Fréchet Sensitivity Analysis and Parameter Estimation in Groundwater Flow Models

Vítor Manuel Leite dos Santos Nunes

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Mathematics

Jeffrey Todd Borggaard, Chair
John Allen Burns
Slimane Adjerid
Lizette Zietsman

2 April 2013
Blacksburg, Virginia

Keywords: Fréchet derivative operators, groundwater flow models, parameter estimation, parameter zonation, sensitivity analysis, uncertainty quantification.

Fréchet Sensitivity Analysis and Parameter Estimation in Groundwater Flow Models

Vítor Manuel Leite dos Santos Nunes

(ABSTRACT)

In this work we develop and analyze algorithms motivated by the *parameter estimation* problem corresponding to a multilayer aquifer/interbed groundwater flow model. The parameter estimation problem is formulated as an optimization problem, then addressed with algorithms based on adjoint equations, quasi-Newton schemes, and multilevel optimization. In addition to the parameter estimation problem, we consider properties of the parameter to solution map. This include invertibility (known as *identifiability*) and differentiability properties of the map. For differentiability, we expand existing results on Fréchet sensitivity analysis to convection diffusion equations and groundwater flow equations. This is achieved by proving that the Féchet derivative of the solution operator is *Hilbert–Schmidt*, under smoothness assumptions for the parameter space. In addition, we approximate this operator by time dependent matrices, where their singular values and singular vectors converge to their infinite dimension peers. This decomposition proves to be very useful as it provides vital information as to which perturbations in the distributed parameters lead to the most significant changes in the solutions, as well as applications to uncertainty quantification. Numerical results complement our theoretical findings.

Acknowledgments

I would like to thank my parents, Suzete and Cecilio, my sisters Soraia and Clara, my Godson Afonso and my cousin Ricardo, for being supportive all these years in my decision on being a mathematician; my wife to be, Nicole, for being by my side in graduate school; my advisor Professor Jeff Borggaard for challenging me throughout my PhD and helping me grow as a mathematician, both analytically and numerically; Professor Lizette Zietsman for mentoring me in the SIAM student chapter; Professor Burns for introducing me to sensitivity analysis; Professor Adjerid for teaching finite elements; Professor Elgart for teaching me fun and interesting mathematics; and Meijing Zhang and Professor Burbey for helping me comprehend groundwater flow modeling. I also would like to thank Mariana and the Polanah family for adopting me as their own, as well as my ICAM friends Weiwei, Hans, Zhu, Boris, Erich, Chris, and Dave.

Contents

List of Figures

Notation

$[D_q[z(t;q)]]_{q=q_0}$	Fréchet derivative of the solution operator $z(t;q)$ with respect to the parameter q at q_0
$[D_q[z(q)]]_{q=q_0}$	Fréchet derivative of the solution operator $z(q)$ with respect to the parameter q at q_0
H	preconsolidation head
h	water head
h^a	water head in the aquifer
h^I	water head in the interbed
T	transmissivity
$\mathbf{K} = (K^x, K^y, K^z)$	hydraulic conductivity
W	groundwater flow source term
k_v	vertical hydraulic conductivity of the interbed
S	specific storage
S_{ske}	elastic skeletal storage
S_{skv}	inelastic skeletal storage
λ, μ	Lamé constants (when used together)
α	Biot-Willis constant
μ_f	fluid viscosity
c_0	constrained specific storage
κ	symmetric permeability tensor
$\mathbf{u}$	displacement
p	pressure
J	cost functional
q	general parameter
$A(q)$	differential operator (depending on parameter q)
Q	parameter space
Q^{ad}	admissible parameter set
Ω	spatial domain, subset of $\mathbb{R}^n$
$\partial\Omega$	boundary of Ω
$\vec{n}$	unit boundary normal vector
λ	Lagrange multiplier
Π_k	interpolation operator of degree k
H^N	finite element space of dimension N
S^N	finite element space of dimension N (solution space)
Q^M	finite element space of dimension M (parameter space)

Chapter 1

Introduction

Since the early stages of human history, the search for water supplies has been a fundamental problem. This alone justifies the importance of modeling and identifying model parameters (parameter estimation) of groundwater flows. Since most of the water is stored beneath the surface; it is very challenging to obtain the information on those parameters that model the distribution the water in reservoirs. Most scholars approach the problem within the framework of poroelastic media models, where the fact that the water flow has a direct impact on the shape and storage properties of the reservoir in question is considered. For instance the processes of dryness by pluming or change in the volume for natural reasons on an aquifer leads to subsidence. The data used in parameter estimation has two distinct forms. The subsidence data is provided by satellite, whereas the volume of water is by measurements on wells. Here the major problem is that the number of wells is relatively small when compared to the size of the land in question. For example, one can have wells that they are dozens of miles apart. Therefore one will need to interpolate the data spatially, it will also be necessary to interpolate the data in time since it is not collected constantly.

In this work water supplies from the Las Vegas Valley is considered. The objective is to improve the existence model and to estimate relevant parameters. The forward problem is modeled numerically by MODFLOW [13] , [22] and [44], [55], as well as a solver that we developed. The parameters, specific storage and transmissivity can be estimated by UCODE [40], or the adjoint quasi Newton optimization solver that we developed. The major problem of UCODE is that it can only estimate the parameter values once their distribution in space is known (zonation). Therefore we developed an algorithm that coupled with UCODE can estimate the parameter values without prior knowledge of the zonation. Both UCODE and MODFLOW use a finite volume scheme to approximate the partial differential equation (PDE).

Once the parameters are identified, the study take another direction, namely determining how those parameters impact the solution or the system's output. This is process is called sensitivity analysis. Sensitivity analysis gives insight into the impact of changes on the distribution of the porous media, wells and even the existence of a construction site on the water reservoir. This is done by considering the data or the solution of a PDE as a function of the distributed parameters, boundary conditions and/or forcing term, since this functional

is well defined on the set of parameters which leads to a unique solution. Under certain assumptions it is possible to differentiate the operator with respect to the parameters of interest. This derivative is an operator called the Fréchet derivative. Motivated by groundwater sensitivity analysis, we developed results on the steady advection-diffusion equation and convection-diffusion equation. These results include the existence of Fréchet derivatives in Hilbert spaces, their spectral decomposition and numerical approximation. In previous works by Herdman and Spies [21] and Seubert and Wade [44], the differentiation is considered in Banach spaces. Although these spaces are more general than Hilbert spaces, they lack some essential mathematical structure for considering approximations. Furthermore, the additional Hilbert space structure allows us to prove that if one assumes some smoothness in the PDE parameters and source term, the Fréchet operator is Hilbert-Schmidt, the infinite dimensional equivalent to the singular value decomposition (SVD). This decomposition allows us to identify which parameters are the most important, reduce the dimension of the Fréchet derivative operator, and create an adaptive meshing strategy. All of these results are developed within an infinite dimensional framework. There are several advantages of such a decomposition.For example, in parameter estimation the gradient is explicitly computed and a lower dimensional representation of it can be computed by truncating the Hilbert-Schmidt expansion.

Outline

This work starts by discussing groundwater flow (GWF) models and their particularities as their approximation technics. The numerical approximation of the flow is not straight forward. One must take aspects such as mass conservation, subsidence and the low values of the storage parameter into account. The main model is based on Leake's model [30] and is presented in Chapter 2. This is followed by the discussion of the existence and uniqueness of the solution of the inverse problem and the Lagrangian associated with it in Chapter 3. The next step is the development of techniques to estimate the parameters and include techniques such as djoint methods and zonation algorithms. The parameter estimation study is followed by theoretical and numerical sensitivity analysis results. Chapters 6 and 7 are dedicated to numerical results for sensitivity analysis and parameter estimation.

Chapter 2

Groundwater Flow Models

2.1 PDE Models

2.1.1 Introduction

In this chapter we present three groundwater flow (GWF) models that are used to study large aquifers. The first is known as the classical model, developed in Section 2.1.2, and is the basis of the U.S. Geological Survey (USGS) package MODFLOW [24]. This model is generalized to a multilayer model in Section 2.1.3 and accounts for land subsidence and is implemented in the USGS SUB Package [30]. The multilayer model is a simplification of Biot's general theory of three-dimensional consolidation [7] and is given in Section 2.1.4. While we do not use the resulting poroelastic model in any of our work, this model of the pressure head and the elastic displacement of the aquifer serves to enhance the understanding of the approximations made in the multilayer model. The extension of the results of this dissertation to the poroelastic equations is outlined Chapter 9. Following the introduction of these models, we review numerical discretization approaches in Section 2.2. This includes finite difference and finite element approximations of the classical model in Sections 2.2.1 and 2.2.2, respectively. The finite element approximation is used for our work on Fréchet sensitivity analysis. A combined finite volume/finite difference approximation for the multilayer model is provided in Section 2.2.3. This combined approximation is used by MODFLOW.

2.1.2 The Classical Model

The classical PDE model for describing the piezometric head h for a saturated flow combines mass balance with Darcy's law. The result (assuming anisotropic hydraulic conductivity, cf. [5, p. 180]) is

$$\begin{cases} S\frac{d}{dt}h(t,\mathbf{x}) = (K^x h_x(t,\mathbf{x}))_x + (K^y h_y(t,\mathbf{x}))_y + (K^z h_z(t,\mathbf{x}))_z + W(t,\mathbf{x}), \\ h(0) = h_0, \qquad h_0 \in H_0^1(\Omega), \end{cases} \tag{$\mathcal{C}_q$}$$

for all $t \in [0,T]$, and $\mathbf{x} \in \Omega$, a convex subset of $\mathbb{R}^3$. The spatially varying functions K^x, K^y and K^z represent the hydraulic conductivity in the x, y and z directions respectively. THe storativity is denoted by $S = S(\mathbf{x})$ and W is the source term. The source term models effects such as rain and well pumping.

Theorem 2.1.1 (Existence and uniqueness of solutions to the classical model)**.** *Let $\Omega \subset \mathbb{R}^3$ be bounded, $S \in Q^s := \{q \in C(\overline{\Omega}) : \ \exists \alpha_s > 0$ such that $q(\mathbf{x}) > \alpha_s \ \forall \mathbf{x} \in \Omega\}$, $K^x, K^y, K^z \in Q^k := \{q \in C(\overline{\Omega}) : \exists \alpha_k > 0$ such that $q(\mathbf{x}) > \alpha_k \ \forall \mathbf{x} \in \Omega\}, W(t,\cdot) \in L^2(\Omega)$ for all $t > 0$, and W satisfies the condition that there exist constants $v, \beta > 0$ such that*

$$\|W(t,\cdot) - W(s,\cdot)\|_{L^2(\Omega)} \leq \beta |t-s|^v \qquad \text{for all } s,t > 0.$$

Then there exists an unique solution $h \in C^1(]0,\infty[\ ; H^2(\Omega) \cap H_0^1(\Omega))$ to $(\mathcal{C}_q)$.

Proof. The proof of this theorem is a direct consequence of Corollary 2.8 in Chapter 7, Pazy [37] since the operator

$$A(\mathbf{x})(\cdot) = \frac{1}{S(\mathbf{x})}[(K^x(\mathbf{x})(\cdot)_x)_x + (K^y(\mathbf{x})(\cdot)_y)_y + (K^z(\mathbf{x})(\cdot)_z)_z]$$

is strongly elliptic. The strong ellipticity follows immediately from our assumptions on S and (K^x, K^y, K^z). □

The main limitation of the classical model is that it does not account for the fact that an aquifer is a poroelastic medium. In other words, as flow occurs in the aquifer, the domain shape changes through time. This displacement leads to a change in the storativity parameter, S.

The standard correction to the classical model subdivides the domain into layers, some of which are incompressible (the *aquifer*) and others are compressible (the *interbeds*). These layers each have different physical properties: thickness, storativity, and conductivity. The layers are approximated by coupling layers of two dimensional models. The individual layer models are constructed by averaging the quantity of water in the vertical direction, this is discussed in the next section.

2.1.3 The Multilayer Model

The previous model is not practical for real aquifers since it does not incorporate subsidence and its effect on the specific storage. Furthermore, most aquifers are well approximated by layers of uniform material. Numerical approximations of $(\mathcal{C}_q)$ do not take specific advantage of this distribution of the aquifer specific storage and transmissivity. Thus a model that divides the domain into a union of horizontal layers is common in the literature. The are two type of layers, *aquifer* layers and *interbed* layers, see Figure 2.1.

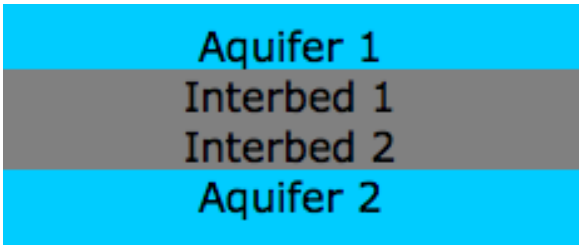

Figure 2.1: Groundwater model layers.

Their physical properties are significantly different and this leads to the following assumptions:

Aquifer Layer

- High permeability that allows water to flow in ordinary conditions.
- The water only flows horizontally within the layer.
- It is incompressible (for example, sand or gravel).

Interbed Layer

- It transmits water at a much lower rate compared to aquifer layers.
- The horizontal conductivity is assumed to be zero or almost zero.
- It is compressible (for example, clay).

The layers are constructed distinctly in order to better fit the previous assumptions. The average quantity of water in a layer l that is bounded at each point (x, y) by $b_1^{a_l}(x, y)$ (bottom) and $b_1^{a_l}(x, y)$ (top) is given by

$$\tilde{h}^{a_l}(x, y) = \frac{1}{b_2^{a_l}(x, y) - b_1^{a_l}(x, y)} \int_{b_1^{a_l}(x,y)}^{b_2^{a_l}(t,x,y)} h(x, y, z)\, dz.$$

It is relevant to point out that aquifer layers are averaged over an interval in which the *relative* location of its endpoints do not change over time. This is in contrast with the similar definition for interbed layers where the relative locations of endpoints may evolve over time because they are assumed to be compressible layers. Within interbed layers, we assume that $K^x(x, y, z) \equiv 0$ and $K^y(x, y, z) \equiv 0$. Thus, the flow within interbed layers is vertical, and contributes a source/sink to adjoining layers.

Aquifer Layer Model Equations If we consider transmissivity to be equal in each direction ($\mathcal{C}_q$) can be reduced to the following PDE:

$$\begin{cases} S(\mathbf{x})\frac{d}{dt}h(t,\mathbf{x}) = \nabla\cdot(K(\mathbf{x})\nabla h(t,\mathbf{x})) + W(t,\mathbf{x}), & h(t,\cdot)\in H^1(\Omega)\ ; \\ K(\mathbf{x})\nabla h(t,\mathbf{x})\cdot\vec{n} = 0 & \mathbf{x}\in\partial\Omega \\ h(0) = h_0 & h_0\in L^2(\Omega),\ t\in[0,T]. \end{cases} \tag{2.1}$$

The source term W is now more complex since it incorporates the discharge/recharge from adjacent (aquifer/interbed or interbed/interbed) layers as well as all other possible sources or sinks. Integrating both sides of (2.1) in the vertical direction z over the interval $[b_1^{a_l}(t,x,y), b_2^{a_l}(t,x,y)]$ (and ignoring the arguments of $b_i^{a_l}$) results in:

$$\int_{b_1^{a_l}}^{b_2^{a_l}} S(\mathbf{x})\frac{d}{dt}h(t,x,y,z)dz = \int_{b_1^{a_l}}^{b_2^{a_l}} \nabla\cdot(K(\mathbf{x})\nabla h(t,\mathbf{x})) + W(t,\mathbf{x})\ dz \tag{2.2}$$

for the aquifer layers and, similarly for the interbed layers (keeping the t argument to emphasize that the interval is time dependent)

$$\int_{b_1^{I_l}(t)}^{b_2^{I_l}(t)} S(\mathbf{x})\frac{d}{dt}h(t,x,y,z)\ dz = \int_{b_1^{I_l}(t)}^{b_2^{I_l}(t)} \nabla\cdot(K(\mathbf{x})\nabla h(t,\mathbf{x})) + W(t,\mathbf{x})\ dz \tag{2.3}$$

for the interbed layers. We now apply the layer modeling assumptions to the $\nabla\cdot(K\nabla h)$ terms in the equations above. In the aquifer layers,

$$\begin{aligned} \int_{b_1^{a_l}}^{b_2^{a_l}} \nabla\cdot(K\nabla h(x,y,z))\,dz &= \int_{b_1^{a_l}}^{b_2^{a_l}} \nabla_{(x,y)}\cdot\left(K\nabla_{(x,y)}h(x,y,z)\right) + \frac{\partial}{\partial z}\left(K\frac{\partial}{\partial z}h(x,y,z)\right)dz \\ &= \int_{b_1^{a_l}}^{b_2^{a_l}} \nabla_{(x,y)}\cdot\left(K\nabla_{(x,y)}h(x,y,z)\right)dz + \left(K\frac{\partial}{\partial z}h(x,y,z)\right)\Big|_{z=b_1^{a_l}}^{z=b_2^{a_l}}. \end{aligned}$$

Interchanging the $\nabla_{(x,y)}\cdot$ and $\int_{b_1^{a_l}}^{b_2^{a_l}}$ operations leads to

$$= \nabla_{(x,y)}\cdot\int_{b_1^{a_l}(x,y)}^{b_2^{a_l}(x,y)} \left(K\nabla_{(x,y)}h(x,y,z)\right)dz + K\nabla_{(x,y)}h(x,y,b_2^{a_l}(x,y))\cdot\nabla_{(x,y)}b_2^{a_l}(x,y)$$

$$-K\nabla_{(x,y)}h(x,y,b_1^{a_l}(x,y))\cdot\nabla_{(x,y)}b_1^{a_l}(x,y) + \left(K\frac{\partial}{\partial z}h(x,y,z)\right)\Big|_{z=b_1^{a_l}(x,y)}^{z=b_2^{a_l}(x,y)}.$$

By assuming there is no subsidence in the aquifer and horizontal flow, $h(t,x,y,z) = h^a(t,x,y)$, one can define, for each aquifer layer l, the transmissivity, storativity, and effective source as

$$T^l(x,y) = \int_{b_1^{a_l}(x,y)}^{b_2^{a_l}(x,y)} K(x,y,z)\ dz, \qquad \tilde{S}(x,y) = \int_{b_1^{a_l}(x,y)}^{b_2^{a_l}(x,y)} S(x,y,z)\ dz,$$

and

$$P(x,y,t) = \int_{b_1^{a_l}(x,y)}^{b_2^{a_l}(x,y)} W(x,y,z,t)\ dz.$$

Each aquifer layer then is modeled by a two dimensional PDE in $\tilde{\mathbf{x}} = (x, y)$.

