(g).

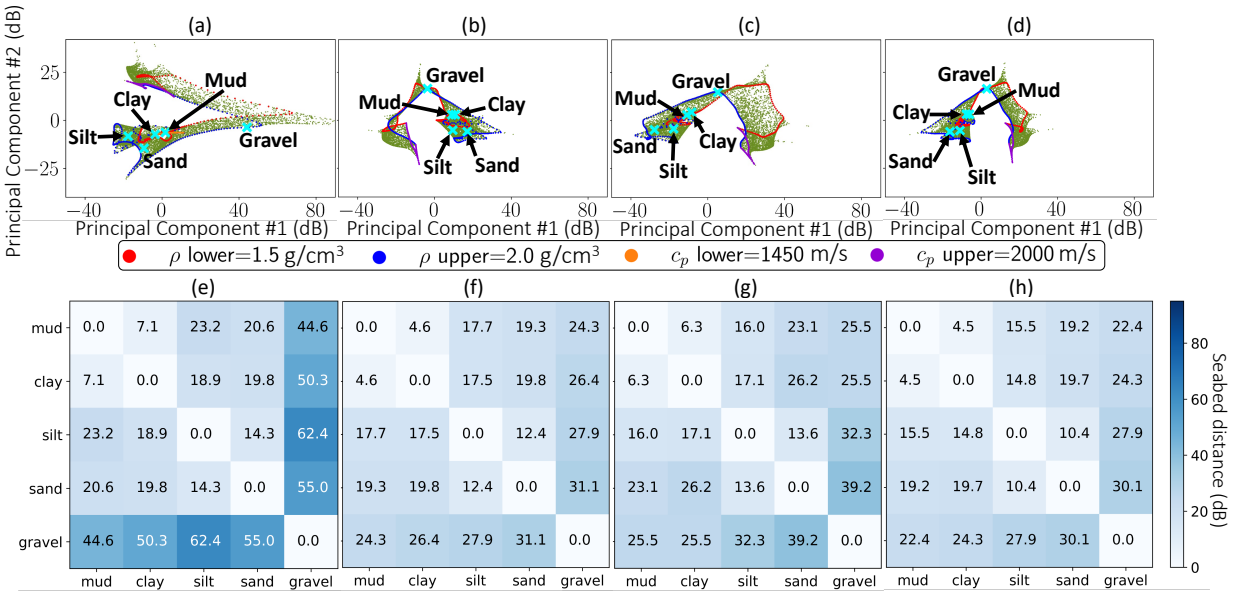

Figure 3.3 2D PCA projections of relative TL model manifolds. Relative TL is calculated by subtracting TL at reference depths of (a) 20 m, (b) 30 m, and (c) 65 m, while in (d) relative TL is calculated by subtracting the mean TL prediction across all 15 VLA elements. Corresponding seabed distance matrices are below each manifold in (e)-(h).

Additionally, Appendix C contains model manifolds and seabed distance matrices relative to all 15 receiver depths on the VLA, as well as a table with average and median seabed distances, a plot presenting the PCA singular values, and a table indicating the percentage of variance explained by the 2D PCA visualizations.

Several notable changes in model geometry and parameter identifiability occur when using relative TL on the VLA. First, as seen in the relative TL model manifold in Fig. 3.1(d), the relative TL model manifolds in Fig. 3.3(a)-(d) are more folded and compact than the absolute TL manifold in Fig. 3.2(a). For example, the model predictions in the first principal component in Fig. 3.3(b)-(d) vary by two-thirds the dB range as Fig. 3.2(a). This compression of the relative TL model manifold is due both to additional folding and the overall reduced scale of the model manifold. Comparison of seabed distances in the full data space can provide a quantitative measure of the compression of the model manifold.

Relative TL model manifolds seabed distances, shown in Fig. 3.3(e)-(h), are smaller than the absolute TL seabed distances shown in Fig. 3.2(b). First, the maximum relative TL seabed distances (darkest blue elements of Fig. 3.3(e)-(h)) are 30 - 70% smaller than the absolute TL maximum seabed distance in Fig. 3.2(b). Additionally, the average seabed distances for relative TL are less than for absolute TL. The average seabed distance in Fig. 3.2(b) (excluding self-correlation) is 56.6 dB. In contrast, the average seabed distances in Fig. 3.3(e)-(h) for relative TL obtained by subtracting TL at reference depths 20 m, 30 m, and 65 m, and by subtracting the mean TL, are 31.6 dB, 20.1 dB, 22.5 dB, and 18.9 dB, respectively. Smaller seabed distances indicate that, in this example, data associated with different seabed parameters are less distinguishable than when using absolute TL.

Most importantly, the respective *ordering* of the seabed distances has changed. For absolute TL, the seabed distances follow the overall trend of increased reflectivity of the seabed, where gravel is furthest from, and, therefore, most distinguishable from, mud and clay, with seabed distances of $\sim$90 dB (Fig. 3.2). However, for the relative TL manifolds in Fig. 3.3(a)-(d), mud and clay are much closer to gravel in data space, and closer than gravel is to sand and silt, due to folding of the model manifold. This trend becomes clear upon inspection of the seabed distance matrices in Fig. 3.3(e)-(h): mud and clay are anywhere between 5 dB and 15 dB closer to gravel than gravel is to sand and silt. Thus, the overall identifiability of the seabed parameters is reduced when using relative TL, due to the subtracting out of common features, which increases the folding of the model manifold. In context of Eq. 3.1, it is not clear exactly what structure in the data is lost when using relative TL; however, qualitatively it appears that some overall bottom loss due to the sediment has been removed. The practical implication of this seabed reordering is that geoacoustic inversions to determine seabed parameters may be more challenging for relative than absolute TL, depending on how much noise reduction is obtained by use of relative TL.

The reference hydrophone used to calculate relative TL changes the model manifolds, effectively increasing or decreasing the information contained in the relative TL data. For example, contrast using TL relative to 30 m (Fig. 3.3(b) and (f)) to using TL relative to 20 m (Fig. 3.3(a) and (e)). The average seabed distance in Fig. 3.3(f) is 20.1 dB, while the average seabed distance in Fig. 3.3(e) is 31.6 dB. Therefore, parameter inference using TL relative to 30 m may be more challenging, with larger uncertainty in parameter estimates, than if a reference depth of 20 m was used. Specifically, the seabed distance between mud and clay for absolute TL is 7.3 dB, as shown in Fig. 3.2(b). The seabed distances between mud and clay for reference depths of 30 m and 20 m, respectively, are 4.6 dB and 7.1 dB (Fig. 3.3(f) and (e)). Thus, using a reference depth of 30 m yields a model manifold in which mud and clay are less distinguishable than when using absolute TL, while using a reference depth of 20 m yields a distinguishability of mud and clay nearly identical to the absolute TL case. Seabed distances for averaged relative TL in Fig. 3.3(h) show similar compression as Fig. 3.3(f), with the distance between mud and clay being 4.5 dB, and an even smaller average seabed distance of 18.9 dB. Thus, if relative TL is used, distances between seabed locations on the model manifolds can be examined to indicate which method and which reference depth maximizes the information with respect to the parameter of interest.

While more studies need to be done looking at relative TL for changing different experimental design parameters, this work provides insights into the loss of information content that may occur when using relative TL for a VLA. In this example, because of decreased and mixed up seabed distances from using relative TL, care should be taken when using parameter values inferred from relative TL. The increased uncertainties from using relative TL may propagate as these values are subsequently used in, for example, source ranging. Thus, the advantages of relative TL in reducing uncertainty due to noise in the data have to be balanced with the potential decrease in information content about the seabed parameters.

3.4 Conclusions

This work provides an example of how model manifolds can be used for optimal experimental design. The model manifold, an information geometry tool for parameter identifiability analysis, has been constructed for TL from the ORCA normal-mode sound propagation model for a shallow water case with a 15 element VLA at a single frequency, source depth and range. The wide dimension of the “ribbon-like" model manifolds correspond to the identifiable sound speed parameter c_p, while the thin dimension of the model manifolds corresponds to density ρ, in some situations, indicating a sloppy parameter. The Principal Component Analysis (PCA) method provides a way to obtain a 2D model manifold visualization projected from a 15D data space, while retaining maximal variation of the model output.

Comparisons of the model manifolds for absolute and relative TL reveals significant insights into information loss and changes in the model (manifold) structure that come from using relative TL, in the case of receivers on a VLA relative to one receiver depth, and for the case of subtracting the mean TL across VLA elements. These relative TL model manifolds are more compact than absolute TL manifolds and have smaller seabed distances in data space. Smaller seabed distances indicate that relative TL contains less information to distinguish certain seabed environments, with implications for geoacoustic inversion for seabed parameters. The compression of the relative TL model manifold can be seen as a symptom of having some overall influence of the seabed environment removed from the acoustical data. Substantially, when using relative TL, the ordering of seabed distances also changes such that seabed distance does not follow the overall trend of seabed reflectivity observed when using absolute TL. With relative TL, mud and clay appear closer to gravel than gravel is to sand and silt, implying that mud and clay are statistically more similar to gravel. With an understanding of the limitations of relative TL, the model manifold approach

allows for selection of the reference depth that maximizes the information content in the relative TL data. Future work is required to explore the implications of relative TL for different experimental design choices, for example considering TL relative to different source depths, frequencies, and source ranges (such as on a horizontal line array).

Other future work includes considering higher dimensional models (and therefore model manifolds), such as including more sediment layers and varying sediment attenuation in addition to sound speed and density. These higher dimensional models will not be amenable to direct visualization, but similar analyses can be performed by visualizing 2D slices of the model manifold. Additionally, future work will utilize differential geometry tools such as geodesics to obtain reduced order models in these more complex cases.

Acknowledgements

The authors acknowledge the US Office of Naval Research for their support of this work under grant number N00014-21-S-B001. We also thank the reviewers and the Associate Editor of JASA-EL for their helpful suggestions.

Chapter 4

Conclusion and future work

4.1 Summary and Discussion

In conclusion, my primary contributions to the field of underwater acoustics have been to write a review paper which introduces information geometry tools in context of parameter sensitivity analysis methods previously utilized in underwater acoustics and to demonstrate how the model manifold, an information geometry tool, can be used to answer questions of experimental design and model selection. These contributions have been demonstrated by two submitted papers included in this thesis as Chapters 2 and 3.

Chapter 2 demonstrates model manifold construction for the Pekeris waveguide, interpreting insights into parameter identifiability both locally, using the Fisher information matrix, and globally, analyzing boundaries of the model manifold. The Fisher information matrix and model manifold are introduced in detail in Sections 2.3 and 2.6, respectively, including a discussion of their connections; these sections provide a resource for readers to reference. The review of papers applying parameter sensitivity analysis methods to ocean acoustics will also be a valuable resource for readers. Chapter

3 demonstrates model manifold construction for the ORCA sound propagation model, exploring the differences in information content between using absolute transmission loss and different choices for relative transmission loss. This paper is a template for how to utilize information geometry and model manifolds in optimizing experimental design in other scenarios.

The following section (Sec. 4.2) discusses some preliminary work conducted constructing neural network (NN) surrogates of the model manifold, enabling differential geometry methods for model selection. Then, Sec. 4.3 gives the outlook for future work in ocean acoustics using information geometry, building upon the work conducted in this thesis.

4.2 Preliminary work on NN surrogates

It was demonstrated in Chapter 2 how information geometry can investigate parameter identifiability both locally and globally. Locally, the Fisher information matrix is a distance metric on the model manifold and provides a measure of parameter identifiability. Globally, the boundaries of the model manifold are simpler models corresponding to the limit of some parameter combination. Identifying the manifold boundaries and the associated reduced-order model closest to the data provide an algorithm for model selection in which the sloppiest parameters are removed from the model; this algorithm is known as the manifold boundary approximation method (MBAM) [183]. To identify manifold boundaries in practice, when there are many parameters and data space dimensions, requires calculation of geodesics on the model manifold. A geodesic is the generalization of a straight line on a curved surface, and gives the shortest distance between two points on a surface. To calculate geodesics, however, requires reliable calculation of accurate derivatives of the model with respect to model parameters. Finite differencing methods, while simple, can yield inaccurate derivatives, in particular for models with various parameter scales. Modern automatic differentiation (AD) methods also provide a method for obtaining model derivatives; however, implementation of

AD directly on a sound propagation model is challenging due to inaccessible or "legacy" model code and numerical computations (e.g., numerical integrals). We propose a method to train a neural network as a surrogate to approximate the model manifold from which derivatives can be rapidly and accurately calculated using AD.

The NN architecture used (implemented in Pytorch [188]) is a simple network with two hidden layers of size 128 and 64, with the inputs being the model parameter values and the labels being TL at the data channels (e.g. receiver depths) specified. Both inputs and outputs are standardized by subtracting the mean and dividing by the standard deviation, using the training data provided. For the example given here, 20000 simulated points were utilized, with a 95/5 train-test split, and the model was trained for 20000 epochs while the learning rate was varied using cosine annealing. The optimizer and activation function used were Adam and Mish, respectively.

