Adaptive optimal experimental design and inversion of a coupled fluid flow and geophysical imaging model for reservoir monitoring

by

Jennifer Fohring

B.Sc., The University of British Columbia, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Geophysics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

May 2016

© Jennifer Fohring 2016
Abstract

Imaging and prediction of fluid flow within the subsurface provides information crucial to decision making processes in fields such as groundwater management and enhanced oil recovery. The flow of a fluid through a reservoir depends primarily on the permeability of the subsurface rock; a quantity that is often unknown throughout the entire domain of the reservoir. One method for predicting flow is to estimate the permeability of the reservoir and simulate flow through a mathematical subsurface flow model. Given the model, flow data can be inverted to estimate the permeability. However, this inversion approach can lead to inaccurate results due to the sparse sampling of flow data, and thus inaccurate predictions.

To acquire a higher sampling of data, geophysical survey techniques are applied in order to efficiently collect a higher density of data sampled at the surface. These data are sensitive to changes to the geophysical properties of the reservoir due to flow. Inversion of geophysical data then provides images of changes to the geophysical properties of the reservoir. In order to estimate the flow parameters using geophysical data, the two mathematical models require coupling.

The thesis therefore proposes two approaches to improve the imaging and prediction of flow. First, a novel coupled inverse problem for estimating the fluid velocity field and the initial geophysical property model from geophysical data is developed. Second, a new method of optimally designing the geophysical survey for the coupled inverse problem is developed. The new adaptive design approach builds on traditional A-Optimal design methods such that historic data are included in the design algo-
Abstract

This produces designs that adapt with flow in the subsurface and reduce the collection of unnecessary data. Both the coupled inverse problem and adaptive survey design method are demonstrated using a seismic tomography geophysical survey and a tracer advection fluid flow model. Numerical examples show that the coupled approach yields an improved flow estimate as well as improved image quality, while the adaptive optimal designs provide sufficient geophysical data.
Preface

This thesis contains original research conducted while studying at the University of British Columbia, resulting in two publications and one expanded conference proceeding.

The idea of the coupled inverse problem presented in Chapter 4 came originally from conversations with Dr. Eldad Haber. The subsequent derivations, code implementation, numerical tests, and manuscript preparation were carried out in collaboration with Dr. Lars Ruthotto, a post doctoral researcher with Dr. Eldad Haber at the time. The work was published in Fohring, J., Haber, E., and Ruthotto, L. (2014). Geophysical imaging of fluid flow in porous media. *SIAM Journal on Scientific Computing*, 36(5):218–236, and various parts of were adapted from an SEG conference proceeding (Fohring, J., Ruthotto, L., and Haber, E. (2013). Geophysical Imaging, Reservoir History Matching and Forecasting. In 2013 SEG Annual Meeting, Houston. Society of Exploration Geophysicists).

The idea of adaptive experimental design also resulted from conversations with Dr. Eldad Haber. The development of this idea is presented in Chapter 5. The subsequent derivations, code implementation, numerical tests, and manuscript preparation for adaptive optimal design were carried out by myself with contributions, input, and advice from Dr. Haber. The bulk of the text in Chapter 5 is adapted from the paper (Fohring, J. and Haber, E. (2016). Adaptive A-optimal experimental design for linear dynamical systems. *SIAM Journal on Uncertainty Quantification*, xx:1–19), which is still at the time of writing, going through the review process.
# Table of Contents

Abstract ................................................................. ii

Preface ................................................................. iv

Table of Contents ..................................................... v

List of Figures ......................................................... viii

Acknowledgements ..................................................... xii

Dedication ............................................................. xiv

1 Introduction .......................................................... 1
  1.1 Reservoir characterization ........................................ 1
  1.2 Geophysics for reservoir characterization ......................... 4
  1.3 Optimal survey design ............................................. 10
  1.4 Thesis objective and outline ..................................... 12

2 A model reservoir monitoring experiment ............................. 15
  2.1 Geophysics: seismic tomography .................................. 17
  2.2 Dynamics: tracer advection ....................................... 18
  2.3 Physical parameter relations .................................... 20
  2.4 Discretization ..................................................... 22
    2.4.1 Discretization of the tomography equation ............... 22
Table of Contents

2.4.2 Discretization of the mass balance .................................. 23
2.4.3 Discretization of the flux-balance conservation ................. 24
2.4.4 A Particle-In-Cell method for flow simulation .................... 25
2.5 Summary .......................................................................... 28

3 Error characterization and model estimation .......................... 30
3.1 Linear inversion .............................................................. 30
3.2 Coupled inversion and error characterization ......................... 34
  3.2.1 Exact dynamics ......................................................... 34
  3.2.2 Inexact dynamics ...................................................... 36
3.3 Summary .......................................................................... 37

4 Coupled inversion for velocity and initial slowness .................. 39
4.1 Flow constrained geophysical imaging .................................. 40
4.2 Discretization of regularization functionals ............................ 42
4.3 Numerical optimization .................................................... 44
4.4 Initialization ..................................................................... 49
4.5 Experimental setup .......................................................... 50
4.6 Sequential reconstruction as initialization ........................... 50
4.7 Joint reconstruction .......................................................... 54
4.8 Forecasting ..................................................................... 55
4.9 Summary .......................................................................... 56

5 Dynamic optimal experimental design ...................................... 57
5.1 Adaptive experimental design .............................................. 58
5.2 Design for dynamical systems ............................................. 64
  5.2.1 Exact dynamics ........................................................ 64
  5.2.2 Inexact dynamics ....................................................... 65
# Table of Contents

5.3 Numerical optimization ............................................ 66  
5.3.1 Exact dynamics .................................................. 66  
5.3.2 Inexact dynamics ................................................. 69  
5.4 Numerical examples .................................................. 70  
5.4.1 Regularization .................................................... 72  
5.4.2 Monitor function .................................................. 74  
5.5 Design results: exact dynamics .................................... 74  
5.6 Results: inexact dynamics .......................................... 78  
5.7 Summary ............................................................. 78  

6 Conclusions ........................................................... 83  

Bibliography ............................................................ 89
List of Figures

1.1 Model of a reservoir with minimal wells. Estimation of the permeability \( k(\vec{x}) \) using only pressure \( (p) \) and hydraulic head \( (z) \) data of a fluid moving with velocity \( (\vec{u}) \), collected from sparsely distributed wells will generate highly inaccurate results. .................... 3

1.2 Model of a reservoir with geophysical survey measurement locations depicted in orange. Changes to the geophysical model parameters \( (m(\vec{x})) \) can be estimated through inversion of the densely sampled geophysical data \( (d_i) \). ................................. 5

2.1 Tomography setting: sources \( (s_i) \) and receivers \( (r_j) \) are placed in two boreholes and the travel times of an acoustic wave traveling along adjoining ray paths \( (\Gamma_{i,j}) \) are measured. ...................... 17

2.2 Tracer advection: Tracer of concentration \( (c) \) is advected in a fluid of density \( (\rho) \) with velocity field \( (\vec{u}) \). The hydraulic conductivity \( (K) \) determines the velocity field given a pressure gradient within the reservoir. ............................... 20

2.3 A grid cell with length \( h_1 \) and height \( h_2 \), slowness value \( (m_j) \) located at the cell center, and components of the fluid velocity flowing in \( (\vec{u} = [u_1^+, u_2^+]) \) and out \( (\vec{u} = [u_1^-, u_2^-]) \) on the cell edges. .................... 24
List of Figures

2.4 Particle-In-Cell method for the advection and mass-conservation equation. A particle located in the cell-centered point $\mathbf{x}_A$ on the regular mesh is pushed forward with velocity $\mathbf{u}$ along the path $\mathbf{u}\Delta t$, to the non-grid point $\mathbf{x}_B$. The value (mass) associated with the particle is then transferred to the cell centered grid points adjacent to $\mathbf{x}_B$ using interpolation weights. ................................................. 27

4.1 Visualizations of the ground true and reference hydraulic conductivity models and associated flow fields. The true and reference models are discretized on a $100 \times 200$ cell grid on a domain of 100m depth by 200m across. A static source, at 55m depth on the left border, and a sink, at 65m depth on the right border of the domain, are indicated by red. The reference model is a simplified model which interpolates borehole data between the two wells. .......................................................... 51

4.2 Visualization of ground truth data for the domain with dimensions 100m (y direction) depth and 200m (x direction) across (first column), individual reconstruction (second column) and results of the joint reconstruction approach. The first row visualizes the initial slowness ($m_0$). The second shows a quiver plot of the recovered velocity field ($\mathbf{u}$). The components of the velocity field $[u_x, u_y]$ are pictured in the third and fourth row and the last row shows the curl of the velocity field ($\nabla \times \mathbf{u}$). Note that there is almost no flow in the y direction, and that at the boundaries of the reservoir there are non-zero values of the curl of $\mathbf{u}$ ............................................................ 52
4.3 The evolution of the tracer is pictured for the ground-truth (left column), individual reconstruction used for initialization (middle column), and the proposed joint reconstruction. The tracer evolution is simulated for 40 days with the respective initial slowness and flow field. The front of the ground truth tracer is visualized by dashed lines for the first 3 rows. Note that the joint method results match the ground truth tracer front to a greater degree than the predictions from the initialization estimates. ................................................. 53

5.1 Initial experiment setup. Tomography sources are pictured in green on the left and receivers in blue on the right. The initial setup covers the entire flow domain with 20 sources and 30 receivers. There is one source of fluid, pumped at a constant rate from the top green point, and a sink at the bottom (red) of the domain. Flow is in the downward direction. ................................................................. 71

5.2 Plot of the relative error ($\|m_1^{\text{true}} - \hat{m}_1\|/\|m_1^{\text{true}}\|$) versus the regularization parameter $\alpha$ for both the linear regularization and smoothed total variation. The TV regularization gives a better estimate of the model for the case of a sharp target. ................................. 73

5.3 Exact dynamics: plots of the adapted mean squared error amse vs. the number of nonzero weights (left column), and the weights used to conduct experiments 1, 3, and 5, at times $t_0$, $t_2$, $t_4$. ................................. 75

5.4 Exact dynamics: relative error plots. Model error is compared for each experiment for both the linear and TV regularization. ................................. 77
List of Figures

5.5 Exact dynamics: optimal designs and recovered models for 9 experiments. The top row shows the rays and the recovered model, with the number of rays ($\#d = 297$) given below, followed by the true models in the second row, the models recovered using total variation in the third row, and finally the models recovered using the linear gradient regularization. ................................................................. [79]

5.6 Inexact dynamics: relative error plots. Model error is compared for each experiment for both the linear and TV regularization ............... [80]

5.7 Inexact dynamics: optimal designs and recovered models for 9 experiments. The top row shows the rays and the recovered model, with the number of rays given below. The second row pictures true models, the third shows the models recovered using total variation, and finally the bottom shows models recovered using the linear gradient regularization. [82]
Acknowledgements

This thesis would never have been possible without the guidance, assistance and support of a large group of individuals.

First and foremost, I would like to express my deepest gratitude to my supervisor Eldad Haber for all of the guidance, advice, and arguments he has provided. I have learned a great deal working with you over the past 6 years, not only about computational geophysics, but also about myself. Toda raba.

During my PhD I had the privilege of working at IBM with Lior Horesh. Thank Lior for your kindness, encouragement, and for providing me with so many different explanations and perspectives on inverse problems and optimization. I am truly grateful for the time spent with you and your lovely family in New York.

And to Lars Ruthotto, the paper wouldn’t have happened without you. I learned so much working with you. Thank you.

Many thanks to all of my office mates and colleagues; particularly Klara and Luz for your much appreciated reading and editing of my thesis, and for lending an ear to many questions, concerns, and complaints! And also to Gudni, Sarah, Justin, Kris, Mike, Mike, Dikun, Rowan, Seogi, and Lindsey, for help, encouragement, editing, and fun!

A very special thanks to my dear friends Marianne and JF: for campus lunches, hikes, dogs, skiing, beer, wine, rowing, science, and games.

And finally, thank you to my best friend and partner Dave Marchant. For everything.
Acknowledgements

The funding of this work was generously provided from a variety of sources. Thanks to University of British Columbia, the Department of Earth and Ocean Sciences, the Society of Exploration Geophysics, the Natural Sciences and Engineering Research Council of Canada, and MITACS for their support.
To my parents.
Chapter 1

Introduction

Subsurface reservoirs are defined as regions of fluid saturated porous rock and are commonly depleted through fluid extraction. Maintenance of groundwater aquifers (a reservoir of water) or oil producing reservoirs as they are depleted, often requires efficient methods of monitoring fluid flow within the subsurface. The ability to track and predict the flow is important for both environmental and economical reservoir management. For example, in an oil recovery scenario, where a fluid such as water is injected into a depleted oil reservoir to enhance recovery, it is crucial to know where the oil and water are at a given time. Monitoring this process provides information crucial to future fluid management decisions. For example, decisions such as from which wells to extract, which wells to shut in, and from which wells to inject can affect production rates, fluid loss, and overall efficiency.

One particular aspect of reservoir management programs involves predicting or forecasting fluid flow within the subsurface. To do so, flow is simulated by building a mathematical model of the subsurface, often referred to as reservoir characterization.

1.1 Reservoir characterization

In most cases, reservoirs consist of isolated regions of porous rock surrounded by regions of non-porous rock. See Figure [1.1] below for a diagram. Thus, the dynamics of a reservoir are commonly governed by mathematical models describing fluid flow in a porous media, which come in many forms with varying degrees of complexity.
1.1. Reservoir characterization

Some of the more common models include: single phase flow, two-phase flow, black oil models, advection-diffusion models, and reactive contaminant transport models. For example, the model describing single phase flow in a porous media as presented in (Chen et al., 2006), given the spatial and temporal variables $\bar{x}$ and $t$, is as follows

$$\frac{\partial (\phi \rho)}{\partial t} = \nabla \cdot (\rho \bar{u}) + q, \tag{1.1a}$$
$$\bar{u} = -\frac{k}{\mu} (\nabla p - \rho g \nabla z). \tag{1.1b}$$

Equation 1.1a is derived by tracking mass conservation of a fluid with viscosity $\mu$ flowing through a rectangular three dimensional differential volume, see (Chen et al., 2006) for the complete derivation. The porosity of the rock (the fraction of a representative elementary volume available for fluid) is denoted by $\phi$, and the density of the fluid by $\rho(t, \bar{x})$. The velocity of the fluid $\bar{u}(\bar{x})$ is defined as the superficial Darcy velocity, and described by Equation 1.1b, the conservation of momentum equation also known as Darcy’s law. The flow in this model is generated by a pressure gradient, $\nabla p(t, \bar{x})$, resulting from the sources and sinks, $q$, and from a gravitational gradient, $\rho g \nabla z$, where $\nabla z$ is the difference in fluid heights, and $g$ is the magnitude of the gravitational acceleration.

If there is no compressibility in the rock then $\frac{\partial (\phi \rho)}{\partial t} \to 0$, such that no fluid is stored within the rock pore space. Additionally, the velocity depends on the permeability tensor $k$ of the porous rock and the fluid viscosity $\mu$. The permeability is an average property of the reservoir that measures the ability of a porous media to transmit fluid. In most cases, $k$ is a tensor, unless it is assumed that $k(\bar{x})_{11} = k(\bar{x})_{22} = k(\bar{x})_{33} = k(\bar{x})$. In this case the porous media is known as isotropic, where the flow is the same in all directions. It is otherwise anisotropic.

In order to simulate flow using a reservoir model and solve for the pressure $p$, the permeability $k(\bar{x})$, often referred to as a reservoir parameter function, must be
1.1. Reservoir characterization

Figure 1.1: Model of a reservoir with minimal wells. Estimation of the permeability \( k(\vec{x}) \) using only pressure \( p \) and hydraulic head \( z \) data of a fluid moving with velocity \( \vec{u} \), collected from sparsely distributed wells will generate highly inaccurate results.

known. However, this function is often unknown and thus must be estimated from borehole rock samples. In most cases boreholes are expensive to drill, and are thus sparsely distributed throughout the reservoir. This leads to inaccurate estimates of \( k(\vec{x}) \) for the entire region, as values must be extrapolated between boreholes over large distances (Roubinet et al., 2013).

One approach to alleviating this problem is to use historic flow data, measurements of the pressure \( p \) and hydraulic head \( z \), to estimate the unknown parameter function \( k(\vec{x}) \). This process is alternately known as parameter estimation, inversion, and in the oil and gas industry is referred to as history matching (Oliver and Chen, 2010; Oliver et al., 2008). The basic idea of history matching is to estimate the reservoir parameters such that simulated flow data fits measured data.

Technically, the inversion process is rather challenging as it involves the discretization of a system of partial differential equations, the solution of the forward problem (flow simulation), the computation of the gradients of the simulator with respect to the parameters, and the solution of an optimization problem. An excellent review of
1.2. Geophysics for reservoir characterization

Although it is possible to estimate the reservoir parameters through inversion and predict flow data, these estimates can be a highly inaccurate representation of the reservoir. This inaccuracy stems from the large null space of the inverse problem associated with the highly sparse spatial sampling of reservoir flow data.