Since water flows more freely in the aquifer layers than the interbed layers, we assume that the flow in the interbed is only vertical, thus $K^x = K^y = 0$. Furthermore, assuming it is an isotropic porous medium, (2.1) simplifies to a series of one dimensional PDEs (at each $\tilde{\mathbf{x}}$). The multilevel model then becomes

$$\begin{aligned}\tilde{S}\frac{\partial h_l^a(t,\tilde{\mathbf{x}})}{\partial t} =& \nabla_{\tilde{\mathbf{x}}} \cdot (T^l(\tilde{\mathbf{x}})\nabla_{\tilde{\mathbf{x}}} h_l^a(t,\tilde{\mathbf{x}})) \\ &+ K^{(l+1)}\frac{\partial h^{I_{l+1}}}{\partial z}\Big|_{z=b_1^{I_{l+1}}(t,(x,y))} - K^{(l-1)}\frac{\partial h^{I_{l-1}}}{\partial z}\Big|_{z=b_2^{I_{l-1}}(t,(x,y))} + P_l(t,\tilde{\mathbf{x}}) \qquad (\mathcal{M}_q^a)\end{aligned}$$

for aquifer layers, and

$$S_o^{(l)}\frac{\partial h_l^I}{\partial t}(z) = K^{(l)}\frac{\partial^2 h_l^I}{\partial z^2}(z) \qquad (\mathcal{M}_q^I)$$

(at each $\tilde{\mathbf{x}}$) for interbed layers. These equations are coupled through the continuity of the flux $K\frac{\partial h}{\partial z}$ across the layers. Thus, the aquifer equation $(\mathcal{M}_q^a)$ for layer l has a source term from the interbed layer above $(l+1)$ and below $(l-1)$. These terms vanish at the top and bottom boundaries and for adjacent aquifer layers.

We assume that the shape of Ω does not change vertically, in other words,

$$\Omega = \left\{(x, y, z) : (x, y) \in \Omega_2 \subset \mathbb{R}^2, z \in [\theta_{bottom}, \theta_{top}]\right\}.$$

The actual compression of the interbed is accounted for through the time-varying specific storage $S_0^{(l)}$. More details of the modeling of this term are provided in Section 2.2.3.

The boundary conditions are then inherited from (2.1) therefore:

$$\left\{\begin{array}{rcll} T^l(\tilde{\mathbf{x}})\nabla_{\tilde{\mathbf{x}}} h^{a_l}(t,\tilde{\mathbf{x}}) \cdot \vec{n}(\tilde{\mathbf{x}}) &=& 0 & \tilde{\mathbf{x}} \in \partial\Omega_2 \\ h^{I_l}(t,\tilde{\mathbf{x}}, b_1^{I_l}(t,\tilde{\mathbf{x}})) &=& h^{a_{l-1}}(t,\tilde{\mathbf{x}}) & t > 0 \\ h^{I_l}(t,\tilde{\mathbf{x}}, b_2^{I_l}(t,\tilde{\mathbf{x}})) &=& h^{a_{l+1}}(t,\tilde{\mathbf{x}}) & t > 0 \\ \tilde{h}^{a_l}(0,\tilde{\mathbf{x}}) &=& \frac{1}{b_2^{a_l}(\tilde{\mathbf{x}}) - b_1^{a_l}(\tilde{\mathbf{x}})}\int_{b_1^{a_l}(\tilde{\mathbf{x}})}^{b_2^{a_l}(\tilde{\mathbf{x}})} h^{a_l}(0,\mathbf{x})\, dz & \\ \tilde{h}^{I_l}(0,\tilde{\mathbf{x}}) &=& \frac{1}{b_2^{I_l}(0,\tilde{\mathbf{x}}) - b_1^{I_l}(0,\tilde{\mathbf{x}})}\int_{b_1^{I_l}(0,\tilde{\mathbf{x}})}^{b_2^{I_l}(0,\tilde{\mathbf{x}})} h^{I_l}(0,\mathbf{x})\, dz. & \end{array}\right. \qquad (\mathcal{BC})$$

These boundary conditions insure that the flow is continuous over the layers. The next section discusses a more fundamental PDE model based on explicitly modeling the elastic effects of the storage media which we will refer to as the *poroelastic model*, [46].

2.1.4 The Poroelastic Media Model

Although we do not delve into too much detail, a more general framework for modeling the elastic deformations in groundwater flows was analyzed by Showalter [46] and numerically simulated by Phillips [38, 39]. The poroelastic model considers the displacement produced

by the flow. This leads to a coupled system for both the pressure p and displacement $\mathbf{u}$,

$$
\begin{array}{lll}
-(\lambda+\mu)\nabla(\nabla\cdot\mathbf{u}(t))-\mu\nabla^2\mathbf{u}(t)+\alpha\nabla p(t)=\mathbf{h}(t) & on\ \Omega & \\
\frac{d}{dt}(c_o p+\alpha\nabla\cdot\mathbf{u}(t))+\frac{1}{\mu_f}\nabla\cdot\kappa(\nabla p(t)+\rho_f\mathbf{g})=h(t) & on\ \Omega & \\
p(t)=p_0 & on & \Gamma_p, \\
\kappa(\nabla p+\rho_f\mathbf{g})\cdot\nu=q & on & \Gamma_f, \\
\mathbf{u}(t)=\mathbf{u}_D & on & \Gamma_o, \\
\tilde{\sigma}(\mathbf{u})\nu=\mathbf{t}_N & on & \Gamma_t, \\
p(0)=p^o, \quad on\ \Omega & & \\
\mathbf{u}(0)=\mathbf{u}^o, \quad on\ \Omega & &
\end{array}
\tag{$\mathcal{P}_q$}
$$

where $\partial\Omega=\Gamma_p\bigcup\Gamma_f$ and $\partial\Omega=\Gamma_t\bigcup\Gamma_o$, λ,μ are the positive Lamé constants, α is the Biot-Willis constant, μ_f the fluid viscosity, c_0 the constrained specific storage, and κ is the symmetric permeability tensor. For the problems considered here, $\mathbf{u}:[0,T]\times\mathbb{R}^3\to\mathbb{R}^3$, $p:[0,T]\times\mathbb{R}^3\to\mathbb{R}_0^+$, $V:=\{v\in H^1(\Omega)\text{ subject to }v|_{\Gamma_S}=0\}$, and $\mathbf{V}:=\{\mathbf{v}\in H^1(\Omega)\times H^1(\Omega)\times H^1(\Omega)\text{ subject to }v|_{\Gamma_o}=0\}$. The physical interpretation of the operator $\mathcal{C}(\mathbf{u})-(\lambda+\mu)\nabla(\nabla\cdot\mathbf{u})-\mu\nabla^2\mathbf{u}$ is the elastic operator and $A(p)=-\partial_j(k\partial_j p)$ is the diffusion operator. By the considering $\mathbf{h}(t)=\mathbf{0}$ the first equation on 2.11 can be rewritten as $\mathbf{u}(t)=\mathcal{C}^{-1}(\nabla p(t))$ which then reduces the problem to one variable, since the second equation is then written as:

$$
\frac{d}{dt}(c_0 P p(t))+\vec{\nabla}\cdot\mathcal{C}^{-1}(\nabla p(t))+A(p(t))=h(t) \tag{2.4}
$$

Which is an implicit evolution equation, this is thoroughly studied by on chapter IV of R.E. Showalter [46]. Since it can be written as:

$$
\frac{d}{dt}\mathcal{B}(p(t))=\mathcal{A}(p(t))+h(t) \tag{2.5}
$$

Obviously the $\mathcal{K}er$ of these operators will play an important role in the uniqueness of a solution. For instance suppose that $p(t)$ is a solution and $r(t)$ is a nontrivial element of $\mathcal{K}er(\mathcal{B})\cap\mathcal{K}er(\mathcal{A})$ then $p(t)+r(t)$ will be a solution of 2.5 as long $r(t)$ satisfies initial and boundary conditions. Thus assumptions on the operators $\mathcal{A}$ and $\mathcal{B}$ must be made in order to guarantee uniqueness. Which in our case in then $\mathcal{K}er(c_0P-\vec{\nabla}\cdot\mathcal{C}^{-1}\vec{\nabla})\cap\mathcal{A}=0$. This motivates the following theorem:

Theorem 2.1.2. *If $\mathcal{K}er(c_0P-\vec{\nabla}\cdot\mathcal{C}^{-1}\vec{\nabla})\cap\mathcal{A}=0$, $T>0$, $v_0\in L^2(\Omega)$, $v_1\in L^2(\Gamma_S)$, $h_0(.)\in C^\alpha([0,T],L^2(\Omega))$ and $h_1(.)\in C^\alpha([0,T],L^2(\Gamma_S))$ then there is an unique pair of functions $p(.):(0,T)\to V$ and $\mathbf{u}(.):(0,T]\to V$ such as:*

$$
(\mathcal{P})_q\begin{cases}
(\lambda+\mu)\nabla(\nabla\cdot\mathbf{u}(t))+\mu\nabla^2\mathbf{u}(t)+\alpha\nabla p(t)=\mathbf{0} & on\ \Omega \\
\frac{d}{dt}(c_o p+\alpha\nabla\cdot\mathbf{u}(t))+\frac{1}{\mu_f}\nabla\cdot\kappa(\nabla p(t)+\rho_f\mathbf{g})=h(t) & on\ \Omega \\
p(t)=0 \quad on \quad \Gamma_1, & \\
\kappa(\nabla p+\rho_f\mathbf{g})\cdot\nu=q \quad on \quad \Gamma_f, & \\
\mathbf{u}(t)=\mathbf{0} \quad on \quad \Gamma_o, & \\
\frac{\partial}{\partial t}((1-\beta)\mathbf{u}(t)\cdot\vec{\mathbf{n}})\chi_S+k\frac{\partial p(t)}{\partial n}=h_1(t)\chi_s \quad on \quad \Gamma_f, & \\
\lim\limits_{t\to0^+}(c_o p+\alpha\nabla\cdot\mathbf{u}(t))=v_0, \quad in\ L^2(\Omega) & \\
\lim\limits_{t\to0^+}(1-\beta)\mathbf{u}\cdot\vec{\mathbf{n}}=v_1, \quad in\ L^2(\Gamma_S) &
\end{cases}
\tag{2.6}
$$

The next last section on theoretical models, will show that the classical equation $(\mathcal{C}_q)$ is a particular case of (2.5).

Groundwater as a particular case

Since the specific storage over $(\mathcal{C}_q)$ in not time dependent one can claim that $S(\mathbf{x})\frac{d}{dt}h(t) = \frac{d}{dt}S(\mathbf{x})h(t)$ thus one can define $\mathcal{B}(h) = h$ and $\mathcal{A}(h) = S(\mathbf{x})^{-1}[(K^x(\mathbf{x})h_x)_x + (K^y(\mathbf{x})h_y) + (K^h(\mathbf{x})h_z)_z]$. Then $(\mathcal{C}_q)$ is of the type:

$$\frac{d}{dt}(h(t)) = \mathcal{A}(h(t)) + W(t) \tag{2.7}$$

It is trivial to prove that $\mathcal{A} : H_0^1(\Omega) \to L^2(\Omega)$ is regular accretive and that under the assumption that $S \in Q^s$and that $K^x, K^y, K^z, \in Q^k$:

$$\lim_{\|v\|_{H_0^1(\Omega)}\to\infty} \frac{(A(v), v)_{L^2(\Omega)} + \|v\|^2_{H_0^1(\Omega)}}{\|v\|_{H_0^1(\Omega)}} = \infty \tag{2.8}$$

This fulfills all the assumptions of Proposition 5.1 on [45] the there is an unique solution to (2.7) and consequently to $(\mathcal{C}_q)$ as next theorem says:

Theorem 2.1.3. *For each $h_0 \in H^2(\Omega)$ and $W \in W^{1,1}(0,T; H_0^1(\Omega))$ then there is an unique $h \in W^{1,\infty}(0,T; H_0^1(\Omega))$ such as $h(0) = h_0$ and $\forall t \in [0,T]$ we have:*

$$\frac{d}{dt}(h(t)) = \mathcal{A}(h(t)) + W(t) \tag{2.9}$$

Which can be translated as there is an unique $h \in W^{1,\infty}(0,T; H_0^1(\Omega))$ such as $h(0) = h_0$ and $\forall t \in [0,T]$ we have:

$$(\mathcal{P})_q \begin{cases} S\frac{d}{dt}h(t,\mathbf{x}) = (K^x h_x(t,\mathbf{x}))_x + (K^y h_y(t,\mathbf{x}))_y + (K^z h_z(t,\mathbf{x}))_z + W(t,\mathbf{x}) \ ; \\ h(0) = h_0 \qquad\qquad h_0 \in H_0^1(\Omega). \end{cases} \tag{2.10}$$

For the non homogeneous case, Showalter uses second consolidation system:

Theorem 2.1.4 (Existence of a strong solution)**.** *Under the assumption that $\lambda^* > 0$, $c_0 > 0$, $v_0 \in {V'}_a$, $\mathbf{w}_0 \in \mathbf{V}$, and*

$$\mathbf{H}(.) \in C^\alpha\left([0,T], L^2(\Omega) \oplus L^2(\Gamma_S)\right)$$
$$h(.) \in C^\alpha([0,T], {V'}_a)$$

There is an unique pair of functions $p(.) : (0,T] \to V$, $\mathbf{u}(.) \to \mathbf{V}$ such as:

$c_0 Pp(.) + \vec{\nabla} \cdot \mathbf{u}(.) \in C^0([0,T], {V'}_a) \cap C^1((0,T], {V'}_a)$
$[\nabla \cdot \mathbf{u}(.), \mathbf{u}(.) \cdot \mathbf{n}] \in C^0([0,T], L^2(\Omega) \oplus L^2(\Gamma_S)) \cap C^1([0,T], L^2(\Omega) \oplus L^2(\Gamma_S))$ *that satisfy the*

initial-value problem

$$(\mathcal{P})_q \begin{cases} -\lambda^*\nabla\left(\frac{d}{dt}\nabla\mathbf{u}(t)\right) - (\lambda+\mu)\nabla(\nabla\cdot\mathbf{u}(t)) - \mu\nabla^2\mathbf{u}(t) + \alpha\nabla p(t) = \mathbf{H}(t) & in\ \mathbf{V}' \\ \frac{d}{dt}(c_o p + \alpha\nabla\cdot\mathbf{u}(t)) + \frac{1}{\mu_f}\nabla\cdot\kappa(\nabla p(t) + \rho_f\mathbf{g}) = h(t) & in\ V_a'\ and\ t\in(0,T] \\ \lim_{t\to 0^+}\left(c_0 P p(t) + \vec{\nabla}\cdot\mathbf{u}(t)\right) = v_0\ in\ V'_a\ , \\ \lim_{t\to 0^+}\vec{\nabla}\cdot\mathbf{u}(t) = \vec{\nabla}\cdot\mathbf{w_0}\ in\ L^2(\Omega)\oplus L^2(\Gamma_S) \end{cases} \tag{2.11}$$

Where $\mathcal{A}(p) = [-\partial_j(k\partial_j p), k\nabla p\cdot\vec{n}]$, $V_a := Rg(\mathcal{A}) \subset L^2(\Omega)\oplus L^2(\Gamma_S)$, $\vec{\nabla} = [\nabla, -\beta p\vec{n}]$, $P: L^2(\Omega)\oplus L^2(\Gamma_d) \to L^2(\Omega)\oplus\{0\}$ *and* $\Gamma_S = \Gamma_f \cap \Gamma_t$

This ends this topic of groundwater and poroelastic media modeling. In the next section we discuss strategies to compute numerical approximations of the models discussed above.

2.2 Discretization

The discretization of a model depends on historical choices, accuracy requirements, enforcing physical constraints such as mass conservation, and how the model will be used. For example a particular discretization may be better suited for forward simulations than for parameter estimation algorithms and *vice versa*. In this section we present several methods to discretize the models presented above as well as discuss some of their advantages and disadvantages.

2.2.1 Finite Difference Discretization of the Classical Model

This is a commonly used method due to its fairly simple implementation, the down side is its convergence rates are not optimal and depend on the smoothness of the problem. We subdivide the spatial domain into a grid of rectangles of dimension Δx by Δy with N_x points on the x-axis and N_y points on the y-axis as shown in Figure 2.2. The time domain is discretized into subintervals of length Δt. Let $h_{i,j}^m$ denote the value of the water head at (x_i, y_j) at time t^m.