For this application, where we are interested in the derivatives of the NN surrogate model manifold, it is important to use a smooth activation function, such as Mish, to yield smoothly varying derivatives. This is illustrated using a toy analytic model, a sum of two decaying exponentials, for which the true derivative values are known, with two parameters, $\boldsymbol{\theta} = [\theta_1, \theta_2]$:

$$y_\theta(\boldsymbol{t}) = e^{-\theta_1 \boldsymbol{t}} + e^{-\theta_2 \boldsymbol{t}}. \tag{4.1}$$

Similar to the sound propagation model considered, this model has two inputs corresponding to the model parameters $\boldsymbol{\theta}$ and three outputs corresponding to the model predictions at $\boldsymbol{t} = [1,2,3]$. A NN surrogate was trained for this model using three different activation functions, ReLU, Tanh, and Mish, which are shown in Fig. 4.1(a)-(c). Note that the ReLU activation function has a discontinuity in its derivative. After training these NN surrogates, Jacobian matrices were calculated from the NN surrogates for a variety of parameter values using automatic differentiation. These predicted Jacobian matrices are then compared to the true Jacobian matrices. For the six elements in the Jacobian for this example, the mean absolute error of the Jacobian elements was calculated to give a

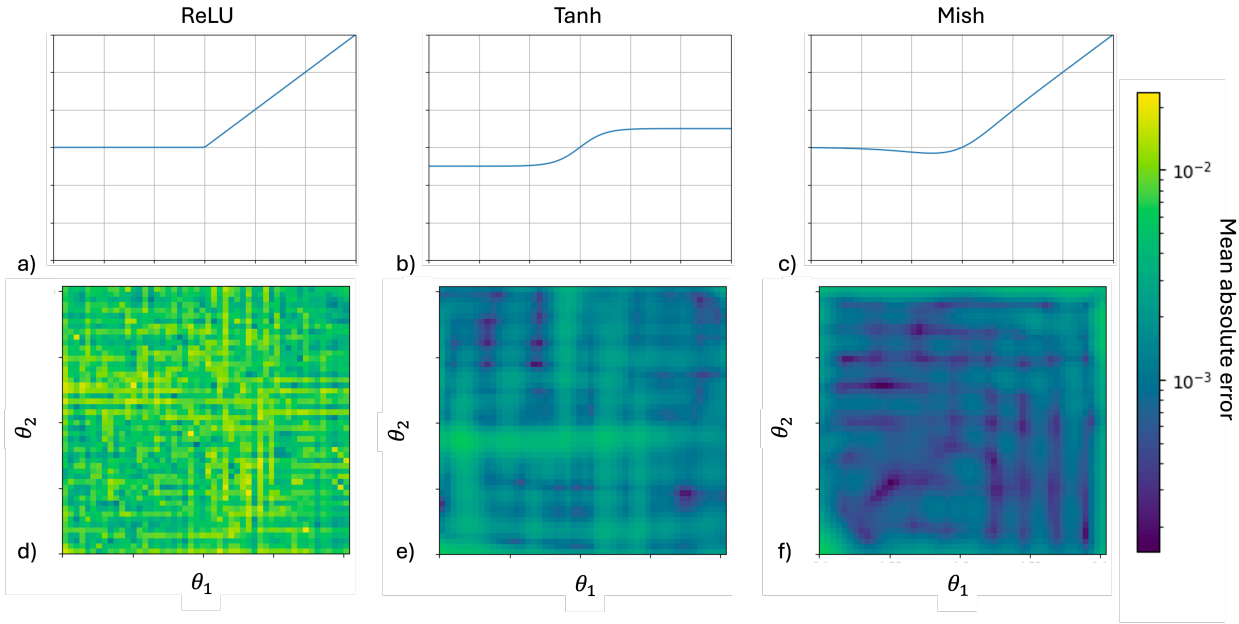

Figure 4.1 NN surrogates were trained using three different activation functions, ReLU, Tanh, and Mish, which are shown in panels (a) - (c). Panels (d) - (f), show the mean absolute error calculated between the NN surrogate Jacobian matrices and the true Jacobian for the same three activation functions, for a variety of parameter values. The error is smaller and the derivatives are more smoothly varying for smooth activation functions such as Tanh and Mish than for ReLU.

single number representing how close the true and predicted Jacobians are at that parameter choice. Plots showing how the mean absolute error between the true and predicted Jacobians changes over parameter space are given in Fig. 4.1(d)-(f). These plots show that for the smooth acivation functions, Tanh and Mish, the error is smaller and the derivatives are more smoothly varying than for the ReLU activation function. Thus, for the neural networks trained here, the Mish activation function is used.

For the case of the Pekeris sound propagation model, a NN surrogate manifold was trained using the Mish activation function for a region of parameter space corresponding to higher values of density and sound speed (c=[4500, 35000] m/s; ρ=[2.5,10000.0] g/cm^3). The NN surrogate (green) is overlaid on the true manifold (blue) in Fig. 4.2 and appears to match very well. Another way to

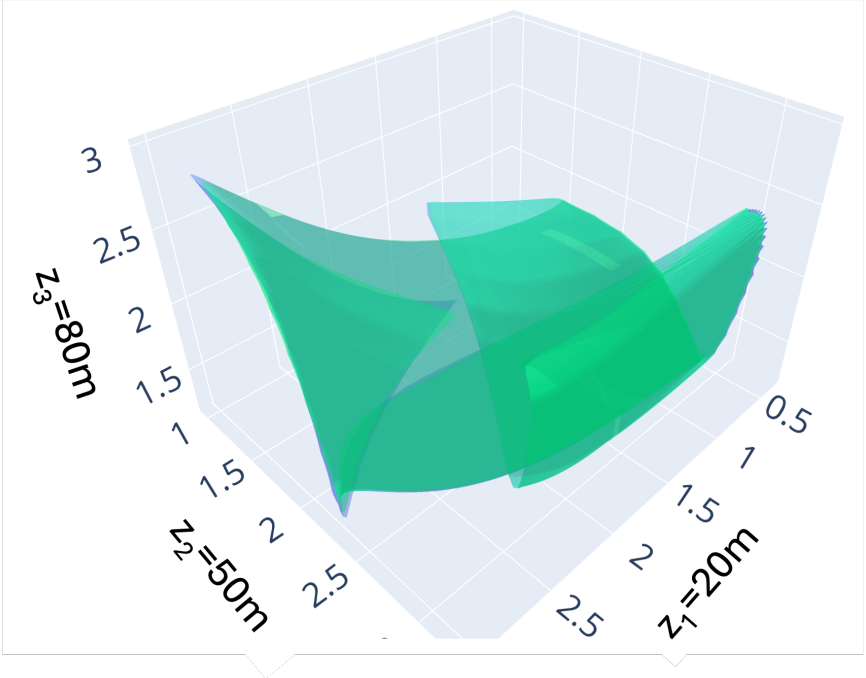

Figure 4.2 Trained NN surrogate (green) overlaid on the true model manifold (blue).

illustrate the accuracy of the NN surrogate compared to the true manifold is by using prediction plots, which plot true model values on the y-axis and predicted model values on the x-axis, as shown in Fig. 4.3, for training and testing data. Perfect predictions will thus lie on on the diagonal black dashed lines in Fig. 4.3. For this NN surrogate, the surrogate model predictions are quite accurate.

With this trained NN surrogate manifold, the model derivatives are obtained from the NN surrogate via automatic differentiation and can be used to calculate geodesics. An example of a geodesic calculated on this NN surrogate model manifold is shown in Fig. 4.4. The geodesic path in parameter space, shown in Fig. 4.4(a), starts at the black point, follows the green path, and ends at the red point; the geodesic path on the model manifold in data space is shown in Fig. 4.4(b), with the same coloring.

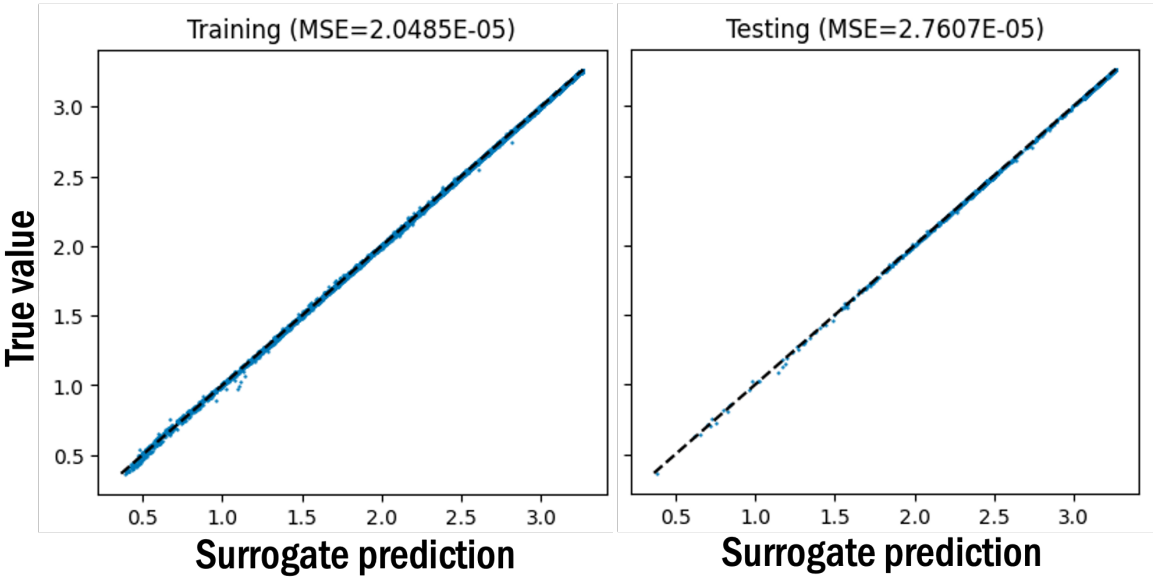

Figure 4.3 Prediction plots for training and test datasets. True value is plotted on the y-axis, predicted value on the x-axis so that a prediction on the black dotted diagonal line is a perfect prediction.

An alternative way to visualize the geodesic path combines the parameter space and data space geodesic paths to help identify when a geodesic encounters model manifold boundaries. The parameter space geodesic path distance (Fig. 4.4(a)) can be parameterized by some variable, say t, and the data space geodesic path distance (Fig. 4.4(b)) can be parameterized by another variable, τ. The plot of τ vs. t in Fig. 4.4(c) is observed to "plateau", indicating that in this parameter limit, changing the parameter value t has ceased changing the model output because that parameter combination has become sloppy. This is the definition of a boundary of the model manifold; thus, plotting τ vs. t can be used to identify model manifold boundaries using geodesics even for complex modeling scenarios in which the model manifold cannot be directly visualized, such as for models with many parameters and and data channels. Two other geodesics on the model manifold, with different initial conditions, are shown in Fig. 4.5 and Fig. 4.6, in addition to the accompanying τ vs. t plots.

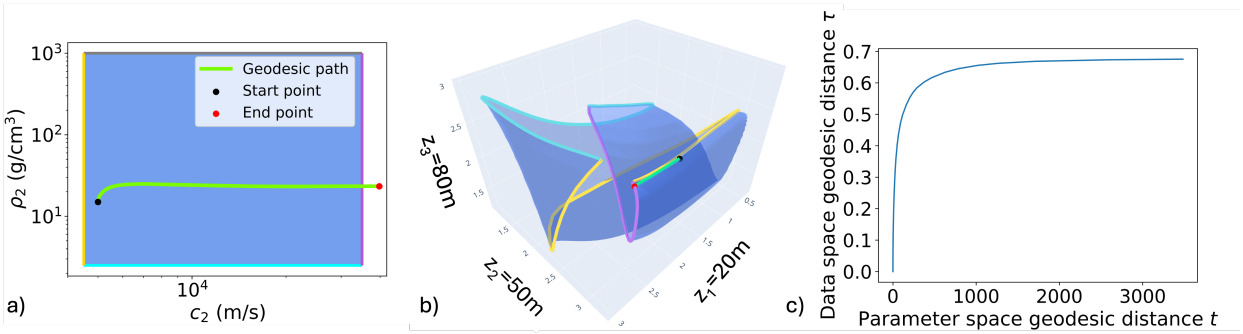

Figure 4.4 (a) Geodesic path in parameter space which starts at the black point, follows the green line, and ends at the red dot, intersecting the high sound speed boundary. (b) Same geodesic, but shown on the model manifold in data space. (c) Plot of the data space geodesic path distance, parameterized by τ, versus the parameter space geodesic path distance, parameterized by t. The plateau-like behavior signals that a manifold boundary has been reached, and some parameter combination has become unidentifiable.

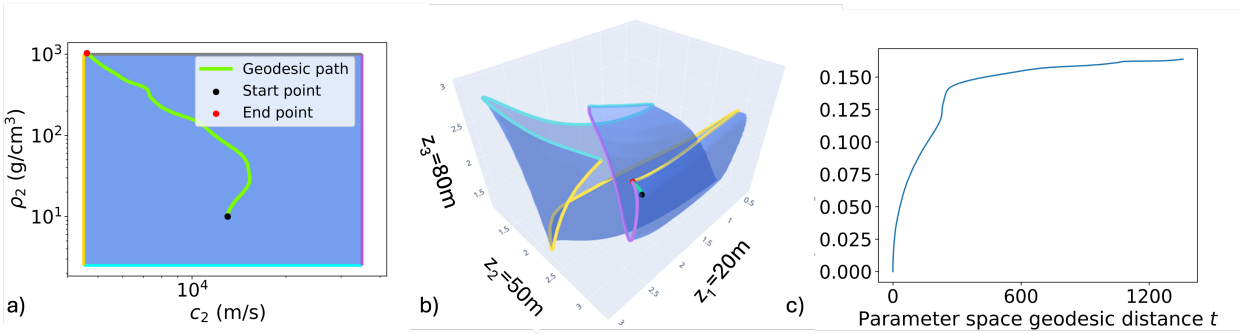

Figure 4.5 (a) Geodesic path in parameter space which starts at the black point, follows the green line, and ends at the red dot, intersecting the high density boundary. (b) Same geodesic, but shown on the model manifold in data space. (c) Plot of the data space geodesic path distance, parameterized by τ, versus the parameter space geodesic path distance, parameterized by t.