To overcome the null space issue subspace techniques have been applied, where the reservoir parameters are restricted to “live” in a small subspace spanned by a (relatively) small number of vectors \( \text{Abacioglu et al., 2001; Gerritsen and Durlofsky, 2005; Oliver and Chen, 2010; Sarma et al., 2013} \). This approach decreases the variance in the recovered estimates of \( k(\mathbf{x}) \) by increasing the bias of the estimated reservoir parameters \( \text{Tenorio, 2001} \).

A different approach to reduce the variance of the recovered reservoir parameter models is to simply add additional flow data and further sample the reservoir. While this is clearly one of the better ways to reduce the uncertainty in the estimates, it is not practical as it is unlikely that many more wells will be drilled just to improve the simulation capability. However, assuming that the geophysical rock properties of the reservoir (density or electrical conductivity for example) will change with fluid flow, implies that geophysical survey techniques can potentially be used to map these changes. This makes geophysics for reservoir characterization a practical method to increase data collection coverage without having to drill additional wells.

1.2 Geophysics for reservoir characterization

Geophysical survey techniques measure variation in the geophysical properties of the subsurface by systematically collecting geophysical data over a large spatial surface grid and within existing boreholes. Geophysical data are not only valuable on their own for detecting variation in the subsurface, but can also be inverted to obtain an
1.2. Geophysics for reservoir characterization

Image of the underlying geophysical property distribution. Thus, geophysical data can complement reservoir flow data by providing a denser grid of measurements than from subsurface flow data obtained from point locations at well sites alone (Hubbard and Rubin, 2000). For example, Figure 1.2 provides an example of a geophysical survey with data, \( d_i \), collected at the surface (orange dots). Sensors could additionally be placed in the existing boreholes to further increase data numbers and depth of investigation.

Figure 1.2: Model of a reservoir with geophysical survey measurement locations depicted in orange. Changes to the geophysical model parameters \((m(\bar{x}))\) can be estimated through inversion of the densely sampled geophysical data \((d_i)\).

The spatial resolution and depth of investigation of geophysical data depends predominantly on the choice of technique. Applied geophysical surveys are a widely used method of imaging the subsurface in resource exploration. Therefore, there are many well-developed survey types available for imaging subsurface flow. In particular, time-lapse seismic surveys, ground penetrating radar, time-lapse gravity measurements, and electromagnetic imaging techniques have been used for oil reservoir and aquifer characterization for some time (Alumbaugh and Morrison, 1995; Archie, 1942; Hubbard and Rubin, 2000; Lumley, 2001a; Vasco et al., 2004; Vesnavaer et al., 2003).
1.2. Geophysics for reservoir characterization

As an example, gravity methods, which are sensitive to shallow changes to the density of the subsurface rock, are a common choice in hydrology applications. Water aquifers tend to be small and close to the surface and thus gravity methods can detect small variations in density as water is depleted, see for example (Alumbaugh and Morrison, 1995; Slater et al., 2000; Wilt et al., 1995). In this case measured geophysical data consist of small changes to the local gravitational field above the aquifer.

Oil reservoirs in comparison, are often large, flat, not necessarily close to the surface, and commonly discovered using seismic techniques. Thus seismic surveys are frequently the first choice for monitoring programs (Lumley, 2001b). Additionally, electromagnetic methods are attractive for monitoring injected fluids because the electrical conductivity of the injected fluid can be controlled. For example, the injected fluid can be selected such that it’s conductivity maximizes the conductivity contrast between the existing media and the introduced fluid. This then generates a greater signal in the measured electric and magnetic potential data.

Although geophysical surveys are an excellent method for imaging changes in the geophysical properties of the subsurface, inferring reservoir parameter functions from geophysical data or recovered images of the geophysical property models is not so straightforward. The use of geophysics for inferring reservoir parameters, sometimes known as geophysical history matching, rests on the assumption that a mathematical relationship (petrophysical relationship) between the geophysical properties and reservoir parameter function(s) exists and is known. Historically, much work has been done to determine these relationships (Archie 1942; Mavko et al. 2009). Although these relations tend to be empirical and site specific, they have been successful in progressing geophysical history matching workflows.

A second point of concern in the geophysical history matching process is that
most often the mathematical geophysical model and the reservoir fluid model are
decoupled. Thus, the inversion process is decoupled. This decoupled inversion process
has a major shortcoming; since geophysical imaging is almost always ill-posed, the
accuracy of the imaging is typically low. Estimating the reservoir model parameters
from the geophysical images alone, or in combination with low resolution models
obtained from inverting fluid flow data, can yield inaccurate and biased estimates.

To address this shortcoming, (Hubbard and Rubin 2000) propose several parameter
estimation approaches using statistical geophysical-hydrological techniques to
characterize an aquifer with geophysical and hydrological data. One particular tech-
nique takes a Bayesian iterative approach where the flow log-permeability (a com-
monly measured reservoir parameter) is first estimated through inversion of sparse
hydrological data. This estimate is considered to be a random variable with a proba-
bility distribution, and is then used as a prior distribution for the Bayesian inversion
of geophysical data. Geophysical data along with the known parameter relationships
are then inverted to recover an update to the log permeability while including the
prior permeability estimate. Alternately, several geostatistical approaches have been
proposed to match structures within images of differing parameters to geologic data
from core samples, or maximize correlations. These approaches suffer the same pa-
rameter function requirements and lower resolution biased estimates (Cassiani et al.,

In oil reservoir monitoring practice, seismic history matching is the predomina-
t method (Abul, 2010; Emerick and Reynolds, 2012; Gosselin et al., 2003; Lumley,
2001b; Mezghani et al., 2013; Nenna et al., 2011; Oliver and Chen, 2010; Sarma et al.,
2013; Trani, 2012; Vasco et al., 2004). Similar to the geophysical-hydrological tech-
niques, seismic history matching involves estimating geophysical properties through
inversion of geophysical data, mapping the estimated geophysical property model to
a reservoir parameter model through a physical property relation, and finally simu-
1.2. Geophysics for reservoir characterization

lating fluid flow data and matching it with observed flow data collected. If the data match is poor, either the reservoir parameter model is adjusted, or new geophysical inversions are carried out and the process is repeated. If the simulations using the reservoir model predict the historic flow data then forecasts can be generated with some certainty. In general, the seismic history matching process suffers as it requires parameter mappings and the mathematical models are decoupled in the inversions.

The concept of coupling the dynamic reservoir model with a geophysical model in order to gain a better estimate of the unknown reservoir parameters is not new. For example, the Kalman filter was developed in the early 1960’s to incorporate new observation data into the estimation of the position of a moving object and became a common method for applications in guidance, navigation and control of vehicles, particularly aircraft and spacecraft.

Classical Kalman filtering is a Bayesian inversion method that iteratively uses a time-series of noisy observed measurements to estimate unknown model parameters. If the noise is assumed to be characterized by a Gaussian distribution then the filter yields the exact conditional probability following Bayes rule (recursive Bayesian estimation), which can also be interpreted as the solution to a linear regularized inverse problem (Kalman, 1960). In the context of geophysical history matching, Kalman filtering is derived from an inverse problem that simultaneously minimizes the geophysical measurement error and the error in the flow model simulation.

Kalman filtering, and ensemble Kalman filtering (EnKF) (Evensen, 1994), have recently been proposed as alternative approaches to seismic history matching as they couple the reservoir model and geophysical model (Abul, 2010; Nævdal et al., 2002; Nenna et al., 2011; Vauhkonen et al., 1998). The Kalman filter (also known as the Kalman gain matrix) couples the geophysics, the dynamics, and the prior error covariance matrices. However, its use for geophysical imaging of flow requires a physical parameter relationship. Additionally, the large dense covariance matrix which
1.2. Geophysics for reservoir characterization

characterizes the prior distribution must be known, and is propagated in time. For
problems such as seismic imaging, where the data set and the parameter space are
large, Kalman filtering becomes computationally infeasible as a result of the storage
and computation with the dense covariance matrices. To alleviate this, stochastic
methods such as the EnKF were developed to approximate the covariance matrices.

The EnKF method is a Monte Carlo type sequential Bayesian inversion method
first proposed by Evensen (1994) to reduce computations and memory. The EnKF
method introduces ensembles of state (parameter models) and measurement vectors
(geophysical data for example) sampled from Gaussian distributions such that the
prior state covariance matrix can be approximated by computing the covariance of
the ensemble set. This reduces the inverse problem to the dimension of the ensem-
ble. Abul (2010) found that there is promise in EnKF methods for reservoir history
matching, however, it was limited by uncertainty in the initial condition and error
estimates, and by the computation required for integration of large amounts of da-
ta and the non-linearity introduced through the relationship between reservoir and
seismic parameters.

In a similar flavour to Kalman filtering, there has been some recent research
to couple non-linear geophysical imaging with flow models in the hydrogeophysics
community through regularization of the geophysical inverse problem with the flow
models, with promising results (Cockett and Haber, 2016; Hoversten et al., 2006;
Steklova and Haber, 2015). These approaches are different in that they were developed
for specific flow models, require Archie’s law, are not solved iteratively, and have
specifically tailored regularization terms. Additionally, the inversions are carried out
for all times, therefore incorporating all historic data in one inversion.

In light of current reservoir monitoring practice, there is room to improve up-
on coupled geophysical imaging and subsurface flow inversion methods for reservoir
monitoring.
1.3 Optimal survey design

In addition to coupling geophysical survey methods to estimate flow model parameters, there is reason to consider optimizing the geophysical survey design to adapt to the flow dynamics, such that the best estimates of the reservoir parameters are obtained from a coupled inversion.

1.3 Optimal survey design

While the coupled flow and geophysical inverse problem can potentially provide better estimates of reservoir parameter functions through inversion of large geophysical data sets, the geophysical data collection can also be designed in an optimal way. For example, if one is interested in monitoring the progression of a fluid front as it moves through a reservoir, collecting geophysical data far downstream may not contribute any additional information to the inversion, and thus those measurements could be excluded. In the context of reservoir monitoring, an optimal experimental design might amount to adding or eliminating a well location as demonstrated by Gharamti et al. (2015), or reducing the number of surface measurements from a maximum allowable set.

In many realistic scenarios, the flow dynamics are sufficiently slow to allow for improved designs after a first experiment has been conducted. This is particularly true when monitoring flow in reservoirs where fluid velocities can be so slow that changes occur over months or even years (Vasco et al., 2004). Additionally, the available historic experimental data and inversion results will contain valuable information about the subsurface. This information should thus be included when determining an optimal survey design for a future survey.

The mathematical method of determining an experimental design is known as optimal experimental design (OED), and has existed for some time for both the well-posed and ill-posed parameter estimation cases (Aggarwal et al., 2015; Ajo-Franklin, 2015).
1.3. Optimal survey design

Typically, optimal experimental designs are computed by minimizing the error in the estimated parameter model according to some specified criteria. In this case the model estimate is assumed to be the mean of a distribution with a covariance matrix characterizing the error. The type of design method depends on the optimality criteria applied to the covariance matrix. There are several standard optimality criteria; for example, A-optimal design minimizes the trace of the covariance matrix or average variance, D-optimality minimizes the determinant of the covariance matrix, and E-optimality minimizes the eigenvalues of the inverse covariance matrix.

There are some optimal design methods which consider dynamic systems, yet there is little in the way of optimal design algorithms for ill-posed inverse problems where the model is governed by a dynamical system. Although there has been work on sequential design for the well-posed case, no work is known for the ill-posed case, (see the following for examples: (Chernoff, 1972; Khodja et al., 2010; Papalambros and Wilde, 2000; Wilkinson et al. 2015) and references within). In some cases, static design methods can be applied to time dependent partial differential equations (PDEs) (Alexanderian et al., 2014; Haber et al., 2011), this approach results in an a-priori estimate of the designs for “all times” which depend only on information provided in the form of a regularization (or prior) to the ill-posed problem. For example, (Alexanderian et al., 2014) compute optimal designs for a tracer advection-diffusion reservoir model. This example however, did not include any geophysical imaging, nor was any historic data included in the experimental design computation.

There has been some work applying optimal design methods to the Kalman filter (Gharamti et al., 2015; Sagnol and Harman, 2014); for example, Sagnol and Harman (2014) propose a D-optimal design method for the Kalman filter, which results
in designs dependent on the posterior state estimation covariance matrix. Alternately, Gharanti et al. (2015) present an optimal method for selecting aquifer well locations to track the flow of a contaminant plume. In both cases, the prior covariance matrix to be minimized includes the transition matrix (a discretization of the dynamics), but the designs still do not depend on historic data, and thus these approaches are no different then the application of static methods.

In fact, as far as can be discerned from the literature, apart from sequential design approaches (Chernoff, 1972; Khodja et al., 2010), which use the posterior model estimate and its posterior covariance as the prior for the design at the next time step, current design techniques for linear inverse problems do not utilize information that is collected at early times in order to design experiments at later times. This is a consequence of optimizing over the posterior variance in the model estimates.

For any linear ill-posed inverse problem, static or dynamic, the posterior covariance matrix will only depend on the physics of the problem, and the linear regularization operator (or prior distribution covariance matrix for the Bayesian perspective), regardless of how the inverse problem is formulated. Thus, in order to include historic data to a standard optimal design problem, the posterior covariance matrix must be reconsidered.

### 1.4 Thesis objective and outline

The main objective of the research presented in this thesis is to improve reservoir monitoring and forecasting by coupling geophysical survey techniques with a dynamic fluid flow reservoir model and optimally designing the data collection in the geophysical survey.

Coupling the geophysics and flow model in one inverse problem allows for the estimation of fluid parameter functions such as the permeability $k(\mathbf{x})$ from geophysical
data. To address the coupling problem, careful consideration of the choice of the reservoir fluid flow model and its discretization is taken such that knowledge of the physical parameter relationship is not necessary. Additionally, error in the reservoir model is assumed to be zero. This is contrary to the usual assumptions. In this way, the geophysical inversion problem will be reformulated as an ill-posed inverse problem with appropriate regularization terms that is constrained by the fluid flow model. Since the geophysical survey provides a much denser sampling of data, the uncertainty in the recovered flow model parameter function is reduced and thus provides better accuracy when forecasting flow.

Optimally designing the geophysical survey for the coupled problem, while including historic survey data, produces a survey that adapts with the dynamics of the reservoir while reducing the number of measurements required. This research builds on A-optimal sparsity constrained design algorithms presented in [Haber et al., 2011] for large-scale ill-posed inverse problems, by introducing a monitor function to the design optimization problem to include historic data.

The main body of the thesis consists of five chapters. Since the same geophysical survey and flow model are used to demonstrate the ideas in the proposed research, the thesis first outlines these models, and then presents the new ideas.

To this end, Chapter 2 introduces the seismic tomography geophysical imaging survey and the tracer advection flow model chosen to represent a simplified reservoir monitoring scenario, presents a mathematical argument for not requiring a physical property relationship between models, and details the discretization of the two model PDEs.

Chapter 3 outlines the decoupled linear inversion process and discusses the issues with coupling the two differing physical models. In particular, consideration must be taken when characterizing the error in each of the models, as this leads to very
different formulations of the coupled inverse problems.

Chapter 4 presents the seismic tomography tracer advection coupled inverse problem formulated to solve for the fluid velocity field and initial tracer concentration within the reservoir. Because of the ill-posed nature of the problem, special consideration of the regularization terms required to solve the problem are discussed. In particular this chapter highlights the construction of a unique regularization function for the fluid velocity field. Due to the non-linearity of the regularization functions, solutions to the inverse problem are not so straightforward. This chapter therefore details the discretization of the regularization and optimization of the inverse problems, and finally presents results for the tomography tracer advection model. The objective of the example is to show that estimates of the initial tracer distribution and fluid velocity field are improved as new geophysical data are obtained, and that forecasting improves as a consequence.

Chapter 5 details the formulation of the adaptive A-optimal experimental design method for two cases of the error characterization in the dynamic fluid flow model. The new formulation proposes the addition of a monitor function in the definition of the posterior covariance matrix of the estimated model parameters. The non-linearity of the resulting optimization problem requires iterative solution techniques, which are outlined in Section 5.3. The adaptive design method is then demonstrated for the coupled seismic tomography tracer advection reservoir monitoring example for both cases of the error characterization. The results show that experimental designs computed using this method adapt to the flow within the reservoir while still providing sufficient data for optimal recovery of tracer distributions.

Chapter 6 summarizes the key findings of the research presented in this thesis, discusses the practical contributions, and comments on areas of future work.
Chapter 2

A model reservoir monitoring experiment

To demonstrate the research ideas presented in this thesis a simple model of a reservoir monitoring scenario is used throughout. The model consists of a seismic tomography survey to image the reservoir, and a tracer advection flow model to characterize the subsurface fluid flow dynamics.