One classical example for this type of discretization is the implicit scheme:

$$S_{i,j}\frac{h_{i,j}^{m+1}-h_{i,j}^m}{\Delta t} = \left[\frac{\partial K^x}{\partial x}\right]_{i,j}\frac{h_{i+1,j}^{m+1}-h_{i,j}^{m+1}}{\Delta x} + K_{i,j}^x\frac{h_{i-1,j}^{m+1}-2h_{i,j}^{m+1}+h_{i+1,j}^{m+1}}{(\Delta x)^2} +$$
$$\left[\frac{\partial K^y}{\partial y}\right]_{i,j}\frac{h_{i,j+1}^{m+1}-h_{i,j}^{m+1}}{\Delta y} + K_{i,j}^y\frac{h_{i,j-1}^{m+1}-2h_{i,j}^{m+1}+h_{i,j+1}^{m+1}}{(\Delta y)^2} + W_{i,j}^{m+1}. \tag{2.12}$$

This is equivalent to the system

$$A(S,K)\mathbf{h}^{m+1} = R(S,K)\mathbf{h}^m \text{ for } m\geq 0 \tag{2.13}$$

where $A(S,K)$ is a matrix of dimension $N\times N$ where $N = N_xN_y - 2N_x - 2(N_y-2)$ and h^m a vector of dimension $N\times 1$. It is relevant to mention that the dimension of A is $N\times N$

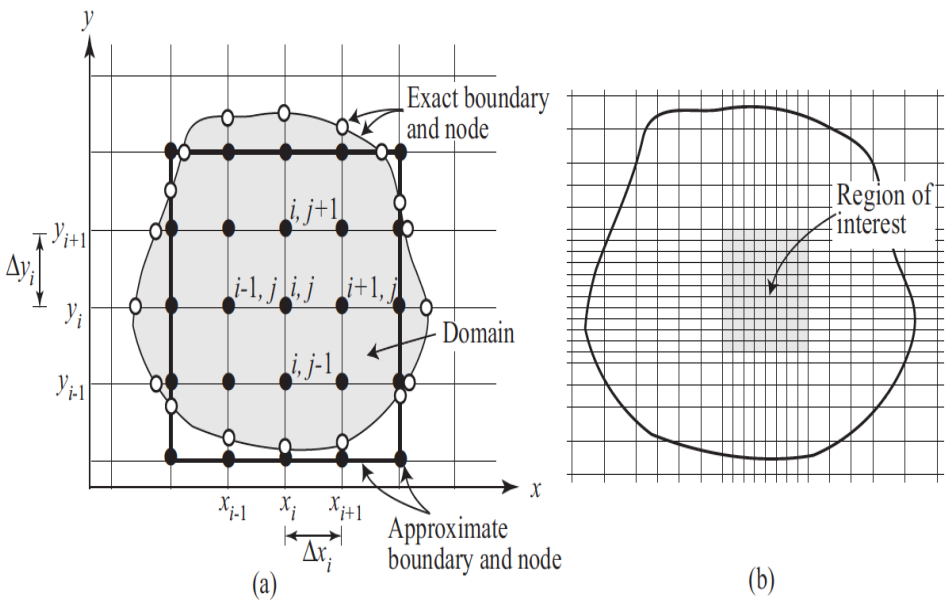

Figure 2.2: Discretization of an aquifer layer (left) around a region of interest (right), [6]

.

rather then $N_xN_y \times N_xN_y$ due to the boundary conditions which results in a reduction of the number of variables.

With the error bounded by ([2] K. E. Atkinson)

$$M[(\Delta x)^4 + (\Delta y)^4 + (\Delta t)^2] \tag{2.14}$$

where M is a constant that depends on the domain and the coefficients S and K.

Note that this traditional finite difference approach does not have the property of conservation of mass.

2.2.2 Finite Element Discretization of the Classical Model

The finite element method was introduced in the early 1940's. The underlying functional analysis used in developing the method enabled convergence results by taking advantage of its Hilbert space structure. In our case consider a Hilbert space $H = H_0^1(\Omega)$ and a finite dimensional subspace H^N with orthonormal basis $\{\phi_i(\mathbf{x})\}_{i=1}^{i=N}$. Multiplying $(\mathcal{C}_q)$ by ϕ_i and integrating results in

$$\int_\Omega \phi_i S \frac{d}{dt} h(t) d\mathbf{x} = \int_\Omega \phi_i [(K^x h_x)_x + (K^y h_y) + (K^y h_z)_z + W(t)] d\mathbf{x}.$$

The weak form of $(\mathcal{C}_q)$ holds for all $\phi_i \in H^N$. Using the divergence theorem gives

$$\int_\Omega \phi_i S \frac{d}{dt} h(t) d\mathbf{x} = -\int_\Omega \nabla \phi_i \cdot [K^x h_x, K^y h_y, K^y h_z] d\mathbf{x} + \int_{\partial\Omega} \phi_i [K^x h_x, K^y h_y, K^y h_z] \cdot \vec{n} d\Gamma + \int_\Omega \phi_i W(t)\ d\mathbf{x}. \tag{2.15}$$

Since we are looking into $H_0^1(\Omega)$ solutions, (2.15) is equivalent to

$$\int_\Omega \phi_i S \frac{d}{dt} h(t) d\mathbf{x} = -\int_\Omega \nabla \phi_i \cdot [K^x h_x, K^y h_y, K^y h_z] d\mathbf{x} + \int_\Omega \phi_i W(t) d\mathbf{x}. \tag{2.16}$$

Let

$$h^N(t, \mathbf{x}) := \sum_{i=0}^{N} \phi_i(\mathbf{x}) h_i(t).$$

Then (2.16) is equivalent to the following system of differential equations

$$\mathbf{S} \frac{d}{dt} \mathbf{h}^N(t) = [\mathbf{K}^x + \mathbf{K}^y + \mathbf{K}^z] \mathbf{h}^N(t) + \mathbf{F}(t) \tag{2.17}$$

where

$$\mathbf{h}^N(t) = [h_1(t), \ldots, h_N(t)]^T, \qquad S_{i,j} = \int_\Omega \phi_i(\mathbf{x}) \phi_j(\mathbf{x}) s(\mathbf{x}) d\mathbf{x},$$

$$K^x_{i,j} = \int_\Omega \phi_i(\mathbf{x})_x \phi_j(\mathbf{x})_x K^x(\mathbf{x}) d\mathbf{x}, \qquad K^y_{i,j} = \int_\Omega \phi_i(\mathbf{x})_y \phi_j(\mathbf{x})_y K^y(\mathbf{x}) d\mathbf{x}$$

$$K^z_{i,j} = \int_\Omega \phi_i(\mathbf{x})_z \phi_j(\mathbf{x})_z K^z(\mathbf{x}) d\mathbf{x}, \qquad \text{and} \qquad F_i(t) = \int_\Omega \phi_i(\mathbf{x}) W(t) d\mathbf{x}.$$

Any implicit method can be used to solve (2.17). For example, Crank-Nicolson or Runge Kutta methods. Using the Crank-Nicolson scheme with a time step Δt leads to

$$\mathbf{S} \frac{\mathbf{h}^N(t + \Delta t) - \mathbf{h}^N(t)}{\Delta t} = [\mathbf{K}^x + \mathbf{K}^y + \mathbf{K}^z] \frac{\mathbf{h}^N(t + \Delta t) + \mathbf{h}^N(t)}{2} + \frac{\mathbf{F}(t + \Delta t) + \mathbf{F}(t)}{2}.$$

This can be factored as

$$\left[\frac{\mathbf{S}}{\Delta t} - \frac{1}{2}[\mathbf{K}^x + \mathbf{K}^y + \mathbf{K}^z]\right] \mathbf{h}^N(t+\Delta t) = \left[\frac{\mathbf{S}}{\Delta t} + \frac{1}{2}[\mathbf{K}^x + \mathbf{K}^y + \mathbf{K}^z]\right] \mathbf{h}^N(t) + \frac{\mathbf{F}(t + \Delta t) + \mathbf{F}(t)}{2}.$$

The error bounds and rate of convergence depend inherently on the type of domain, type of mesh and of course, the type of basis, and the time discretization. The advantage of this method is that it is very similar to the weak formulation of the PDE. Therefore all the tools from PDE analysis and functional analysis can be applied to analyze its convergence. The details of these properties will be discussed in Appendix A.

2.2.3 Finite Volume Discretization of the Multilayer Model

The most common discretization methods in the groundwater flow literature are based on finite volume approximations. This is true for both forward and inverse problems. The main motivation for these methods is better mass conservation. Therefore, finite volume discretizations do not seek to estimate the quantity at each point but rather the volume average of a quantity over small geometric regions. The domain is discretized into rectangular prisms or cubes in three dimensions and into rectangles or squares in two dimensions. Due to the different nature of both layers their discretization changes as indicated in Figure 2.3. The image on the left illustrates the fact the flow moves horizontally in the aquifer layer, while the image on the right relies on the fact that flow is only vertical in the interbed as discussed in Section 2.1.3.

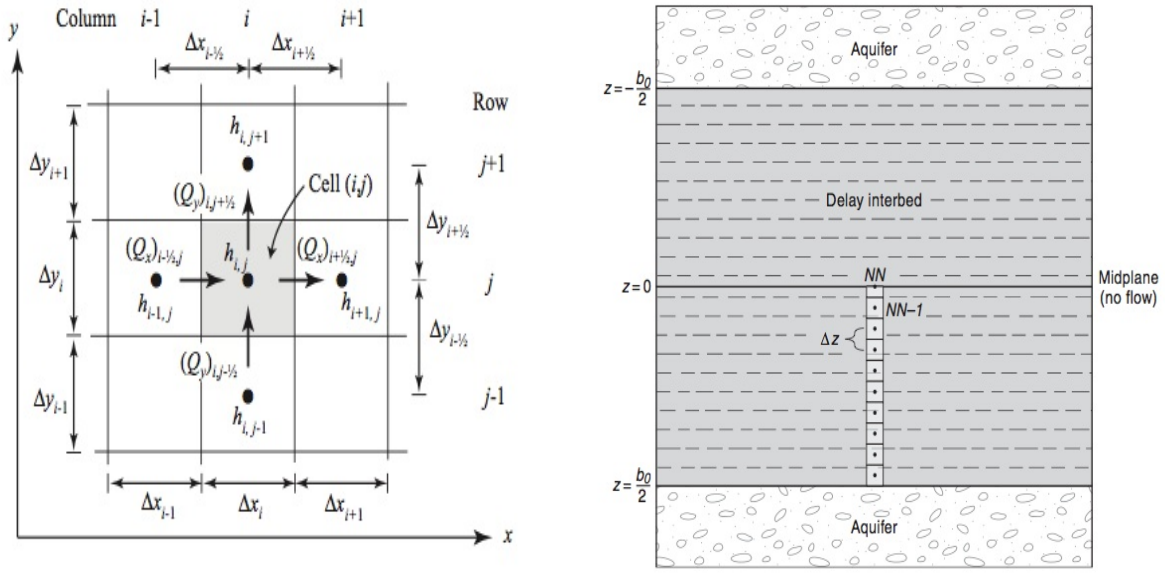

Figure 2.3: Finite volume discretizations.. Aquifer layer (left) and interbed layer (right), [6].

Discretizing the Equations for Aquifer Layers

In Section 2.1.3, we showed that the flow in aquifer layer l could be modeled by $(\mathcal{M}_q^a)$

$$\begin{aligned}\tilde{S}(\tilde{\mathbf{x}})\frac{\partial h^{a_l}(t,\tilde{\mathbf{x}})}{\partial t} =& \nabla_{\tilde{\mathbf{x}}}\cdot(T^l(\tilde{\mathbf{x}})\nabla_{\tilde{\mathbf{x}}}h^{a_l}(t,\tilde{\mathbf{x}})) \\ &+ K^{(l+1)}\frac{\partial h^{I_{l+1}}}{\partial z}\Big|_{z=b_1^{I_{l+1}}(t,(x,y))} - K^{(l-1)}\frac{\partial h^{I_{l-1}}}{\partial z}\Big|_{z=b_2^{I_{l-1}}(t,(x,y))} + P_l(t,\tilde{\mathbf{x}}). \quad (\mathcal{M}_q^a)\end{aligned}$$

Let the mesh of Ω_2 be composed of disjoint rectangles $\Omega_{i,j}$ of dimension Δx_i by Δy_j and centered at the point (x_i, y_j). The finite volume method is based on converting equation

$(\mathcal{M}_q^a)$ to an integral equation on each rectangle. This leads to

$$\int_{\Omega_{i,j}} \tilde{S}(\tilde{\mathbf{x}}) \frac{\partial h^{a_l}(t,\tilde{\mathbf{x}})}{\partial t} d\tilde{\mathbf{x}} = \int_{\Omega_{i,j}} \nabla_{\tilde{\mathbf{x}}} \cdot (T^l(\tilde{\mathbf{x}}) \nabla_{\tilde{\mathbf{x}}} h^{a_l}(t,\tilde{\mathbf{x}})) d\tilde{\mathbf{x}} + \int_{\Omega_{i,j}} K^{(l+1)} \frac{\partial h^{I_{i+1}}}{\partial z}|_{z=b_1^{I_{i+1}}(t,(x,y))} \, d\tilde{\mathbf{x}}$$
$$- \int_{\Omega_{i,j}} K^{(l-1)} \frac{\partial h^{I_{i-1}}}{\partial z}|_{z=b_2^{I_{i-1}}(t,(x,y))} d\tilde{\mathbf{x}} + \int_{\Omega_{i,j}} P_l(x,y) \, d\tilde{\mathbf{x}}. \tag{2.18}$$

Using Stokes' theorem on the term $\int_{\Omega_{i,j}} \nabla_{\tilde{\mathbf{x}}} \cdot (T^l(\tilde{\mathbf{x}}) \nabla_{\tilde{\mathbf{x}}} h_l^a(t,\tilde{\mathbf{x}})) d\tilde{\mathbf{x}}$, we obtain

$$\int_{\Omega_{i,j}} \nabla_{\tilde{\mathbf{x}}} \cdot (T^l(\tilde{\mathbf{x}}) \nabla_{\tilde{\mathbf{x}}} h^{a_l}(t,\tilde{\mathbf{x}})) \, d\tilde{\mathbf{x}} = \int_{\partial\Omega_{i,j}} T^l(\tilde{\mathbf{x}}) \nabla h^{a_l}(t,\tilde{\mathbf{x}}) \cdot \vec{n} \, d\partial\Omega_{i,j}. \tag{2.19}$$

Substituting this expression into (2.18) yields

$$\int_{\Omega_{i,j}} \tilde{S} \frac{\partial h^{a_l}}{\partial t} d\tilde{\mathbf{x}} = \int_{\partial\Omega_{i,j}} T^l \nabla h^{a_l} \cdot \vec{n} \, d\partial\Omega_{i,j} + \int_{\Omega_{i,j}} K^{(l+1)} \frac{\partial h^{I_{i+1}}}{\partial z}|_{z=b_1^{I_{i+1}}(t,(x,y))} d\tilde{\mathbf{x}}$$
$$- \int_{\Omega_{i,j}} K^{(l-1)} \frac{\partial h^{I_{l-1}}}{\partial z}|_{z=b_2^{I_{l-1}}(t,(x,y))} d\tilde{\mathbf{x}} + \int_{\Omega_{i,j}} P(x,y) d\tilde{\mathbf{x}}. \tag{2.20}$$

Replacing h^{a_l}, T, and $\tilde{S}$ by their cell averages, with $h_{a_l}(t_k, x_i, y_j) \equiv h_{i,j}^k$, the left hand side (2.20) is discretized as

$$\int_{\Omega_{i,j}} \tilde{S} \frac{\partial h(t^k,\tilde{\mathbf{x}})}{\partial t} d\tilde{\mathbf{x}} \approx S_{i,j} \frac{h_{i,j}^{k+1} - h_{i,j}^k}{\Delta t} \Delta x_i \Delta y_j.$$

The discretization of the right hand side of (2.20) requires more care since the derivative at the edge of the rectangle must be approximated from the surrounding cell averages and T^l must also be approximated there. This can be performed by either a simple average of T^l in neighboring cells or by the harmonic average. Each cell is connected to four adjacent cells. Thus a cell can receive or lose volume of water through each of those. Let $\partial_1\Omega_{i,j}, \partial_2\Omega_{i,j}, \partial_3\Omega_{i,j}$ and $\partial_4\Omega_{i,j}$ denote the left, bottom, right and top edges of the rectangle $\partial\Omega_{i,j}$ respectively. By considering an implicit scheme, we have the following approximations for the boundary integrals:

$$\int_{\partial_1\Omega_{i,j}} T\nabla h \cdot \vec{n} d\partial\Omega_{i,j} \approx (Q_x)_{i-\frac{1}{2},j} := -T_{i-\frac{1}{2},j} \Delta y_j \frac{h_{i,j}^{k+1} - h_{i-1,j}^{k+1}}{\Delta x_{i-\frac{1}{2}}},$$

$$\int_{\partial_3\Omega_{i,j}} T\nabla h \cdot \vec{n} d\partial\Omega_{i,j} \approx (Q_x)_{i+\frac{1}{2},j} := T_{i+\frac{1}{2},j} \Delta y_j \frac{h_{i+1,j}^{k+1} - h_{i,j}^{k+1}}{\Delta x_{i+\frac{1}{2}}},$$

$$\int_{\partial_2\Omega_{i,j}} T\nabla h \cdot \vec{n} d\partial\Omega_{i,j} \approx (Q_y)_{i,j-\frac{1}{2}} := -T_{i,j-\frac{1}{2}} \Delta x_i \frac{h_{i,j}^{k+1} - h_{i,j-1}^{k+1}}{\Delta y_{j-\frac{1}{2}}},$$

$$\int_{\partial_4\Omega_{i,j}} T\nabla h \cdot \vec{n} d\partial\Omega_{i,j} \approx (Q_y)_{i,j+\frac{1}{2}} := T_{i+\frac{1}{2},j} \Delta x_i \frac{h_{i,j+1}^{k+1} - h_{i,j}^{k+1}}{\Delta y_{j+\frac{1}{2}}},$$

where

$$\Delta x_{i-\frac{1}{2}} := \frac{\Delta x_{i-1} + \Delta x_i}{2},\ \Delta x_{i+\frac{1}{2}} := \frac{\Delta x_{i+1} + \Delta x_i}{2},\ \Delta y_{j-\frac{1}{2}} := \frac{\Delta y_{j-1} + \Delta y_j}{2}$$

$$\text{and } \Delta y_{j+\frac{1}{2}} := \frac{\Delta y_{j+1} + \Delta y_j}{2}.$$

The harmonic mean is

$$T_{i-\frac{1}{2},j} = \frac{\Delta x_{i-1} + \Delta x_i}{\frac{\Delta x_{i-1}}{T_{i-1,j}} + \frac{\Delta x_i}{T_{i,j}}}, \qquad T_{i+\frac{1}{2},j} = \frac{\Delta x_{i+1} + \Delta x_i}{\frac{\Delta x_{i+1}}{T_{i+1,j}} + \frac{\Delta x_i}{T_{i,j}}}$$

and $T_{i,j+\frac{1}{2}}$ and $T_{i,j+\frac{1}{2}}$ are analogously defined. Then

$$\int_{\partial\Omega_{i,j}} T\nabla h \cdot \vec{n} d\partial\Omega_{i,j} \approx (Q_x)_{i-\frac{1}{2},j} + (Q_x)_{i+\frac{1}{2},j} + (Q_y)_{i,j-\frac{1}{2}} + (Q_y)_{i,j+\frac{1}{2}}.$$