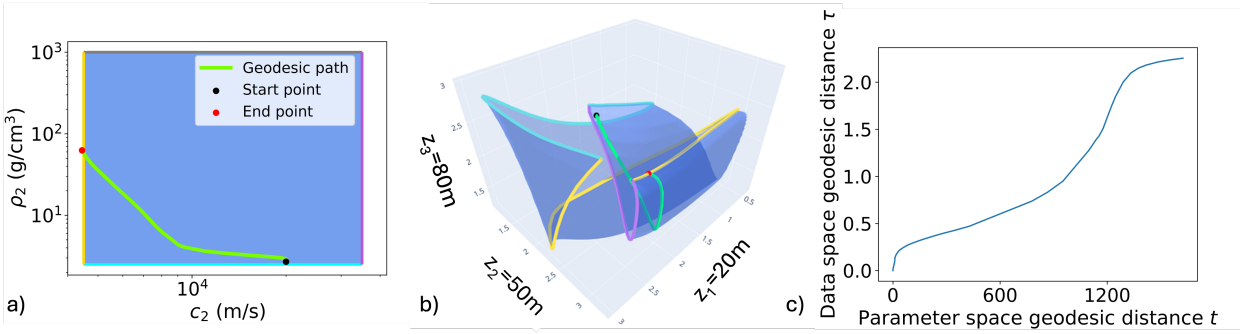

Figure 4.6 (a) Geodesic path in parameter space which starts at the black point, follows the green line, and ends at the red dot, intersecting the low sound speed boundary. (b) Same geodesic, but shown on the model manifold in data space. (c) Plot of the data space geodesic path distance, parameterized by τ, versus the parameter space geodesic path distance, parameterized by t.

4.3 Outlook

The work conducted in this thesis represents some of the first steps in applying information geometry methods to problems in ocean acoustics, but more progress needs to be made in these efforts. This section gives ideas for future work that build upon the research in this thesis, specifically directed towards the areas of experimental design and model selection, the latter which includes continued development of the method for training NN surrogate model manifolds. The reader may also reference Sec. 2.8 for a summary of other ways that information geometry has successfully led to advances in other fields, which may provide ideas that can be applied in underwater acoustics.

In the area of experimental design, more studies can be conducted utilizing model manifolds and information geometry. Chapter 3 inspects model manifolds for both relative and absolute transmission loss (TL) for hydrophones on a vertical array, giving insights into seabed distinguishability when using relative TL, and how an optimal reference depth can be chosen. To draw more general conclusions about how using relative TL changes information content, additional studies would have to be performed. Ideas include inspecting model manifolds and seabed distance matrices for relative TL for a horizontal array of hydrophones (towed array) and different source ranges, and

for different source frequencies. More generally (i.e., not looking at relative TL), model manifolds for different ranges and frequencies can be compared, providing insight into optimal experimental design choices for a variety of different scenarios. The same studies could be conducted for coherent or incoherent mode summations in the sound propagation model, where an incoherent summation provides a more realistic model of a broadband signal, whereas a coherent mode summation is for a single frequency [166]. Additional work could also include considering different sediment type differentiations in constructing seabed distance matrices.

Continued development of methods for training NN surrogates and calculating geodesics on the surrogate manifold provide many other opportunities for future work, related to model selection. Near term future work includes continuing to develop NN surrogate code to make it adaptable to a wide variety of modeling scenarios. Additionally, the geodesic code can be integrated with the NN code so that geodesics can be calculated on the model manifold dynamically; if the geodesic leaves the trained region of parameter space, the geodesic is paused and the NN can be further trained on additional data in that direction of parameter space.

Some of the most exciting opportunities for future work revolve around using geodesics to better discover and understand physically significant parameter combinations and limits in underwater sound propagation. In modeling, it can often be shown that there are parameter combinations which are more relevant to characterizing the model behavior than the original model parameters– for example, the ratio of sediment sound speed over water sound speed c_2/c_1 instead of the exact sound speed values. The viscous grain shearing (VGS) parameterization is a good example of an alternate parameterization of an ocean sediment in which its parameters are physically interpretable combinations of more traditional seabed sediment parameters– density, sound speed, attenuation,

etc. Geodesics can be calculated for model manifolds of a variety of modeling scenarios, and different parameter limits explored to look for these different ways to think about model parameters. This relates to other work being pursued to create model manifolds in terms of the VGS parameters, which can also be used with this NN surrogate manifold method.

Another idea for future work is to calculate geodesics on the Pekeris waveguide model manifold and characterize the global boundary structure. Chapter 2 contains examples of several Pekeris waveguide model manifolds. Understanding the boundary structure of the Pekeris waveguide model manifold will allow for a better understanding of the limit in which an entire sediment layer is removed from the model ($c_2 = c_1$ and $\rho_2 = \rho_1$); this may lead to more general insights into how sediment layers can be removed from more complex environmental configurations, for obtaining reduced-order models for geoacoustic inversion.

Additionally, development of these methods for obtaining geodesics on the model manifold will enable application of the MBAM algorithm to more complex sound propagation model scenarios, with several sediment layers and many data channels, identifying reduced-order models that represent the given data with reasonable accuracy. This reduced model can then be used in geoacoustic inversion and the performance compared to the original, more complex model. Future developments of the MBAM algorithm may also lead to more opportunities for applying information geometry to sound propagation modeling and geoacoustic inversion.

Appendix A

Pekeris waveguide with sediment attenuation implementation details

In this section, additional details regarding the implementation of the Pekeris waveguide from [5] are given. In a future work we plan to interpret the full boundary structure of the Pekeris waveguide, requiring that we take parameter values to their extremes (e.g. $\rho_2 = \infty$). Some checks in the code handle numerical intricacies at these limits but are unnecessary for most parameter choices.

First, Eq. 2.35 in Sec. 2.7.1 gives the number of total proper modes required in the solution. M is constrained to be greater than zero because sometimes the equation yields a number less than zero. Additionally, as the sediment attenuation parameter is increased ($\alpha_2 \to \infty$), M increases exponentially, but the values for higher mode numbers become effectively zero. In the implementation, a parameter of MAX_MODES=5000 is used, along with a check to verify that the modal contributions have gone to zero before the cutoff.

In [5], a source strength parameter Q with units of [length]^2 is used, but it does not specify how this value is chosen. This parameter has a net effect of shifting the TL curve. For the simulations in this paper, a value of Q =9E-5 is used because it is found to give TL that aligned well with similar simulations performed with the ORCA model [189] with a reference pressure $P_{\text{ref}} = 1$.

Numerical calculation of the EJP integral was done using SCIPY.INTEGRATE.QUAD. The absolute error parameter ABSERR was chosen to be its default value of 1.49E-8. This same value, which is referred to as ε, is used in several places throughout the algorithm in determining a threshold for meaningful contributions to the final answer.

Several specific adjustments were required to make this integral evaluation feasible and accurate.

First, there is an e^{-ipr} term in the integrand of Eq. 2.40. In this implementation, p must be chosen to be the positive or negative root $p = \pm\sqrt{k_2^2 - \eta_2^2}$ depending on what will give a decaying exponential in the infinite integral limits. Specifically, for the case with zero sediment attenuation, k_2 is real, and η_2 (as the integration variable) is assumed to be real. Therefore, when $|\eta_2| < |k_2|$, $\sqrt{k_2^2 - \eta_2^2}$ is real so the default positive root is used, making the exponential a propagating wave between $[-k_2, k_2]$ due to the i in the exponent. When $|\eta_2| > |k_2|$, however, $\sqrt{k_2^2 - \eta_2^2}$ yields a positive imaginary number, so the negative root of p must be chosen to produce a decaying exponential outside of the region $[-k_2, k_2]$. This check is also necessary in the code for when attenuation is introduced and k_2 is complex, because it can change the positive or negative root chosen from the zero attenuation situation. The decaying exponential when $|\eta_2| > |k_2|$ makes the integrand drop off quickly with source range r.

Next, notice in Eq. 2.40 that the bounds of the EJP integral go from $-\infty$ to ∞. While SCIPY.INTEGRATE.QUAD can handle infinite bounds, as discussed above the "effective range" of this integral is much smaller due to the decaying exponential in the integrand. This numerical integral is evaluated much more accurately if the integration region is restricted. For this implementation, the integral bounds over η_2 were chosen to be $[-1.11k_1, 1.11k_1]$, as discussed below.

For the case of zero sediment attenuation (which means k_2 is real), the region $[-k_2, k_2]$ is where the integral is a propagating wave and not a decaying exponential. This is where the vast majority of integral contributions are, due to source range r usually being chosen to be relatively large (e.g., 3 km in this work). The largest possible values of k_2 is when $c_2 = c_1$ and $k_1 = k_2$, so the range of $[-k_1, k_1]$ can be used to include all possible propagating wave contributions.

In a few cases, the decaying exponential does not drop off as quickly and meaningful contributions exist outside the $[-k_2, k_2]$ range. To slightly expand the range, a multiplier of 1.11 is added. For $c_2 = c_1$, the range $[-1.11k_1, 1.11k_1]$ was tested and found to sufficiently capture any features larger than ε in the decaying exponential portion for a source range of 100 m with a 100 m ocean depth.

Inclusion of sediment attenuation leads to an extra imaginary term, and a decaying exponential term is introduced to the integral overall. This decay term causes the integral to get smaller with increasing attenuation. Because calculating the integral increases the overall run time, the integral is considered zero and not calculated when the integrand gets small enough (smaller than ε).

Additionally, some edge cases have singularities in the integrand, which can be handled by specifying them in the POINTS argument of SCIPY.INTEGRATE.QUAD. First, recall that the integrand changes from a propagating wave to a decaying exponential when $|\eta_2| > |k_2|$. At this point, $\eta_2 = k_2$ so $p = 0$ leading to a singularity due to the $\frac{\eta_2}{p}$ term in the integrand. Singularities can also occur in the denominator of F_1. Due to the sine and cosine terms in the denominator, in most circumstances

the denominator is not identically zero. However, when a large density ρ_2 is chosen, b_{12} becomes very small and so the zeros of the cosine function are the zeros of the denominator, causing singularities. And lastly, in the edge case when c_2 is chosen to be equal to c_1, $\eta_1 = \eta_2$. This means that when $\eta_2 = 0$, $\eta_1 = 0$. F_1 has an η_1 in the denominator, so when $c_2 = c_1$ there is a singularity at $\eta_2 = 0$.

Appendix B

Infinite density hard bottom model derivation

This derivation shows how for the Pekeris waveguide in the limit of infinite sediment density, both ρ_2 and c_2 drop out of the model. Recall that the solution of the Pekeris waveguide consists of a summation of proper modes plus the EJP branch line integral. The proper modes summation is shown to go to a constant value (in c_2 and ρ_2), and the EJP integral goes to zero. This derivation is specifically for the case of zero sediment attenuation.

The summation of proper modes is given by Eq. 2.37. This summation is given again here, but with the value for $\text{Res}[F_1(\eta_{1m},\eta_{2m})e^{ip_m r}]$ in Eq. 2.39 inserted so that

$$\phi_{\text{modes}}(r,z,\boldsymbol{\theta}) = \frac{Q}{2\pi}\left[2\pi i \sum_{m=1}^{M} \frac{\eta_{1m}\sin(\eta_{1m}z)\sin(\eta_{1m}z')}{p_m[\eta_{1m}h - \sin(\eta_{1m}h)\cos(\eta_{1m}h) - b_{12}^2\sin(\eta_{1m}h)\tan(\eta_{1m}h)]} e^{ip_m r}\right] \tag{B.1}$$

The variables that depend on c_2 and ρ_2 are b_{12}, which depends on density as $b_{12} = \rho_1/\rho_2$, and η_{1m} which is solved for using a Newton-Raphson procedure that depends on both c_2 and ρ_2. And then because p_m is calculated from η_{1m} as $p_m = -\sqrt{k_1^2 - \eta_1^2}$, p_m also depends on c_2 and ρ_2. First, as $\rho_2 \to \infty$, the ratio $b_{12} \to 0$. Therefore the term with b_{12}^2 in the denominator of Eq. B.1 goes to zero.

Next, the Newton-Raphson procedure to find η_{1m} that is outlined in [5] Eqs. 37-42 is given here:

$$f(X) = X - \left(m - \frac{1}{2}\right)\pi - \tan^{-1}[g(X)] = 0, \quad m = 1, 2, \ldots, M \tag{B.2}$$

where

$$X = \eta_1 h \tag{B.3}$$

$$g(X) = b_{12}\frac{\sqrt{A^2 - X^2}}{X} \tag{B.4}$$

and

$$A = \sqrt{k_1^2 - k_2^2}. \tag{B.5}$$

X is solved for by iterating the equation

$$X_{n+1} = X_n - \frac{f(X_n)}{f'(X_n)} \tag{B.6}$$

where

$$f'(X) = \frac{df}{dX} = 1 + \frac{1}{X(1+g^2)}\left\{g + \frac{b_{12}^2}{g}\right\}. \tag{B.7}$$

As $\rho_2 \to \infty$, b_{12} goes to zero, and therefore $g(X)$ in Eq. B.4 also goes to zero. All c_2 dependence is in A, and therefore disappears. When $g(X) \to 0$, because the $\tan^{-1}(0)$ equals zero, Eq. B.2 becomes

$$X - \left(m - \frac{1}{2}\right)\pi = 0$$

$$\eta_{1m} = \left(m - \frac{1}{2}\right)\frac{\pi}{h}. \tag{B.8}$$

Now η_{1m} doesn't have any dependence on c_2 or ρ_2, so Eq. B.1 also doesn't depend on c_2 or ρ_2. Additional simplifications of the expression are possible because $\sin(\eta_{1m}h) = (-1)^{n+1}$ and $\cos(\eta_{1m}h) = 0$, so that

$$\phi_{\text{modes}}(r,z,\boldsymbol{\theta}) = \frac{Q}{2\pi}\left[2\pi i \sum_{m=1}^{M} \frac{\sin(\eta_{1m}z)\sin(\eta_{1m}z')}{p_m h} e^{ip_m r}\right] \tag{B.9}$$

which has no c_2 or ρ_2 dependence.