Although there are several choices when considering geophysical surveys for monitoring subsurface flow, such as gravity surveys or electrical methods, borehole seismic tomography surveys provide a good starting point for monitoring a reservoir. In particular, the tomography survey utilizes already existing boreholes for source and receiver locations, while receivers can be placed on the surface to expand the measurement area. It has also been established that seismic wave velocity is sensitive to changes in rock fluid properties (Alumbaugh and Morrison, 1995; Nolet, 1987; Oliver and Chen, 2010), and thus the data are sensitive to flow. An additional benefit is the linearity of the governing equations. The measured tomography data are a linear mapping of the subsurface rocks seismic properties, which can be inverted to recover a spatial image of the properties. Thus imaging the seismic wave velocity involves solving a fairly straightforward linear inverse problem.

There are several well established and researched models describing flow within a reservoir, such as two-phase flow models or a black oil model, among many others.
Chapter 2. A model reservoir monitoring experiment

See (Chen et al., 2006) for an excellent outline of flow models. However, many of these models are complicated and can be discretized with a limited number of techniques with time step and spatial discretization length restrictions. Since the goal is to couple the flow model and the geophysical model to form an already computationally expensive ill-posed inverse problem, it is not beneficial at this stage of proof of concept to incorporate more physically accurate, and computationally difficult models.

Although the tracer advection model might seem like an unlikely choice for simulating flow within a reservoir, there is some justification; the tracer itself can be thought of as an analogy to the second phase (a differing fluid than the already present fluid) moving in a reservoir and mathematically resembles two-phase flow models (Chen et al., 2006). In reality, the tracer advection model is a single phase flow model. Also, at very long time scales, where an injected fluid front might take months to move even a few meters, it is not necessary for the purposes of this research to incorporate a model that characterizes minute changes. Additionally, in hydrology, tracer experiments are common practice for estimating aquifer parameters such as hydraulic conductivity, and have been used in conjunction with seismic tomography surveys in the past (Hyndman et al., 1994).

Practically, from a computational perspective, the tracer advection models linearity lends itself to a variety of discretization techniques, including unconditionally stable Lagrangian particle in cell methods. A second benefit becomes clear in Section 2.3 when discussing physical parameter relationships and coupling the flow model with the geophysical imaging.

This chapter introduces the governing partial differential equations (PDEs) for the tomography survey experiment and tracer advection model, discusses the relationship between the different physical properties of the two PDEs, and finally presents a discretization of the resulting governing system of equations.
2.1 Geophysics: seismic tomography

Classical travel time tomography uses ray paths to model travel times and so assumes that the data are of high frequency, meaning the wavelength is at least several times smaller than the spatial variations of the velocity model (Yilmaz, 2013).

A typical 2D seismic tomography experiment places $n_s$ sources on one side of the region to be imaged and $n_r$ receivers on the other. Travel times of an acoustic wave traveling from source to receiver constitute the measured data, where the ultimate goal is to recover an image of changes to the distribution of physical properties generating changes in these data. A schematic of the seismic experiment is illustrated in Fig. 2.1.

![Figure 2.1: Tomography setting: sources ($s_i$) and receivers ($r_j$) are placed in two boreholes and the travel times of an acoustic wave traveling along adjoining ray paths ($\Gamma_{i,j}$) are measured.](image)

Mathematically, the travel time data are calculated by integrating the inverse of the acoustic wave velocity, also know as the slowness, $m(t, \vec{x})$, over a ray path $\Gamma_{i,j}$. 
2.2. Dynamics: tracer advection

The travel time along the ray path from the $i$th source to the $j$th receiver is given by

\[ d_{i,j}(t) = \int_{\Gamma_{i,j}} m(t, \bar{x}) d\ell. \]  \hspace{1cm} (2.1)

where $m \in \mathcal{M}$ is the slowness model belonging to the subspace, $\mathcal{M}$. Assuming a layered earth with homogeneous isotropic porous rock layers, then the ray paths the waves follow will be linear and independent of the slowness (this is true for small perturbations in the slowness field). Thus, the problem can be cast as a linear inverse problem for $m(t, \bar{x})$ \cite{Jones2010}.

Given noisy data vectors $d(t)$ measured at times $(t_0, \ldots, t_n)$ Equation (2.1) can be written as,

\[ \mathcal{F}m(t_k, \bar{x}) + \epsilon_k = d(t_k), \]  \hspace{1cm} (2.2)

where $d$ is a $n_s \times n_r$ vector of real numbers, $\mathcal{F} : \mathcal{M} \to \mathbb{R}^{n_s \times n_r}$ is the forward mapping operator, and measurement error $\epsilon_k$. The receivers, or modern seismometers, measure tiny displacements in the earth. These instruments use electronic sensors, amplifiers, and recording devices, which can cover a wide range of frequencies. In general, the first arrival time and initial displacement of earth is measured successively many times over a very short interval and stacked such that the measurement is given as a continuous random variable, with error $\epsilon_{i,j} \pm \sigma_{i,j}$.

2.2 Dynamics: tracer advection

The tracer advection equations as presented in \cite{Chen2006}, describe the transport of a solute in a fully saturated fluid phase with the assumption that fluxes due to dispersion and diffusion are small relative to the advective transport of the solute. The motion of a tracer with conserved concentration $c$ (volumetric fraction in
the fluid phase), in a fluid of density \( \rho(t, \vec{x}) \), with spatial and temporal variables \( \vec{x}, t \) respectively, is described by the following relation,

\[
\frac{\partial (\rho c)}{\partial t} + \nabla \cdot c \rho \vec{u} = 0, \quad (2.3a)
\]

s.t. \( c(0, \vec{x}) = c_0(\vec{x}) \),

where the flow field \( \vec{u} \) satisfies

\[
\nabla \cdot \vec{u} = q, \quad (2.3b)
\]

\[
\vec{u} = -K(\vec{x})\nabla p, \quad (2.3c)
\]

\[
p(0, \vec{x}) = p_0(\vec{x}). \quad (2.3d)
\]

accompanied with either Dirichlet or Neumann boundary conditions for the pressure, \( p \). The fluid velocity field \( \vec{u} \) is determined by Darcy flow in this approximation, where only the hydraulic conductivity \( K = k(\vec{x})/\mu \), a function of the permeability \( k \) and fluid viscosity \( \mu \), and pressure gradient describe the flow field. Recalling Equations (1.1b), the gravitational term has been excluded on the assumption that it is very small in comparison to the pressure gradient. This implies that pumping rates are high enough to generate significant pressure within the reservoir. Note that Equation (2.3b) implies that what flows in flows out within the entire domain, such that the source pumping rate in is equal to the pumping rate out, \( q_{in} = q_{out} \). Thus the rock is assumed to be incompressible, such that no fluid is stored. The tracer advection model is illustrated in Figure 2.2.

In a typical reservoir characterization setting one would collect pressure and concentration measurements and invert these data to recover \( K(\vec{x}) \). However, assuming the presence of the tracer generates a change in the seismic wave velocity, then the idea is to first estimate \( \vec{u} \), and later estimate \( K \). To do this, the petrophysical parameter
2.3 Physical parameter relations

In order to use geophysical techniques to estimate fluid flow parameters through inversion of geophysical data, there must be a relationship between the flow parameters and the geophysical rock properties, also known as a petrophysical relationship. One famous petrophysical function, known as Archie’s law, relates the electrical conductivity of a sedimentary rock to its porosity and brine saturation (Archie 1942). In this way, changes in saturation will result in changes to the electrical conductivity, a property that governs the flow of electric current in the presence of an electric field. However, Archie’s law is an empirical relation that is site dependent and requires significant laboratory work which cannot be generated for the general case (Archie, 1942).
Many empirical relations between seismic velocities and rock porosity have also been developed (Abul, 2010; Mavko et al., 2009). However, these are also rock and site specific, and are often highly nonlinear.

For the seismic tomography experiment it may be that in the presence of the tracer, the seismic slowness changes relative to areas where there is tracer present. Thus, it is assumed that the tracer $c = c_0$ and the seismic slowness $m$ are related such that petrophysical relationship $m(c)$ exists, however, given the choice of flow model it is not necessary to know $m(c)$.

To see this consider the case where the flow is divergence free such that $\nabla \cdot \bar{u} = 0$. The fluid velocity model, Equation (2.3b), typically has a very sparse right hand side with delta functions for injection and extraction points. Therefore, the flow field is divergence free almost everywhere within the reservoir.

Recalling Equation (2.3a), the equation describing the advection of the tracer is

$$\frac{\partial (c)}{\partial t} + \nabla \cdot \bar{c} \bar{u} = 0.$$ 

Now, assuming that the only source of change to the subsurface rock comes from the transport of the tracer, then the “transport” of the seismic slowness is governed by the same conservation law as follows,

$$\frac{\partial m(c)}{\partial t} + \nabla \cdot (m(c) \bar{u}),$$

applying the chain rule to the first term and expanding the divergence term yields,

$$\frac{\partial m(c)}{\partial c} \frac{\partial c}{\partial t} + \bar{u}^T \nabla m(c) + c \nabla \cdot \bar{u}.$$
2.4. Discretization

Expanding the gradient term and recalling the assumption $\nabla \cdot \vec{u} = 0$ gives,

$$\frac{\partial m(\hat{c})}{\partial \hat{c}} \frac{\partial \hat{c}}{\partial t} + \vec{u}^\top \frac{\partial m(\hat{c})}{\partial \hat{c}} \nabla \hat{c},$$

and finally factoring out $\frac{\partial m(\hat{c})}{\partial \hat{c}}$ while recalling that Equation (2.3a) is equal to zero,

$$\frac{\partial m(\hat{c})}{\partial \hat{c}} \left( \frac{\partial \hat{c}}{\partial t} + \vec{u}^\top \nabla \hat{c} \right) = \frac{\partial m(\hat{c})}{\partial \hat{c}} (0) = 0$$

Thus, it can be assumed from this result that the transport equation describing the flow of the tracer can also be used for the “transport” of the slowness. This feature of the tracer advection model is particularly advantageous as it eliminates the necessity of knowing the petrophysical relationship between the tracer and seismic velocity.

Given the two models, the discretization is now outlined in the following sections.

2.4 Discretization

The numerical discretization of the geophysical imaging and the flow dynamics is presented in three parts: the discretization of the tomography experiment (Equation (2.2)), the tracer advection (Equation (2.3a)), and the fluid conservation law (Equation (2.3b)).

For simplicity, the discretization is restricted to a two-dimensional setting where the computational domain $\Omega = [0, L_1] \times [0, L_2]$ is rectangular, and divided into $l_1$ by $l_2$ rectangular cells of edge length $h_1 = L_1/l_1$ and $h_2 = L_2/l_2$.

2.4.1 Discretization of the tomography equation

To discretize the tomography experiment with discrete data $d_k$ measured at time $k = 1, \ldots, n$ the line integrals in (2.1) are approximated by summing over line segments.
2.4. Discretization

That is, the approximation of the integral in Equation 2.1 for the travel time $d_{i,j}$ of a wave traveling along the ray path $\Gamma_{i,j}$ between the $i^{th}$ source and the $j^{th}$ receiver is

$$d_{i,j} = \int_{\Gamma_{i,j}} m(t, \vec{x}) dl \approx \sum_{k=1}^{p} m_k d_{l_k},$$

where $p$ is the number of cells that $\Gamma_{i,j}$ intersects in the mesh, $m_k$ is the value of the slowness at the cell center of the $k^{th}$ cell, and $d_{l_k}$ are the line segments of $\Gamma_{i,j}$ passing through each cell respectively.

Thus, the discrete tomography experiment can be written as a linear system

$$Fm_k + e_k = d_k; \quad k = 1, \ldots, n,$$

where the entry of $F_{ij}$ is the length $d_{l_k}$ of the intersection of $\Gamma_{i,j}$ through the $k^{th}$ cell. The cell centered values of the grid are stored in the vector $m$.

The operator $F : R^M \rightarrow R^N$, maps from the model space to the data space $N = ns \times nr$ (number of sources $\times$ number of receivers), and thus $d_k$ is an $N \times 1$ vector of discrete travel time measurements, and $m$ is an $M \times 1$ vector of seismic slowness values with units $s/m$.

2.4.2 Discretization of the mass balance

Since vector operators such as the curl and divergence will be used later, a Marker and Cell (MAC) (Fletcher 2012) grid is chosen for the discretization of the fluid velocity vector field $\vec{u}$ and the initial seismic slowness $m_0$. The MAC method is discretized on an Eulerian staggered grid system, and is a finite difference solution technique for investigating the dynamics of an incompressible viscous fluid. One key feature is the use of Lagrangian virtual particles, whose coordinates are stored, and which move
2.4. Discretization

Figure 2.3: A grid cell with length $h_1$ and height $h_2$, slowness value $(m_j)$ located at the cell center, and components of the fluid velocity flowing in ($\bar{u} = [u^+_1, u^+_2]$) and out ($\bar{u} = [u^-_1, u^-_2]$) on the cell edges.

from a cell to the next according to the latest computed velocity field (McKee et al., 2008).

Figure 2.3 pictures a cell in the grid where $\bar{u} = [u_1, u_2]$ is discretized on the cell faces, and $m$ at the cell centers using a cell-centered grid function $m_0$. This yields an approximation of the divergence in the cell-centers using short differences.

2.4.3 Discretization of the flux-balance conservation

A finite volume discretization of the flux-balance equation $\nabla \cdot \bar{u} = q$ on a staggered grid is obtained by using the divergence theorem;

$$\int_{V_{cell}} \nabla \cdot \bar{u} \, dV = \int_{S_{cell}} \bar{u} \cdot \bar{n} \, dS,$$

where, $V_{cell}$ is the volume of the cell, and $S_{cell}$ is the cell surface area. The vector $\bar{n}$ is the unit normal vector to the surface. See (Haber and Ascher, 2001) for further details.

Referring to Figure 2.3 the divergence over a cell with dimensions of $h_1 \times h_2$ in
2.4. Discretization

The mesh can be expressed as

\[(\nabla \cdot \bar{u})_{\text{cell}} \approx \frac{h_2(u_1^+ - u_1^-) + h_1(u_2^- - u_2^+)}{h_1 h_2},\]

where the \(u_i^+\) is the magnitude of the flux into the cell and \(u_i^-\) is the magnitude of the flux out of the cell in the \(i\)th coordinate direction. This discretization leads to the following standard discretization of the divergence

\[\text{DIV} \ u = q, \quad \text{where} \ \text{DIV} = \left( I_{l_2} \otimes d_{l_1+1}^{l_2} \quad d_{l_2+1}^{l_1} \otimes I_{l_1} \right), \quad (2.5)\]

where \(\otimes\) is the matrix Kronecker product, \(I_l \in R^{l \times l}\) denotes the identity matrix, \(l\) is the number of cells in a respective direction, and where \(d_{l+1}^{l_2} \in R^{l_2 \times l}\) is a short finite difference matrix; see (Modersitzki, 2009) for more details about the implementation.

The discrete divergence operator maps from faces of the mesh to cell-centers.

2.4.4 A Particle-In-Cell method for flow simulation

The tracer advection equation (Equation (2.3a)) can be discretized in a number of ways. One option is to use explicit Eulerian techniques such as up-winding, Lax-Wendroff or Lax-Friedrichs (Ascher et al., 2006). However, explicit techniques suffer from one main shortcoming; since the velocity of the equation is unknown, one has to monitor the time step making sure stability is maintained. Monitoring and adjusting the time step can add enormous complexity to the optimization technique used if attempting to estimate the velocity. Methods such as up-winding and other flux limiters have another shortcoming; they are not differentiable with respect to the velocity and therefore simple optimization techniques can run into difficulties. Alternately, one has the option of using implicit methods which require the solution of linear systems at every time step, or using semi-Lagrangian techniques.
2.4. Discretization

Semi-Lagrangian techniques are particularly attractive for this application as they are explicit, unconditionally stable, and can be easily made differentiable. Semi-Lagrangian methods can suffer from low accuracy, however for this application where the velocity field is approximated and thus inaccurate, working with highly accurate discretizations is usually computationally wasteful. Therefore a Particle-In-Cell (PIC) method, that can also be interpreted as a semi-Lagrangian technique, is applied for the discretization.

The idea of using PIC methods for the solution of conservation laws is not new and was proposed in [Evans and Harlow (1957)]. Recently, PIC methods have emerged in computer graphics [Edwards and Bridson 2012] and in flow in porous media [Roubinet et al. 2013].