Consequently, the aquifer finite volume representation is given by

$$S_{i,j}\Delta x_i \Delta y_j \frac{h_{i,j}^{k+1} - h_{i,j}^k}{\Delta t} = (Q_x)_{i-\frac{1}{2},j} + (Q_x)_{i+\frac{1}{2},j} + (Q_y)_{i-\frac{1}{2}} + (Q_y)_{i+\frac{1}{2}}$$
$$+ N_{i,j}\Delta x_i \Delta y_j + P_{i,j}\Delta x_i \Delta y_j$$

where $N_{i,j}$ is the connection between the aquifer and interbed. Abstractly, we can write the discrete form of the model as a system with the following structure

$$B_a(S, T, \Delta t)h^{m+1} = A_a(S, T, \Delta t)h^m + f^m. \tag{2.21}$$

In order to introduce Neumann boundary conditions, $T\nabla h \cdot \vec{n} = 0$ one has to set

$$(Q_x)_{1-\frac{1}{2},j} = (Q_x)_{N_x+\frac{1}{2},j} = (Q_y)_{i,1-\frac{1}{2}} = (Q_y)_{i,N_y+\frac{1}{2}} = 0\ \forall i, j.$$

Discretizing the Equations for Interbed Layers

The Interbed equation at a point (x, y) in the horizontal plane is given by

$$S\frac{dh^I(t)}{dt} = K_v \frac{\partial^2 h^I(t)}{\partial z^2}. \tag{2.22}$$

Since the interbed is a compressible medium, there must be some adjustments to its discretization. The subsidence alters the speed of the flow and the specific storage as well. S. A. Leak [30] introduces an artificial variable called the precondition head H which is defined as

Definition 2.2.1. *Precondition Head*
Given $H^0 = H_0$ *then*

$$H_{i,j}^m := \begin{cases} H_{i,j}^{m-1} & \text{if } h_{i,j}^{I,m-1} > H_{i,j}^{m-1} \\ h_{i,j}^{I,m-1} & \text{if } h_{i,j}^{I,m-1} \leq H_{i,j}^{m-1} \end{cases} \tag{2.23}$$

Definition 2.2.2. *Skeletal Specific Storage*

$$\mathcal{S}_{kj}^{m} := \begin{cases} S_{kej} & \text{if } h_j^{I,m} > H_j^{I,m-1} \\ S_{kvj} & \text{if } h_j^{I,m} \leq H_j^{I,m-1} \end{cases} \tag{2.24}$$

where S_{ke} is the elastic storage and S_{kv} is the inelastic storage. Their contribution to the model is that the term $S\frac{dh}{dt}$ can be estimated at time t^m by

$$S\frac{dh(t)}{dt}\bigg|_{t=t_m} \approx \frac{S_k^m}{\Delta t}(h_j^{I,m} - H_j^{I,m-1}) + \frac{S_{ke}}{\Delta t}(H_j^{I,m-1} - h_j^{I,m-1}). \tag{2.25}$$

This implies that the discretization of $(\mathcal{M}_q^I)$ is as follows

$$\frac{S_k^m}{\Delta t}(h_j^{I,m} - H_j^{I,m-1}) + \frac{S_{ke}}{\Delta t}(H_j^{I,m-1} - h_j^{I,m-1}) = K_{vi}\frac{h_{j+1}^m - 2h_j^m + h_{j-1}^m}{2\Delta z} \tag{2.26}$$

which leads to the following scheme

$$B_I(S,T,h^{I,m+1})h^{m+1} = A_I(S,T)h^{I,m} + r^m. \tag{2.27}$$

Aquifer-Interbed Coupled Model

To connect both layers we use Darcy's law. The term $N = K_v\frac{(h^I - h^a)}{\Delta z}$ is added as forcing term to the aquifer and $\frac{K_v}{\Delta z}h^a$ as a boundary condition to the interbed. Since the two equations are now coupled, (2.21) and (2.27) give rise to

$$\begin{bmatrix} B_a & C^1 \\ C^2 & B_I(h_I^{m+1}) \end{bmatrix}\begin{bmatrix} h_a^{m+1} \\ h_I^{m+1} \end{bmatrix} = \begin{bmatrix} A_a & C^3 \\ C^4 & A_I \end{bmatrix}\begin{bmatrix} h_a^m \\ h_I^m \end{bmatrix} + \begin{bmatrix} F^m \\ r^m(h^m, m) \end{bmatrix}. \tag{2.28}$$

This has the form

$$\mathbf{B}(q(h^{m+1}))h^{m+1} = \mathbf{A}(q)h^m + \mathbf{R}(q(h^m)). \tag{2.29}$$

The nonlinearity arising in the interbed model (since the storage is a function of h) is treated using a Newton method. Groundwater models where the domain consists of more than two layers is given by the block tridiagonal system

$$\begin{bmatrix} B_{a_1} & C^2_{I_1,a_1} & 0 & 0 & 0 & 0 \\ C^1_{a_1,I_1} & B_{I_1} & C^2_{I_1,a_2} & 0 & 0 & 0 \\ 0 & C^1_{I_1,a_2} & B_{a_2} & C^2_{a_2,I_2} & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & C^1_{a_{n-1},I_{n-1}} & B_{I_{n-1}} & C^2_{I_{n-1},a_n} \\ 0 & 0 & 0 & 0 & C^1_{I_{n-1},a_n} & B_{a_n} \end{bmatrix}\begin{bmatrix} h_{a_1}^{m+1} \\ h_{I_1}^{m+1} \\ h_{a_2}^{m+1} \\ \vdots \\ h_{I_{n-1}}^{m+1} \\ h_{a_n}^{m+1} \end{bmatrix} =$$

$$\begin{bmatrix} A_{a_1} & C^4_{a_1,I_1} & 0 & 0 & 0 & 0 \\ C^3_{a_1,I_1} & A_{I_1} & C^4_{I_1,a_2} & 0 & 0 & 0 \\ 0 & C^3_{I_1,a_2} & A_{a_2} & C^4_{a_2,I_2} & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & C^3_{a_{n-1},I_{n-1}} & A_{I_{n-1}} & C^4_{I_{n-1},a_n} \\ 0 & 0 & 0 & 0 & C^3_{I_{n-1},a_n} & A_{a_n} \end{bmatrix}\begin{bmatrix} h_{a_1}^m \\ h_{I_1}^m \\ h_{a_2}^m \\ \vdots \\ h_{I_{n-1}}^m \\ h_{a_n}^m \end{bmatrix} + \begin{bmatrix} F^m \\ r_1^m(h^m,m) \\ 0 \\ r_2^m(h^m,m) \\ 0 \\ \vdots \\ r_{n-1}^m(h^m,m) \\ 0 \end{bmatrix}.$$

2.2.4 Four Layer Model

An example of a four layer, aquifer-interbed-interbed-aquifer is shown in Figure 2.1. This is equivalent to the model derived in Section 2.1.4 by appropriately arranging the layers and introducing the proper connection terms. The four layer model will be used in some of our numerical results in Chapter 7.

2.3 General Finite Volumes Methods for Groundwater Flow Models

The main difference between finite volume and finite element schemes is the definition of the flux, more generally the discretization of

$$\int_{\partial_1 \Omega_e} T \nabla h \cdot \vec{n} d\partial\Omega_e.$$

As one can imagine this can be done in several ways with various types/orders of discretization. Although these methods have substantial differences, they all follow the same foundational premise, mass conservation. In this section we only considered one methodology which is used in MODFLOW.

In the Chapter 3 we discuss the well-posedness of the inverse problem as well as necessary conditions for optimality.

Chapter 3

Theoretical Aspects of Parameter Estimation in Groundwater Flow Models

3.1 Introduction

The availability of accurate groundwater flow models (GWFMs) is valuable since they can be used to evaluate a variety of scenarios ranging from drought to residential development. This so-called *forward problem* is described in Chapter 2. It consists of being given a set of parameters, the initial state of the aquifer, and knowledge of source terms then using a GWFM to predict the evolution of the aquifer (piezostatic head and/or displacement). In most cases, the parameters are not available or information known about them is sparse or inaccurate. This chapter is dedicated to answering a different problem, known as the *inverse problem* or *parameter estimation problem*. Given data about the flow, the objective is to recover those parameters that best generate the known flow data when used in a GWFM.

Mathematically, we pose this problem as finding an estimate of $q^* = (S^*, K^{x*}, K^{y*}, K^{z*})$ that solves an optimization problem. Given continuous or discrete time measurements of the piezostatic head h_{data}, we seek the q^* that solves the following parameter estimation problem:

$$\begin{cases} \min\limits_{q \in \mathcal{Q}} \dfrac{1}{2} \|h(q) - h_{data}\|^2_{\mathcal{H}} + P(q) \\ \\ \text{subject to:} \\ \\ (\mathcal{F}) \begin{cases} \frac{d}{dt} S h(t;q) = A(q)h(q;t) + W(t) & h(t) \in H^1_0(\Omega) \quad ; \\ h(0) = h_0 & h_0 \in H^1_0(\Omega), t \in [0,T]. \end{cases} \end{cases} \tag{$\mathcal{PE}$}$$

Here $q = (S, K^x, K^y, K^z)$, $A(q)h = (K^x h_x)_x + (K^y h_y)_y + (K^z h_z)_z$, $P(q)$ is a penalty term such as the subsidence or the norm of q and finally Ω is a subset of $\mathbb{R}^3$ with C^0 Lipchitz

boundary.

The first instinct of a mathematician is to ask, "is there a solution?", "how dependent is that solution on the data?", "where is the data and the solution defined?" and "is the problem well defined?". All of these questions are reasonable and crucial and one should not attack the problem numerically without having these answered. The next section is dedicated to address this issues.

3.2 Well-Posedness of the Groundwater Flow Inverse Problem

In this section we discuss the well-posedness of $(\mathcal{PE})$. That is, is there a unique parameter associated with each data set? One needs to decide in which spaces the data, the parameters and cost functional are defined. All of those choices will determine the applicability of any numerical method discussed in the next chapter.

3.2.1 Identifiability

The first step in parameter estimation is to assure that the parameter is identifiable and if the process of doing it is stable. The definition of identifiability may vary, in this work we use the Kravaris and Seinfield definition [27]. First we need to clarify some concepts, the goal is to use data on the water heads and/or subsidence to identify the vector $q = (S, K^x, K^y, K^z)$ that reasonably estimates the data. In other words, given a tolerance $\delta > 0$, determine q such that

$$\|h(q) - h_{data}\|_{\mathcal{S}} < \delta \quad \text{where } \mathcal{S} \text{ denotes the solution space.} \tag{3.1}$$

This automatically gives rise to the following question, " Does there exist a q in the parameter space $\mathcal{Q}$ that satisfies this condition for each δ?". The answer is affirmative or negative depending on the choices of $\mathcal{Q}$ and $\mathcal{S}$. To answer this question we define the observation functional.

Definition 3.2.1 (Observation Functional-Infinite Dimensional Spaces)**.** *An observation functional, Φ, is a mapping from the parameter space $\mathcal{Q}$ to the observation space $\mathcal{H}$. That is $\Phi : \mathcal{Q} \to \mathcal{H}$. This functional can be can be considered as $\Phi(q) = \mathcal{C}(h(q))$ where $h(\cdot)$ is the solution operator and $\mathcal{C}$ is the projection from the solution space into the observation space*

$$\Phi(q) : \mathcal{Q} \to \mathcal{S} \to \mathcal{H}. \tag{3.2}$$

Now that the observation functional if defined the definition of identifiability comes naturally.

Definition 3.2.2 (Identifiability, Banks and Kunish, [4])**.** *The parameter q is identifiable at q^* with respect to $\mathcal{Q}$ if for any $q \in \mathcal{Q}$, $\Phi(q) = \Phi(q^*)$ implies $q = q^*$.*

The previous definition is then equivalent to sate that the observation functional is injective. Which is a reasonable property to ask. Since given data one should only recover one and one only parameter.

Theorem 3.2.3 (Identifiability)**.** *If $f(x) > 0$ a.e. and S is constant, and known, over Ω then the system $\mathcal{C}_q$ is identifiable over the spaces $\mathcal{Q} := Q^s \times Q^k \times Q^k \times Q^k$ where $Q^s := \{q \in C(\overline{\Omega}) : \exists \alpha_s > 0$ such that $\forall \boldsymbol{x} \in \Omega$, $q(\boldsymbol{x}) \geq \alpha_s\}$ and $Q^k := \{q \in C(\overline{\Omega}) : \exists \alpha_k > 0$ such that $\forall \boldsymbol{x} \in \Omega$, $q(\boldsymbol{x}) \geq \alpha_k\}$ and $\mathcal{S} = \mathcal{H} = L^2((0,T); H_0^1(\Omega))$ and $\|q\|_{\mathcal{Q}} := \|S\|_\infty + \|K^x\|_\infty + \|K^y\|_\infty + \|K^z\|_\infty$.*

Proof. This is a direct result of Kravaris and Seinfield [27]. □

The existence of a minimum requires a smaller admissible set and a larger solution space both defined in the following theorem.

Theorem 3.2.4 (Cost functional Minimizer)**.** *There is a global minimum $q_\beta^* \in \mathcal{Q}^{ad}$ to the objective functional*

$$J_\beta(h,q) = \frac{1}{2}\|h - h_{data}\|_{\mathcal{H}}^2 + \frac{\beta_S}{2}\|S\|_{L^2(\Omega)}^2 + \frac{\beta_{K^x}}{2}\|K^x\|_{L^2(\Omega)}^2 + \frac{\beta_{K^y}}{2}\|K^y\|_{L^2(\Omega)}^2 + \frac{\beta_{K^z}}{2}\|K^z\|_{L^2(\Omega)}^2 \tag{3.3}$$

where the admissible set is defined by:

$$\mathcal{Q}^{ad} = \{(S, K^x, K^y, K^z) \in \mathcal{Q} : s_1 \geq S(\boldsymbol{x}) \geq s_0,\ K_1^x \geq K^x(\boldsymbol{x}) \geq K_0^x,\ K_1^y \geq K^y(\boldsymbol{x}) \geq K_0^y, \text{ and } K_1^z \geq K^z(\boldsymbol{x}) \geq K_0^z, \forall \boldsymbol{x} \in \overline{\Omega}\} \tag{3.4}$$

where $s_i > 0$, K_i^x, K_i^y and K_i^z for i=1, 2 are fixed positive constants.

The solution space by $\mathcal{H} = W^1((0,T); H_0^1(\Omega))$ and $\beta_S, \beta_{K^x}, \beta_{K^y}$ and β_{K^z} are positive constants.

Proof. By Theorem 1.45 in [23] there is a minimum, under the following assumptions:

H1 $\mathcal{Q}^{ad}$ is convex bounded and closed

H2 (3.3) has a feasible point $h \in \mathcal{H}$

H3 $e(h,q)$ has a bounded solution operator $h(q)$ where $e(h,q) := \begin{pmatrix} \frac{d}{dt}h(t) - A(q)h(t) + W(t) \\ h(0) - h_0 \end{pmatrix}$

H4 $(h,q) \in \mathcal{H} \times \mathcal{Q} \mapsto e(h,q) \in L^2((0,T); L^2(\Omega))$ is continuous under weak convergence.

H5 J_β is sequentially weakly lower semicontinuous

Hypotheses 1 and 2 are easily verified but more detail are needed for Hypotheses 3 to 5.

Consider Hypothesis 3: On the admissible set we have the following inequality by Evans, (Theorem 7.2, [18]):

$$\|h(q)\|_{W^1((0,T);H^1(\Omega))} \leq C(q)\left(\|W\|_{L^2((0,T);L^2(\Omega))} + \|h_0\|_{L^2(\Omega)}\right) \tag{3.5}$$

where $C(q)$ depends continuously on q in the norm on C^0 denoted by $\|\cdot\|_0$. It follows that $C(q)$ is bounded since $\mathcal{Q}^{ad}$ is bounded and this implies that $h(q)$ is bounded.

Consider Hypothesis 4: Note that we only focus on the convergence of the first component of the vector $e(h,q)$. The second component converges by hypothesis. Let $\{(h^n,q^n)\}$ be an arbitrary convergent sequence with $(h^n,q^n) \underset{\mathcal{H}\times\mathcal{Q}}{\to} (h,q)$, and $\phi \in \mathcal{H}$ then one have

$$\begin{aligned}\left|\int_0^T\int_\Omega \phi\left(e(h^n,q^n)-e(h,q)\right)d\mathbf{x}dt\right| &= \left|\int_0^T\int_\Omega \phi\left(\frac{d}{dt}Sh-\frac{d}{dt}S^nh^n+A(q^n)h^n-A(q)h(q;t)\right)d\mathbf{x}dt\right| \\ &\leq \left|\int_0^T\int_\Omega \phi\left(\frac{d}{dt}Sh-\frac{d}{dt}S^nh^n\right)d\mathbf{x}dt\right| \\ &\quad+\left|\int_0^T\int_\Omega \phi\left(A(q^n)h^n-A(q)h(q;t)\right)d\mathbf{x}dt\right|. \end{aligned} \tag{3.6}$$

Now we consider the two previous terms separately. First,

$$\begin{aligned}\left|\int_0^T\int_\Omega \frac{d}{dt}Sh-\frac{d}{dt}S^nh^nd\mathbf{x}dt\right| &\leq \left|\int_0^T\int_\Omega \frac{d}{dt}(Sh-S^nh)d\mathbf{x}dt\right|+\left|\int_0^T\int_\Omega \frac{d}{dt}(S^nh-S^nh^nd\mathbf{x}dt\right| \\ &\leq \|S-S^n\|_0\|\frac{d}{dt}h\|_{L^2(0,T;L^2(\Omega))}\|\phi\|_{L^2(0,T;L^2(\Omega))} \\ &\quad+\|S^n\|_0\|\frac{d}{dt}(h-h^n)\|_{L^2(0,T;L^2(\Omega))}\|\phi\|_{L^2(0,T;L^2(\Omega))} \\ &\leq \|\phi\|_{L^2(0,T;L^2(\Omega))}\left[\|S-S^n\|_0\|\frac{d}{dt}h\|_{\mathcal{H}}+\|S^n\|_0\|(h-h^n)\|_{\mathcal{H}}\right].\end{aligned}$$

Since $\|\phi\|_{L^2(0,T;L^2(\Omega))}<\infty$ and S^n converges uniformly to S we have that for $n>p_1$, $p_1\in\mathbb{N}$ the previous equation is bounded by $\frac{\varepsilon}{2}$.