Additionally, the EJP integral can be shown to go to zero as $\rho_2 \to \infty$. The integral is given in Eq. 2.40, which has $F_1(\eta_1,\eta_2)$ (given in Eq. 2.38) as part of the integrand. The combined equation is

$$\phi_{\text{EJP}}(r,z,\boldsymbol{\theta}) = \frac{Q}{2\pi}\int_{-\infty}^{\infty} \frac{\eta_2}{p}\frac{\sin(\eta_1 z_<)}{\eta_1}\frac{\eta_1\cos[\eta_1(h-z_>)] + ib_{12}\eta_2\sin[\eta_1(h-z_>)]}{[\eta_1\cos(\eta_1 h) + ib_{12}\eta_2\sin(\eta_1 h)]} e^{-ipr} d\eta_2 \tag{B.10}$$

Because $b_{12} \to 0$ as $\rho_2 \to \infty$, in this limit the integral simplifies to

$$\phi_{\text{EJP}}(r,z,\boldsymbol{\theta}) = \frac{Q}{2\pi}\int_{-\infty}^{\infty} \frac{\eta_2}{\eta_1 p}\frac{\sin(\eta_1 z_<)\cos[\eta_1(h-z_>)]}{\cos(\eta_1 h)} e^{-ipr} d\eta_2 \tag{B.11}$$

The crux of the argument that this integral goes to zero is that the integrand can be shown to be an odd function of η_2. First, in calculating the integral, given a value of the integration variable η_2, p can be calculated as $p = \pm\sqrt{k_2^2 - \eta_2^2}$. Because there is an η_2^2, p is therefore an even function of η_2. This means the e^{-ipr} term in B.11 is also an even function. Additionally η_1, which is calculated as $\eta_1 = \sqrt{k_1^2 - p^2}$, is also an even function of η_2 because p is an even function.

Because η_1 is an even function of η_2, all the trigonometric functions in Eq. B.11 which depend on η_1 are also even functions of η_2. Therefore, in Eq. B.11 the only term that is an odd function of η_2 is η_2 itself. This means that the whole integrand is an odd function of η_2, and integrates to zero.

In conclusion, for $\alpha_2 = 0$, as $\rho_2 \to \infty$ both ρ_2 and c_2 drop out of the Pekeris waveguide solution (the proper modes and the EJP integral), giving a reduced parameter model. This is observed as the red corner on the model manifold shown in Fig. 2.4 in the infinite density limit.

Appendix C

Supplementary Material for Information geometry analysis example for absolute and relative transmission loss in a shallow ocean

The supplementary material presented here provides the exact modeling parameters used, the parameter values for the five seabed types, additional PCA information, the figures of relative TL model manifolds and seabed distance matrices with different reference depths, and the average and median seabed distances for all model manifolds. In the main text, absolute TL is compared to relative TL for a few specific choices of reference depth, of the 15 possible reference depths. Here, information is given for all 15 receiver depths spaced evenly throughout the water column.

C.1 ORCA parameters

The sound propagation model used for this work is the range-independent normal mode model ORCA [187]. The ORCA configuration we use in this paper assumes an upper halfspace corresponding to air, the water layer, the main sediment layer, a halfspace layer, and a lower halfspace providing a numerical basement. The sound speed and density of the sediment layer are varied in making the model manifold. Table C.1 lists the parameter values held constant in the ORCA ocean configuration. All parameters are assumed to be uniform across the layer. All shear parameters (shear sound speed, shear attenuation) are zero and, therefore, not listed in the table. The attenuation in the table is compressional attenuation.

Table C.1 Constant parameter values used in the ORCA ocean configuration. There is an upper halfspace (air), a water layer, a sediment layer, a halfspace layer, and then a lower halfspace, which is needed numerically in calculating ORCA.

Layer	Parameter	Value (units)
Upper halfspace	Sound speed	343 m/s
	Density	0.00121 g/cm^3
	Attenuation	0 dB/m/kHz
Water layer	Layer height	75 m
	Sound speed	1500 m/s
	Density	1 g/cm^3
	Attenuation	0 dB/m/kHz
Sediment layer	Layer height	35 m
	Attenuation	0.63 dB/m/kHz
Halfspace layer	Layer height	100 m
	Sound speed	5250 m/s
	Density	2.7 g/cm^3
	Attenuation	0.02 dB/m/kHz
Lower halfspace	Sound speed	6000 m/s
	Density	5.0 g/cm^3
	Attenuation	0.83 dB/m/kHz

C.2 Seabed parameters

Five sediment types–mud, clay, silt, sand, and gravel–are marked on the model manifolds presented in the main text. The exact parameter values used to represent these sediment types are given in Table C.2. The parameter values for clay, silt, sand, and gravel are from Table 1.3 in *Computational Ocean Acoustics* [1]. The value for the sound speed of mud is based on an estimate of the sediment-water sound speed ratio of 0.99, which is an average value from Fig. 7 in Wilson *et al.*, [3] showing 15 inversion results. The density of mud comes from Potty *et al.* [2]

Table C.2 Sound speed and density parameters for five seabed types: mud, clay, silt, sand, and gravel. All parameter values come from *Computational Ocean Acoustics* [1] except the density ρ of mud [2] and the sound speed c_p of mud [3].

Seabed Type	Sound Speed (c_p, m/s)	Density (ρ, g/cm^3)
Mud	1485	1.6
Clay	1500	1.5
Silt	1575	1.7
Sand	1650	1.9
Gravel	1800	2.0

C.3 PCA visualizations of the model manifold

Principal component analysis (PCA) is a method that finds a new basis for data space aligned with the directions in which the model output varies most. PCA is accomplished using a singular value decomposition, in which the singular vectors are referred to as the principle components. The first principle components are associate with the largest singular values and often explain a large degree of the variance across the model manifold. The percentage of variance explained by the first two

principle components (PC) is shown in Table C.3 for all 15 relative TL model manifold visualizations (i.e., the 15 manifolds of TL relative to the individual receiver depths). All explained variance percentages (except $z = 30$ m) are $> 80\%$. This makes sense from inspection of singular values obtained from the PCA, shown in Fig. C.1. The singular values span many orders of magnitude, indicating that the data space is sloppy. Sloppiness is manifest in the model manifold experiencing much larger variation in some directions (related to combinations of modeling parameters) than others. Sloppy data can be represented faithfully by a lower dimensional representation, as seen in Table C.3.

Table C.3 Explained variance (as a %) of using only the first two principle components (PC) from Principle Component Analysis of the relative TL manifolds for 15 reference depths. First two PC's are used for visualization of manifolds.

	Explained variance % (of PC 1 & 2)
Absolute	91.2
Rel. to mean TL	80.3
Rel. to 5 m	86.8
Rel. to 10 m	82.4
Rel. to 15 m	92.7
Rel. to 20 m	88.7
Rel. to 25 m	85.4
Rel. to 30 m	71.5
Rel. to 35 m	82.8

	Explained variance % (of PC 1 & 2)
Rel. to 40 m	83.2
Rel. to 45 m	84.3
Rel. to 50 m	86.5
Rel. to 55 m	84.7
Rel. to 60 m	84.1
Rel. to 65 m	87.8
Rel. to 70 m	92.9
Rel. to 75 m	93.3

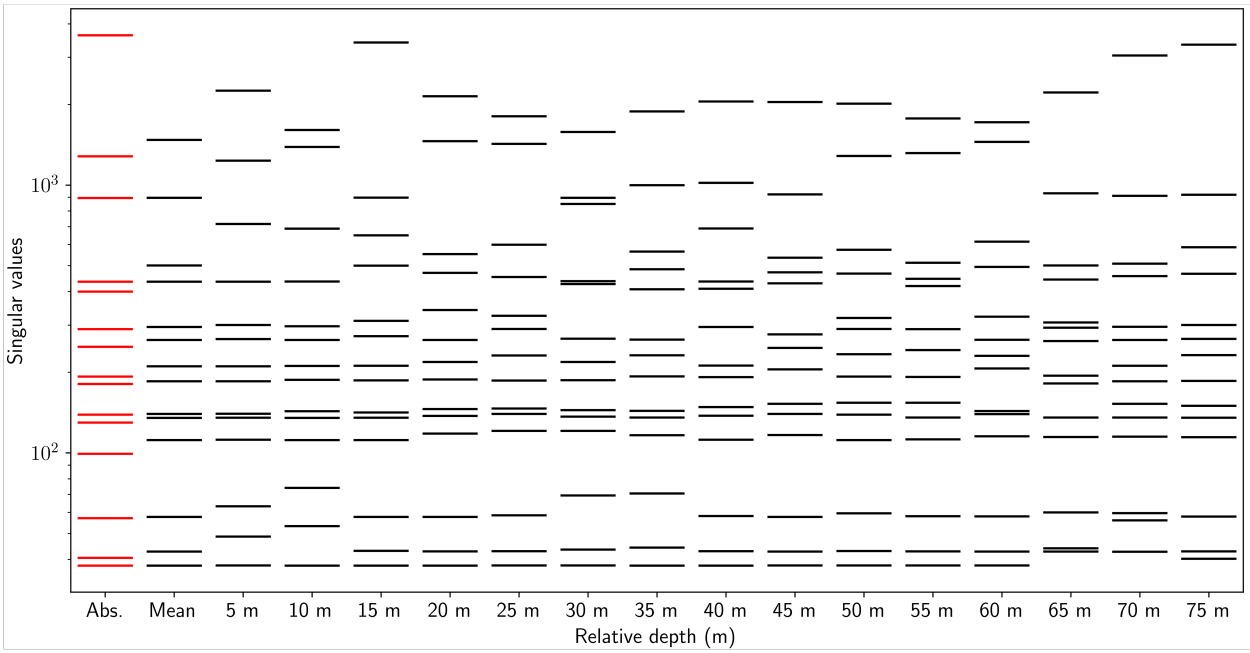

Figure C.1 Singular values for the Principle Component Analysis done for the absolute TL model manifold (left), the relative TL model manifold from subtracting the mean TL across 15 VLAs (second), and the relative TL model manifolds at 15 different reference depths.

C.4 Relative TL model manifolds and seabed distance matrices

Relative TL model manifolds are shown on the left side of Figs. C.2-C.5 using all 15 receiver depths as reference depth. For comparison, the plot limits are the same as those used for visualization in the paper. For a couple cases, these limits cut off portions of the model manifold, but the segments of the manifold that are cut off do not include any of the five labeled sediments. Additionally, seabed distance matrices (explained in the main text) are shown on the right side of Figs. C.2-C.5 for all 15 reference depths.

The average and median seabed distances (excluding self-correlation) for each relative depth are listed in Table C.4, along with the values for absolute TL and averaged relative TL. For reference depths in the middle of the water column, many seabed distances are smaller than those near the top or bottom of the water column. This is most clearly seen by inspection of the median seabed distance values. Apparently use of these mid-depth receivers as the reference channels lessens the distinction between the seabeds. Likely this is related to the combination of source depth, water depth, source-receiver distance, and frequency. Additional studies are needed to further explore more cases and formulate general rules for experimental designs that optimize the seabed distances.

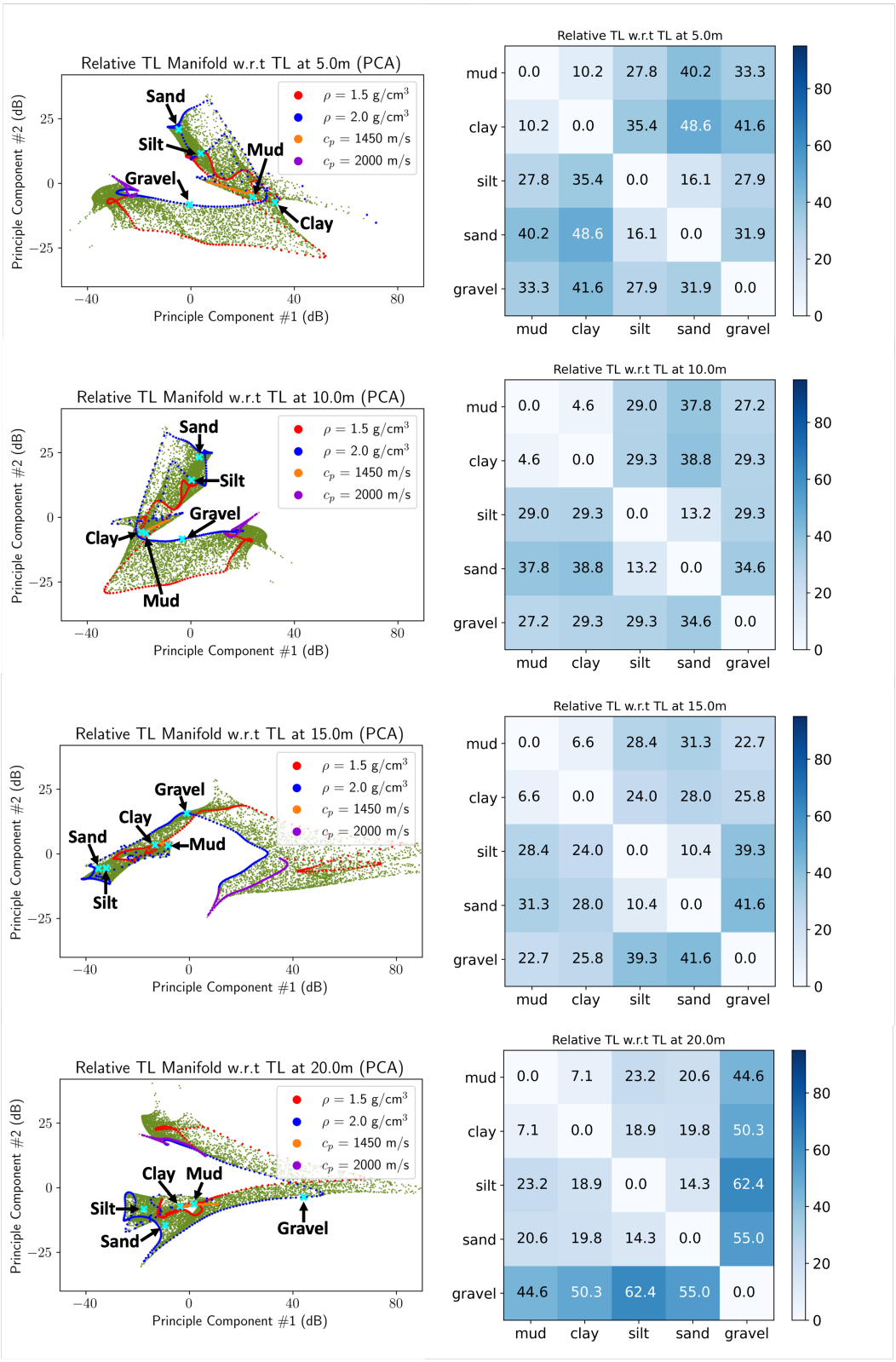

Figure C.2 Relative TL model manifolds (left) and seabed distance matrices (right) for reference depths of 5 m, 10 m, 15 m, and 20 m. Seabed distance matrices quantify the Euclidean distances between sediment types in the full 15D data space, as explained in the main text.