To use PIC methods, a particle $x_j$ is associated with the midpoint of the $j$th grid cell. The particle is assigned a value $m_j = m_0(\mathbf{x}_j)$, and the advection of the particle $x_j$ is described by the following ordinary differential equation (ODE)

$$\frac{\partial x_j(t)}{\partial t} = \mathbf{u}(x).$$  \hspace{1cm} (2.6)

Using a midpoint quadrature rule, the particles position after the first time step is approximated by

$$x_j(t_1) = x_j(0) + \Delta t(A^c_j u)_j + O(\Delta t^2),$$  \hspace{1cm} (2.7)

where $A^c_j$ denotes an averaging operator from cell faces (where the velocity resides) to cell centers. Thus, the position of the particle at a later time is easily computed given a time step and velocity field. Since the new position will in general not be the mid-point of a cell, the value $m_j$ is distributed to the closest surrounding points as illustrated in Figure 2.4. To compute the distribution weights, bilinear basis functions (bilinear interpolation) were used, however, higher order interpolation techniques can and have been applied [Edwards and Bridson 2012]. Thus, the slowness field $m_k$ at
2.4. Discretization

Figure 2.4: Particle-In-Cell method for the advection and mass-conservation equation. A particle located in the cell-centered point $\vec{x}_A$ on the regular mesh is pushed forward with velocity $\vec{u}$ along the path $\vec{u}\Delta t$, to the non-grid point $\vec{x}_B$. The value (mass) associated with the particle is then transferred to the cell centered grid points adjacent to $\vec{x}_B$ using interpolation weights.
2.5. Summary

Time $t_k = t_{k-1} + \Delta t$ is given by

$$m_k = T(\Delta tu)m_{k-1},$$

(2.8)

where $T(\Delta tu)$ denotes the push-forward matrix containing the values of the bilinear hat functions associated with the particles at the grid points. Since the value of each particle is spread between neighboring cell-centers, this construction guarantees exact mass preservation as long as no flux exits at the boundaries of the computational domain. Using the push forward matrix, the time-stepping process is then

$$m_k = T(\Delta tu)m_{k-1} = T^k m_0.$$

(2.9)

The slowness at the $k^{th}$ time step is therefore a pushed-forward version of $m_0$, where the path the particle takes is piece-wise linear.

2.5 Summary

This chapter presented the physical models selected to represent the geophysical imaging survey and tracer advection fluid flow model for the reservoir monitoring experiment, discussed the relationship between the model physical properties, and detailed the discretization.

One novel and important result of the tracer advection model was discussed in Section 2.3, is that under certain conditions the "motion" of the seismic slowness is governed by the advection model. This result eliminates the need to know the petrophysical relationship between the tracer and the slowness, and will be beneficial when formulating the coupled inverse problems later in Chapters 3 and 4.

In addition, the discretization of the physical models was detailed, and in particular, an explicit unconditionally stable Lagrangian technique was applied to discretize
2.5. Summary

the tracer advection equation. Because the method is explicit and unconditionally stable, it will not add additional complexity to the inverse problems to be proposed.
Chapter 3

Error characterization and model estimation

Estimating either the seismic slowness, or the fluid flow parameter functions and later optimally designing the geophysical survey requires the solution of an inverse problem. Before formulating the coupled inverse problem to estimate flow parameters, and developing an optimal experimental design method, it is important to understand the basis of a linear inverse problem from both the Bayesian and frequentist perspectives, and additionally, how the characterization of the error in the dynamic flow model effects the formulations. The main goal of this thesis is to give the reader a quick background in linear inverse theory in preparation for the inverse problem developed in Chapter 4, and the adaptive optimal experimental design method developed in Chapter 5.

This chapter first presents a brief review of Bayesian linear inversion and its relationship to frequentist linear inversion, followed by two formulations of a coupled inverse problem that assumes knowledge of the velocity field $\bar{u}$, for the case where there is no noise in the flow model, and the case where there is random noise.

3.1 Linear inversion

The process of inferring model parameters from a set of noisy observations is commonly known as inversion, or parameter estimation. In this context it is assumed
3.1. Linear inversion

that given a physical model and model parameters, observation data can be simul-
ted by solving the forward problem. Thus, if one has some real observation data, then
an estimate of the model parameters can be obtained by minimizing the difference
between simulated data (predicted data) and the observations.

Inverse problems can be thought of as statistical estimation problems that can
be studied from both Bayesian and frequentist perspectives (Biegler et al., 2011; S-
tark and Tenorio 2010). In either case, a stochastic model for the data is required
and constraints on the unknown model parameters can be incorporated. The major
difference in the two perspectives is that Bayesian methods require that constraints
be formulated as a prior probability distribution, whereas the frequentist method-
ology does not. An important well known observation worth noting, is that for a
linear Tikhonov regularized inverse problem, with some simple assumptions, the two
methodologies reduce to an equivalent problem. This result becomes important when
later in Chapter 5 when the adaptive optimal experimental design of the ill-posed
coupled problem is examined.

To begin a basic review of inversion, consider the dynamic process and measure-
ment technique described by the following set of discrete linear equations describing
the tracer advection (Equation (3.1a)), and the seismic tomography (Equation (3.1b)),
presented in Chapter 2

\[ m_k = Tm_{k-1} + \eta_k, \quad (3.1a) \]
\[ d_k = Fm_k + \epsilon_k, \quad (3.1b) \]

where the noise vectors \( \eta_k \) and \( \epsilon_k \) are assumed to be Gaussian normal, \( \eta_k \sim N(0, Q_k) \)
and \( \epsilon_k \sim N(0, W^{-1}_k) \). If the noise \( \epsilon \) is uncorrelated, then the matrix \( W \) is a diagonal
matrix with entries \( w_i = 1/\sigma_i^2 \), where \( \sigma_i \) is the standard deviation of \( \epsilon_i \) away from
3.1. Linear inversion

zero.

Assuming for the time being that the goal is to recover only the geophysical model properties given geophysical data without consideration of the dynamics, then the geophysical properties, or slowness $m$ for the tomography experiment, given travel time data set $d$, can be estimated by maximizing the Bayesian posterior likelihood of $m$ given $d$, which is proportional to the product of the probability of the data given the model, and the model likelihood:

$$P(m|d) \sim P(d|m)P(m),$$

where $m$ is a random variable with Gaussian prior distribution $m \sim N(\mu, \Sigma)$. Thus the probability of the model given the data is,

$$P(m|d) \propto \exp\left(-\frac{1}{2}(Fm - d)^\top W(Fm - d)\right) \exp\left(-\frac{1}{2}(m - \mu)^\top \Sigma^{-1}(m - \mu)\right).$$

Taking the negative log yields,

$$-\ln(P(m|d)) \propto \frac{1}{2}(Fm - d)^\top W(Fm - d) + \frac{1}{2}(m - \mu)^\top \Sigma^{-1}(m - \mu),$$

and minimizing with respect to $m$ gives

$$F^\top W(Fm + d) + \Sigma^{-1}(m - \mu) = 0.$$  

Finally, solving for $m$, yields the maximum a-posterior estimate (MAP estimate)

$$\hat{m} = (F^\top WF + \Sigma^{-1})^{-1}(F^\top Wd + \Sigma^{-1}\mu).$$

Alternately, a common linear inverse problem for estimating the geophysical pa-
3.1. Linear inversion

Parameters $m$ from the frequentist perspective is given by the following regularized Tikhonov inverse problem (Hansen, 1998; Tikhonov and Arsenin, 1977),

$$
\min_m \frac{1}{2} \|Fm - d\|^2 + \frac{1}{2} \alpha \|L(m - m_p)\|^2 .
$$

(3.2)

The first term in Equation (3.2) penalizes the difference between the simulated data and measured data, and the second term (the regularization term) is chosen to promote smoothness between the estimated model and a reference model $m_0$. The regularization term improves the convexity of the inverse problem as it is ill-posed in the classical sense; without it solutions will fit the measurement noise and are thus meaningless.

The choice of regularization is an interesting problem in itself, and is usually problem specific. There is much research in the area of optimally selecting and constructing regularization functionals, particularly in medical imaging (Hansen, 1998; Huang et al., 2012). Later, when developing the coupled inverse problem, specific care will be taken in constructing an appropriate regularization functional for reservoir parameter estimation.

Minimizing Equation (3.2) yields

$$
\hat{m} = (F^T F + \alpha L^T L)^{-1} (F^T d + \alpha L^T L m_p).
$$

Note that if $L^T L = \Sigma^{-1}$, $W = 1/\alpha I$, and $m_p$ is considered the prior mean, then the two problems are equivalent.

This well known fact will allow for the application of Bayesian optimal experimental design techniques to coupled inverse problems with linear regularization terms later discussed in Chapter 5.

However, in order to couple the Equations (3.1) to form a single inverse problem,
the error $\eta_k$ in the dynamic model must be considered.

## 3.2 Coupled inversion and error characterization

Before combining Equations (3.1a) and (3.1b) there are two instances of the noise in the dynamics that must be addressed which result in different coupled inverse problem formulations.

Assume first that the dynamics are exact, that is $\eta_k = 0, \forall k$. Such a scenario is often assumed in history matching (Oliver and Chen [2010]). In this case the dynamics are determined by the initial model $m_0$, and estimated using the measurements for all times. The second instance refers to the case that $\eta_k \neq 0$, and therefore models $m_1, \ldots, m_k$ for all times are estimated from all available data.

### 3.2.1 Exact dynamics

Setting $\eta_k = 0$ and assuming the velocity field $u$ is known, the model at time $t_k$ can be written as a linear mapping of the initial model $m_0$, such that

$$m_k = \mathbf{T}(\Delta t u)m_{k-1}, \quad \text{or} \quad m_k = \mathbf{T}(\Delta t u)^k m_0,$$

which can be written as the following linear system,

$$
\begin{pmatrix}
-I \\
\mathbf{T}(\Delta t u) & -I \\
\mathbf{T}(\Delta t u) & -I \\
\vdots & \ddots & \ddots \\
\mathbf{T}(\Delta t u) & \cdots & -I \\
\end{pmatrix}
\begin{pmatrix}
m_1 \\
m_2 \\
m_3 \\
\vdots \\
m_n \\
\end{pmatrix}
= 
\begin{pmatrix}
\mathbf{T}(\Delta t u) \\
0 \\
0 \\
\vdots \\
0 \\
\end{pmatrix}
\begin{pmatrix}
m_0 \\
= \mathbf{A}(u) \\
= \mathbf{m} \\
= \mathbf{B}(u) \\
\end{pmatrix}.$$

(3.4)
3.2. Coupled inversion and error characterization

In compact form the system is given by

\[ A(u)\overline{m} + B(u)m_0 = 0. \]

Thus it is possible to write the time evolution of the initial model \( m_0 \) in terms of the future models and the dynamics, such that

\[ \overline{m} = -A(u)^{-1}B(u)m_0. \] (3.5)

The data measurements at each time are then given by

\[ d_k = F_k m_0 + \epsilon_k, \] (3.6)

where \( F_k = FT^k \), and for the large system for all times,

\[ \overline{d} = (I \otimes F) A(u)^{-1}B(u)m_0 + \overline{\epsilon}, \] (3.7)

where \( \otimes \) is the matrix Kronecker product, \( \overline{d} = [d_1,..d_k]^\top \) and \( \overline{\epsilon} = [\epsilon_1,..\epsilon_k]^\top \).

The coupled linear Tikhonov inverse problem to estimate \( m_0 \), assuming that the dynamics are known, is then

\[ \min_{m_0} \frac{1}{2} \left\| (I \otimes F) A(u)^{-1}B(u)m_0 - \overline{d}_{\text{obs}} \right\|^2 + \frac{1}{2} \alpha \| Lm_0 \|^2. \] (3.8)

The case where the parameters governing the dynamics are unknown, that is, the reservoir parameters and velocity field are to be estimated, again results in a different formulation of the inverse problem. A new formulation for estimating these models is presented in Chapter 4.
3.2. Coupled inversion and error characterization

3.2.2 Inexact dynamics

To address the case where the dynamics are inexact the geophysical imaging and the tracer advection of \( m_0 \) are written for all times

\[
\begin{align*}
F m_0 - d_0 &= \epsilon_0, \\
F m_1 - d_1 &= \epsilon_1, \\
&\vdots \\
F m_k - d_k &= \epsilon_k, \\
m_1 - T m_0 &= \eta_1, \\
&\vdots \\
m_k - T m_{k-1} &= \eta_k.
\end{align*}
\]

Collecting all equations in matrix form results in the following system,

\[
\begin{pmatrix}
F & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot \\
\end{pmatrix}
\begin{pmatrix}
m_0 \\
m_1 \\
m_k \\
\end{pmatrix}
- 
\begin{pmatrix}
d_0 \\
d_1 \\
d_k \\
\end{pmatrix} = 
\begin{pmatrix}
\epsilon_0 \\
\epsilon_1 \\
\epsilon_k \\
\end{pmatrix},
\]

\[ (3.9) \]

\[
\begin{pmatrix}
-T & I & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
\end{pmatrix}
\begin{pmatrix}
m_0 \\
m_1 \\
m_k \\
\end{pmatrix} = 
\begin{pmatrix}
\eta_0 \\
\eta_1 \\
\eta_k \\
\end{pmatrix}.
\]

Defining the transport matrix \( \overline{T} \) for all times as

\[
\overline{T} = \begin{pmatrix}
-T & I \\
\cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot \\
\end{pmatrix},
\]

\[
\overline{m}_k = (m_0^T, \ldots, m_k^T)^T, \quad \overline{d} = (d_0^T, \ldots, d_k^T)^T, \quad \overline{\epsilon} = (\epsilon_0^T, \ldots, \epsilon_k^T)^T, \quad \text{and} \quad \overline{\eta} = (\eta_0^T, \ldots, \eta_k^T)^T
\]

the compact system is
3.3 Summary

\[ \tilde{d} = (\mathbf{I} \otimes \mathbf{F}) \tilde{m} + \tilde{\varepsilon}, \]  
(3.11)

\[ \overline{T \tilde{m}} = \overline{\eta}. \]  
(3.12)

Assuming again that the velocity field \( \mathbf{u} \) is known, the most recent model \( \tilde{\mathbf{m}}_k \) is estimated by solving the following regularized minimization problem which includes all historic data and model estimates \( \mathbf{m}_0, \ldots, \mathbf{m}_{k-1} \),

\[
\hat{\mathbf{m}}_k = \arg\min \quad \frac{1}{2} \left\| \left( \mathbf{I} \otimes \mathbf{F} \right) \tilde{\mathbf{m}}_k - \tilde{\mathbf{d}}_k \right\|^2 \overline{\mathbf{W}_k} + \frac{1}{2} \left\| \overline{T \tilde{\mathbf{m}}_k} \right\|^2 \overline{\mathbf{Q}_k}^{-1} \\
+ \frac{\alpha}{2} \left\| \left( \mathbf{I} \otimes \mathbf{L} \right) \tilde{\mathbf{m}}_k \right\|^2.
\]  
(3.13)

A second option is to re-estimate all models given all current data. This is akin to minimizing over \( \tilde{\mathbf{m}}_k \). Writing the problem in this way, where the inverse problem minimizing both error vectors, is also known as Kalman smoothing (Aavkin, 2010; Evensen, 1994).

3.3 Summary

This chapter provides a brief review of linear inverse theory from both the Bayesian and frequentist perspectives, highlights the relationship between the two methodologies for linear problems, and discusses the error characterization in the dynamic flow model.

Given the exact and inexact noise considerations presented, the main questions of the thesis can now be addressed. Chapter 4 considers the exact case and presents the formulation of the inverse problem for estimating both the initial state of the dynamic model, \( \mathbf{m}_0 \) and the velocity field \( \mathbf{u} \). Chapter 5 presents a new optimal design method.
3.3. Summary

developed from the Bayesian perspective for the frequentist coupled inverse problem. The design method is presented for both exact and inexact noise realizations in the dynamic flow model.
Chapter 4

Coupled inversion for velocity and initial slowness

In many cases in reservoir characterization, it is the hydraulic conductivity $K$ that is sought in order to forecast flow. For the tracer advection model, one might argue that this is not necessary, and that it may be a simpler problem to estimate the velocity field $u$. In particular, once the velocity field is known, and given an estimate of the initial slowness, the flow can be predicted by marching the forward in time. This is the motivation for the approach detailed here.

Whereas the previous chapter outlined a method for estimating the initial slowness $m_0$ given seismic travel time data while the velocity field is assumed to be known, this chapter highlights a novel approach for estimating both $m_0$ and the velocity field $u$ from seismic tomography data, through a coupled inverse problem. This formulation assumes that there is no error in the dynamic flow model. That is, $m_k = T(u)m_{k-1} + \eta_k$, where $\eta_k = 0$.

The mathematical formulation developed in this chapter is similar to that of Section 3.2.1, with the exception that the velocity field is unknown. The resulting parameter estimation problem is posed as a flow constrained inverse problem with specifically tailored regularization functionals for both the initial model $m_0$ and the velocity field $u$. Following the discretization of the regularization functionals, the initialization and numerical optimization of the resulting non-linear inverse problem is discussed.
4.1 Flow constrained geophysical imaging

The chapter concludes with a numerical example, where the initial slowness model and velocity field are recovered and used to forecast the flow of the tracer.