Now we focus on the second term of (3.6).

$$\begin{aligned}\left|\int_0^T\int_\Omega \phi\left(A(q^n)h^n-A(q)h(q;t)\right)d\mathbf{x}dt\right| &\leq \left|\int_0^T\int_\Omega K^{xn}\phi_xh_x^n-K^x\phi_xh_x\mathbf{x}dt\right| \\ &\quad+\left|\int_0^T\int_\Omega K^{yn}\phi_yh_y^n-K^x\phi_yh_yd\mathbf{x}dt\right| \\ &\quad+\left|\int_0^T\int_\Omega K^{zn}\phi_zh_z^n-K^z\phi_zh_zd\mathbf{x}dt\right|.\end{aligned} \tag{3.7}$$

The first term can be bounded by:

$$\begin{aligned}\left|\int_0^T\int_\Omega K^{xn}\phi_x h_x^n - K^x\phi_x h_x \mathbf{x}dt\right| &\leq \left|\int_0^T\int_\Omega K^x\phi_x h_x^n - K^x\phi_x h_x \mathbf{x}dt\right| \\ &\quad + \left|\int_0^T\int_\Omega K^x\phi_x h_x - K^{xn}\phi_x h_x^n \mathbf{x}dt\right| \\ &\leq \|K^x\|_0\|h_x^n - h_x\|_{L^2(0,T,H_0^1(\Omega))}\|\phi_x\|_{L^2(0,T,H_0^1(\Omega))} \\ &\quad + \|K^x - K^{xn}\|_0\|\phi_x\|_{L^2(0,T,H_0^1(\Omega))}\|h\|_{L^2(0,T,H_0^1(\Omega))}.\end{aligned} \tag{3.8}$$

There exists a $p_x \in \mathbb{N}$ such that for $n > p_x$, (3.8) is bounded by $\frac{\varepsilon}{6}$. Similarly for the remaining terms in (3.7). Thus for $p := \max\{p_1, p_x, p_y, p_z\}$ and $n > p$ we have that

$$\left|\int_0^T\int_\Omega \phi\left(e(h^n, q^n) - e(h,q)\right)d\mathbf{x}dt\right| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{6} + \frac{\varepsilon}{6} + \frac{\varepsilon}{6} = \varepsilon$$

which proves the weak continuity of $e(h,q)$.

In Hypothesis 5, J_β is the sum of continuous functions and is consequently continuous.

□

3.3 Non Identifiability of the Homogenous Case

Theorem 3.3.1. *By Definition 3.2.2 the system not identifiable*

$$(\mathcal{P})_q \begin{cases} S\frac{d}{dt}h(t,\mathbf{x}) = (K^x h_x(t,\mathbf{x}))_x + (K^y h_y(t,\mathbf{x}))_y + (K^z h_z(t,\mathbf{x}))_z \ \ ; \\ h(0) = h_0 \qquad\qquad h_0 \in H_0^1(\Omega). \end{cases} \tag{3.9}$$

Proof. Let $(S_1, K_1) = \alpha(S,K)$ where $\alpha \in \mathbb{R}_1^+$ and $(S,K) \in \mathcal{Q}^{ad}$. Then $h(S_1,K_1) = h(S,K)$ since $w = h(S_1,K_1) - h(S,K)$ is the solution of

$$(\mathcal{P})_q \begin{cases} (S_1 - S)\frac{d}{dt}h(t,\mathbf{x}) = ((K_1^x - K^x)h_x(t,\mathbf{x}))_x + ((K_1^y - K^y)h_y(t,\mathbf{x}))_y + \\ \qquad\qquad\qquad\qquad\qquad\qquad ((K_1^z - K^z)h_z(t,\mathbf{x}))_z; \\ h(0) = 0. \end{cases} \tag{3.10}$$

This implies that $w = 0$ so the cost functional is not injective and therefore the problem is not identifiable. □

3.4 Optimization with PDE Constraints

In this section we discuss necessary conditions to optimality, the existence of the Lagrange multiplier and the adjoint equation. The framework is quite different from the previous section since $(\mathcal{PE})$ is a constrained optimization problem.

3.4.1 Lagrange Multipliers and Optimality Conditions

Rewriting $(\mathcal{PE})$ we have

$$\begin{cases} \min\limits_{(h,q)\in\mathcal{H}\times Q} \dfrac{1}{2}\|h-h_{data}\|^2_{\mathcal{H}} + \dfrac{\beta_S}{2}\|S\|^2_{L^2(\Omega)} + \dfrac{\beta_{K^x}}{2}\|K^x\|^2_{L^2(\Omega)} + \dfrac{\beta_{K^y}}{2}\|K^y\|^2_{L^2(\Omega)} + \dfrac{\beta_{K^z}}{2}\|K^z\|^2_{L^2(\Omega)} \\ \\ \text{subject to:} \\ \\ (\mathcal{F}) \begin{cases} S\frac{d}{dt}h(t) = A(q)h(t) + W(t) & z(t)\in H^1_0(\Omega) \quad ; \\ h(0) = h_0 & h_0 \in H^1_0(\Omega), t\in[0,T]. \end{cases} \\ \\ \begin{cases} s_1 \geq S(\mathbf{x}) - s_0 \geq 0 \\ K^z_1 \geq K^x(\mathbf{x}) - K^x_0 \geq 0 \\ K^z_1 \geq K^y(\mathbf{x}) - K^y_0 \geq 0 \\ K^z_1 \geq K^z(\mathbf{x}) - K^z_0 \geq 0 \end{cases} \\ \\ \text{where } s_i > 0,\ K^x_i,\ K^y_i \text{ and } K^z_i \text{ for i=1,2 are fixed positive constants.} \end{cases} \tag{3.11}$$

As stated in Theorem 2.1.1 this condition is sufficient for existence and uniqueness of the forward problem. In the next paragraphs we use the following notation

- $J(h,q) := \frac{1}{2}\|h-h_{data}\|^2_{\mathcal{H}} + \frac{\beta_S}{2}\|S\|^2_{L^2(\Omega)} + \frac{\beta_{K^x}}{2}\|K^x\|^2_{L^2(\Omega)} + \frac{\beta_{K^y}}{2}\|K^y\|^2_{L^2(\Omega)} + \frac{\beta_{K^z}}{2}\|K^z\|^2_{L^2(\Omega)}$
- $e(h,q) := \begin{pmatrix} \frac{d}{dt}h(t) - A(q)h(t) + W(t) \\ h(0) - h_0 \end{pmatrix}$
- $\mathcal{Q} = C(\overline{\Omega}) \times C(\overline{\Omega}) \times C(\overline{\Omega}) \times C(\overline{\Omega})$
- $\mathcal{Q}^{ad} = \{(S,K^x,K^y,K^z) \in Q : s_1 \geq S(\mathbf{x}) \geq s_0,\ K^x_1 \geq K^x(\mathbf{x}) \geq K^x_0,\ K^y_1 \geq K^y(\mathbf{x}) \geq K^y_0,$ and $K^z_1 \geq K^z(\mathbf{x}) \geq K^z_0,\ \forall \mathbf{x} \in \overline{\Omega}\}$
- $\mathcal{K} := \{\mathbf{0}\}$
- $\mathcal{H} = L^2((0,T); H^1_0(\Omega))$

Then 3.11 is equivalent to

$$\begin{cases} \min\limits_{(h,q)\in\mathcal{H}\times Q} J(h,q) \\ \\ \text{subject to} \\ e(h,q) \in \mathcal{K} \text{ and} \\ q \in Q^{ad}. \end{cases} \tag{3.12}$$

Definition 3.4.1 (Robinson's Condition)**.** *$w = (h,q)$ satisfies the Robinson's condition if*

$$0 \in int\{G(w) + G'(w)(\mathcal{C} - w) - \mathcal{K}_G\} \tag{3.13}$$

where

$$G(w) = \begin{pmatrix} e(h,q) \\ h \end{pmatrix} \tag{3.14}$$

and $\mathcal{K}_G = \{\mathbf{0}\} \times \mathcal{H}$.

Lemma 3.4.2 (Ulbrich [23])**.** *If $e_y(\overline{w}) \in \mathcal{L}(Y,Z)$ is surjective and if there exists a $\overline{y} \in Y$ and $\overline{q} \in Q^{ad}$such that*

$$e_y(\overline{w})(\tilde{y} - \overline{y}) + e_q(\overline{w})(\tilde{q} - \overline{q}) = 0 \tag{3.15}$$

then the Robinson condition is satisfied.

Theorem 3.4.3. *Equation (3.12) satisfies the Robinson condition.*

Proof. The proof is based on the fact that (3.12) satisfies the hypothesis of Lemma 3.4.2. Obviously $e_h(\overline{w})$ is a bounded linear operator. The next step is to prove surjectiveness

$$e_h(\overline{w})\delta h = \begin{pmatrix} \frac{d}{dt}\delta h(t) - A(q)\delta h(t) \\ \delta h(0) \end{pmatrix}. \tag{3.16}$$

Given a pair $(f,p) \in L^2((0,T);L^2(\Omega)) \times H_0^1(\Omega)$. Find $\delta h \in \mathcal{H}$ such that

$$e_h(\overline{w})\delta h = \begin{pmatrix} f \\ p \end{pmatrix}. \tag{3.17}$$

This is equivalent to prove that for all pairs (f,p) the following there is a solution $v(t)$ to

$$\begin{cases} S\frac{d}{dt}v(t) = A(q)v(t) + f(t) & v(t) \in H_0^1(\Omega), \forall t \in [0,T] \\ v(0) = p & p \in H_0^1(\Omega). \end{cases} \tag{3.18}$$

The next step is to prove that there is $\tilde{h}$ and $\tilde{q}$ such as

$$e_h(\overline{w})(\tilde{h}(t) - \overline{h}) + e_q(\overline{w})(\tilde{q} - \overline{q}) = \mathbf{0} \Leftrightarrow \tag{3.19}$$

$$\begin{pmatrix} \frac{d}{dt}(h(t) - \tilde{h}(t)) - A(q)(h(t) - \tilde{h}(t)) + (S - \tilde{S})\frac{d}{dt}h(t) - A(q - \tilde{q})h(t) \\ h(0) - \tilde{h}(0) \end{pmatrix} = \mathbf{0}. \tag{3.20}$$

Doing the change of variables $v(t) := h(t) - \tilde{h}(t)$ it is equivalent to prove that there is a $v(t)$ such as

$$\begin{pmatrix} \frac{d}{dt}v(t) - A(q)v(t) + (S - \tilde{S})\frac{d}{dt}h(t) - A(q - \tilde{q})h(t) \\ v(0) \end{pmatrix} = \mathbf{0} \tag{3.21}$$

which is true for all $\tilde{q}$ such as $\tilde{S}(\mathbf{x}) - S \geq 0$, $\tilde{K_x}(\mathbf{x}) - K^x(\mathbf{x}) \geq 0$, $\tilde{K_y}(\mathbf{x}) - K^y(\mathbf{x}) \geq 0$ and $\tilde{K_z}(\mathbf{x}) - K^y(\mathbf{z}) \geq 0$ for all $\mathbf{x} \in \overline{\Omega}$. This completes the proof. □

Since (3.12) satisfies Robinson's condition, J and e are Fréchet differentiable, K is a closed convex cone and moreover $\mathcal{Q}^{ad}$ is closed. The problem fulfills all the hypotheses of Theorem 1.56 in [23] that is taken from Zowe and Kurcyusz, [56].

Theorem 3.4.4 (Karush Kuhn Tucker (KKT)-Optimality Condition Groundwater Flow)**.** *Let $(h^*, q^*) \in \mathcal{H} \times \mathcal{Q} \subset \mathcal{H} \times [H^2(\Omega) \times H^2(\Omega) \times H^2(\Omega)]$ be a the solution of (3.12) then there is a Lagrange multiplier λ that defines the Lagrangian*

$$\mathcal{L}(h, q, \lambda) := J(h, q) + \langle e(h, q), \lambda \rangle_{L^2((0,T);H^{-1}(\Omega)),L^2((0,T);H_0^1(\Omega))} \tag{3.22}$$

.Moreover

$$(\mathcal{OC}) \begin{cases} (\mathcal{F}) \begin{cases} \frac{d}{dt} S h(t) = A(q)h(t) + W(t) & h(t) \in H_0^1(\Omega), t \in [0, T]; \\ h(0) = h_0 & h_0 \in H_0^1(\Omega). \end{cases} \\ (\mathcal{A}) \begin{cases} S\frac{d}{dt}\lambda(t) = -A(q)\lambda(t) - (z - z_{data}) & \lambda(t) \in H_0^1(\Omega), t \in [0, T]; \\ \lambda(T) = 0. \end{cases} \\ \langle \beta_S, S - S^* \rangle + \beta_{K^x} \langle K^x, K^x - K^{x*} \rangle + \beta_{K^y} \langle K^y, K^y - K^{y*} \rangle + \\ + \beta_{K^z} \langle K^z, K^z - K^{z*} \rangle + \langle e_q(h, q), q - q^* \rangle \geq 0 \ \forall q \in \mathcal{Q}^{ad}. \end{cases} \tag{3.23}$$

3.4.2 Identifiability Under Finite Dimensional Approximations

In the real world the optimization is done in the finite dimensional setting. This means that the adjoint equation has to be discretized. This implies an approach to first discretize then optimize. This gives rise to two questions. Is there a solution for each discretization and does the sequence of solutions converge to a solution if the $(\mathcal{PE})$? This problem is addressed by [29]. They define q^{*N} the solution of the following inverse problem

$$\min_{q \in \mathcal{Q}} \frac{1}{2} \| h^N(q) - h_{data} \|_{\mathcal{H}}^2$$

subject to: $((\mathcal{PE})^N)$

$$(\mathcal{F}) \begin{cases} \frac{d}{dt} S^N h^N(t; q) = A^N(q) h^N(q; t) & h^N(t) \in H_0^1(\Omega) \quad ; \\ h^N(0) = h_0^N & h_0^N \in H_0^1(\Omega), t \in [0, T]. \end{cases}$$

This gives the rise to the question: "Does $q^{*N} \to q^*$ solution of $((\mathcal{PE})^N)$ as $N \to \infty$?". The answer is: not necessarily. Some assumptions must made in the cost functional regularity, definition of $\mathcal{H}$ regularity of the initial condition, the definition of the admissible set and finally the type of discretization of $\mathcal{H}$. In order to have a precise and clear idea of the problem Kunish and White on [29] provided the following definition:

Definition 3.4.5 (PEC- Kunish and White [29])**.** *A sequence $(H^N, A^N(q), \mathcal{C})$ is called parameter estimation convergent (PEC) scheme for (1.2) if $(\mathcal{PE}^N)$ has a solution q^{*N} for N=1,2..., if there exist a convergent subsequence $\tilde{q}^{N_k} \to q^*$ a solution of $(\mathcal{PE})$ such that*

$$h^N(t, \cdot; q^{*N}) \to h(t, \cdot; q^*) \in H^0(\Omega), t \in [0, T], \qquad \text{as } N \to \infty \tag{3.24}$$

$$J^N(q^{*N}) \to J(q^*), \qquad \text{as } N \to \infty. \tag{3.25}$$

As one can see this is a very reasonable propriety to ask, specially if one intents to find a numerical approximation of the $(\mathcal{PE})$ solution. Under some regularity assumptions this is true, the next paragraph is dedicated to the introduction of those assumptions and to their applicability to our problem, $((\mathcal{PE})^N)$. The problem of identifiability is ill-conditioned in the homogenous case and that hapens when it is pumping season. Therefore one should aim to identify the quotient.

$$S\frac{dh}{dt} = (K^x(h(t))_x)_x + (K^y(h(t))_y)_y + (K^z(h(t))_z)_z \Leftrightarrow \tag{3.26}$$

$$\frac{dh}{dt} = \left(\frac{K^x}{S}(h(t))_x\right)_x + \left(\frac{K^y}{S}(h(t))_y\right)_y + \left(\frac{K^z}{S}(h(t))_z\right)_z -$$

$$\left(\frac{K^x}{S}\right)_x (h(t))_x - \left(\frac{K^y}{S}\right)_y (h(t))_y - \left(\frac{K^z}{S}\right)_z (h(t))_z \Leftrightarrow \tag{3.27}$$

$$\frac{dh}{dt} = \sum_{i=1}^{3} \left(D^i(a_i D^i(h(t))\right) - b \cdot \nabla h(t). \tag{3.28}$$

Where $a_i = \frac{K^{x_i}}{S}$, $b_i = D^i(\frac{K^{x_i}}{S})D^i(h(t))$, $x_1 = x$, $x_2 = y$, $x_3 = z$ and $D_i = \frac{\partial}{\partial_{x_i}}$. In the paper written by Kunish and White there are sufficient conditions to the $(\mathcal{PE})$ previous problem to be PEC. For that (3.26) must fulfill the following hypothesis:

Theorem 3.4.6 (PEC-Groundwater Flow Inverse problem)**.** *If $z_0 \in H_0^1$ and the cost functional is defined as $\mathcal{C}(u) = \sum_{i=1}^{M} \int_\Omega |u(t_i, q, \mathbf{x}) - z(t_i, \mathbf{x})|$ and the admissible set is*

$$Q := \overline{\{(q_1, q_2) \in C^1(\Omega) : q_i \geq \gamma,\ \forall x \in \Omega \text{ and } \|q_i\|_1 \leq \eta \text{ for } i = 1, 2\}}^{W^{1,2}(\Omega,\mathbb{R}^{3,3})\times L^2(\Omega,R^3)}$$

$$Q^{ad} := \mathcal{P}^s \cap Q$$

where $\mathcal{P}^s$ is the set of all polynomials of degree s, the discrete version of $\mathcal{H}$ is the space H^N and is the set of all functions which are linear with respect to some triangulation of Ω, with diameter of the triangles bounded by $\frac{1}{N}$.