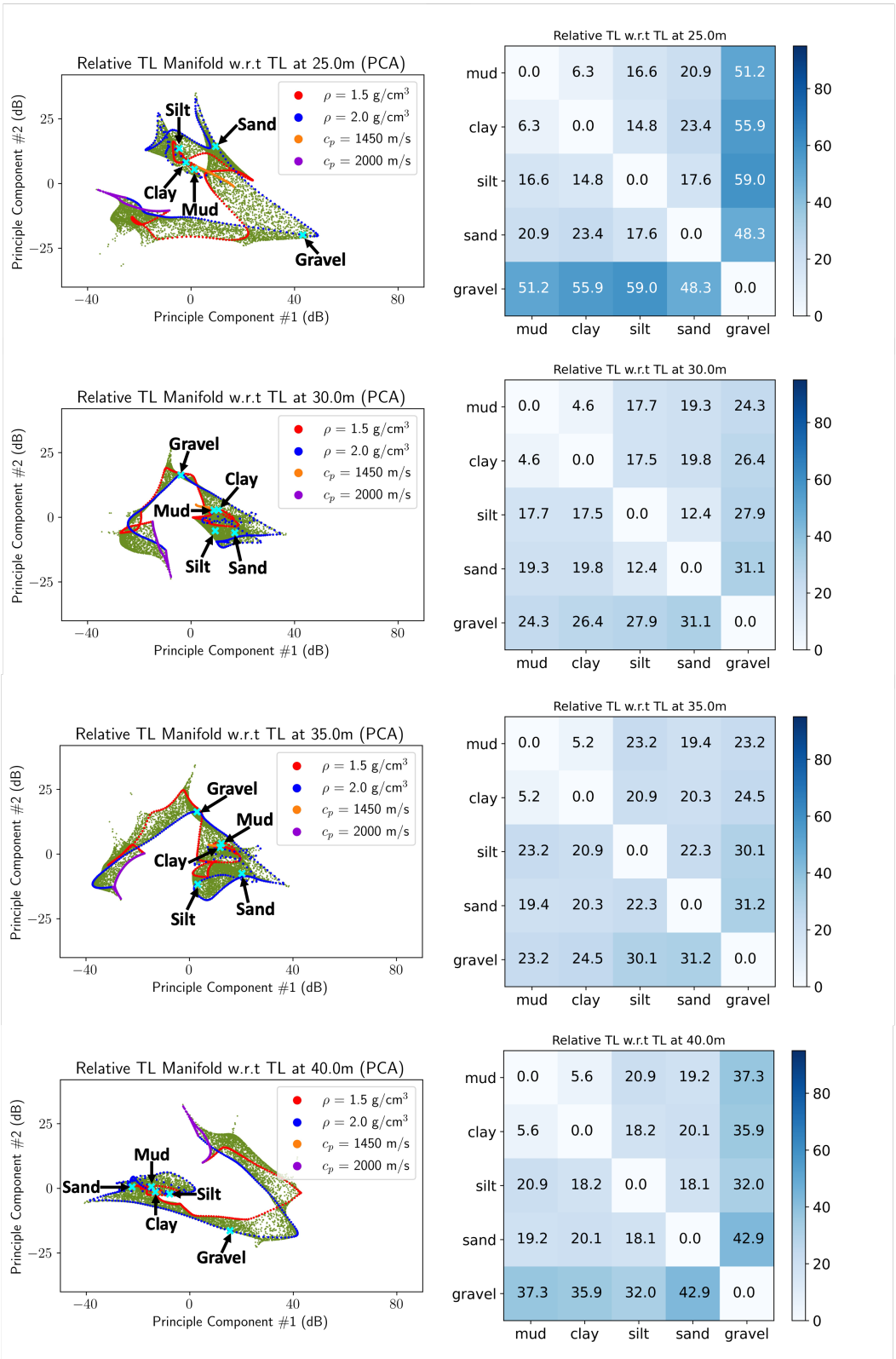

Figure C.3 Similar to Fig. C.2, relative TL model manifolds (left) and seabed distance matrices (right) are shown for reference depths of 25 m, 30 m, 35 m, and 40 m.

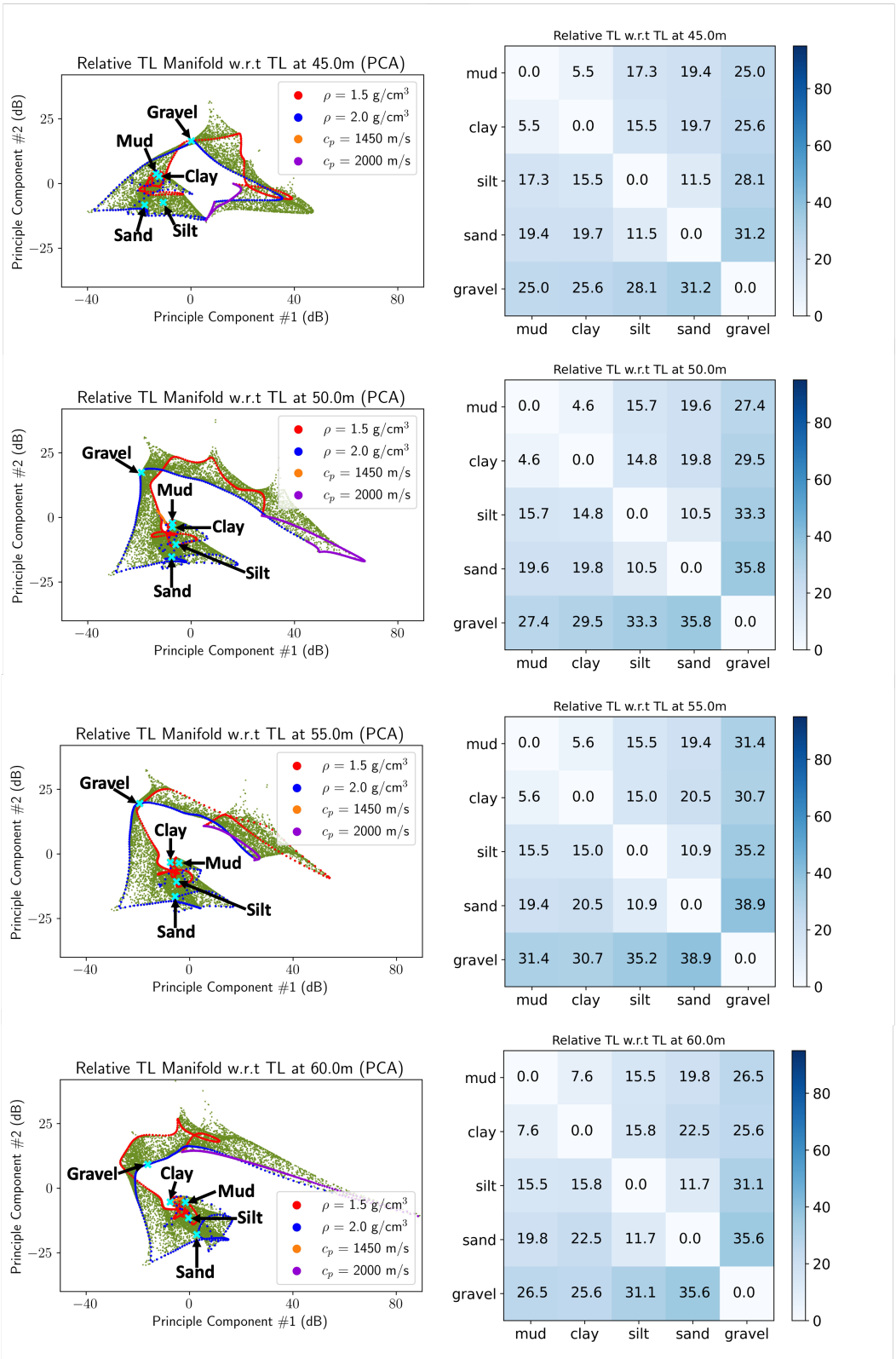

Figure C.4 Similar to Fig. C.2, relative TL model manifolds (left) and seabed distance matrices (right) are shown for reference depths of 45 m, 50 m, 55 m, and 60 m.

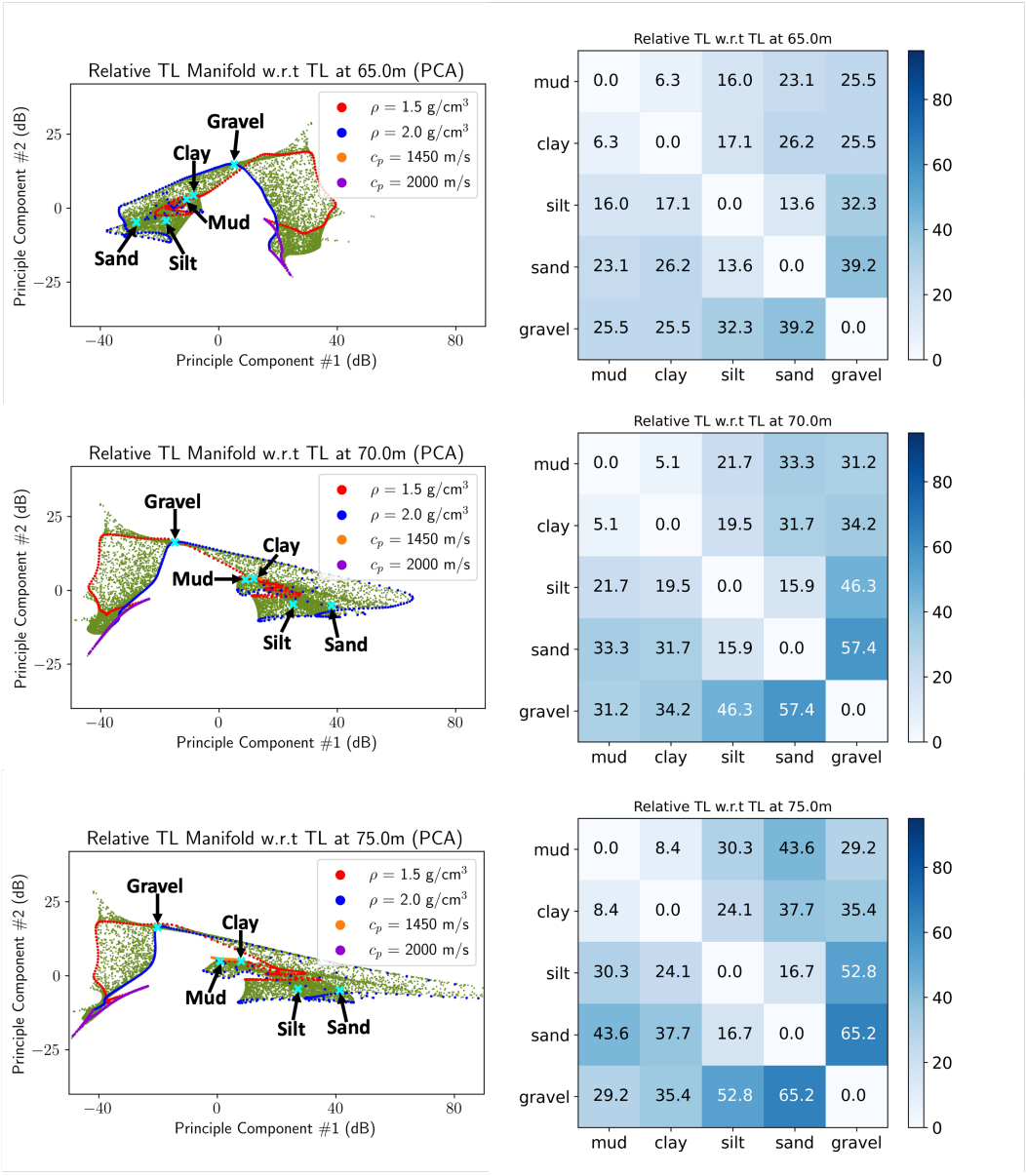

Figure C.5 Similar to Fig. C.2, relative TL model manifolds (left) and seabed distance matrices (right) are shown for reference depths of 65 m, 70 m, and 75 m.