4.1 Flow constrained geophysical imaging

The formulation presented in Section 3.2.1 can be thought of as a flow constrained inverse problem for the case where the velocity field $u$ is known. In the case where one does not know the flow field, the following constrained problem can be solved to obtain estimates of both $m_0$ and $u$,

$$\min_{u,m_0} \frac{1}{2} \sum_{j=0}^{k} \|Fm_j - d_j\|^2$$ (4.1a)

subject to $m_k = T(\Delta t u)m_{k-1}$, $m(0,x) = m_0$, $\text{DIV} u = q$. (4.1b)

Recall Equation (3.4) such that $m_k = T(\Delta t u)m_{k-1}$ can be written in matrix form for all times in terms of the initial condition $m_0$ as $\overline{m} = -A(u)^{-1}B(u)m_0$. With the assumption that the dynamics are exact, i.e. the noise vector $\eta_k = 0, \forall k$, and the data for all times are given by Equation (3.7), then Equation (4.1) reduces to

$$\min_{u,m_0} \frac{1}{2} \|(I \otimes F)A(u)^{-1}B(u)m_0 - d\|^2$$ (4.2a)

subject to $\text{DIV} u = q$. (4.2b)

Minimizing Equation (4.2) under the flow constraints yields a velocity field and initial slowness distribution that fits the tomography data. However, since both the tomography and flow estimation problems are ill-posed, regularization is needed. Here $R^{\text{flow}}(u)$ is defined as a regularization functional on the velocity field, and $R^{m}(m_0)$ as a regularization functional for the initial slowness.

To regularize the flow field estimation the following continuous regularization func-
4.1. Flow constrained geophysical imaging

tional for \( \tilde{u} \) is defined:

\[
\mathcal{R}_{\text{flow}}(\tilde{u}) = \int \alpha_1 \phi(\nabla \times \tilde{u}) + \frac{\alpha_2}{2} w(\tilde{x}) |\tilde{u} - \tilde{u}_{\text{ref}}|^2 d\tilde{x},
\]  

(4.3)

where the first term penalizes the curl of the flow field and the second term seeks to minimize the difference between the recovered velocity, \( \tilde{u} \), and a reference velocity, \( \tilde{u}_{\text{ref}} \). The parameters \( \alpha_1, \alpha_2 > 0 \) balance the contribution of both terms.

The choice to control the curl of the flow in the regularization was made since the divergence and the curl complement each other. Given that the divergence of the flow field is set by the constraint (4.1b) it makes sense to regularize only over its orthogonal complement. The penalty function \( \phi \) was chosen by noting that sharp contrasts in hydraulic properties of different rock units are common; for example see the visualization of the \( \nabla \times \tilde{u} \) in the left column in Figure 4.2. Taking the curl of the velocity field \( \tilde{u} = K \nabla p \) with \( \nabla \cdot \tilde{u} = q \) gives

\[
\nabla \times \tilde{u} = \nabla \times (K \nabla p) = \nabla p \times \nabla K;
\]

thus, the curl of the flow field has tangential discontinuities where \( K \) has jumps between rock types. For the most part the pressure gradient is parallel to \( \nabla K \) except along the boundaries of rock types. For example, an aquifer might be surrounded by non porous rock with zero hydraulic conductivity, and thus the cross product along that boundary will not necessarily vanish. Thus, assuming that \( K \) is piecewise constant, the curl of the velocity is expected to be sparse.

The following convex approximation to the \( \ell_1 \)-norm is applied to promote the
4.2 Discretization of regularization functionals

sparsity of the curl of $u$,

$$\phi(c) = \sqrt{c^2 + \epsilon}. \quad (4.4)$$

The reference model, $\breve{u}_{\text{ref}}$, is included in order to incorporate prior information about the subsurface. The weighting function $w(\breve{x})$ quantifies the confidence in $\breve{u}_{\text{ref}}$; see (Oldenburg and Pratt, 2002). One option to compute $\breve{u}_{\text{ref}}$ is by solving (2.3b) and (2.3c) using a reference conductivity model $K_0$ constructed from a priori borehole data. In this case, $w(\breve{x})$ is large close to the boreholes and grows smaller as the distance from the known drill site grows.

It was mentioned in Section 2.2 that the diffusion term is not included in the advection-diffusion model. Because of this assumption, it is unlikely that the initial tracer model will have diffuse edges. Thus, smoothed Total Variation (TV) is used as the regularization to promote sharp edges in the estimated models (Ascher et al., 2006),

$$\mathcal{R}^m(m_0) = \beta \int \phi(|\nabla m_0|) \, d\breve{x}. \quad (4.5)$$

Finally, the regularization parameters $\alpha_1$, $\alpha_2$ and $\beta$ should be selected such that the data misfit is approximately equal to the norm of the noise; a quantity that is in general unknown (Parker, 1994). Therefore, a cooling strategy starting with large regularization parameters that are then decreased incrementally until a reasonably small data misfit is achieved is applied, (see Section 4.6).

4.2 Discretization of regularization functionals

The discretization of the regularization functionals applies standard finite difference approximations of the partial differential operators on orthogonal staggered grids, following the discretization of the tomography experiment and the flow equations.
4.2. Discretization of regularization functionals

outlined in Chapter 2.

To discretize the curl of the flow field \( \mathbf{u} \), Stoke’s theorem is used,

\[
\int_S \nabla \times \mathbf{u} \cdot \mathbf{n} dS = \oint_{\Gamma} \mathbf{u} \cdot d\Gamma,
\]

where, \( \mathbf{n} \) is the unit normal vector to the surface \( dS \), \( \Gamma \) is the path around the surface, and \( d\Gamma \) is the infinitesimal path length.

The discretization of the curl operator follows (Haber and Ascher, 2001; Modersitzki, 2009), such that the discrete curl operator is defined by

\[
\text{CURL} = \left( d_{t+1}^l \otimes I_{t_1} - I_{t_2} \otimes d_{t_1}^l \right).
\]

(4.7)

Here \( d_{t+1}^l \in R^{l \times l+1} \) is again a short finite difference matrix. Note that the CURL operates from the cell faces to the nodes. For a complete derivation of the matrices see (Haber and Ascher, 2001; Modersitzki, 2009).

The integral in Equation (4.3) is discretized using a midpoint rule, resulting in the discrete regularization,

\[
R_{\text{low}} (\mathbf{u}) = \alpha_1 h^2 \mathbf{e}^\top A_n^c \phi (\text{CURL } \mathbf{u}) + \frac{\alpha_2 h^2}{2} (\mathbf{u} - \mathbf{u}_{\text{ref}})^\top \mathbf{W} (\mathbf{u} - \mathbf{u}_{\text{ref}}),
\]

(4.8)

where the matrix \( A_n^c \) averages from nodes to cell-centers, \( h \) is the cell length, and \( \mathbf{e} \in R^M \) is a vector of ones.

Next, the regularization functional for the initial slowness is discretized. Noting that \( \mathbf{m}_0 \) is discretized on a cell-centered grid, a standard discrete approximation for the smoothed total variation regularization is applied (Ascher et al., 2006)

\[
R^n (\mathbf{m}_0) = h^2 \mathbf{e}^\top \sqrt{A_j^c ((\text{GRAD } \mathbf{m}_0) \odot (\text{GRAD } \mathbf{m}_0))} + \epsilon,
\]

(4.9)
4.3 Numerical optimization

where \( \circ \) is the Hadamard product, and \( \text{GRAD} \) is a standard 2-point discretization of the gradient of a cell-centered variable which maps from cell-centers to faces, as described in (Ascher et al., 2006; Haber and Ascher, 2001). \( \mathbf{A}_f^c \) is an averaging matrix from cell-faces to cell-centers. Note that the notation is somewhat abused, and that here the square root of a vector is the point-wise square root.

Now that the regularization functionals have been established and discretized, the optimization methods are discussed in the following section.

4.3 Numerical optimization

In this section the approach to solving the flow constrained discrete optimization problem is outlined. A variable projection method is chosen to solve for the initial slowness and velocity field in turn. Because the TV regularization is used for \( m_0 \) the objective function is non-linear with respect to \( m_0 \), and thus a primal-dual Newton method (Chan et al., 1999) is applied. To estimate \( u \) an approximate Sequential Quadratic Programming (SQP) method is utilized, (see Nocedal and Wright (2000) for further details).

The discrete form of the coupled variational optimization problem (4.2) for \( u \) and \( m_0 \) with the discrete regularization functionals is given by,

\[
\begin{align*}
\min_{u, m_0} & \quad \frac{1}{2} \| (\mathbf{I} \otimes \mathbf{F}) \mathbf{A}(u)^{-1} \mathbf{B}(u) m_0 - \mathbf{d} \|^2 + R^{\text{flow}}(u) + R^{m}(m_0) \\
\text{s.t.} & \quad \text{DIV} \ u = \mathbf{q}.
\end{align*}
\]

(4.10a)

(4.10b)

There are a number of options for the solution of such problems. One option is to solve the problem directly with respect to \( u \) and \( m_0 \). This approach has a number of disadvantages. First, it requires solving a large coupled problem where the
4.3. Numerical optimization

parameters may have different scales. Second, it does not take advantage of the existing ability to solve the decoupled problems efficiently. Finally, solving the coupled problem requires the simultaneous evaluation of two regularization parameters. An alternate, more attractive option is to use a variant of the variable projection method (Golub and Pereyra, 2003). This method was applied successfully in (Chung et al., 2006) for solving the related super-resolution problem. Furthermore, it has been shown in (Chung, 2009) that it is possible to choose regularization parameters for the different variables in the algorithm separately, thus decoupling the problem of selecting regularization parameters. The variable projection method as applied to the tomography-flow optimization problem is as follows.

First, assuming $u$ to be fixed, the conditions for a minimum with respect to $m_0$ are

$$G(u)^\top(G(u)m_0 - d) + \nabla_{m_0} R^m(m_0) = 0,$$  \hspace{1cm} (4.11)

where $G(u) = (I \otimes F)A(u)^{-1}B(u)$. In the standard variable projection method, where quadratic regularization is used, Equation (4.11) is linear with respect to $m_0$ and therefore can be solved directly. In this application the equation is nonlinear due to the TV regularization applied to $m_0$. Nonetheless, it is possible to solve this problem rather quickly using, for example, a primal-dual Newton method (Chan et al., 1999).

Assuming that $m_0$ solves Equation (4.11), then $m_0 = m_0(u)$ and the optimization problem can be rewritten as a problem for $u$ alone

$$\min_u \frac{1}{2} \|G(u)m_0(u) - d\|^2 + R^{\text{flow}}(u) + R^m(m(u)) \hspace{1cm} (4.12a)$$

$$\text{s.t.} \quad \nabla V u = q.$$  \hspace{1cm} (4.12b)
4.3. Numerical optimization

Introducing the Lagrange multiplier $\lambda$, the Lagrangian $L(u, \lambda)$ of the problem is

$$L(u, \lambda) = \frac{1}{2} \|G(u)m_0(u) - d\|^2 + R^{\text{flow}}(u) + R^m(m_0(u)) + \lambda^\top(D\nabla u - q). \quad (4.13)$$

An important observation made in (Golub and Greif, 2003) was that if $m_0$ solves the system Equation (4.11) then $\nabla_u L(u, \lambda, m_0(u)) = \frac{\partial L}{\partial u} + \frac{\partial L}{\partial m_0} \frac{\partial m_0}{\partial u} = \frac{\partial L}{\partial u}$, because Equation (4.11) implies that $\frac{\partial L}{\partial m_0} = 0$ and thus $m_0$ can be treated as a constant in the objective function.

This observation leads to the following conditions for a minimum

$$J(u)^\top(G(u)m_0(u) - d) + \nabla_u R^{\text{flow}}(u) + D\nabla^\top \lambda = 0 \text{ and } D\nabla u = q, \quad (4.14)$$

where $J(u)$ is the Jacobian (sensitivity) of the data misfit term with respect to $u$. To compute the Jacobian the misfit is rewritten as,

$$D(u) = \frac{1}{2} r(u)^\top r(u),$$

where $r(u) := (I \otimes F)A(u)^{-1}B(u)m_0 - d$ and $J(u) = \frac{\partial r}{\partial u}$.

Using (3.4) to simplify notation, and recalling the definition $\overline{m} = A(u)^{-1}B(u)m_0$, the Jacobian is then

$$\frac{\partial r(u)}{\partial u} = (I \otimes F) \frac{\partial \overline{m}(u)}{\partial u}. \quad (4.15)$$

In order to differentiate $\overline{m}$ with respect to $u$, Equation (3.4) is implicitly differentiated,
4.3. Numerical optimization

yielding

\[
\frac{\partial (A(u)\bar{m})}{\partial u} + A(u) \frac{\partial \bar{m}(u)}{\partial u} = \frac{\partial (B(u)m_0)}{\partial u}.
\] (4.16)

For the computation of the derivatives of \(A(u)\bar{m}\) and \(B(u)m_0\) only the push forward operator times the \(i^{th}\) model, \(T(u)m_i\), needs to be differentiated for indexes \(i = 1, \ldots, n\). These derivatives depend on the employed interpolation basis functions.

Differentiating the product \(T(u)m_i\) with respect to \(u\) has been done in (Chung et al., 2006). Furthermore, since the interpolation matrix is sparse, the derivative matrices \(\frac{\partial A(u)\bar{m}}{\partial u}\) and \(\frac{\partial B(u)m_0}{\partial u}\) are sparse.

To summarize, Equation (4.16) was solved for the derivative of \(\bar{m}\) with respect to \(u\) and substituted into Equation (4.15) to obtain the derivative of the residual function \(r(u)\) as

\[
J(u) = \frac{\partial r(u)}{\partial u} = (I \odot F)A(u)^{-1}\left(\frac{\partial (B(u)m_0)}{\partial u} - \frac{\partial (A(u)\bar{m})}{\partial u}\right).
\] (4.17)

An important observation needs to be made here. While the matrix \(J(u)\) is dense in general, its product with a vector can be computed efficiently using sparse matrix techniques. To calculate \(J(u)\) times an arbitrary vector \(z\) one first computes the matrix vector product

\[
y = \left(\frac{\partial (B(u)m_0)}{\partial u} - \frac{\partial (A(u)\bar{m})}{\partial u}\right)z.
\]

Since this matrix is sparse the computation can be done efficiently. Next, the solution \(x = A(u)^{-1}y\) is obtained by solving the linear system

\[
A(u)x = y.
\]
4.3. Numerical optimization

Solving this linear system is equivalent to solving a single flow forward problem, that is, advection in time. Finally, the matrix $F$ is also sparse and thus computing the matrix vector product, $(I \otimes F)x$ can also be done quickly and efficiently.

To complete the computation of the derivatives the differentiation of the regularization term is now discussed. It is straightforward to verify that

$$\frac{\partial F_{\text{flow}}}{\partial u} = \alpha_1 \text{CURL}^T \text{diag}(1/\phi(\text{CURL}u)) \text{CURL}u,$$

where the notation $1/\phi$ is a vector that is divided point wise.

Given all the components described above the goal is to now solve the discrete optimization problem. There are a number of options for the solution of the problem. Here an approximation to the Sequential Quadratic Programming approach (SQP) is applied. By linearizing the system in (4.12) and using a Gauss-Newton approximation of the Hessian with respect to $u$ the following linear system is obtained,

$$
\begin{pmatrix}
    J^T J + \alpha_1 \text{CURL}^T \text{diag}(1/\phi(\text{CURL}u)) \text{CURL} & \text{DIV}^T \\
    \text{DIV} & 0
\end{pmatrix}
\begin{pmatrix}
    \delta u \\
    \delta \lambda
\end{pmatrix}
= -
\begin{pmatrix}
    I_u \\
    I_\lambda
\end{pmatrix}.
$$

This system is solved for $\delta u$ and $\delta \lambda$ and a soft (backtracking) line search is used for the update of $u$ and $\lambda$ (Nocedal and Wright 2000).

To solve the linear system (4.19) note that the system has many similarities to the Stokes problem with a positive-semi definite (1,1) block, as outlined in Golub and Greif (2003). Such systems can be solved by a combination of the augmented Lagrangian method and an approximation to the Schur complement (Benzi et al., 2005).
4.4 Initialization

The optimization problem given by Equation (4.10) is a nonlinear non convex problem that without an appropriate initialization will result in solutions of low quality. Therefore an initialization methodology is used that cheaply yields a “reasonable” starting point.