Proof. The theorem's proof requires that the problems fulfills three hypothesis, but first one must rewrite the problem in the following fashion:

H$_1$ $\exists \varepsilon > 0$ such as $\varepsilon\|x\|^2 \leq \sum_{i=1}^{3} a_i x_i^2$

H$_2$ There exist a constants α such as Q is a closed convex subset of

$$X := \left\{((a_i), (b_i)) : \|a_i\|_{W^{1,2}(\Omega)} \leq \alpha,\ \|b_i\|_{L^2(\Omega)} \leq \alpha \text{ and } \|z_0\|_{H^0} \leq \alpha\right\}$$

H$_3$ $\mathcal{Q}^{ad}$ is a compact subset of $W^{1,2}(\Omega, \mathbb{R}^{3,3}) \times L^2(\Omega, R^3)$

Hypothesis 1 and 2 are fulfilled since $\mathcal{Q} = C^1(\Omega) \times C^1(\Omega) \times C^1(\Omega) \times C^1(\Omega)$ and by the definition if an element on the admissible set.
Compactness, since $\mathcal{Q}^{ad}$ is the intersection of a finite dimensional space $\mathcal{P}^s$, therefore compact, with a closed subspace $\mathcal{Q}$ it makes it compact. □

This section ends then this chapter. In the following chapter we will discuss strategies to estimate q^* more exactly how to estimate q^{*N}.

Chapter 4

Numerical Groundwater Inverse Problems

4.1 Introduction

Contrasting with the previous chapter, this chapter is dedicated to the numerical aspect of inverse problems and the choice of the cost functional. Inverse problems are inexorably linked to a cost functional and the consequent problem of minimizing it. The choice of the cost functional is a discipline on it is own, one can not derive a priori the cost functional. Its form, the space where is defined, the convexity or lack of it, the choice between the strategy of discretize and then minimize versus minimize and then discretize, the type of discretization, all depends exclusively in the problem in hands. In Groundwater flow modeling it is very common to discretize using Finite Volumes and then perform the minimization on the discrete level. The main obstacle on inverse problems in GWF is that the data is very sparse in the space domain, for instance information about the water heads might be kilometers apart and there is almost no information in depth, this leads to use an auxiliary type of data by satellite, which consists on knowing how much the land has subsided. The chosen strategy by my the research group was to use Leaks model since that is the one available in MODFLOW. After that choice was made the next step is to define the how to invert the data in hands. We chose to use a least square cost functional with a regularization term the penalized bounded variation norm. This last one was chosen because leads to a piecewise constant solution, which is ideal for rocks modeling. As gradient estimation we chose the adjoint methods. Adjoint methods have been used to compute gradients of functionals for several years, their advantage comes from the fact their computation is fast and they don't require to much memory. One of the problems of adjoint methods is that the only outcome is the gradient, one should then decide afterwards which path to take on the optimization method. For instance a linear approximation of the cost functional, as the steepest descendent method only requires the computation of the gradient but it has a small rate of convergence. On the other hand a better accuracy and faster convergence can be reached by the quadratic approximation of the functional, Newton methods, here the problem is that computing the Hessian is often not sustainable in terms of computation time and memory space. The first

problem is solved by the Quasi-Newton methods, among this class of methods a very popular one is the BFGS method, were the Hessian is approximated by small rank matrix easy to compute. The second problem by a line search. Sensitivity analysis is the other hand is post parameter estimation study. The main purpose of of sensitivity analysis is to module the impact and influence that the parameters have on the PDE. For instance one can rank the importance of the different parameters on the model and which kind of effect they produce on the PDE, form dissipativity or delay in time propagation.

4.2 The Cost Functional and Optimization Algorithm

In this subsection we will derive the adjoint equations and an iterative method to estimate the parameters. For instance one should assure that every formulation and step on the estimation process the parameters are in the conditions of 2.1.1.

4.2.1 Cost Functional

The main goal is to find the set of parameter such the forward problem generated by them fits the best the data. The definition of best is very dubious and dependent on the problem itself, for instance depends on the definition of the parameter space and solution space as well. For well-posedness purposes the common technic is to add a penalty term and a regularization term. The objective of the penalty term is to introduce numerical stability and improve convergence rates in the other hand the regularization terms are responsible to add regularity to the solution and the problem it self, for instance requiring the solution to C^2 or piecewise constant as adding the term $\|\nabla q\|$ on the cost functional, by regularizing the cost functional is always common to refer it as adding convexity to the problem.
A classical Cost Functional for the ground water flow would be the following:

$$J(h,q) := \frac{1}{2}\|h-h_{data}\|_Z^2+\frac{\beta_S}{2}\|S-S_{data}\|_{L^2(\Omega)}^2+\frac{\beta_{K^x}}{2}\|K^x-K^x_{data}\|_{L^2(\Omega)}^2+\frac{\beta_{K^y}}{2}\|K^y-K^y_{data}\|_{L^2(\Omega)}^2 \tag{4.1}$$

Where $q = (S, K)$ the solution space Z is $L^2((0,T); H_0^1(\Omega))$ and the time space domain $\Omega_T = (0,T)\times\Omega$ and h is the solution of:

$$(\mathcal{P})_q \begin{cases} S\frac{d}{dt}h(t) = (K^x h_x(t))_x + (K^y h_y(t))_y + W(t) & h(t)\in H_0^1(\Omega) \quad ; \\ h(0) = h_0 & t\in[0,T]. \end{cases} \tag{4.2}$$

The dependence of h from the the parameters trough the forward problem leads to the definition os the set of $q's$ such as h is well defined, the admissible set of parameters Q^{ad}, a subset of Q a Banach or Hilbert space. So the parameter estimation set up is:

$$\min_{(h,q)\in Q^{ad}} \frac{1}{2}\|h-h_{data}\|_Z^2+\frac{\beta_S}{2}\|S-S_{data}\|_{L^2(\Omega)}^2+\frac{\beta_{K^x}}{2}\|K^x-K^x_{data}\|_{L^2(\Omega)}^2+\frac{\beta_{K^y}}{2}\|K^y-K^y_{data}\|_{L^2(\Omega)}^2 \tag{4.3}$$

Subject to

$$(\mathcal{P})_q \begin{cases} S\frac{d}{dt}h(t) = (K^x h_x(t))_x + (K^y h_y(t))_y + W(t) & h(t) \in H^1(\Omega); \\ h(0) = h_0 & t \in [0,T]. \end{cases} \tag{4.4}$$

This is the classical approach to a general inverse problem, due to the differences in the physics and model of each problem the choice of the perturbation on regularization term is tremendously important. In groundwater flow modeling often the distribution of the parameters is considered to be piecewise constant, then the most adequate regularization term is the bounded variation norm, $\beta_q \int_\Omega \sqrt{\|\nabla q\|^2 + \gamma_q}$, since this penalizes the non piecewise constant functions. Changing then the cost functional to:

$$J(h,q) = \frac{1}{2}\|h - h_{data}\|_Z^2 + \frac{\beta_S}{2}\|S - S_{data}\|_{L^2(\Omega)}^2 + \frac{\beta_{K^x}}{2}\|K^x - K^x_{data}\|_{L^2(\Omega)}^2 + \frac{\beta_{K^y}}{2}\|K^y - K^y_{data}\|_{L^2(\Omega)}^2 +$$

$$\alpha_S \int_\Omega \sqrt{\|\nabla S\|_2^2 + \gamma_S} d\mathbf{x} + \alpha_{K^x} \int_\Omega \sqrt{\|\nabla K^x\|_2^2 + \gamma_{K^x}} d\mathbf{x} + \alpha_{K^y} \int_\Omega \sqrt{\|\nabla K^y\|_2^2 + \gamma_{K^y}} d\mathbf{x} \tag{4.5}$$

4.2.2 Steepest Descent and Newton's Method Using the Adjoint Equation

Now the problem becomes a constrained optimization problem, therefore one should set the Lagrangian and the adjoint equation associated with problem. Looking to local optimal solutions, the Lagrangian associated with this problem:

$$L : Z \times Q^S \times Q^K \times H^{-1} \to \mathbb{R}$$

Defined by:

$$L(h,S,K,\lambda) = J(h,q) - \int_{\Omega_T} \lambda(t)(S\frac{d}{dt}h(t) - (K^x h_x(t))_x - (K^y h_y(t))_y - W(t))dtd\mathbf{x} \tag{4.6}$$

Integrating by parts in time the Lagrangian becomes:

$$J(h,q) - \int_\Omega \lambda(T)Sh(T) - \lambda(0)Sh_0 d\mathbf{x} + \int_{\Omega_T} S\lambda_t(t)h(t)dtd\mathbf{x} -$$

$$\int_{\Omega_T} \lambda(t)(-(K^x h_x(t))_x - (K^y h_y(t))_y - W(t)))dtd\mathbf{x} \tag{4.7}$$

Then the adjoint equation is then defined by:

$$\frac{\partial L}{\partial z}(\delta h) = 0 \qquad\qquad \forall \delta h \tag{4.8}$$

Which is equivalent to:

$$\int_{\Omega_T} \delta h(h - h_{data}) dt d\mathbf{x} + \int_{\Omega_T} \lambda_t(t) S(\delta h) dt d\mathbf{x} -$$

$$- \int_{\Omega} \lambda(T) S \delta h(T) d\mathbf{x} + \int_{\Omega_T} \lambda (K^x \delta h_x(t))_x + (K^y \delta h_y(t))_y dt d\mathbf{x} = 0 \tag{4.9}$$

Since (4.8) has to hold for all directions δh the Lagrangian multiplier must be a solution of:

$$(\mathcal{A})_q \begin{cases} S\frac{d}{dt}\lambda(t) = -(K^x \lambda_x)_x - (K^y \lambda_y)_y - (h - h_{data}) & \lambda(t) \in H_0^1(\Omega); \\ \lambda(T) = 0 & t \in [0, T]. \end{cases} \tag{4.10}$$

The strategy now is to use the adjoint equation to find the gradient of J if $\lambda(t)$ is the solution of the adjoint equation at a state $(h(t; S, K))$ then $\nabla_{S,K} J$ is given by:

$$J'(h(q), q)\delta h = J_q(h, q)\delta q + \langle D_q[S\frac{d}{dt}h(t) - (K^x h_x(t))_x - (K^y h_y(t))_y - W(t)]\delta q, \lambda(t)\rangle_Z \tag{4.11}$$

Where

$$J_q(h, q)\delta q = \beta_S \int_{\Omega} (S - S_{data})\delta S d\mathbf{x} + \beta_{K^x} \int_{\Omega} (K^x - {K^x}_{data})\delta K^x d\mathbf{x} + \beta_{K^y} \int_{\Omega} (K^y - {K^y}_{data})\delta K^y d\mathbf{x} +$$

$$\alpha_S \int_{\Omega} \frac{\nabla S \cdot \nabla \delta S}{\sqrt{\|\nabla S\|^2 + \gamma_{K^y}}} d\mathbf{x} + \alpha_{K^x} \int_{\Omega} \frac{\nabla K^x \cdot \nabla \delta K^x}{\sqrt{\|\nabla K^x\|^2 + \gamma_{K^x}}} d\mathbf{x} + \alpha_{K^y} \int_{\Omega} \frac{\nabla K^y \cdot \nabla \delta K^y}{\sqrt{\|\nabla K^y\|^2 + \gamma_{K^y}}} d\mathbf{x} \tag{4.12}$$

and

$$\langle D_q[S\frac{d}{dt}h(t) - (K^x h_x(t))_x - (K^y h_y(t))_y - W(t)]\delta q, \lambda(t)\rangle_Z =$$

$$= \langle \delta S \frac{d}{dt}h(t), \lambda(t)\rangle_Z + \langle \delta K^x h_x(t), \lambda_x(t)\rangle_Z + \langle \delta K^y h_y(t), \lambda_y(t)\rangle_Z \tag{4.13}$$

Steepest Descent Methods

This is a gradient method base approach since it only uses gradient information . The main idea is to use the fact that once one is close to the the local minimum the gradient is a descent direction therefore the optimal steepest descent would be by starting with an initial guess q_0

$$\begin{cases} \mu^* := \arg\min_{\mu} J(q_k - r\nabla J(q_k)) \\ q_{k+1} = q_k - \mu^* \nabla J(q_k) \end{cases} \tag{4.14}$$

The main problem of this method that it has a low rate of convergence. For this same reason the Newton method is more popular and that is the one used in this work.

Newton's method

Now the following approximation

$$m(\delta q) = J(q_0 + \delta q) = J(q_0) + \langle J'(q_0), \delta q\rangle + \frac{1}{2}\langle J''(q_0)\delta q, \delta q\rangle \tag{4.15}$$

Then locally the maximum is the solution of:

$$m'(\delta q) = 0 \tag{4.16}$$

$$\langle J'(q_0), \delta q\rangle + \langle J''(q_0)\delta q, \delta q\rangle = 0 \tag{4.17}$$

For all possible variations of δq.
The main goal now is to find the minima for that the most appropriate is the Newton's method. One start with an initial guess q^0 and recursively update the new q in the following fashion:

$$(\mathcal{N})_q \begin{cases} q^{k+1} = q^k + \delta q; \text{ where h is the solution of:} \\ \langle J'(q^k), \delta q\rangle + \langle J''(q^k)\delta q, \delta q\rangle = 0. \end{cases} \tag{4.18}$$

The evaluation of the Hessian is generally costly, a vast of literature of ways of approximate the Hessian in an iterative fashion. The one we use is the so called BFGS, this method, named after its creators, Broyden, Fletcher, Goldfarb, and Shanno. Theoreitical aspects of the algorithm are discussed at [8]. It is defined by the recursive algorithm:

BFGS-Algorithm

- start with an initial guess of x_0 and an Hessian inverse H_0
- $x_1 = x_0 - H_0 x_0$

Repeat the steps 1 to 5 until convergence

1. $s_k = x_{k+1} - x_k$
2. $y_k = \nabla f_{k+1} - \nabla f_k$
3. $H_{k+1} = (I - \rho_k s_k y_k^T) H^k (I - \rho_k y_k s_K^T) + \rho_k s_k s_k^T$
4. $x_{k+2} = x_{k+1} - H_{k+1}\nabla f_{k+1}$
5. stop if $\|\nabla f_{k+2}\| \leq tol$

4.2.3 KKT Conditions

As the KKT conditions are necessary conditions to optimality, this section presents two numerical methods to find a local optimal solution. The constrained optimization problem (4.3)–(4.4) can be rewritten in the following fashion.

$$\left\{ \begin{array}{l} \min\limits_{(h,q)\in Q^{ad}} J(h,q) \text{ subject to:} \\ e(h,q) = S\frac{d}{dt}h(t) - (K^x h_x(t))_x - (K^y h_y(t))_y - W(t) = 0, \; for \text{ all } t \in [0,T]. \end{array} \right. \tag{$(\mathcal{O})_q$}$$

Then once one write the Lagrangian, $J(h,q) + \int_0^T \langle e(h,q), \lambda(t) \rangle_{L^2(\Omega)} dt$ the KKT condition is true at an extremal point (q^*, λ^*):

$$\begin{cases} L_h = \nabla_h L(q^*, \lambda^*) = 0 \\ L_q = \nabla_q L(q^*, \lambda^*) = 0 \\ L_\lambda = \nabla_\lambda L(q^*, \lambda^*) = 0 \end{cases} \quad (\mathcal{KKT})$$

Which equivalent to the system of five equations

$$(L_h, L_{K^x}, L_{K^y}, L_S, L_\lambda) = 0 \tag{4.19}$$

That is translated to:

1.
$$L_h(\delta q) = \int_{\Omega_T} \delta h(h(t) - h(t)_{data}) dt d\mathbf{x} - \int_{\Omega_T} \frac{d\lambda(t)}{dt} S(\delta h) dt d\mathbf{x} +$$
$$\int_\Omega \lambda(T) S \delta h(T) d\mathbf{x} + \int_{\Omega_T} \lambda(t)(K^x \delta h_x(t))_x + \lambda(t)(K^y \delta h_y(t))_y dt d\mathbf{x} = 0 \tag{4.20}$$

2.
$$L_S(\delta S) = \beta_S \int_\Omega \delta S(S - S_{data}) d\mathbf{x} + \int_{\Omega_T} \lambda(t) \delta S \frac{dh(t)}{dt} dt d\mathbf{x} = 0 \tag{4.21}$$

3.
$$L_{K^x}(\delta K^x) = \beta_{K^x} \int_\Omega \delta K^x (K^x - K^x_{data}) d\mathbf{x} - \int_{\Omega_T} \lambda(t)(\delta K^x h_x(t))_x dt d\mathbf{x} = 0 \tag{4.22}$$

4.
$$L_{K^y}(\delta K^y) = \beta_{K^y} \int_\Omega \delta K^y (K^y - K^y_{data}) dy - \int_{\Omega_T} \lambda(t)(\delta K^y h_y(t))_x dt d\mathbf{x} = 0 \tag{4.23}$$

5.
$$L_\lambda(\delta\lambda) = \int_{\Omega_T} (\delta\lambda) S \frac{dh(t)}{dt} dt d\mathbf{x} +$$
$$\int_{\Omega_T} \delta\lambda(-K^x \delta h_x(t))_x - (K^y \delta h_y(t))_y - W(t))) dt d\mathbf{x} = 0 \tag{4.24}$$

Using integration by parts in time and Green's formula, the previous integrals are equivalent to

1.
$$\int_{\Omega_T} \delta h(h - h_{data}) dt d\mathbf{x} - \int_{\Omega_T} \frac{d\lambda(t)}{dt} S(\delta h) dt d\mathbf{x} + \int_\Omega \lambda(T) S \delta h(T) -$$
$$\int_{\Omega_T} \delta h(K^x \lambda_x(t))_x + \delta h(K^y \lambda_y(t))_y dt d\mathbf{x} = 0 \tag{4.25}$$