Table C.4 Average and median sediment distances for the absolute TL model manifold, the relative TL model manifold from subtracting the mean TL across 15 VLAs, and the relative TL model manifolds for all 15 reference depths.

	Average seabed distance (dB)	Median seabed distance (dB)
Absolute	56.6	60.0
Rel. to mean TL	18.9	19.4
Rel. to 5 m	31.3	32.6
Rel. to 10 m	27.3	29.3
Rel. to 15 m	25.8	26.9
Rel. to 20 m	31.6	21.9
Rel. to 25 m	31.4	22.2
Rel. to 30 m	20.1	19.6
Rel. to 35 m	22.0	22.7
Rel. to 40 m	25.0	20.5
Rel. to 45 m	19.9	19.6
Rel. to 50 m	21.1	19.7
Rel. to 55 m	22.3	20.0
Rel. to 60 m	21.2	21.2
Rel. to 65 m	22.5	24.3
Rel. to 70 m	29.6	31.4
Rel. to 75 m	34.3	32.9

Bibliography

[1] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, in *Computational Ocean Acoustics*, 2nd ed. (Springer, 2011), Chap. 2, pp. 118–133.

[2] G. R. Potty and J. H. Miller, "Effect of shear on modal arrival times," IEEE Journal of Oceanic Engineering **45,** 103–115 (2020).

[3] P. S. Wilson, D. P. Knobles, and T. B. Neilsen, "Guest Editorial An Overview of the Seabed Characterization Experiment," IEEE Journal of Oceanic Engineering **45,** 1–13 (2020).

[4] N. R. Chapman and E. C. Shang, "Review of Geoacoustic Inversion in Underwater Acoustics," Journal of Theoretical and Computational Acoustics **29,** 2130004 (2021).

[5] M. J. Buckingham and E. M. Giddens, "On the acoustic field in a Pekeris waveguide with attenuation in the bottom half-space," J. Acoust. Soc. Am. **119,** 123–142 (2006).

[6] W. Munk and C. Wunsch, "Ocean acoustic tomography: Rays and modes," Reviews of Geophysics **21,** 777–793 (1983).

[7] L. A. Brooks and P. Gerstoft, "Ocean acoustic interferometry," The Journal of the Acoustical Society of America **121,** 3377–3385 (2007).

[8] J. A. Colosi *et al.*, "A review of recent results on ocean acoustic wave propagation in random media: Basin scales," IEEE Journal of Oceanic Engineering **24,** 138–155 (1999).

[9] K. D. LePage, "Modeling propagation and reverberation sensitivity to oceanographic and seabed variability," IEEE Journal of Oceanic Engineering **31,** 402–412 (2006).

[10] P. F. Lermusiaux, J. Xu, C.-F. Chen, S. Jan, L. Y. Chiu, and Y.-J. Yang, "Coupled ocean–acoustic prediction of transmission loss in a continental shelfbreak region: Predictive skill, uncertainty quantification, and dynamical sensitivities," IEEE Journal of Oceanic Engineering **35,** 895–916 (2010).

[11] J. R. Apel, L. A. Ostrovsky, Y. A. Stepanyants, and J. F. Lynch, "Internal solitons in the ocean and their effect on underwater sound," The Journal of the Acoustical Society of America **121,** 695–722 (2007).

[12] M. Siderius *et al.*, "Effects of ocean thermocline variability on noncoherent underwater acoustic communications," The Journal of the Acoustical Society of America **121,** 1895–1908 (2007).

[13] S. D. Rajan, "Determination of geoacoustic parameters of the ocean bottom—data requirements," The Journal of the Acoustical Society of America **92,** 2126–2140 (1992).

[14] R. T. Kessel, "The variation of modal wave numbers with geoacoustic parameters in layered media," The Journal of the Acoustical Society of America **102,** 2690–2696 (1997).

[15] D. P. Knobles, T. W. Yudichak, R. A. Koch, P. G. Cable, J. H. Miller, and G. R. Potty, "Inferences on seabed acoustics in the East China Sea from distributed acoustic measurements," IEEE Journal of Oceanic Engineering **31,** 129–144 (2006).

[16] A. Thode, "The derivative of a waveguide acoustic field with respect to a three-dimensional sound speed perturbation," The Journal of the Acoustical Society of America **115,** 2824–2833 (2004).

[17] O. A. Godin, V. U. Zavorotny, A. G. Voronovich, and V. V. Goncharov, "Refraction of sound in a horizontally inhomogeneous, time-dependent ocean," IEEE Journal of Oceanic Engineering **31,** 384–401 (2006).

[18] K. Heaney and H. Cox, "A Tactical Approach to Environmental Uncertainty and Sensitivity," IEEE Journal of Oceanic Engineering **31,** 356–367 (2006).

[19] J. A. Colosi and W. Zinicola-Lapin, "Sensitivity of mixed layer duct propagation to deterministic ocean features," The Journal of the Acoustical Society of America **149,** 1969–1978 (2021).

[20] A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky, "An overview of matched field methods in ocean acoustics," IEEE Journal of Oceanic Engineering **18,** 401–424 (1993).

[21] A. Tolstoy, *Matched field processing for underwater acoustics* (World Scientific, 1993).

[22] D. R. Del Balzo, C. Feuillade, and M. M. Rowe, "Effects of water-depth mismatch on matched-field localization in shallow water," The Journal of the Acoustical Society of America **83,** 2180–2185 (1988).

[23] D. F. Gingras, "Methods for predicting the sensitivity of matched-field processors to mismatch," The Journal of the Acoustical Society of America **86,** 1940–1949 (1989).

[24] B. F. Harrison, "An inverse problem in underwater acoustic source localization: robust matched-field processing," Inverse Problems **16,** 1641 (2000).

[25] J. Bonnel, C. Gervaise, B. Nicolas, and J. I. Mars, "Single-receiver geoacoustic inversion using modal reversal," The Journal of the Acoustical Society of America **131,** 119–128 (2012).

[26] C. E. Lindsay and N. R. Chapman, "Matched field inversion for geoacoustic model parameters using adaptive simulated annealing," IEEE Journal of Oceanic Engineering **18,** 224–231 (1993).

[27] A. Tolstoy, "Using matched-field processing to estimate shallow-water bottom properties from shot data taken in the mediterranean sea," IEEE Journal of Oceanic Engineering **21,** 471–479 (1996).

[28] M. Taroudakis and M. Markaki, "Bottom geoacoustic inversion by matched field processing-a sensitivity study," Inverse Problems **16,** 1679 (2000).

[29] Z.-H. Michalopoulou and U. Ghosh-Dastidar, "Tabu for matched-field source localization and geoacoustic inversion," The Journal of the Acoustical Society of America **115,** 135–145 (2004).

[30] S. A. Stotts, R. A. Koch, S. M. Joshi, V. T. Nguyen, V. W. Ferreri, and D. P. Knobles, "Geoacoustic inversions of horizontal and vertical line array acoustic data from a surface ship source of opportunity," IEEE Journal of Oceanic Engineering **35,** 79–102 (2010).

[31] C. Park, W. Seong, and P. Gerstoft, "Geoacoustic inversion in time domain using ship of opportunity noise recorded on a horizontal towed array," The Journal of the Acoustical Society of America **117,** 1933–1941 (2005).

[32] P. Santos, O. Rodríguez, P. Felisberto, and S. Jesus, "Seabed geoacoustic characterization with a vector sensor array," The Journal of the Acoustical Society of America **128,** 2652–2663 (2010).

[33] J. Lipor, J. Gebbie, and M. Siderius, "On the limits of distinguishing seabed types via ambient acoustic sound," The Journal of the Acoustical Society of America **154,** 2892–2903 (2023).

[34] S. E. Dosso, "Quantifying uncertainty in geoacoustic inversion. I. A fast Gibbs sampler approach," The Journal of the Acoustical Society of America **111,** 129–142 (2002).

[35] S. E. Dosso and P. L. Nielsen, "Quantifying uncertainty in geoacoustic inversion. II. Application to broadband, shallow-water data," The Journal of the Acoustical Society of America **111,** 143–159 (2002).

[36] S. E. Dosso, P. M. Giles, G. H. Brooke, D. F. McCammon, S. Pecknold, and P. C. Hines, "Linear and nonlinear measures of ocean acoustic environmental sensitivity," The Journal of the Acoustical Society of America **121,** 42–45 (2007).

[37] S. D. Rajan, J. F. Lynch, and G. V. Frisk, "Perturbative inversion methods for obtaining bottom geoacoustic parameters in shallow water," The Journal of the Acoustical Society of America **82,** 998–1017 (1987).

[38] R. T. Kessel, "A mode-based measure of field sensitivity to geoacoustic parameters in weakly range-dependent environments," The Journal of the Acoustical Society of America **105,** 122–129 (1999).

[39] B. J. DeCourcy, Y.-T. Lin, and W. L. Siegmann, "Estimating the parameter sensitivity of acoustic mode quantities for an idealized shelf-slope front," The Journal of the Acoustical Society of America **143,** 706–715 (2018).

[40] S. E. Dosso, M. G. Morley, P. M. Giles, G. H. Brooke, D. F. McCammon, S. Pecknold, and P. C. Hines, "Spatial field shifts in ocean acoustic environmental sensitivity analysis," The Journal of the Acoustical Society of America **122,** 2560–2570 (2007).

[41] L. Jaschke and N. R. Chapman, "Matched field inversion of broadband data using the freeze bath method," The Journal of the Acoustical Society of America **106,** 1838–1851 (1999).

[42] N. R. Chapman and L. Jaschke, in *Inverse Problems in Underwater Acoustics* (Springer New York, 2001), p. 15–35.

[43] C.-F. Huang, P. Gerstoft, and W. S. Hodgkiss, "Uncertainty analysis in matched-field geoacoustic inversions," The Journal of the Acoustical Society of America **119,** 197–207 (2006).

[44] P. Gerstoft, "Inversion of seismoacoustic data using genetic algorithms and a posteriori probability distributions," The Journal of the Acoustical Society of America **95,** 770–782 (1994).

[45] P. Gerstoft, "Inversion of acoustic data using a combination of genetic algorithms and the Gauss–Newton approach," The Journal of the Acoustical Society of America **97,** 2181–2190 (1995).

[46] P. Gerstoft and C. F. Mecklenbräuker, "Ocean acoustic inversion with estimation of a posteriori probability distributions," The Journal of the Acoustical Society of America **104,** 808–819 (1998).

[47] J. Dettmer, S. E. Dosso, and C. W. Holland, "Uncertainty estimation in seismo-acoustic reflection travel time inversion," The Journal of the Acoustical Society of America **122,** 161–176 (2007).

[48] C.-F. Huang, P. Gerstoft, and W. S. Hodgkiss, "Effect of ocean sound speed uncertainty on matched-field geoacoustic inversion," The Journal of the Acoustical Society of America **123,** EL162–EL168 (2008).

[49] G. Steininger, J. Dettmer, S. E. Dosso, and C. W. Holland, "Trans-dimensional joint inversion of seabed scattering and reflection data," The Journal of the Acoustical Society of America **133,** 1347–1357 (2013).

[50] J. Dettmer and S. E. Dosso, "Trans-dimensional matched-field geoacoustic inversion with hierarchical error models and interacting Markov chains," The Journal of the Acoustical Society of America **132,** 2239–2250 (2012).

[51] S. E. Dosso, C. W. Holland, and M. Sambridge, "Parallel tempering for strongly nonlinear geoacoustic inversion," The Journal of the Acoustical Society of America **132,** 3030–3040 (2012).

[52] J. Dettmer, S. E. Dosso, and C. W. Holland, "Trans-dimensional geoacoustic inversion," The Journal of the Acoustical Society of America **128,** 3393–3405 (2010).

[53] J. Dettmer, C. W. Holland, and S. E. Dosso, "Transdimensional uncertainty estimation for dispersive seabed sediments," Geophysics **78,** WB63–WB76 (2013).

[54] S. E. Dosso, J. Dettmer, G. Steininger, and C. W. Holland, "Efficient trans-dimensional Bayesian inversion for geoacoustic profile estimation," Inverse Problems **30,** 114018 (2014).

[55] J. Dettmer, S. E. Dosso, and C. W. Holland, "Sequential trans-dimensional Monte Carlo for range-dependent geoacoustic inversion," The Journal of the Acoustical Society of America **129,** 1794–1806 (2011).

[56] P. Ratilal, P. Gerstoft, and J. T. Goh, "Subspace approach to inversion by genetic algorithms involving multiple frequencies," Journal of Computational Acoustics **6,** 99–115 (1998).

[57] S. E. Dosso, M. J. Wilmut, and A.-L. Lapinski, "An adaptive-hybrid algorithm for geoacoustic inversion," IEEE Journal of Oceanic Engineering **26,** 324–336 (2001).

[58] M. D. Collins and L. Fishman, "Efficient navigation of parameter landscapes," The Journal of the Acoustical Society of America **98,** 1637–1644 (1995).

[59] J. R. Buck, J. C. Preisig, and K. E. Wage, "A unified framework for mode filtering and the maximum a posteriori mode filter," The Journal of the Acoustical Society of America **103,** 1813–1824 (1998).

[60] J. R. Buck, "Information theoretic bounds on source localization performance," In *Sensor Array and Multichannel Signal Processing Workshop Proceedings, 2002*, pp. 184–188 (2002).

[61] X. Fang, W. Yan, Y. Wang, and Z. Wang, "Observability analysis of underwater localization based on fisher information matrix," In *Proceedings of the 32nd Chinese Control Conference*, pp. 4866–4869 (2013).