To this end, consider the decoupling of the imaging of the slowness for all times. That is, consider the individual problems

\[ \mathbf{F} \mathbf{m}_k + \epsilon = \mathbf{d}_k; \quad k = 1, \ldots, n. \]

Inverting each data set to obtain \( n \) initial estimates of \( \mathbf{m}_k \) is accomplished by solving \( n \) decoupled optimization problems;

\[ \hat{\mathbf{m}}_k = \arg \min_{\mathbf{m}_k} \frac{1}{2} \| \mathbf{F} \mathbf{m}_k - \mathbf{d}_k \|^2 + \mathcal{R}^m(\mathbf{m}_k); \quad k = 1, \ldots, n. \quad (4.20) \]

Given the estimates of \( \hat{\mathbf{m}}_k \) an estimate of the initialization velocity is computed by solving the optimization problem

\[ \hat{\mathbf{u}} = \arg \min_{\mathbf{u}} \frac{1}{2} \| \mathbf{A}(\mathbf{u}) \hat{\mathbf{m}} - \mathbf{B}(\mathbf{u}) \mathbf{m}_0 \|^2 + \mathcal{R}^{\text{flow}}(\mathbf{u}). \quad (4.21) \]

This estimate is equivalent to obtaining a flow estimate without improving \( \mathbf{m} \), which in this experiment leads to a good initial guess for both the slowness and velocity.

In the following sections the approach to jointly estimate the initial slowness and the velocity field are presented for a small numerical example. The results show that the recovered flow field and reconstructed initial slowness can be used to predict the flow of an injected tracer within the subsurface.
4.5 Experimental setup

The computational domain, \( \Omega = [0, 100] \times [0, 200] \) meters, is divided into \( l = [200, 100] \) cells of width \( h = [1, 1] \). The ground truth conductivity model consists of five layers, with conductivities ranging between 10 and 1000 m\(^3\)·day/kg. The ground truth model is pictured in the top panel of Figure 4.1. Two boreholes are used for the injection and extraction of fluids. A reference conductivity model was constructed by linear interpolation of the borehole data. The reference conductivity model is a simple layer of high conductivity \( (k = 1000 \text{ m}^3\cdot\text{day/kg}) \) surrounded by a background layer \( (k = 10 \text{ m}^3\cdot\text{day/kg}) \). Both the true and reference conductivity models are pictured in Figure 4.1. The ground-truth initial tracer \( m_{gt} \) is a piece-wise constant model with two regions of concentration 0.5 and 1; see top left plot in Figure 4.2. Fluid is injected from the left well at 50 m depth and extracted on the right well at 60 m depth with a static pumping rate of \( \pm 100 \text{ m}^3/\text{day} \).

Given the hydraulic conductivity model and the source term, the ground-truth flow field \( u_{gt} \) and the reference flow field \( u_{ref} \) are obtained by solving Equation (2.3); see quiver plot in Figure 4.1. Tomography travel-time data are simulated by solving the forward problem given by Equation (3.7) for \( m_{gt} \) and \( u_{ref} \) and adding Gaussian white noise with a standard deviation \( \sigma = 0.5 \). The tomography experiment consists of \( n_s = n_r = 35 \) transmitters and receivers equally spaced from 20m to 100m depth along the left and right boundary of the domain, respectively. Data were simulated for 15 days with a time step \( \Delta t = 1 \) day. Snapshots of the advection of the tracer can be seen in the first column of Figure 4.3.

4.6 Sequential reconstruction as initialization

To obtain a starting guess of the flow field and the initial slowness \( m_0 \), the procedure outlined in Section 4.4 is applied. First, the slowness evolution is individually
4.6. **Sequential reconstruction as initialization**

Figure 4.1: Visualizations of the ground true and reference hydraulic conductivity models and associated flow fields. The true and reference models are discretized on a $100 \times 200$ cell grid on a domain of 100m depth by 200m across. A static source, at 55m depth on the left border, and a sink, at 65m depth on the right border of the domain, are indicated by red. The reference model is a simplified model which interpolates borehole data between the two wells.
4.6. Sequential reconstruction as initialization

Figure 4.2: Visualization of ground truth data for the domain with dimensions 100m (y direction) depth and 200m (x direction) across (first column), individual reconstruction (second column) and results of the joint reconstruction approach. The first row visualizes the initial slowness ($m_0$). The second shows a quiver plot of the recovered velocity field ($\vec{u}$). The components of the velocity field $[u_x, u_y]$ are pictured in the third and fourth row and the last row shows the curl of the velocity field ($\nabla \times \vec{u}$). Note that there is almost no flow in the y direction, and that at the boundaries of the reservoir there are non-zero values of the curl of $\vec{u}$.
4.6. **Sequential reconstruction as initialization**

Figure 4.3: The evolution of the tracer is pictured for the ground-truth (left column), individual reconstruction used for initialization (middle column), and the proposed joint reconstruction. The tracer evolution is simulated for 40 days with the respective initial slowness and flow field. The front of the ground truth tracer is visualized by dashed lines for the first 3 rows. Note that the joint method results match the ground truth tracer front to a greater degree than the predictions from the initialization estimates.
reconstructed for all 15 days solving Equation (4.20) using a standard Gauss-Newton method with a regularization parameter $\beta = 0.4$. Subsequently, a flow estimate $\hat{u}$ was obtained by solving Equation (4.21) using the described SQP method with regularization parameters $\alpha_1 = 10^{-3}$ and $\alpha_2 = 4 \cdot 10^{-3}$. The regularization parameters were found using a cooling strategy which starts with large parameters that are decreased iteratively until the measurement noise became visually too prominent in both the reconstructed images and the estimated flow fields.

A larger bias towards the reference flow field, $u_{\text{ref}}$, is given close to the boreholes by using the weighting function $w(x_1, x_2) = (x_2 - 100)^2 / 100^2$.

Estimates of the initial slowness and flow field are visualized in the second column of Figure 4.2. The first row shows the reconstruction of the initial slowness, estimated by inverting the tomography data of the first day. The flow field estimate, visualized in the second, third, forth, and fifth rows; in a quiver plot, component wise and by its curl, captures the main characteristics of the ground truth. However, the magnitude of the velocity field is small compared to the ground truth. This can be seen by looking at the plots of $u_y$ and in the second column of Figure 4.3 where the front of the high concentration plume is behind the ground truth. This demonstrates the importance of using the flow to better guide the imaging which in turn yields more accurate flow predictions.

4.7 Joint reconstruction

To improve the estimates of $m_0$ and the velocity field, the coupled tracer tomography inverse problem (Equation (4.1a)) is solved. The regularization parameters were adjusted to $\alpha_1 = 5 \times 10^{-3}, \alpha_2 = 7 \times 10^{-4}$ and $\beta = 200$. The result of the individual estimate of $\hat{u}$, was used as a starting guess for the optimization and $u_{\text{ref}}$ for the reference model.
4.8. Forecasting

Since errors in the flow field accumulate in time, an outer loop is used to introduce the tomography data sequentially as the reconstruction of the velocity field improves. To be precise, only the tomography data from the first day is used in the estimation problem for $m_0$, and then flow field is estimated by solving Equation (4.12) using the SQP method outlined earlier. Afterwards the estimate $m_0$ is updated by introducing tomography data from one additional day and updating the estimate of $u$. In total, five outer iterations are performed using tomography data from the first six days. Note that all tomography data is used in the flow estimation at all iterations. It is important to note that computational time increases with the coupled recovery method since it requires solving the System (4.19) in addition to the imaging step.

The initial slowness and velocity field reconstruction results are visualized in the third column of Figure 4.2. Note the significant improvement in the reconstruction of $m_0$, particularly the sharpening of the edges.

4.8 Forecasting

The numerically estimated flow fields and initial slowness are used to forecast the evolution of the tracer beyond the first 15 days by solving the forward flow model in Equation (2.9) for 35 days. The numerical solutions of the initialization and the joint history matching problems are compared to the ground-truth in Figure 4.3.

In the considered application, the most interesting quantity is the prediction of the arrival time. The front of the tracer is thus visualized by a dashed line in the subplots. It can be seen that the coupled estimate of $u$ and $m$ improves the prediction of the arrival time when compared with the estimates obtained in the initialization.
4.9 Summary

This chapter presented a new method for the estimation and prediction of tracer flow in porous media given seismic tomography geophysical data. The new coupled method was formulated as a flow constrained inverse problem to estimate both the initial seismic slowness and the velocity field in which tracer is injected. As a consequence of the coupled ill-posed inverse problem, a novel regularization for the velocity field was additionally constructed to promote discontinuities at rock property interfaces.

The coupled method was demonstrated for the model borehole seismic tomography survey of tracer advection. The numerical results were compared with numerical results obtained through a decoupled process and demonstrate that this new coupled approach yields not only accurate reconstructions of the initial slowness and flow field, but, especially, accurate predictions of the fluid flow velocity and tracer evolution.

In addition to coupling the flow and geophysics, the next approach to improving reservoir monitoring is to consider what data are collected as the tracer moves through the subsurface. This optimal experiment design question is investigated in the following chapter.
Chapter 5

Dynamic optimal experimental design

The previous chapter presented a new joint method for estimating the initial slowness $m_0$, the velocity field $u$, given all previous seismic tomography data sets for the reservoir monitoring experiment outlined in Chapter 2. Although this method alleviates some of the uncertainty and expense involved in reservoir monitoring programs by making use of geophysical data instead of sparse fluid measurements, it is possible that the tomography data set could be reduced while maintaining an optimal estimation of the initial slowness. In this chapter, the idea is therefore to reduce the number of necessary tomography data measurements in an optimal way. This is commonly known as optimal experimental design.

The optimal design method will generate an experiment with a reduced number of measurements while maintaining the best estimate of the initial slowness, incorporate the historic tomography data sets, and couple the flow dynamics such that the experimental design adapts with the moving tracer. For the formulation of the design method, the assumption is made that dynamics are well known, that is, it is assumed that $\bar{u}$ is known, or estimated to reasonable accuracy for forecasting.

The new approach, defined as adaptive optimal experimental design, applies a method similar to classic A-optimal design criteria, by minimizing the adapted mean square error (amse) of a regularized model estimate, instead of the mean square error.
error (Atkinson and Donev 1992, Fedorov 1972). The adapted mean square error is defined by introducing a monitor function which scales the mean squared error according to historic model estimates.

This chapter introduces the adaptive design method from first principals, the numerical optimization of the resulting design objective function, and discusses the results for the tomography tracer advection reservoir model example.

5.1 Adaptive experimental design

To introduce the adaptive experimental design method it is necessary to understand A-optimal design criteria from first principals. To begin, consider only two measurement vectors of the form

\[ F_1 m + \epsilon_1 = d_1 \quad \text{and} \quad F_2 m + \epsilon_2 = d_2, \quad (5.1) \]

where \( F_1 \) is the forward tomography operator at time \( t_1 \) and \( F_2 \) is the forward operator at time \( t_2 \). As before, the measurement error vectors \( \epsilon_k \) are assumed to be normally distributed uncorrelated noise, \( \epsilon_k \sim N(0, W_k^{-1}) \), with zero mean and a diagonal covariance matrix \( W_k^{-1} \), for \( k = 1, 2 \).

At this point consider only the static case, where the goal is to design the experiments at times \( t_1 \) and \( t_2 \) such that the “best” recovery of \( m \) is obtained by some criteria, such as minimizing the Tikhonov optimization problem. This scenario fits the problem where the dynamical system is “exact” (containing no noise as in Section 3.2.1) and \( m \) represents the dynamical system parameters, such as the initial condition of the tracer in the reservoir modeling experiment.

Two different designs can be considered. First, it is possible to perform a-priori design, that is, to design the experiments \( F_1 \) and \( F_2 \) prior to the data collection.
5.1. Adaptive experimental design

This is the case when the time between \( t_1 \) and \( t_2 \) is much shorter than the time for the processing of the data. A second approach is to use post-priori design. The idea here is to use the results obtained from the first experiment in order to design a “better” second experiment. The goal is to first obtain an a-priori design for the first experiment and then, after the data is collected and processed, use the estimator obtained at \( t_1 \) in order to design the data collection at time \( t_2 \).

Before discussing the design of the data acquisition for time \( t_2 \), it is necessary to review the a-priori design of the experiment at time \( t_1 \). Consider the regularized estimation of \( m \) given \( d_1 \), accomplished by solving the following penalized-weighted least squares (Tikhonov regularized) optimization problem

\[
\hat{m}_1 = \arg \min_m \frac{1}{2} \|F_1 m - d_1\|_W^2 + \frac{\alpha}{2} \|L(m - \hat{m}_0)\|_2^2,
\]

(5.2)

where \( W_1 = \text{diag}(w_1) \) is a matrix of inverse variances, that is, the inverse covariance of the noise, \( L \) is a smoothing penalty matrix, \( \hat{m}_0 \) is the current estimate of the model prior to having data, and \( \alpha \) is a regularization parameter. Recall from Chapter 3 that from the Bayesian perspective, \( \hat{m}_0 \) can be considered the mean of a prior probability distribution with covariance matrix \((L^TL)^{-1}\). Minimizing Equation (5.2) yields the estimator

\[
\hat{m}_1 = (F_1^TW_1F_1 + \alpha L^TL)^{-1}(F_1^TW_1d_1 + \alpha L^TL\hat{m}_0).
\]

(5.3)

Defining the matrix \( C = (F_1^TW_1F_1 + \alpha L^TL) \), and recalling that \( F_1 m + \epsilon_1 = d_1 \) the error in the recovery can be written as

\[
\hat{m}_1 - m = C^{-1}F_1^TW_1F_1 m + C^{-1}F_1^TW_1\epsilon_1 + \alpha C^{-1}L^TL\hat{m}_0
\]

\[
+ (\alpha C^{-1}L^TLm - \alpha C^{-1}L^TLm) - m.
\]

(5.4)
5.1. Adaptive experimental design

Collecting terms and using the definition of $C$ gives,

$$ \hat{m}_1 - m = C^{-1} F_1^\top W_1 \epsilon_1 + \alpha C^{-1} L^\top L (\hat{m}_0 - m). $$  

(5.5)

Squaring and taking the expectation over $\epsilon_1$ and recalling that $\epsilon_1 \sim N(0, W_1^{-1})$, the mean square error is then given by

$$ \text{mse}(w_1, m) = \mathbb{E} \| \hat{m}_1 - m \|_2^2 = \text{trace}[F_1 C^{-2} F_1^\top W_1] + \alpha^2 \mathbb{E} \| C^{-1} L^\top L (\hat{m}_0 - m) \|_2^2. $$  

(5.6)

Taking the Bayesian point of view assumes that $m - \hat{m}_0$ is Gaussian with a zero mean and a covariance matrix $(\alpha L^\top L)^{-1}$ such that the mean squared error is equivalent to the Bayesian risk.

$$ \text{mse}(w_1) = \phi_1(w_1) = \text{trace}[C^{-2} F_1^\top W_1 F_1] + \alpha \text{trace}[C^{-2} L^\top L]. $$  

(5.7)

Finally, using the linearity of the trace and the definition of $C$ results in

$$ \phi_1(w_1) = \text{trace} \left[ (F_1^\top W_1 F_1 + \alpha L^\top L)^{-1} \right]. $$  

(5.8)

That is, the Bayesian risk is the trace of the inverse precision matrix, or trace of the posterior covariance matrix; a well known result.

The goal of the design problem seeks to obtain a better estimate for $m$ by assuming that the (inverse) variances, $w_1$, of the collected data can be controlled in some way.

By controlled, it is assumed that a data measurement is accompanied by a standard deviation determined by the instrumentation. If this is not the case, several measurements at a particular location can be conducted to estimate the error. Similar treatment is given in (Alexanderian et al., 2014; Haber et al., 2008, 2010, 2011) where
5.1. Adaptive experimental design

this point is further discussed. In this way, through optimization, the weights $w_i$ are estimated prior to the actual data collection, and a measurement $d_i$ that is assigned with infinite standard deviation ($w_i = 0$), or zero variance, by the optimization is thus not collected.

To estimate the variances, the Bayesian risk $\phi_1(w_1)$ is minimized with respect to $w$. Additionally, since the goal is to obtain a sparse design (that is, collect only a few measurements) an additional cost on $w_1$ is added. This cost is equivalent to the 1-norm of $w$ and promotes its sparsity (see (Alexanderian et al., 2014; Haber et al., 2008) for further discussion), resulting in the following optimization problem which balances the minimization of the mean squared error (mse) and the cost of the experiment (Haber et al., 2008),

$$
\phi_1^2(w_1) = \text{trace} \left[ (F_1^T W_1 F_1 + \alpha L^T L)^{-1} \right] + \beta e^T w_1; \quad 0 \leq w_{1,i},
$$

(5.9)

where $e^T$ is a vector of ones.