2.
$$\beta_S \int_\Omega \delta S(S - S_{data})d\mathbf{x} + \int_{\Omega_T} \lambda(t)\delta S \frac{dh(t)}{dt} dt d\mathbf{x} = 0 \tag{4.26}$$

3.
$$\beta_{K^x} \int_\Omega \delta K^x (K^x - K^x_{data})d\mathbf{x} + \int_{\Omega_T} \delta K^x \lambda_x(t) h_x(t) dt d\mathbf{x} - \int_{\partial\Omega_T} \lambda(t)\delta K^x h_x(t) n_x dt d\Gamma = 0 \tag{4.27}$$

4.
$$\beta_{K^y} \int_\Omega \delta K^y (K^y - K^y_{data})dy +$$
$$\int_{\Omega_T} \delta K^y \lambda_y(t) h_y(t) dt d\mathbf{x} - \int_{\partial\Omega_T} \lambda(t)\delta K^y h_y(t) n_y dt d\Gamma = 0 \tag{4.28}$$

5.
$$L_\lambda(\delta\lambda) = \int_{\Omega_T} (\delta\lambda) S \frac{dh(t)}{dt} dt d\mathbf{x} -$$
$$\int_{\Omega_T} \delta\lambda(-K^x h_x(t))_x - (K^y h_y(t))_y - W(t)))dt d\mathbf{x} = 0 \tag{4.29}$$

Finally this can be described as the weak solution of:

$$\begin{cases} (\mathcal{F})_q \begin{cases} S\frac{d}{dt}h(t) = (K^x h_x(t))_x + (K^y h_y(t))_y - W(t) \\ h(0) = h_0 \end{cases} \\ \\ (\mathcal{A})_q \begin{cases} S\frac{d}{dt}\lambda(t) = -(K^x \lambda(t)_x)_x - (K^y \lambda(t)_y)_y - (h(t) - h_{data}(t)) \\ \lambda(T) = 0 \end{cases} \\ \\ \beta_{K^x}(S - S_{data}) + \int_0^T \frac{d\lambda(t)}{dt} h(t) dt = 0 \\ \beta_{K^x}(K^x - K^x_{data}) + \int_0^T \lambda_x(t) h_x(t) dt = 0 \\ \beta_{K^y}(K^y - K^y_{data}) + \int_0^T \lambda_y(t) h_y(t) dt = 0 \end{cases} \qquad ((\mathcal{KKT})_q)$$

Gauss-Seidel Algorithm

Given an initial guess (S^n, K^n) one can solve for $h(t)$:

$$S^n \frac{d}{dt} h(t) = (K^{xn} h_x(t))_x + (K^{yn} h_y(t))_y + W(t) \tag{4.30}$$

which makes it possible to solve for $\lambda(t)$:

$$S^n \frac{d}{dt} \lambda(t) = -(K^{xn} \lambda(t)_x)_x - (K^{yn} \lambda(t)_y)_y - (h(t) - h_{data}(t)) \tag{4.31}$$

This makes it possible to update S^{n+1}, K^{xn+1} and K^{yn+1} through:

$$\beta_S(S^{n+1} - S_{data}) + \int_0^T \lambda_t(t) h(t) dt = 0 \tag{4.32}$$

$$\beta_{K^x}(K^{xn+1} - K^x_{data}) + \int_0^T \lambda_x(t)h_x(t)dt = 0 \tag{4.33}$$

$$\beta_{K^y}(K^{yn+1} - K^y_{data}) + \int_0^T \lambda_y(t)h_y(t)dt = 0 \tag{4.34}$$

Theorem 4.2.1. *If $\beta_S = \beta_{K^x} = \beta_{K^y} = \beta > 0$ and q_0 is sufficiently close to q^* then the previous algorithm converges to q^*. Moreover the method is equivalent to a steepest descendent method with an update of $\frac{1}{\beta}J_q(h(q), q)$.*

Proof. By definition of the Lagrange multiplier:

$$J_q(h(q), q) = J_q(h, q) + \langle e_q(h, z), \lambda \rangle \tag{4.35}$$

which is equivalent to :

$$J_q(h(q), q) = \begin{bmatrix} \beta(S^n - S_{data}) + \int_0^T \lambda_t(t)h(t)dt \\ \beta(K^{xn} - K^x_{data}) + \int_0^T \lambda_x(t)h_x(t)dt \\ \beta(K^{yn} - K^y_{data}) + \int_0^T \lambda_y(t)h_y(t)dt \end{bmatrix} \tag{4.36}$$

Therefore one can write the updates (4.32), (4.33) and (4.34) as

$$q^{n+1} = q^n - \frac{1}{\beta}\begin{bmatrix} -\beta S_{data} + \int_0^T \lambda_t(t)h(t)dt \\ -K^x_{data} + \int_0^T \lambda_x(t)h_x(t)dt \\ -K^y_{data} + \int_0^T \lambda_y(t)h_y(t)dt \end{bmatrix} = q^n - \frac{1}{\beta}J_q(h(q), q)(q^n + 1) \tag{4.37}$$

which is the steepest descendent update. □

The natural path to follow is to derive the finite dimensional equivalent of the previous algorithm and this is discussed in the next subsection.

Finite Dimensional Approximation Using Gauss-Seidel

Using finite elements, the finite dimensional version of the previous algorithm is:

Given an initial guess of the vectors S_0^N, K^{xN}_0 and K^{yN}_0 solve:

$$S_n^N \otimes \frac{d}{dt} h^N(t) = A^{N,N}(K_n^N) h^N(t) + W^N(t) \tag{4.38}$$

given that $h^N(0) = h_0^N$. Solve for $\lambda^N(t)$:

$$S_n^N \otimes \frac{d}{dt} \lambda^N(t) = -A^{N,N}(K_n^N) \lambda^N(t) - (h^N(t) - h_{data}^N(t)) \tag{4.39}$$

Where $\lambda^N(0) = \mathbf{0}^N$. Update S_{n+1}^N, K^{xN}_{n+1} and K^{yN}_{n+1} by:

$$S_{n+1}^N = S_{data}^N - \frac{1}{\beta_S} \int_0^T \lambda_t^N(t) h^N(t) dt \tag{4.40}$$

$$K^{xN}_{n+1} = K^{xN}_{data} - \frac{1}{\beta_{K^x}} \int_0^T \lambda_x^N(t) h^N(t)_x dt \tag{4.41}$$

$$K^{yN}_{n+1} = K^{yN}_{data} - \frac{1}{\beta_{K^y}} \int_0^T \lambda_y^N(t) h^N(t)_y dt \tag{4.42}$$

It is relevant to point out that the data has the same space discretization and time step as the approximation. In the event that the data is sparser one can interpolate in time and space. The previous scheme is still infinite dimensional in time, therefore it is a differential equation. To solve the differential equations (4.38) and (4.39) the Crank-Nicolson method and *ode*45 in Matlab are used. The comparison is shown in Chapter 6.

All at Once

This is a common method to solve (4.19). Instead of solving the set of equations ($\mathcal{KKT}$) sequentially, this method tries to find the root of $(L_h, L_{K^x}, L_{K^y}, L_S, L_\lambda)$ by firstly finding the numerical approximation $(L_h^N, L_{K^x}^N, L_{K^y}^N, L_S^N, L_\lambda^N)$ and then finding the root. Which by defining the functional $F(h^N, \lambda^N, S^N, K^{xN}, K^{yN})$ as:

$$F(h^N, \lambda^N, S^N, K^{xN}, K^{yN}) := \begin{bmatrix} S^N \otimes \frac{d}{dt} h^N(t) - A^{N,N}(K^N) h^N(t) - W^N(t) \\ S^N \otimes \frac{d}{dt} \lambda^N(t) + A^{N,N}(K^N) \lambda^N(t) + h^N(t) - h_{data}^N(t) \\ S_{data}^N - \frac{1}{\beta_S} \int_0^T \lambda_t^N(t) h^N(t) dt \\ K^{xN}_{data} - \frac{1}{\beta_{K^x}} \int_0^T \lambda_x^N(t) h^N(t)_x dt \\ K^{yN}_{data} - \frac{1}{\beta_{K^y}} \int_0^T \lambda_y^N(t) h^N(t)_y dt \end{bmatrix}$$

Then all at once method is based on finding $(h^{N*}, \lambda^{N*}, S^{N*}, K^{xN*}, K^{yN*})$ such that

$$F(h^{N*}, \lambda^{N*}, S^{N*}, K^{xN*}, K^{yN*}) = 0. \tag{4.43}$$

Any root solver can be used.

4.3 Las Vegas Model

As stated before, the main goal is to model the Las Vegas Valley as described in Section 2.1.3. The methods used will now exclusively be in the finite dimensional space. Due to the lack of depth data, subsidence data becomes crucial to groundwater flow parameter estimation. This leads to the introduction of a new regularization term into the cost functional, the subsidence data fitting. This incorporates the subsidences data obtained by satellite and the Leak's simulation of subsidence. The model will then discretized in the following fashion: The Aquifer is divided into N_A cells, the interbed into $N_I = \sum_{i=1}^{N_A} N_{Ii}$ cells and finally the time interval $[0, T]$ into M intervals, where 0 is when the first measurements take place. Since the wells pump water periodically, distributed into pumping cycles and non-pumping cycles. The non-pumping cycles lead to the non identifiability discussed in Section 3.3 therefore one should try to estimate the data over the periods of pluming. The following paragraphs define subsidence and its relationship with the specific storage and the water head.

Definition 4.3.1 (Subsidence-Interbed)**.** *Let i be a column as described in Figure 2.3 where b_i its thickness and N_{Ii} the number of cells in that same column. Then the subsidence at each time t^m in the column i is given by*

$$[\Delta b^m]_i = \sum_{j=1}^{N_{Ii}} \Delta b_{i,j}^m \tag{4.44}$$

where $\Delta b_{i,j}^m$ is given by:

$$\Delta b_{i,j}^m = -\Delta z[S_{sk}^m(h_{i,j}^m - H_{i,j}^{m-1}) + S_{ske}(H_{i,j}^{m-1} - h_{i,j}^{m-1}),\ j < N_{Ii} \tag{4.45}$$

$$\Delta b_{i,j}^m = -\frac{\Delta z}{2}[S_{sk}^m(h_{i,j}^m - H_{i,j}^{m-1}) + S_{ske}(H_{i,j}^{m-1} - h_{i,j}^{m-1}),\ j = N_{Ii} \tag{4.46}$$

and S_{kj}^m:

$$S_{kj}^m = \begin{cases} S_{kej} & if \quad h_i^{I,m} > H_j^{I,m-1} \\ S_{kvj} & if \quad h_i^{I,m} \leq H_j^{I,m-1} \end{cases} \tag{4.47}$$

Due to the nonlinearity a Newton step must be introduced at each iteration.

Definition 4.3.2.

$$S_{sk}(h^{I,m}; \alpha) := \frac{S_{ke}}{\pi}[\tan^{-1}(\alpha(h_j^{I,m} - H_j^{I,m})) + \frac{\pi}{2}] + \frac{S_{kv}}{\pi}[\tan^{-1}(\alpha(H_j^{I,m} - h_j^{I,m})) + \frac{\pi}{2}]. \tag{4.48}$$

As $\alpha \to \infty$ (4.3.2) converges to (4.47) so one should use a $\alpha > 10^4$ to approximate the parameter S_{sk}, or an α that is large comparatively difference between the precondition-head and the the drawdown.

4.3.1 Cost Functional

Definition 4.3.3. *Let $X \subset \mathbb{R}^n$ then*

$$PWCT(X) := \{f : X \to \mathbb{R}^+ \textit{such as } f \textit{ is piecewise constant over } X\}$$

Admissible set:

$Q^{ad} := \{q = (S, K^x, K^y, K_v, S_{ke}, S_{kv}),\ S, K^x, K^y \in PWCT(\Omega_a) \text{ and } K_v, S_{ke}, S_{kv} \in PWCT(\Omega_I)\}$
Where Ω_a (Ω_b) is a connected finite union of rectangles (rectangular prisms) in $\mathbb{R}^2(\mathbb{R}^3)$. The cost functional is then defined by

$$J^\beta(h, q) = \frac{1}{2}\sum_{m\leq M} \|h^a(t(m)) - h_{data}(t(m))\|^2_{l^2_w} + \frac{1}{2}\sum_{m\leq M} \|\Delta b^m - \Delta b^m_{data}\|^2_{l^2_w} + r(q, \beta) \quad (4.49)$$

and the penalty/regularization term r, is defined as:

$$r(q,\beta) = \frac{\theta_S}{2}\|S - S_{data}\|^2_{L^2(\Omega)} + \frac{\theta_{K^x}}{2}\|K^x - K^x_{data}\|^2_{L^2(\Omega)} + \frac{\theta_{K^y}}{2}\|K^y - K^y_{data}\|^2_{L^2(\Omega)} +$$

$$\alpha_S \int_\Omega \sqrt{\|\nabla S\|_2^2 + \gamma_S} d\mathbf{x} + \alpha_{K^x} \int_\Omega \sqrt{\|\nabla K^x\|_2^2 + \gamma_{K^x}} d\mathbf{x} + \alpha_{K^y} \int_\Omega \sqrt{\|\nabla K^y\|_2^2 + \gamma_{K^y}} d\mathbf{x}$$

and subject to the constraints:

(a) $e^a(h^a, h^I, q) = S\frac{d}{dt}h^a - (K^x h^a_x)_x - (K^y h^a_y)_y) - \sum P_i(t)\delta(\mathbf{x} - \mathbf{x}^i) - \sum K_{vi}\frac{dh^l}{dz}\delta(\mathbf{x} - \mathbf{x}^j) = 0$

(b) $e^I_j(h^a, h^I, q) = \frac{S^m_k}{\Delta t}(h^{I,m}_j - H^{I,m-1}_j) + \frac{S^m_{ke}}{\Delta t}(H^{I,m-1}_j - h^{I,m-1}_j) - K_{vi}h^{I^m}_{j\,zz} = 0 \qquad \forall m, j$

(c) $e^I_{Bj}(h^a, h^I, q) = h^I{}_j(\mathbf{x}^i) - h^a(\mathbf{x}^i) = 0$

The constraint (a) is responsible for the aquifer, (b) for the interbed and finally (c) for the link between the aquifer and the interbed and l^2_w is the weighted l^2 inner product.

4.3.2 Optimization Approach

This is an the opposite approach form the previous methods since it is discretize and then optimize. The book Lagrange Multiplier Approach to Variational Problems and Applications from It and Kunisch [25] this is discussed very thoroughly.This approach is the most intuitive since the data lives in a discrete setting therefore lifting it to a continuos space might enable the existence of convergence results it won't necessary make the problem easier. As it was discussed on 3.
The previous constraints in finite dimensional on 2.29 approximation using finite volumes can be written as the augmented system:

$$e(h^m, q^m) = \mathbf{B}(q(h^{m+1}))h^{m+1} - \mathbf{A}(q)h^m - \mathbf{R}(q(h^m)) \quad (4.50)$$

Where $h^m = [h^{a,m}, h^{I,m}]$ and M,N, N_I are the time, horizontal and vertical dimensions respectively.

The next step is to construct the finite dimensional Lagrangian:

$$L^{\beta}(h^a, h^I, q, \lambda^a, \lambda^I, \lambda^B) = J^{\beta,N,N_I,M}(h^a, h^I, q)+$$

$$\sum_{m\leq M} \langle \mathbf{B}(q(h^{m+1}))h^{m+1} - \mathbf{A}(q)h^m - \mathbf{R}(q(h^m)), \lambda^m \rangle_{l^2} \tag{4.51}$$

Solving for $\lambda^{M,N}$ the adjoint equation (backwards in time with $\lambda^{M,N}_{M,\cdot} = \mathbf{0}$):

$$\frac{\partial}{\partial h} J^{\beta,N,N_I,M}(h^a, h^I, q) + \sum_{m\leq M} \langle \frac{\partial}{\partial h}[\mathbf{B}(q(h^{m+1}))h^{m+1} - \mathbf{A}(q)h^m - \mathbf{R}(q(h^m))], \lambda^{M,N}_{m,\cdot} \rangle_{l^2} = \mathbf{0} \tag{4.52}$$

then we can use $\lambda^{M,N}$ to compute:

$$\frac{d}{dq} J^{\beta,N,N_I,M}(h(q), q)\delta q = \frac{\partial}{\partial h} J^{\beta,N,N_I,M}(h^a, h^I, q)\delta q+$$

$$\sum_{m\leq M} \langle \frac{\partial}{\partial q}[\mathbf{B}(q(h^{m+1}))h^{m+1} - \mathbf{A}(q)h^m - \mathbf{R}(q(h^m))]\delta q, \lambda^{M,N}_{m,\cdot} \rangle_{l^2} \tag{4.53}$$

Since the Lagrangian 4.51 the partial derivatives in (4.52) and (4.53) and the total derivatives (4.52) are all finite dimensional, it is important to point the fact that δq has the dimension of the discrete version of $(S, K^x, K^y, K_v, S_{ke}, S_{kv})$ consequently the dimension of the vector is large this requires an intelligent assembly and the use of parallel programing rather than sequential. As in (4.15) we approximate $J(h(q), q)$ around q_0 by a quadratic function:

$$m(q) = J(q_0) + J_q(q_0)q + q^T J_{qq}(q_0)q \tag{4.54}$$

As before to find the minimum we have to solve the $m_q(q) = 0$ and use the $BFGS$ method to approximate $J_{qq}(q_0)$. Since the optimization problem is hard and costly one should use the so popular multilevel optimization. Starting then with a corse mesh and agglutinating then zones that have nearly the same values as the following algorithm illustrates:

This approach can be summarized in the following pseudo code:

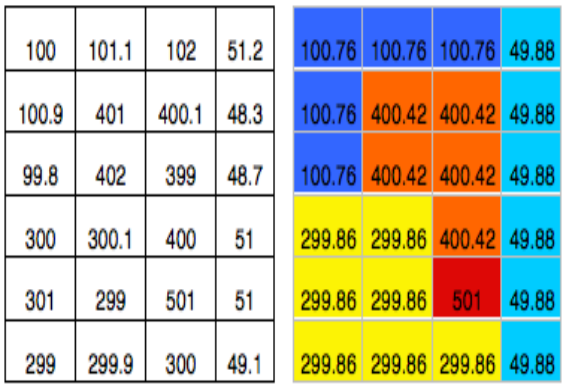

Figure 4.1: On the left is the mesh before the agglutination where on the right after the agglutination

4.3.3 Zonation Algorithm

- Given $T^{0,0}$, $S^{0,0}$, tol, ${Z_T}^0$, ${Z_S}^0$,ϵ, N^{out}, r^T, r^S, $k = 0$, $j = 0$.
  1. Optimize using Quasi-Newton while 1a and 1b are false
     (a) $\|J(T^{k,j}, S^{k,j})\| < \epsilon$ *stop*
     (b) $\|J(T^{k,j+1}, S^{k,j+1}) - J(T^{k,j}, S^{k,j})\| < tol$ move to 2
  2. Regroup from ${Z_T}^k$, ${Z_S}^k$, to ${Z_T}^{k+1}$, ${Z_S}^{k+1}$, zones as follows:
     – for $2 \leq n \leq N_y - 1$
     – for $2 \leq m \leq N_x - 1$
       (a) For all neighbors $T^{k,j}_{a,b}$ of $T^{k,j}_{n,m}$, if $|T^{k,j}_{n,m} - T^{k,j}_{a,b}| < r^T$ then set $Z^k_{Ta,b} = Z^k_{Tn,m}$
       (b) For all neighbors $S^{k,j}_{a,b}$ of $S^{k,j}_{n,m}$, if $|S^{k,j}_{n,m} - S^{k,j}_{a,b}| < r^S$ then set $Z^k_{Sa,b} = Z^k_{Sn,m}$
     – end for
     – end for
  3. Change the variables change from $T^{k,j}$ to $T^{k+1,0}$ and $S^{k,j}$ to $S^{k+1,0}$ by the transformations
     – $T^{k+1,0}_l = average\{T^{k,j}$, such as $T^{k,j} \in Z(l)\}$
     – $S^{k+1,0}_l = average\{S^{k,j}$, such as $S^{k,j} \in Z(l)\}$
     – $k = k + 1$
  4. Repeat 1,2 and 3 until convergence.