[62] S. Finette and P. C. Mignerey, "Stochastic matched-field localization of an acoustic source based on principles of Riemannian geometry," The Journal of the Acoustical Society of America **143,** 3628–3638 (2018).

[63] S. R. Tuladhar and J. R. Buck, "Optimum array design to maximize Fisher information for bearing estimation," The Journal of the Acoustical Society of America **130,** 2797–2806 (2011).

[64] L. N. Kloepper, J. R. Buck, Y. Liu, and P. E. Nachtigall, "Off-axis targets maximize bearing Fisher Information in broadband active sonar," The Journal of the Acoustical Society of America **143,** EL43–EL48 (2018).

[65] X. Zhu and J. R. Buck, "Designing nonuniform linear arrays to maximize mutual information for bearing estimation," The Journal of the Acoustical Society of America **128,** 2926–2939 (2010).

[66] M. D. Tidwell and J. R. Buck, "Designing linear fm active sonar waveforms for continuous line source transducers to maximize the fisher information at a desired bearing," In *2019 Sensor Signal Processing for Defence Conference (SSPD)*, pp. 1–5 (2019).

[67] A. ElMoslimany, M. Zhou, T. M. Duman, and A. Papandreou-Suppappola, "An underwater acoustic communication scheme exploiting biological sounds," Wireless Communications and Mobile Computing **16,** 2194–2211 (2016).

[68] Z. Wang, J. Xu, Y. Feng, Y. Wang, G. Xie, X. Hou, W. Men, and Y. Ren, "Fisher-Information-Matrix-Based USBL Cooperative Location in USV–AUV Networks," Sensors **23,** 7429 (2023).

[69] S. Finette, "Non-Euclidean estimation of acoustic cross-spectral density matrices: Evaluation by matched-field localization in an ocean waveguide," JASA Express Letters 2 (2022).

[70] D. J. Brooker, S. Finette, and P. C. Mignerey, "Detection on a distributed network using information geometry," JASA Express Letters **1,** 036001 (2021).

[71] C.-T. J. Hsu, G. J. Brouhard, and P. François, "Numerical Parameter Space Compression and Its Application to Biophysical Models," Biophysical Journal **118,** 1455–1465 (2020).

[72] M. K. Transtrum and P. Qiu, "Bridging Mechanistic and Phenomenological Models of Complex Biological Systems," PLOS Computational Biology **12,** e1004915 (2016).

[73] B. K. Mannakee, A. P. Ragsdale, M. K. Transtrum, and R. N. Gutenkunst, in *Uncertainty in Biology: A Computational Modeling Approach*, L. Geris and D. Gomez-Cabrero, eds., (Springer International Publishing, Cham, 2016), pp. 271–299.

[74] R. Guzzi, T. Colombo, and P. Paci, in *Systems Biology* (Springer New York, 2017), p. 69–94.

[75] H. Hass, C. Loos, E. Raimúndez-Álvarez, J. Timmer, J. Hasenauer, and C. Kreutz, "Benchmark problems for dynamic modeling of intracellular processes," Bioinformatics **35,** 3073–3082 (2019).

[76] D. V. Raman, J. Anderson, and A. Papachristodoulou, "Delineating parameter unidentifiabilities in complex models," Physical Review E 95 (2017).

[77] K. Erguler and M. P. H. Stumpf, "Practical limits for reverse engineering of dynamical systems: a statistical analysis of sensitivity and parameter inferability in systems biology models," Molecular BioSystems **7,** 1593 (2011).

[78] D. M. Lombardo and W.-J. Rappel, "Systematic reduction of a detailed atrial myocyte model," Chaos: An Interdisciplinary Journal of Nonlinear Science 27 (2017).

[79] K. S. Brown and J. P. Sethna, "Statistical mechanical approaches to models with many poorly known parameters," Physical Review E 68 (2003).

[80] K. S. Brown, C. C. Hill, G. A. Calero, C. R. Myers, K. H. Lee, J. P. Sethna, and R. A. Cerione, "The statistical mechanics of complex signaling networks: nerve growth factor signaling," Physical Biology **1,** 184–195 (2004).

[81] R. N. Gutenkunst, J. J. Waterfall, F. P. Casey, K. S. Brown, C. R. Myers, and J. P. Sethna, "Universally Sloppy Parameter Sensitivities in Systems Biology Models," PLoS Computational Biology **3,** e189 (2007).

[82] H. Boigenzahn, L. D. González, J. C. Thompson, V. M. Zavala, and J. Yin, "Kinetic Modeling and Parameter Estimation of a Prebiotic Peptide Reaction Network," Journal of Molecular Evolution **91,** 730–744 (2023).

[83] G. M. Monsalve-Bravo, B. A. J. Lawson, C. Drovandi, K. Burrage, K. S. Brown, C. M. Baker, S. A. Vollert, K. Mengersen, E. McDonald-Madden, and M. P. Adams, "Analysis of sloppiness in model simulations: Unveiling parameter uncertainty when mathematical models are fitted to data," Science Advances 8 (2022).

[84] D. G. Whittaker, J. Wang, J. G. Shuttleworth, R. Venkateshappa, J. M. Kemp, T. W. Claydon, and G. R. Mirams, "Ion channel model reduction using manifold boundaries," Journal of The Royal Society Interface 19 (2022).

[85] A. White, M. Tolman, H. D. Thames, H. R. Withers, K. A. Mason, and M. K. Transtrum, "The Limitations of Model-Based Experimental Design and Parameter Estimation in Sloppy Systems," PLOS Computational Biology **12,** e1005227 (2016).

[86] C. Schillings, M. Sunnåker, J. Stelling, and C. Schwab, "Efficient Characterization of Parametric Uncertainty of Complex (Bio)chemical Networks," PLOS Computational Biology **11,** e1004457 (2015).

[87] P. Subsoontorn, J. Kim, and E. Winfree, "Ensemble Bayesian Analysis of Bistability in a Synthetic Transcriptional Switch," ACS Synthetic Biology **1,** 299–316 (2012).

[88] C. N. K. Anderson and M. K. Transtrum, "Sloppy model analysis identifies bifurcation parameters without normal form analysis," Physical Review E **108,** 064215 (2023).

[89] E. Braun, in *Phenotypic Switching* (Elsevier, 2020), p. 305–334.

[90] A. P. Browning and M. J. Simpson, "Geometric analysis enables biological insight from complex non-identifiable models using simple surrogates," PLOS Computational Biology **19,** e1010844 (2023).

[91] M. H. Hennig, "The sloppy relationship between neural circuit structure and function," The Journal of Physiology **601,** 3025–3035 (2022).

[92] A. Ponce-Alvarez, G. Mochol, A. Hermoso-Mendizabal, J. de la Rocha, and G. Deco, "Cortical state transitions and stimulus response evolve along stiff and sloppy parameter dimensions, respectively," eLife **9,** e53268 (2020).

[93] T. Jabri and J. N. MacLean, "Large-Scale Algorithmic Search Identifies Stiff and Sloppy Dimensions in Synaptic Architectures Consistent With Murine Neocortical Wiring," Neural Computation **34,** 2347–2373 (2022).

[94] D. Panas, H. Amin, A. Maccione, O. Muthmann, M. van Rossum, L. Berdondini, and M. H. Hennig, "Sloppiness in Spontaneously Active Neuronal Networks," The Journal of Neuroscience **35,** 8480–8492 (2015).

[95] A. Hartoyo, P. J. Cadusch, D. T. J. Liley, and D. G. Hicks, "Parameter estimation and identifiability in a neural population model for electro-cortical activity," PLOS Computational Biology **15,** e1006694 (2019).

[96] R. M. Warner, K. S. Brown, J. D. Benson, A. Eroglu, and A. Z. Higgins, "Multiple cryoprotectant toxicity model for vitrification solution optimization," Cryobiology **108,** 1–9 (2022).

[97] K. Ghosh, S. Vernuccio, and A. W. Dowling, "Nonlinear Reactor Design Optimization With Embedded Microkinetic Model Information," Frontiers in Chemical Engineering 4 (2022).

[98] B. J. Befort, A. Garciadiego, J. Wang, K. Wang, G. Franco, E. J. Maginn, and A. W. Dowling, "Data science for thermodynamic modeling: Case study for ionic liquid and hydrofluorocarbon refrigerant mixtures," Fluid Phase Equilibria **572,** 113833 (2023).

[99] M. Quaglio, C. Waldron, A. Pankajakshan, E. Cao, A. Gavriilidis, E. S. Fraga, and F. Galvanin, "An online reparametrisation approach for robust parameter estimation in automated model identification platforms," Computers & Chemical Engineering **124,** 270–284 (2019).

[100] A. Goshtasbi, J. Chen, J. R. Waldecker, S. Hirano, and T. Ersal, "Effective Parameterization of PEM Fuel Cell Models—Part I: Sensitivity Analysis and Parameter Identifiability," Journal of The Electrochemical Society **167,** 044504 (2020).

[101] M. K. Transtrum, A. T. Saric, and A. M. Stankovic, "Measurement-Directed Reduction of Dynamic Models in Power Systems," IEEE Transactions on Power Systems **32,** 2243–2253 (2017).

[102] M. K. Transtrum, B. L. Francis, A. T. Saric, and A. M. Stankovic, "Simultaneous Global Identification of Dynamic and Network Parameters in Transient Stability Studies," In *2018 IEEE Power & Energy Society General Meeting (PESGM)*, (IEEE, 2018).

[103] V. Svenda, M. Transtrum, B. Francis, A. Saric, and A. Stankovic, "State Estimation Model Reduction Through the Manifold Boundary Approximation Method," IEEE Transactions on Power Systems **37,** 272–281 (2022).

[104] A. T. Saria, A. A. Saric, M. K. Transtrum, and A. M. Stankovic, "Symbolic Regression for Data-Driven Dynamic Model Refinement in Power Systems," IEEE Transactions on Power Systems **36,** 2390–2402 (2021).

[105] S. A. Vollert, C. Drovandi, G. M. Monsalve-Bravo, and M. P. Adams, "Strategic model reduction by analysing model sloppiness: A case study in coral calcification," Environmental Modelling & Software **159,** 105578 (2023).

[106] G. L. Marschmann, H. Pagel, P. Kügler, and T. Streck, "Equifinality, sloppiness, and emergent structures of mechanistic soil biogeochemical models," Environmental Modelling & Software **122,** 104518 (2019).

[107] Y. Kurniawan, C. L. Petrie, K. J. Williams, M. K. Transtrum, E. B. Tadmor, R. S. Elliott, D. S. Karls, and M. Wen, "Bayesian, frequentist, and information geometric approaches to parametric uncertainty quantification of classical empirical interatomic potentials," The Journal of Chemical Physics **156,** 214103 (2022).

[108] M. K. Transtrum, B. B. Machta, and J. P. Sethna, "Geometry of nonlinear least squares with applications to sloppy models and optimization," Physical Review E **036701,** 1–35 (2011).

[109] B. B. Machta, R. Chachra, M. K. Transtrum, and J. P. Sethna, "Parameter Space Compression Underlies Emergent Theories and Predictive Models," Science **342,** 604–607 (2013).

[110] J. Mao, I. Griniasty, H. K. Teoh, R. Ramesh, R. Yang, M. K. Transtrum, J. P. Sethna, and P. Chaudhari, "The training process of many deep networks explores the same low-dimensional manifold," Proceedings of the National Academy of Sciences **121,** e2310002121 (2024).

[111] K. N. Quinn, C. B. Clement, F. De Bernardis, M. D. Niemack, and J. P. Sethna, "Visualizing probabilistic models and data with Intensive Principal Component Analysis," Proceedings of the National Academy of Sciences **116,** 13762–13767 (2019).

[112] K. Naumann-Woleske, M. S. Knicker, M. Benzaquen, and J.-P. Bouchaud, "Exploration of the Parameter Space in Macroeconomic Agent-Based Models,", 2021.

[113] D. S. Brée, D. Challet, and P. P. Peirano, "Prediction accuracy and sloppiness of log-periodic functions," Quantitative Finance **13,** 275–280 (2013).

[114] G. Curato, J. Gatheral, and F. Lillo, "Optimal execution with non-linear transient market impact," Quantitative Finance **17,** 41–54 (2016).

[115] M. Sarovar, J. Zhang, and L. Zeng, "Reliability of analog quantum simulation," EPJ Quantum Technology 4 (2017).

[116] V. Petzold, T. Bligaard, and K. W. Jacobsen, "Construction of New Electronic Density Functionals with Error Estimation Through Fitting," Topics in Catalysis **55,** 402–417 (2012).

[117] J. J. Waterfall, F. P. Casey, R. N. Gutenkunst, K. S. Brown, C. R. Myers, P. W. Brouwer, V. Elser, and J. P. Sethna, "Sloppy-Model Universality Class and the Vandermonde Matrix," Physical Review Letters **97,** 150601 (2006).

[118] M. Imbrišak and K. Nomura, "Classical and Bayesian error analysis of the relativistic mean-field model for doubly magic nuclei," Physical Review C **108,** 024321 (2023).

[119] M. Imbrišak and K. Nomura, "Stability of the manifold boundary approximation method for reductions of nuclear structure models," Physical Review C **107,** 034304 (2023).

[120] M. Dias, J. Frazer, and M. D. Marsh, "Seven lessons from manyfield inflation in random potentials," Journal of Cosmology and Astroparticle Physics **2018,** 036–036 (2018).

[121] M. C. Mortenson, T. B. Neilsen, M. K. Transtrum, and D. P. Knobles, "Accurate broadband gradient estimates enable local sensitivity analysis of ocean acoustic models," Journal of Theoretical and Computational Acoustics p. 2250015 (2023).