Consider now using the design problem of estimating $m$ given the estimated model $\hat{m}_1$. One could rewrite the recovery problem in a similar way, that is

$$
\hat{m}_2 = \arg \min \frac{1}{2} \| F_1 m - d_1 \|^2_{W_1} + \frac{1}{2} \| F_2 m - d_2 \|^2_{W_2} + \frac{\alpha}{2} \| L(m - \hat{m}_0) \|^2_2,
$$

(5.10)

which leads to the estimate

$$
\hat{m}_2(w_2) = (F_1^T W_1 F_1 + F_2^T W_2 F_2 + \alpha L^T L)^{-1} (F_1^T W_1 d_1 + F_2^T W_2 d_2 + \alpha L^T L \hat{m}_0).
$$

(5.11)

Note that $w_1$ is assumed to be fixed and therefore the new estimate is a function of $w_2$ alone. If the steps above are repeated then the A-optimal design for time $t_2$ yields
5.1. Adaptive experimental design

the minimization of the function

$$\phi_2^\beta(w_2) = \text{trace} \left[ (F_1^\top W_1 F_1 + F_2^\top W_2 F_2 + \alpha L^\top L)^{-1} \right] + \beta e^\top w_2; \quad 0 \leq w_{2,i}. \quad (5.12)$$

At this point it is worth pausing for a moment and noting an important feature of the experimental design criteria. The design criteria do not depend on the data. This observation is true for any of the design criteria (that is, C,D and E designs) (Fedorov, 1972). Furthermore, it is easy to verify that any linear estimator of the data yields a covariance matrix that is independent of $d_1$. This implies that using current design criteria does not make use of the estimated model obtained at time $t_1$ to obtain a better estimate at time $t_2$.

In order to use the information obtained at time $t_1$ for the design of the experiment at time $t_2$ the concept of adaptive design is presented.

Assume that the model, $\hat{m}_1$ has some “interesting” features and some “boring” features. To be more specific, assume that the difference

$$\delta_1 = |\hat{m}_1 - \hat{m}_0| \quad (5.13)$$

is small in some norm over a region and large in others. The goal then is to better estimate the new features that appear in the model.

To this end, the monitor function, $\tau(\delta)$ is introduced to measure the change in the model. For example, to start, consider the function

$$\tau = \delta, \quad (5.14)$$

that simply measures the change in the estimator compared to the previously known estimator. A different function that was found to be useful in numerical experiments
5.1. Adaptive experimental design

is

$$\tau = \chi_\theta(\delta),$$  \hfill (5.15)$$

where $\chi_\theta$ is a smoothed characteristic function. In this case the optimization focuses only on areas where $\delta$ is large and does not sample areas where $\delta$ is small.

Given the monitor function, rather than minimizing the mean square error, the idea is to minimize the adaptive mean square error (amse), defined by

$$\text{amse}(w_2, m) = E[\|\hat{m}_2 - m\|_1^2] = E[(\hat{m}_2 - m)^\top \text{diag}(\tau_1)(\hat{m}_2 - m)].$$  \hfill (5.16)$$

The idea behind the amse is to obtain a tighter bound on the model where the estimator exhibits large changes compared to the known a-priori estimator.

Starting from Equation (5.4), modifying it to deal with the amse, and repeating the calculation above, the expectation over the amse is

$$\phi_2(w_2) = \text{trace} \left[ \text{diag}(\tau_1)(F_1^\top W_1 F_1 + F_2^\top W_2 F_2 + \alpha L^\top L)^{-1} \right].$$  \hfill (5.17)$$

Similar to the non-adaptive case, the adaptive design is computed by minimizing the (penalized) function $\phi_2(w_2)$.

The above concept can easily be extended to any number of time steps. Consider $k - 1$ experiments that are conducted at times $t_1, t_{k-1}$ using the forward operators $F_1, \ldots, F_{k-1}$ and assume that the goal is to design an experiment for the problem at time $t_k$. The design criteria in this case reads

$$\phi_2^\alpha(w_k) = \text{trace} \left[ \text{diag}(\tau_{k-1}) \left( \sum_{j=1}^{k} F_j^\top W_j F_j + \alpha L^\top L \right)^{-1} \right] + \beta e^\top w_k; \quad 0 \leq w_{k,i},$$  \hfill (5.18)$$
5.2. Design for dynamical systems

The previous section presented a formulation for adaptive A-optimal experimental design that assumes the model, \( m \) is static. This section therefore addresses the case where the model is governed by a (potentially noisy) dynamical system.

Recalling Chapter 3 and the system of Equations (3.1), the adaptive optimal design method is applied for both cases, where the dynamic tracer advection model is assumed to be exact or inexact. Note that in these cases the velocity field \( u \) is assumed to be known.

5.2.1 Exact dynamics

Considering the data vector given by Equation (3.6) in Chapter 3 and noting that this case is exactly the case described in the previous section, the anse criteria can by directly applied.

Given the assumption that there is no error, the estimation of the initial model \( m_0 \) given \( k \) data sets is

\[
\hat{m}_{0,k} = \text{argmin} \frac{1}{2} \sum_{j=1}^{k} \| F_j m_0 - d_j \|_W^2 + \frac{\alpha}{2} \| L(m_0 - \hat{m}_0) \|_W^2, \quad (5.19)
\]

and therefore the design function is identical to Equation (5.18)

\[
\phi_k^\beta(w_k) = \text{trace} \left[ \text{diag}(\tau_{k-1}) \left( \sum_{j=1}^{k} F_j^T W_j F_j + \alpha L^T L \right)^{-1} \right] + \beta e^T w_k; \quad 0 \leq w_{k,i}. \quad (5.20)
\]
5.2. Design for dynamical systems

5.2.2 Inexact dynamics

To address the case where the dynamics are inexact recall Equations (3.11) and (3.12), which correspond to the noise vectors for the tomography imaging experiment and the tracer dynamics model.

In order to apply the amse method the problem is formulated as in Kalman smoothing (Aravkin et al., 2013; Kalman, 1960). A similar approach that maximizes the information gain was presented in (Gharamti et al., 2015) for the ensemble Kalman filter.

Recalling that the noise vector \( \eta_k \sim N(0, Q_k) \), and assuming that all \( Q_k \) are known, that is, that the dynamics are not changing in time, all models \( \hat{\mathbf{m}}_k = (\hat{\mathbf{m}}_0^T, \ldots, \hat{\mathbf{m}}_k^T)^T \) are estimated by minimizing Equation (3.13) presented in Section 3.2.2

\[
\hat{\mathbf{m}}_k = \arg\min_{\mathbf{m}_k} \frac{1}{2} \left\| (\mathbf{I} \otimes \mathbf{F}) \mathbf{m}_k - \mathbf{d}_k \right\|_{\mathbf{W}_k}^2 + \frac{1}{2} \left\| \mathbf{T} \mathbf{m}_k \right\|_{Q_k^{-1}}^2 + \frac{\alpha}{2} \left\| (\mathbf{I} \otimes \mathbf{L}) \mathbf{m}_k \right\|_2^2,
\]

such that,

\[
\hat{\mathbf{m}}_k = \left( (\mathbf{I} \otimes \mathbf{F}^\top) \mathbf{W}_k (\mathbf{I} \otimes \mathbf{F}) + \mathbf{T}^\top \mathbf{Q}_k^{-1} \mathbf{T} + \alpha \mathbf{I} \otimes \mathbf{L}^\top \mathbf{L} \right)^{-1} (\mathbf{I} \otimes \mathbf{F}^\top) \mathbf{W}_k (\mathbf{I} \otimes \mathbf{F}) \mathbf{d}_k,
\]

(5.21)

where \( \mathbf{W}_k = \text{diag}(w_0, \ldots, w_k) \), and \( \mathbf{Q}_k^{-1} = \text{blkdiag}(Q_1^{-1}, \ldots, Q_k^{-1}) \).

It is assumed in this case that the first experiment \( \mathbf{d}_0 \), will either be conducted such that all data are collected, or will be conducted in a naive way. In either case it is assumed that the initial experimental design, \( \mathbf{W}_0 = \text{diag}(\mathbf{w}_0) \), is known or can be designed for as was suggested in previous work (Haber et al., 2011).

In a straightforward extension of the previous section, the amse design optimiza-
5.3. Numerical optimization

The objective function for a design at time $t_k$ is then

$$\Phi_k^\beta(w_k) = \text{trace} \left[ \text{diag}(\tilde{\tau}_k) \left( (I \otimes F)\top \mathbf{W}_k (I \otimes F) + T\top \tilde{Q}_k^{-1} T + \alpha I \otimes (L\top L) \right)^{-1} \right] + \beta e\top w;$$

$$0 \leq w_{k,i}. \quad (5.22)$$

where $\tilde{\tau}_k = (\tau_0(\hat{m}_0), \ldots, \tau_k(T\hat{m}_{k-1}))$.

It is assumed here that all previous experiments $0, \ldots, k - 1$ have already been conducted. Therefore, solutions are only for the current time point $t_k$. At this point, $\hat{m}_k$ is not known. To predict $\hat{m}_k$ at time $t_k$ the previous estimator is propagated ahead in time, $\hat{m}_k \approx T\hat{m}_{k-1}$.

The two formulations of the adaptive design method for a dynamical system yield two different numerical optimization problems. However, in the limit that all $Q_k \rightarrow 0$, the inexact optimization problem reduces to the exact formulation, see (Nocedal and Wright, 2000) for a discussion of penalty methods. The numerical optimization techniques for the solution of these problems are discussed in the following section.

5.3 Numerical optimization

This section presents the calculation of the gradients required to solve the design optimization problems for both the exact and inexact tracer dynamics.

5.3.1 Exact dynamics

Finding an optimal design for time step $k$ requires the minimization of the objective functional given by Equation (5.20), which involves computing the trace of a large dense matrix. Typically, when solving geophysical problems, the size of matrices can be on the order of millions. Thus, forming and storing these matrices is memory and computationally intense.
5.3. Numerical optimization

In the non-linear optimization algorithm, at each iteration the computation of the objective function and its gradient is required. In order to avoid numerous expensive calculations while storing and manipulating large dense matrices, a stochastic Hutchinson trace estimator is applied to approximate the trace (Haber et al., 2011; Hutchinson, 1990). By doing so, the objective function reduces to matrix vector products which can be carried out much more efficiently.

If the vectors \( v_1, \ldots, v_m \) are independent and each with independent entries taking the values of 1 and \(-1\) with equal probability, then the stochastic trace estimator of the trace of a matrix \( H \) is given by

\[
\text{trace}(H) \approx \frac{1}{m} \sum_{i=1}^{m} v_i^T H v_i.
\]

Applying the trace estimator for \( m = 1 \), Equation (5.20) becomes

\[
\phi_k^\beta(w_k) = v^T \text{diag}(\tau) \left( F_k^T W_k F_k + G \right)^{-1} v + \beta e^T w_k,
\]

where \( e \) is a vector of ones, \( G = \sum_{j=1}^{k-1} F_j^T W_j F_j + \alpha L^T L \), and \( v \) is a random vector with equally distributed values of 1 and \(-1\).

Note that computing the the gradient of \( \phi_k^\beta(w_k) \) involves computing the gradient of an inverse matrix. This computation is carried out by defining \( z \) such that

\[
z = (F_k^T W_k F_k + G)^{-1} v \iff (F_k^T W_k F_k + G)z = v, \quad (5.24)
\]

recalling that \( W_k = \text{diag}(w_k) \), and differentiating implicitly to obtain that

\[
F_k^T \text{diag}(F_k z) + (F_k^T \text{diag}(w_k) F_k + G)J = 0. \quad (5.25)
\]
5.3. Numerical optimization

Solving for the Jacobian $J = \frac{\partial z}{\partial w}$ results in,

$$J = -(\mathbf{F}_k^\top \text{diag}(\mathbf{w}_k) \mathbf{F}_k + \mathbf{G})^{-1} \mathbf{F}_k^\top \text{diag}(\mathbf{F}_k \mathbf{z}) .$$

(5.26)

Substituting $J$ back into $\nabla \phi^\beta_k$, transposing, defining the matrix $\mathbf{C} = \mathbf{F}_k^\top \mathbf{W}_k \mathbf{F}_k + \mathbf{G}$, and the vectors, $\mathbf{y}$ and $\mathbf{z}$ by,

$$\mathbf{Cz} = \mathbf{v} \quad \text{and} \quad \mathbf{Cy} = \tau \odot \mathbf{v},$$

yields the gradient of the design function,

$$\nabla \phi^\beta_k(\mathbf{w}_k) = -(\mathbf{F}_k \mathbf{z}) \odot (\mathbf{F}_k \mathbf{y}) + \beta \mathbf{e},$$

(5.27)

where $\odot$ is the Hadamard product.

Unfortunately due to the inclusion of the monitor function $\tau$, the computation of the gradient is slightly more complicated compared with the classic A-optimal design case. Here the optimization requires solutions to two linear systems, $\mathbf{Cz} = \mathbf{v}$ and $\mathbf{Cy} = \tau \odot \mathbf{v}$ compared with a single linear system for the classic A design. However, since only matrix vector products are required, the conjugate gradient method is used to solve these systems avoiding ever forming $\mathbf{C}$, since it requires only matrix vector products.

Due to the non-linearity of the design problem, iterative solution methods are required. Here, a gradient descent method is used for the solution of the problem with a backtracking line search to solve for $\mathbf{w}_k$. 
5.3.2 Inexact dynamics

To minimize the noisy design function $\Phi_k^\beta(w_k)$, the Hutchinson trace estimator is again used to approximate the trace operation in Equation (5.22),

$$
\Phi_k^\beta(w_k) = \overline{v}^T \text{diag}(\overline{\tau}) \left( (I \otimes F)^T W_k (I \otimes F) + \overline{T}^T Q_k^{-1} \overline{T} + \alpha I \otimes (L^T L) \right)^{-1} \overline{v},
$$

(5.28)

where $\overline{\tau} = (\tau_0, \tau_1, \ldots, \tau_k)^T$ and $\overline{v}$ is a $(k+1)N \times 1$ vector of evenly distributed values of $[-1, 1]$. As in the noiseless case, again note that the computation of the gradient of $\Phi_k^\beta$ involves the computation of the gradient of a large dense inverse matrix. Therefore, the vector $\overline{z} = (z_0, \ldots, z_k)^T$ is defined such that

$$
\left( (I \otimes F)^T W_k (I \otimes F) + \overline{T}^T Q_k^{-1} \overline{T} + \alpha I \otimes (L^T L) \right) \overline{z} = \overline{v}.
$$

(5.29)

Differentiating with respect to $w_k$, and solving for $\overline{J} = \frac{\partial \overline{z}}{\partial \overline{v}}$ results in

$$
\overline{J} = -\left( (I \otimes F)^T W_k (I \otimes F) + \overline{T}^T Q_k^{-1} \overline{T} + \alpha I \otimes (L^T L) \right)^{-1} B.
$$

(5.30)

where

$$
B = \begin{pmatrix}
0 \\
0 \\
\vdots \\
F^T \text{diag}(Fz_k)
\end{pmatrix}.
$$

(5.31)
Substituting $\mathbf{J}$ back into the gradient of the design objective function and transposing gives,

$$\nabla \Phi_k^\beta (\mathbf{w}_k) = -\mathbf{B}^\top \mathbf{y} + \beta \mathbf{e}, \quad (5.32)$$

where

$$\mathbf{y} = \left( (\mathbf{I} \otimes \mathbf{F})^\top \mathbf{W}_k (\mathbf{I} \otimes \mathbf{F}) + \mathbf{T}^\top \mathbf{Q}_k^{-1} \mathbf{T} + \alpha \mathbf{I} \otimes (\mathbf{L}^\top \mathbf{L}) \right)^{-1} (\mathbf{r} \otimes \mathbf{v}). \quad (5.33)$$

To solve for $\mathbf{z}$ and $\mathbf{y}$, the preconditioned conjugate gradient method was used to avoid forming the large matrices. The steepest descent method was used for the non-linear design optimization.

### 5.4 Numerical examples

The adaptive design method is demonstrated for the reservoir monitoring simulation using the tomography experiment and tracer advection. The survey was designed to fully demonstrate the ability of the method to track motion and areas of interest in the model. The example consists of a circular volume of tracer moving vertically. The dynamics governing the motion of the tracer are described by the tracer advection model, and imaged using the borehole seismic tomography survey.

The partial differential equations governing both the seismic tomography experiment (Equation (2.1)) and the tracer dynamics, (Equation (2.3)) are discretized on a 2D computational domain $\Omega = [0, 400] \times [0, 100]$m, divided into $[100, 50]$ cells of width $h = [4, 2]$m. The fluid velocity field $\mathbf{u}$ was computed by solving Equations (2.3b) and (2.3c) with a constant hydraulic conductivity for $10 \text{m}^3 \cdot \text{day/kg}$, and pumping rate of $150 \text{m}^3 / \text{s}$. 
5.4. Numerical examples

Figure 5.1: Initial experiment setup. Tomography sources are pictured in green on the left and receivers in blue on the right. The initial setup covers the entire flow domain with 20 sources and 30 receivers. There is one source of fluid, pumped at a constant rate from the top green point, and a sink at the bottom (red) of the domain. Flow is in the downward direction.
5.4. Numerical examples

5.4.1 Regularization

Two different inversions were carried out to recover models $m_k$ by applying two different regularization functions. First a linear inversion was performed where the regularization operator $L = I$ was assigned to promote smallness in the recovered models. Alternately, $L = \text{GRAD}$ is a common choice to promote smooth edges in the models. However, since the original model has sharp edges, using a smooth recovery may lead to more inaccurate results.