[122] K. N. Quinn, M. C. Abbott, M. K. Transtrum, B. B. Machta, and J. P. Sethna, "Information geometry for multiparameter models: new perspectives on the origin of simplicity," Reports on Progress in Physics **86,** 035901 (2022).

[123] M. K. Transtrum, B. B. Machta, K. S. Brown, B. C. Daniels, C. R. Myers, and J. P. Sethna, "Perspective: Sloppiness and emergent theories in physics, biology, and beyond," Journal of Chemical Physics 143 (2015).

[124] W. Bergan, I. Bazarov, C. Duncan, D. Liarte, D. Rubin, and J. Sethna, "Online storage ring optimization using dimension-reduction and genetic algorithms," Physical Review Accelerators and Beams 22 (2019).

[125] M. K. Transtrum and P. Qiu, "Model Reduction by Manifold Boundaries," Physical Review Letters **113,** 098701 (2014).

[126] C. Petrie, C. Anderson, C. Maekawa, T. Maekawa, and M. K. Transtrum, "The supremum principle selects simple, transferable models,", 2021.

[127] B. L. Francis, M. K. Transtrum, A. T. Saric, and A. M. Stankovic, "Piecemeal Reduction of Models of Large Networks," In *2021 60th IEEE Conference on Decision and Control (CDC)*, (IEEE, 2021).

[128] K. Theriault and A. Baggeroer, "Structure Estimation from Acoustic Reflection Measurements," In *OCEANS '76*, (IEEE, 1976).

[129] K. Theriault and A. Baggeroer, "Structure estimation accuracy in discrete layered media," IEEE Journal of Oceanic Engineering **4,** 4–13 (1979).

[130] J. Moura and A. Baggeroer, "Passive systems theory with narrow-band and linear constraints: Part I-Spatial diversity," IEEE Journal of Oceanic Engineering **3,** 5–13 (1978).

[131] Y. Rockah and P. Schultheiss, "Array shape calibration using sources in unknown locations–Part II: Near-field sources and estimator implementation," IEEE Transactions on Acoustics, Speech, and Signal Processing **35,** 724–735 (1987).

[132] Y. Rockah and P. Schultheiss, "Array shape calibration using sources in unknown locations–Part I: Far-field sources," IEEE Transactions on Acoustics, Speech, and Signal Processing **35,** 286–299 (1987).

[133] A. B. Baggeroer, W. A. Kuperman, and H. Schmidt, "Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem," The Journal of the Acoustical Society of America **83,** 571–587 (1988).

[134] A. B. Baggeroer, "The stochastic Cramér-Rao bound for source localization and medium tomography using vector sensors," The Journal of the Acoustical Society of America **141,** 3430–3449 (2017).

[135] H. Song, "Cramer-Rao bound for passive localization of a moving source in ocean acoustics," In *IEEE International Conference on Acoustics Speech and Signal Processing*, (IEEE, 1993).

[136] A. Baggeroer and H. Schmidt, "Cramér-Rao bounds for matched field tomography and ocean acoustic tomography," In , **5,** 2763–2766 (IEEE, 1995).

[137] M. J. D. Rendas and J. M. Moura, "Cramer-Rao bounds for passive range and depth in a vertically inhomogeneous medium," In *International Conference on Acoustics, Speech, and Signal Processing*, pp. 2779–2782 (1990).

[138] M.-J. D. Rendas and J. M. Moura, "Cramer-Rao bound for location systems in multipath environments," IEEE Transactions on Signal Processing **39,** 2593–2610 (1991).

[139] B. Friedlander and A. Weiss, "Direction finding using noise covariance modeling," IEEE Transactions on Signal Processing **43,** 1557–1567 (1995).

[140] J. Gebbie, M. Siderius, and J. S. Allen III, "A two-hydrophone range and bearing localization algorithm with performance analysis," The Journal of the Acoustical Society of America **137,** 1586–1597 (2015).

[141] A. Koochakzadeh and P. Pal, "Cramér–Rao Bounds for Underdetermined Source Localization," IEEE Signal Processing Letters **23,** 919–923 (2016).

[142] P. Pal and P. P. Vaidyanathan, "Parameter identifiability in Sparse Bayesian Learning," In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (IEEE, 2014).

[143] T. Li, A. Ekpenyong, and Y.-F. Huang, "Source Localization and Tracking Using Distributed Asynchronous Sensors," IEEE Transactions on Signal Processing **54,** 3991–4003 (2006).

[144] A. Weiss, T. Arikan, H. Vishnu, G. B. Deane, A. C. Singer, and G. W. Wornell, "A semi-blind method for localization of underwater acoustic sources," IEEE Transactions on Signal Processing **70,** 3090–3106 (2022).

[145] M. Siderius and J. Gebbie, "Environmental information content of ocean ambient noise," The Journal of the Acoustical Society of America **146,** 1824–1833 (2019).

[146] S. C. Walker, C. Yardim, A. Thode, and E. Arias-Castro, "Using Fisher information to quantify uncertainty in environmental parameters estimated from correlated ambient noise," The Journal of the Acoustical Society of America **133,** EL228–EL234 (2013).

[147] S. Narasimhan and J. L. Krolik, "Fundamental limits on acoustic source range estimation performance in uncertain ocean channels," The Journal of the Acoustical Society of America **97,** 215–226 (1995).

[148] P. K. Tam, K. T. Wong, and Y. Song, "An hybrid Cramer-Rao bound in closed form for direction-of-arrival estimation by an "acoustic vector sensor" with gain-phase uncertainties," IEEE transactions on signal processing **62,** 2504–2516 (2014).

[149] W. Xu, A. B. Baggeroer, and C. D. Richmond, "Bayesian bounds for matched-field parameter estimation," IEEE Transactions on Signal Processing **52,** 3293–3305 (2004).

[150] J. Gebbie and M. Siderius, "Optimal environmental estimation with ocean ambient noise," The Journal of the Acoustical Society of America **149,** 825–834 (2021).

[151] J.-P. Hermand, M. Meyer, M. Asch, and M. Berrada, "Adjoint-based acoustic inversion for the physical characterization of a shallow water environment," The Journal of the Acoustical Society of America **119,** 3860–3871 (2006).

[152] M. Hawkes and A. Nehorai, "Wideband source localization using a distributed acoustic vector-sensor array," IEEE Transactions on Signal Processing **51,** 1479–1491 (2003).

[153] R. J. Cederberg and M. D. Collins, "Application of an improved self-starter to geoacoustic inversion," IEEE Journal of Oceanic Engineering **22,** 102–109 (1997).

[154] D. K. Dacol, M. D. Collins, and J. F. Lingevitch, "An efficient parabolic equation solution based on the method of undetermined coefficients," The Journal of the Acoustical Society of America **106,** 1727–1731 (1999).

[155] J. F. Lingevitch and M. D. Collins, "Estimating elastic sediment properties with the self-starter," Wave Motion **31,** 157–163 (2000).

[156] L. T. Fialkowski, J. F. Lingevitch, J. S. Perkins, D. K. Dacol, and M. D. Collins, "Geoacoustic inversion using a rotated coordinate system and simulated annealing," IEEE Journal of Oceanic Engineering **28,** 370–379 (2003).

[157] T. B. Neilsen and D. P. Knobles, "Geoacoustic inversion of range-dependent data with added Gaussian noise," IEEE Journal of Oceanic Engineering **28,** 446–453 (2003).

[158] T. B. Neilsen, "An iterative implementation of rotated coordinates for inverse problems," The Journal of the Acoustical Society of America **113,** 2574–2586 (2003).

[159] T. B. Neilsen, "Localization of multiple acoustic sources in the shallow ocean," The Journal of the Acoustical Society of America **118,** 2944–2953 (2005).

[160] M. Nicholas, J. S. Perkins, G. J. Orris, L. T. Fialkowski, and G. J. Heard, "Environmental inversion and matched-field tracking with a surface ship and an L-shaped receiver array," The Journal of the Acoustical Society of America **116,** 2891–2901 (2004).

[161] M. J. Isakson and T. B. Neilsen, "The viability of reflection loss measurement inversion to predict broadband acoustic behavior," The Journal of the Acoustical Society of America **120,** 135–144 (2006).

[162] L. T. Fialkowski, M. D. Collins, J. S. Perkins, and W. Kuperman, "Source localization in noisy and uncertain ocean environments," The Journal of the Acoustical Society of America **101,** 3539–3545 (1997).

[163] S. E. Dosso and J. Dettmer, "Bayesian matched-field geoacoustic inversion," Inverse Problems **27,** 055009 (2011).

[164] Z.-H. Michalopoulou and P. Gerstoft, "Multipath broadband localization, bathymetry, and sediment inversion," IEEE Journal of Oceanic Engineering **45,** 92–102 (2019).

[165] C. L. Pekeris, "Theory of propagation of explosive sound in shallow water," in *Geological Society of America Memoirs* (Geological Society of America, 1948), pp. 1–116.

[166] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, in *Computational Ocean Acoustics*, 2nd ed. (Springer, 2011), Chap. 2, pp. 118–133.

[167] F. Garofalo, L. Iannelli, and F. Vasca, "Participation Factors and their Connections to Residues and Relative Gain Array," IFAC Proceedings Volumes **35,** 125–130 (2002).

[168] F. Casey, R. Gutenkunst, C. Myers, D. Baird, K. Brown, J. Waterfall, Q. Feng, R. Cerione, and J. Sethna, "Optimal Experimental Design in an Epidermal Growth Factor Receptor Signalling and Down-Regulation Model," IET Systems Biology **1,** 190–202 (2007).

[169] T. Nikšić and D. Vretenar, ""Sloppy" nuclear energy density functionals: Effective model reduction," Physical Review C **94,** 024333 (2016).

[170] T. Nikšić, M. Imbrišak, and D. Vretenar, ""Sloppy" nuclear energy density functionals. II. Finite nuclei," Physical Review C **95,** 054304 (2017).

[171] P. E. Pare, D. Grimsman, A. T. Wilson, M. K. Transtrum, and S. Warnick, "Model Boundary Approximation Method as a Unifying Framework for Balanced Truncation and Singular Perturbation Approximation," IEEE Transactions on Automatic Control **64,** 4796–4802 (2019).

[172] X. Zhao, Y. Hou, D. Song, and W. Li, "A Confident Information First Principle for Parameter Reduction and Model Selection of Boltzmann Machines," IEEE Transactions on Neural Networks and Learning Systems **29,** 1608–1621 (2018).

[173] A. Holiday, M. Kooshkbaghi, J. M. Bello-Rivas, C. William Gear, A. Zagaris, and I. G. Kevrekidis, "Manifold learning for parameter reduction," Journal of Computational Physics **392,** 419–431 (2019).

[174] H. K. Teoh, K. N. Quinn, J. Kent-Dobias, C. B. Clement, Q. Xu, and J. P. Sethna, "Visualizing probabilistic models in Minkowski space with intensive symmetrized Kullback-Leibler embedding," Physical Review Research **2,** 033221 (2020).

[175] M. K. Transtrum, B. B. MacHta, and J. P. Sethna, "Why are nonlinear fits to data so challenging?," Physical Review Letters **104,** 2–5 (2010).

[176] M. Wen, J. Li, P. Brommer, R. S. Elliott, J. P. Sethna, and E. B. Tadmor, "A KIM-compliantpotfitfor fitting sloppy interatomic potentials: application to the EDIP model for silicon," Modelling and Simulation in Materials Science and Engineering **25,** 014001 (2016).

[177] D. S. Berman, M. S. Klinger, and A. G. Stapleton, "Bayesian renormalization," Machine Learning: Science and Technology **4,** 045011 (2023).

[178] M. Girolami and B. Calderhead, "Riemann Manifold Langevin and Hamiltonian Monte Carlo Methods," Journal of the Royal Statistical Society: Series B (Statistical Methodology) **73,** 123–214 (2011).

[179] H. H. Mattingly, M. K. Transtrum, M. C. Abbott, and B. B. Machta, "Maximizing the information learned from finite data selects a simple model," Proceedings of the National Academy of Sciences **115,** 1760–1765 (2018).

[180] C. H. LaMont and P. A. Wiggins, "Correspondence between thermodynamics and inference," Physical Review E **99,** 052140 (2019).

[181] S. ichi Amari, "Information geometry," Japanese Journal of Mathematics **16,** 1–48 (2021).

[182] F. Nielsen, "An Elementary Introduction to Information Geometry," Entropy **22,** 1100 (2020).

[183] M. K. Transtrum and P. Qiu, "Model Reduction by Manifold Boundaries," Physical Review Letters 113 (2014).

[184] K. S. Brown and J. P. Sethna, "Statistical mechanical approaches to models with many poorly known parameters," Physical Review E 68 (2003).

[185] M. K. Transtrum, B. B. Machta, K. S. Brown, B. C. Daniels, C. R. Myers, and J. P. Sethna, "Perspective: Sloppiness and emergent theories in physics, biology, and beyond," The Journal of Chemical Physics **143,** 010901 (2015).

[186] K. N. Quinn, M. C. Abbott, M. K. Transtrum, B. B. Machta, and J. P. Sethna, "Information geometry for multiparameter models: new perspectives on the origin of simplicity," Reports on Progress in Physics **86,** 035901 (2022).

[187] E. K. Westwood, C. T. Tindle, and N. R. Chapman, "A normal mode model for acousto-elastic ocean environments," The Journal of the Acoustical Society of America **100,** 3631–3645 (1996).

[188] J. Ansel *et al.*, "PyTorch 2: Faster Machine Learning Through Dynamic Python Bytecode Transformation and Graph Compilation," In *Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2*, ASPLOS '24 (ACM, 2024).

[189] E. K. Westwood, C. T. Tindle, and N. R. Chapman, "A normal mode model for acousto-elastic ocean environments," The Journal of the Acoustical Society of America **100,** 3631–3645 (1996).