Second, a non-linear inversion was performed with smoothed total variation (TV) as the regularization (Ascher et al., 2006). The total variation regularization promotes discontinuous boundaries, or sharp edges. This is particularly valid for this experiment since the tracer advection model does not include a diffusion term. The tracer model is concentrated in the circle and zero everywhere else. In this ideal case there should be no diffusion of the tracer, and thus the boundaries should remain sharp as the tracer moves along in the fluid velocity field.

In continuous space, smoothed total variation is described by the following relation,

$$R(m) = \int \phi(|\nabla m|) \, d\mathbf{x}, \quad (5.34)$$

where the convex function $\phi(c) = \sqrt{c^2 + \epsilon}$, is an approximation to the $\ell_1$ norm.

A standard discrete approximation for the smoothed total variation regularization found in (Ascher et al., 2006) is,

$$R(m) = h^2 e^\top \sqrt{A_f^c ((\text{GRAD} \, m) \odot (\text{GRAD} \, m))} + \epsilon, \quad (5.35)$$

where the gradient operator, \text{GRAD}, maps from cell centers to cell faces, and the averaging matrix $A_f^c$ maps from faces to cell centers.
5.4. Numerical examples

Although the optimal design method was formulated based on a linear estimation of \( m_k \), it is not unreasonable to estimate \( m_k \) using the optimal design for the linear model estimation problem with a non-linear regularization, provided that this will contribute to recovering the best estimates of \( m_k \). As can be seen in Figure 5.2 below, the overall relative model error for estimates recovered with the total variation regularization is significantly lower than that of the models estimated using the linear regularization. Thus the estimates obtained using the TV regularization are more accurate with respect to the true model for this particular problem. Since the monitor function has a very large impact on the future optimal designs, it is desirable to get the best possible estimate of \( m_k \).

![Figure 5.2: Plot of the relative error \( \frac{\| m_1^{\text{true}} - \hat{m}_1 \|}{\| m_1^{\text{true}} \|} \) versus the regularization parameter \( \alpha \) for both the linear regularization and smoothed total variation. The TV regularization gives a better estimate of the model for the case of a sharp target.](image-url)
5.4.2 Monitor function

To design only for the tracer and not the entire reservoir, the monitor function delta is defined by 
\[ \delta = | \mathbf{m}_k - \mathbf{m}_b | , \]
where \( \mathbf{m}_b \) is the background model (that is, the model without the dynamic target). Because neither the estimated model nor the background model are perfect, a threshold value \( \theta \) was chosen to remove unwanted or erroneous information from \( \delta \) where the background model \( \mathbf{m}_b \), is the model used to compute the velocity field in Equation (2.3c).

5.5 Design results: exact dynamics

For the exact case, data were simulated by marching the tracer along in time for time steps of 25 days and measured by conducting a tomography survey at each time with 4% Gaussian noise added to each data set. In total 9 experiments were conducted for times \( t_0, t_1, ..., t_8 \).

The initial design for time point \( t_0 \) did not include the dynamics. This amounts to setting \( \tau_0 = 1 \) everywhere. The design optimization problem in this case is identical to the static case and produces an optimal design based only on the physics and information from the linear regularization operator \( \mathbf{L} = \mathbf{I} \). This is apparent in both the plot of the weights for experiment 1 at time \( t_0 \) in Figure 5.3 and the image of the rays in Figure 5.5, where the design specifies fewer rays that cover the entire domain.

For each experiment the amse was plotted versus the number of nonzero weights for a set of penalty parameters \( \beta \), see Figure 5.3 for examples for times \( t_0, t_2, t_4 \). From these curves, the best set of weights were chosen such that the weights were sparse, but also so the amse was kept reasonably small. After a set of weights was chosen, both \( \mathbf{m}_0 \) and the current model \( \mathbf{m}_k \) were reconstructed from the reduced set of data \( \mathbf{d}_k \), in addition to reconstruction obtained from the original A-optimal design from \( t_0 \), and using all data for both the linear regularization and TV. The values of \( \mathbf{d} \) which
5.5. Design results: exact dynamics

Figure 5.3: Exact dynamics: plots of the adapted mean squared error amse vs. the number of nonzero weights (left column), and the weights used to conduct experiments 1, 3, and 5, at times $t_0$, $t_2$, $t_4$. 
5.5. **Design results: exact dynamics**

Contribute to the reduced data set correspond to the non-zero weights of $w_k$. The total number of data required to image the models is significantly reduced. In most cases the number of data are on average 60 of a possible 600.

The model error was then calculated for each case with respect to the ground truth, true model for each experiment, and is plotted in Figure 5.4. It is clear from the figure that models estimated using all data have the lowest mse compared with the true model. However, those constructed with the reduced set, particularly when using the TV regularization, are not far off. This implies that one can gain a reasonable estimate of the model from a significantly reduced set of data, without giving up too much on accuracy.

In Figure 5.5, the rays corresponding to $d_k$ used to estimate the current models are pictured in blue over an image of the model. Note that the rays tend to follow the target as it moves through the domain and pass through the tracer. There are some spurious rays that do not pass through the tracer which are included in the optimal design. However, since the design depends on the monitor function $\tau(\hat{m}_0)$, and the reconstructions of $\hat{m}_0$ are never exact, these spurious rays are expected.

It is also apparent that the number of spurious rays increases further in time even though estimates of $m_0$ improve as more data are collected. In particular, at time $t_6$ it appears that the design algorithm has a harder time generating a design that captures the target well. There are many more rays which do not pass through the target. This is partially due to the increasing number of multiplications by the transport matrix $T$ as time progresses, compounding errors, but also the loss of mass as the tracer exits the domain out of the sink. This is also apparent in the error calculations in Figure 5.4, where even while estimating models using all data, the error increases over time. However, the designs still provide enough information to recover models which indicate the location of the tracer. The reconstructions of $m_k$ with smoothed total variation as the regularization are much closer to the true models, even when the
5.5. Design results: exact dynamics

Figure 5.4: Exact dynamics: relative error plots. Model error is compared for each experiment for both the linear and TV regularization.
optimal design was computed for the linear regularization model estimation problem. Again, this is to be expected since the true model has discontinuous edges, and since there is no diffusion in the flow model.

5.6 Results: inexact dynamics

For the first experiment $w_0$, the initial experiment computed in the noiseless example was used in the subsequent design estimations. The covariance matrices $Q_k$ were assigned a scalar value for the variance for each time step to represent independently and identically distributed (iid) noise in the dynamics, such that $Q_k = \sigma_k I$. A constant value was assigned for the standard deviation, $\sigma_k = 0.2$, for all experiments. Models were reconstructed using the non-linear smoothed total variation regularization, and also with the linear regularization. The model estimation error is plotted in Figure 5.6.

Models were estimated using all data, the initial A-optimal design for each experiment, and finally using the adaptive method reduced data set. It is again clear that inversion carried out using all data generated models with the lowest mse, however, the benefit of reducing the number of measurements to approximately 60 from 600 might be much greater than the loss in accuracy in the model estimates, depending on the cost of data acquisition. The designs, and models estimates are pictured in Figure 5.7 below. Notice that in this case there are fewer spurious rays. It is apparent that adapted A-optimal designs were again able to track the motion of the tracer with a significantly reduced set of measurements.

5.7 Summary

In this chapter, a new method for the design of experiments for dynamical systems was presented. The method generates survey designs which adapt to the motion of
Figure 5.5: Exact dynamics: optimal designs and recovered models for 9 experiments. The top row shows the rays and the recovered model, with the number of rays (\(#d = 297\) given below, followed by the true models in the second row, the models recovered using total variation in the third row, and finally the models recovered using the linear gradient regularization.
Figure 5.6: Inexact dynamics: relative error plots. Model error is compared for each experiment for both the linear and TV regularization specific areas of the model while incorporating historic data. The motivation for not simply applying static optimal design methods to the dynamic problem is that the mean squared error of the current estimated model never contains information from previous data. To this end, the design optimization problem “knows” nothing about the changes in the model due to the dynamics and thus designs that are based on classical formulations only reflect information from the physics and the regularization (or prior) of the estimation problem.

The adaptive optimal experimental design approach is based on the introduction of a monitor function to scale the mean squared error of an estimator whose motion is governed by a dynamical system, according to historic data. To determine a design for a future experiment, the adaptive mse (amse) is minimized with the added constraint
5.7. Summary

that the design is sparse. Two model estimation problems were considered, whose construction depends on the characterization of the error in the dynamical system. This leads to experimental designs that track the changes in the model with a reduced set of measurements. The methodology was tested using seismic tomography to image the advection of a tracer in a reservoir.
5.7. Summary

Figure 5.7: Inexact dynamics: optimal designs and recovered models for 9 experiments. The top row shows the rays and the recovered model, with the number of rays given below. The second row pictures true models, the third shows the models recovered using total variation, and finally the bottom shows models recovered using the linear gradient regularization.
Chapter 6

Conclusions

The main objective of the research presented in this thesis was to improve reservoir monitoring and forecasting by coupling geophysical survey techniques with a dynamic fluid flow model and optimally designing the geophysical survey.

Forecasting of flow within a reservoir requires knowledge of fluid flow parameters, such as hydraulic conductivity. Estimates of such parameters can be obtained by direct measurements from subsurface rock samples, or through the inversion of flow data. However, these estimates are often highly inaccurate due to sparse measurements of both the parameters and flow data.

Therefore, one approach investigated in this thesis is to include a higher sampling of measurements without adding a large additional cost to the monitoring program, was to couple a geophysical survey with the dynamic fluid flow model in a single inverse problem, such that the unknown flow parameters could be estimated from geophysical data.

Geophysical surveys tend to cover much larger spatial areas at lower cost since sensors can be deployed at the surface as well as below ground. Collecting fluid flow data requires expensive boreholes, and is thus sparsely sampled. In the context of an ill-posed under-determined inverse problem, it is very difficult to estimate parameter models for large subsurface domains with little data. Thus, the greater number of data and spatial coverage improves the estimates obtained by solving the coupled ill-posed inverse problem.
Chapter 6. Conclusions

The second approach to improve reservoir monitoring proposed in this thesis was to compute optimal geophysical survey designs for data measurements for the coupled subsurface flow inverse problem. In this case, it is inefficient to be collecting data over regions of the subsurface where there is no change to the geophysical properties due to flow. Ideally, one would want to be able to adapt the survey with the changes in the subsurface flow. However, traditional optimal design methods do not utilize historic estimates of the subsurface models in a way that includes them in the design optimization problem. To this end, the idea was to formulate an optimal design estimation method for the coupled geophysics and flow model inverse problem that includes historic model estimates.

To demonstrate the ideas proposed in the thesis, the physical models for the geophysical survey and the fluid flow model were first presented in Chapter 2. The choice of models was particularly important for the coupling problem. The seismic tomography experiment was chosen primarily due to its linearity, but also due the fact that data can be collected from existing boreholes as well as from sensors placed on the surface. Borehole seismic tomography is also commonly used in practice, with a wide variety of applications, making it an ideal survey for the model reservoir monitoring problem.

The tracer advection model was used because of its linearity and because it is very similar mathematically to more complex multiple phase flow models. In addition, as was shown in Section 2.3, the advection of the tracer can also be thought of as the advection of the seismic slowness. This result eliminates the need for a petrophysical function relating the geophysical properties (seismic slowness) to flow parameters (tracer or hydraulic conductivity). Technically, the tracer advection model lends itself to efficient numerical optimization schemes, in particular a stable and differentiable discretization of the physical models was applied. For the discretization of the flow a Particle-In-Cell method was used, whose stability is independent of the time-step
size and therefore can be safely applied, even if the fluid velocity field is unknown. The fluid velocity field was discretized on staggered grids, resulting in a stable discretization of the regularization operators and flow constraints, see Section 2.4 for details.

Following the introduction of the physical models, Chapter 3 provides a brief introduction to linear inversion and highlights the different formulations which result from different characterizations of the noise in the dynamical fluid flow model: zero noise and non-zero noise. The choice of noise characterization is important not only because it results in different formulations of the inverse problem, but also because it is often difficult to estimate a priori.

In Chapter 4 a new methodology for estimating the initial tracer concentration and fluid velocity field from geophysical data was presented for the case where the noise in the dynamical flow model was assumed to be zero for all times. The new estimation problem was formulated as a constrained inverse problem with a specifically tailored regularization for the velocity field that promotes discontinuities in the tangential components of the estimated flow field. Estimates of the velocity were then used for the estimation of the hydraulic conductivity.

Numerical results for a simple layered earth model with a homogeneous isotropic reservoir demonstrated that this new approach yields not only accurate estimates of the initial change to the slowness but, especially, accurate predictions of the fluid flow velocity tracer evolution. The work was published in the SIAM Journal on Scientific Computing (Fohring et al., 2014). However, a number of topics relating to the coupled geophysics flow problem have yet to be studied.

One important aspect of the inverse problem is the evaluation of uncertainty in the flow. Although the error in the dynamic flow model was assumed to be zero, it is likely that measurement and numerical errors were propagated with the dynamics throughout the inverse problem. This was particularly apparent when the number
of historic experiments, and thus time steps, increased. With increased times steps, came an increase in the difficulty in recovering an accurate approximation to the true velocity field. This is counter to what one would expect; that more information should provide better results, and is likely due to the noise propagation. In particular, the particle in cell (PIC) discretization of the tracer advection equation relies on a bilinear interpolation. Bilinear interpolation error is quadratic, and thus with each time step, and therefore each multiplication of the tracer by the interpolation matrix, the error grows. Thus experimenting further with higher order interpolations may help to reduce error propagation. However, higher order interpolation functions may simultaneously add complexity to the inverse problem because of the difficulty associated with computing gradients.

An additional topic of research involves applying the coupled problem with more complex flow models. However, the use of such models might require the knowledge of petrophysical parameter functions and the application of differing discretization techniques that are not necessarily stable for all time step lengths. This would lead to added complexity in the inverse problem, which would require innovative solution techniques.

Finally, to truly demonstrate the ability of the method to recover fluid flow parameters, the conduction of a field or laboratory experiment to obtain a real data set is required to fully evaluate the contribution of the work to reservoir monitoring and geophysical history matching.

To address the second goal of the thesis, Chapter 5 presents a new adaptive optimal experimental design method for the linear coupled reservoir monitoring problem. The new adapted mean square error (amse) method builds on classical A-Optimal design criteria by altering the posterior mean squared error (mse) of a model estimate to include historic data. In the classic case the mse is minimized to recover an optimal set
of measurements. Since the mse for the linear estimator includes only the physics and the prior covariance matrix, designs do not adapt with the motion of the tracer. The amse was therefore defined to include historic data through an introduced monitor function. Results presented in Section 5.4 demonstrate the amse method for the coupled seismic tomography tracer advection. Applying the method produced designs that tracked the motion of the tracer, while requiring significantly fewer measurements to recover images of the tracer compared with using all data available. The amse optimal design method research presented in the thesis is currently in the final review process for publication in SIAM Journal on Uncertainty Quantification (Fohring and Haber, 2016).

The idea to include historic information in the amse through the introduction of a monitor function originates from adaptive mesh methods. In this thesis only two options for the monitor function were chosen and tested. Thus there is significant room for further investigation of how differing monitor functions can affect optimal designs.

The adapted design criteria presented in the thesis, only included reducing the number of measurements from a maximum set. However, in many cases it might be more efficient to eliminate entire borehole locations, instead of just individual sensor locations, and additionally impose a monitory cost on more expensive measurement locations to quantitatively estimate cost reduction.

Another avenue of investigation includes designing for dynamic coupled non-linear inverse problems. Current non-linear design methods are limited by the fact that there is no closed form calculation of the posterior mean squared error. This results in expensive bi-level design optimization problems that require training sets of possible models. However, for the coupled dynamic problem there maybe room to include historic estimates through propagation of earlier time models to generate training sets.
Chapter 6. Conclusions

Recently (Alexanderian et al., 2015) presented a method for A-optimal design for non-linear inverse problems where the posterior covariance matrix is approximated by a stochastic average over a set of data by an approximate Hessian. The method was demonstrated for an advection diffusion flow problem. Building on this method might be an interesting approach to introducing the adaptive design method for non-linear problems.

One might also ask how sensitive the adaptive optimal design is to variations in the fluid velocity field. This could be investigated by estimating and updating the flow dynamics in time, and would amount to developing an adaptive optimal design method for the constrained variable projection problem, that is, for the non-linear flow coupled inverse problem presented in Chapter 4.
Bibliography


Bibliography