The results of the implementation of this algorithm can be find Section 7.1

4.3.4 Multi-Level Optimization

. This technique as been used in medical imaging more specifically in registration. Jan Modersitzki has developed this very elegantly on his book Numerical Methods for Image Registration [33], and Melissa Weber Mendona in her thesis Multilevel Optimization: Convergence Theory, Algorithms and Application to Derivative-Free Optimization [32] has proved relevant results in convergence. As it is known it is a very hard problem to choose the right initial guess for a newton scheme and since their convergence is local rather than global its importance must not be neglected. The Multi Level is then an approach that tries to get a good initial guess by, firstly starting with an initial guess, it can be constant across the whole domain, and a very sparse mesh which is a easier problem to solve. This will converge eventually then the next step to refine the previous mesh and optimize on that same mesh using as the initial guess the latest iteration of the previous Newton optimization, this process is repeated until a minimum or a tolerance is reached. Mesh independence [26], a nonintuitive result, states that the number of iterations for convergence of the Newton method is independent of the size of the discretization. The following pseudo code illustrates this method when applied to the problem in hand.

- Given $T^{0,0}$, $S^{0,0}$, tol, ${Z_T}^0$, ${Z_S}^0$,ϵ, N^{out}, r^T, r^S, $k = 0$, $j = 0$.
  1. Optimize using Quasi-Newton while 1a and 1b are false
     (a) $\|J(T^{k,j}, S^{k,j})\| < \epsilon$ *stop*
     (b) $\|J(T^{k,j+1}, S^{k,j+1}) - J(T^{k,j}, S^{k,j})\| < tol$ move to 2
  2. Partition ${Z_T}^k$, ${Z_S}^k$, into a finer submesh ${Z_T}^{k+1}$, ${Z_S}^{k+1}$,.
  3. Change the variables change from $T^{k,j}$ to $T^{k+1,0}$ and $S^{k,j}$ to $S^{k+1,0}$ by the transformations
     - $T_l^{k+1,0} = \{T^{k,j}$, such as $T^{k,j} \in Z(l)\}$
     - $S_l^{k+1,0} = \{S^{k,j}$, such as $S^{k,j} \in Z(l)\}$
     - $k = k + 1$
  4. Repeat 1,2 and 3 until convergence.

Figure 4.2 illustrates an example of this mesh partition:

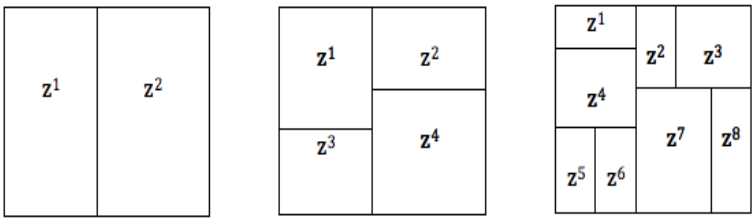

Figure 4.2: From left to right is a sequence of meshes from coarse to fine

In the sequence above, we would have $T_2^0(Z_1^2) := T_1^f(Z_1^1)$. The implementation of this method is discussed in 7.3.

Chapter 5

Sensitivity Analysis and Fréchet Derivative Operators

5.1 Outline

This chapter is dedicated to the study of sensitivity analysis for steady elliptic equations, convection-diffusion, and groundwater flow equations. In particular, it discusses the existence of the Fréchet derivative of the solution of a PDE with respect to distributed coefficients in the PDE as well as boundary conditions. This is followed by results on the existence of a spectral decomposition of this operator, which under suitable conditions is shown to be *Hilbert–Schmidt*. Furthermore, to set up our numerical algorithms, we present results on the approximation of the operator using sensitivity equations. An overview of the finite element spaces is presented and convergence results for the spectral decomposition are thoroughly discussed.

5.2 Background

One of the main problems in parameter estimation is the lack of sensitivity of the cost functional to variations in the parameters. In other words, when large variations in the parameter lead to small variations in the solution and consequently, the cost functional. The study of the effect of the parameters on the solution is known as *sensitivity analysis*. To illustrate its importance we consider perturbing one parameter in the 1D groundwater flow equation. The model equation, including specific parameter values for this study is

$$
\begin{aligned}
(1+sin(\pi x))\frac{d}{dt}z(t,x;q) &= (e^{-4x}z(t,x;q)_x)_x + 1 + x + t, \qquad z(t,\cdot;q) \in H^1([0,1]) \qquad (5.1)\\
z(0,x;q) &= x \qquad\qquad x \in [0,1]\\
z(t,0;q) &= 0 \qquad\qquad t \in [0,1]\\
z(t,1;q) &= e^{-t} \qquad\qquad t \in [0,1].
\end{aligned}
$$

Perturbing the elliptic coefficient $q(x) = e^{-4x}$ by the small variation $\delta q(x) = \frac{1}{100}e^{9(x-1)}$, a relative change of 2.4×10^{-3} in the L^2-norm leads to the system

$$
\begin{aligned}
(1 + sin(\pi x))\frac{d}{dt}z(t,x;q+\delta q) &= ([e^{-4x} + \frac{1}{100}e^{9(x-1)}]z(t,x;q+\delta q)_x)_x + 1 + x + t \\
z(0,x;q+\delta q) &= x && x \in [0,1] \\
z(t,0;q+\delta q) &= 0 && t \in [0,1] \\
z(t,1;q+\delta q) &= e^{-t} && t \in [0,1].
\end{aligned}
$$

These two solutions are simulated and displayed in Figure 5.1. As one can see, a small

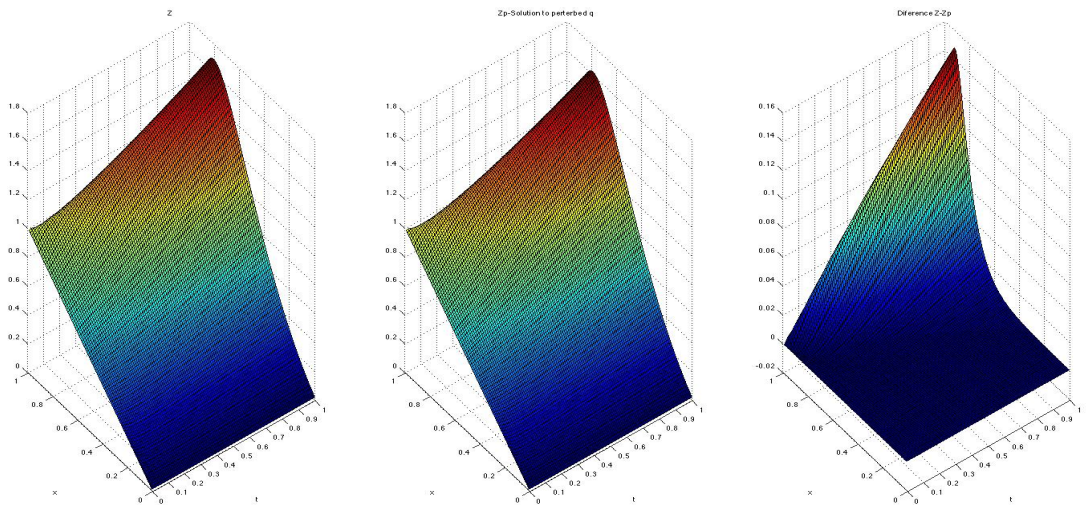

Figure 5.1: On the left is the solution of the original system, on the middle the solution of the perturbed system and finally on the right the difference between both solutions

perturbation in the parameter leads to a dramatic change in the solution (the difference has a maximum value of 1.5×10^{-1}, or a 10% relative difference for a relative change of less than a percent). This effect can be amplified by changing parameter values. Thus, this simple example motivates the study of sensitivity analysis, which is detailed in this chapter. We apply sensitivity analysis on groundwater flow equations, among others.

In mathematics when one needs to study the behavior of a real valued function, such as finding a local maximum or minimum, or study where it is increasing or decreasing, an important tool is the function's derivative, if it exists. This same line of reasoning can be extended to a functional, and leads to, for example, Gâteaux or Fréchet derivatives. We review their definitions below.

Definition 5.2.1. *[Gâteaux differentiable operator] Let X and Y be normed spaces and $A : X \to Y$ be an operator. The operator A is* Gâteaux-differentiable *at x in the direction h if and only if there exists an operator $G_x : X \to Y$ such that*

$$\lim_{t \to 0} \frac{\|A(x+th) - A(x) - G_x(h)\|_Y}{t} \to 0. \tag{5.2}$$

Definition 5.2.2 (Fréchet differentiable operator)**.** *Let X and Y be normed spaces and $A : X \to Y$ be an operator. The operator A is* Fréchet-differentiable *at x if and only if there exists a bounded linear operator F_x such that*

$$\lim_{\|h\|\to 0} \frac{\|A(x+h) - A(x) - F_x(h)\|_Y}{\|h\|_X} \to 0. \tag{5.3}$$

There are several differences between Gâteux and Fréchet differentiability. For one, the Gâteux derivative is not necessarily either continuous or linear, whereas the Fréchet derivative must be a bounded linear operator. However, if the operator is Fréchet differentiable, we know it is also Gâteux differentiable.

Proposition 5.2.3. *Let X and Y be normed spaces and $A : X \to Y$ be a Fréchet-differentiable operator. Then A is also Gâteux-differentiable at x. Furthermore, the directional derivative of A with respect to x in the direction h, written $G_x(h)$ is given as*

$$G_x(h) = F_x(h),$$

where F_x is the Fréchet derivative of A with respect to x.

5.3 Sensitivity Analysis using Fréchet Derivatives

The Fréchet derivative operator is very useful for parameter estimation and performing sensitivity analysis for PDE's or dynamical systems. They provide useful information on parameter variations, including their "energy" and the spatial characterization of the impact they have. In the following paragraphs, we discuss the existence of the Fréchet derivative of the solution to a PDE with respect to the PDE parameters. Motivated by our groundwater flow models, we first consider an elliptic equation then consider advection-diffusion equations.

Consider the following elliptic PDE

$$\begin{aligned} -\nabla \cdot (q \nabla z) &= f, \qquad f \in L^2(\Omega) \\ z &\in H_0^1(\Omega), \qquad \Omega \subset \mathbb{R}^n. \end{aligned} \tag{5.4}$$

One can consider the solution z, as an operator acting on q. Then we can write this operator $z(q)$ with domain $Q^{ad} \subset L^2(\Omega)$, the set of all q such that there exists an unique solution of (5.4). Now that the operator $z(q)$ is defined, we can study the existence of the Fréchet derivative of $z(q)$ with respect to q.

5.3.1 Fréchet Differentiability and the Sensitivity Equation

For this discussion, we assume Ω is bounded with a C^0 boundary and the set of admissible parameters q is

$$Q^{ad} \equiv \{q : q \in C^1(\Omega) \quad \text{and} \quad q(x) \geq \alpha_0 > 0, \quad \forall x \in \Omega\} \tag{5.5}$$

and consider the max norm on Q^{ad}, $\|h\| = max_{x\in\Omega}\{|h(x)|\} + max_{x\in\Omega}\{|h'(x)|\}$.

Theorem 5.3.1 (Fréchet Differentiability of Solutions to Elliptic Equations)**.** *Let $z(q)$ be the solution to (5.4) with $q \in Q^{ad}$. Then the Fréchet derivative of $z(q)$ with respect to q applied to h,*

$$v_h \equiv [D_q z(q)]_{q=q_0}(h) \tag{5.6}$$

is the weak solution of the well-posed elliptic equation

$$\mathcal{S}_{q_0} \begin{cases} -\nabla \cdot (q_0 \nabla v_h) = \nabla \cdot (h \nabla z(q_0)) \\ v_h \in H_0^1(\Omega). \end{cases} \tag{5.7}$$

Proof. Let z_h be the solution to

$$\begin{cases} -\nabla \cdot ((q+h)\nabla z_h) = f, \quad f \in L^2(\Omega) \\ z_h \in H_0^1(\Omega). \end{cases} \tag{5.8}$$

then subtracting (5.4) from (5.8) we obtain

$$-\nabla \cdot (q \nabla (z_h - z(q))) = \nabla \cdot (h \nabla z_h)$$

then subtracting equation (5.7) we find

$$-\nabla \cdot (q \nabla (z_h - z(q) - v_h)) = \nabla \cdot (h \nabla (z_h - z(q))).$$

Multiplying by $(z_h - z(q) - v_h)$, and integrating by parts we obtain:

$$\int q |\nabla (z_h - z(q) - v_h)|^2 = -\int h \nabla (z_h - z(q)) \nabla (z_h - z(q) - v_h).$$

Taking the absolute value and applying the Cauchy-Schwarz inequality leads to

$$\int \alpha_0 |\nabla (z_h - z(q) - v_h)|^2 \leq \|h\| \int |\nabla (z_h - z(q))| \|\nabla (z_h - z(q) - v_h)|.$$

Since Ω is bounded we apply Poincare's inequality on the left side and Holder's inequality on right side. Thus there is a constant C such that

$$\alpha_0 \|z_h - z(q) - v_h\|_{H^1}^2 \leq C \|h\| \|z_h - z(q) - v_h\|_{H^1} \|z_h - z(q)\|_{H^1}. \tag{5.9}$$

However, for all h such that $\|h\| \leq 1$ there exists $C > 0$ such that

$$\|z_h - z(q)\|_{H^1} \leq C \|h\| \|f\|_{L^2}. \tag{5.10}$$

Therefore by combining (5.9) with (5.10) we obtain $\|z_h - z(q) - v_h\|_{H^1} \leq C \|h\|^2 \|f\|_{L^2}$. Dividing both sides by $\|h\|$ and taking limits,

$$\lim_{\|h\| \to 0} \frac{\|z_h - z(q) - v_h\|_{H^1}}{\|h\|} \leq \lim_{\|h\| \to 0} C \|h\| \|f\|_{L^2},$$

which leads to

$$\lim_{\|h\| \to 0} \frac{\|z_h - z(q) - v_h\|_{H^1}}{\|h\|} = 0.$$

We can conclude that $[D_q z(q)]_{q=q_0}(h) = v_h$. In other words, the solution to (5.7) is indeed the Frechét derivative of z with respect to q applied to h. □

5.3.2 Approximation of the Fréchet Derivative Operator

Upon careful inspection of equation (5.7), we see that we need to first evaluate the solution to (5.4) in order to obtain the derivative. The solution to (5.4), z, is required to form the right hand side of (5.7). There are a number of relevant questions to ask when approximating the Fréchet derivatives: when we compute the finite element solution of (5.4) with mesh size Δx_1, $z_{\Delta x_1}$, and substitute it into (5.7) approximated with mesh size Δx_2, does the Finite Element solution, $v_{h\Delta x_2}$, converge as Δx_2 and Δx_1 converge to zero? especially in the typical case $\Delta x_1 = \Delta x_2$? In other words, does $v_{h\Delta x_2}$ converge to $[D_q z(q)]_{q=q_0}(h) = v_h$? This is answered by the following theorem.

Theorem 5.3.2 (Operator convergence)**.** *Let $z(q)$ and $z_{\Delta x_1}$ be the solution and the finite element solution of the elliptic equation (5.4). Let $r = v_h$ be the solution of the sensitivity equation (5.7). Respectively, let $\tilde{r}$ and $\tilde{r}_{\Delta x_2}$ be the solution and finite element solution of*

$$-\nabla \cdot (q\nabla \tilde{r}) = \nabla \cdot (h\nabla z(q)_{\Delta_{x_1}}), \qquad \tilde{r} \in H_0^1(\Omega). \tag{5.11}$$

Then if $\Delta x_1 \to 0$ and $\Delta x_2 \to 0$, we have $\|\tilde{r}_{\Delta x_2} - [D_q z(q)]_{q=q_0}(h)\|_{H^1} \to 0$.

Proof. Adding and subtracting $\tilde{r}$ and applying the triangle inequality to $\|r - \tilde{r}_{\Delta x_2}\|_{H^1}$, we obtain

$$\|r - \tilde{r}_{\Delta x_2}\|_{H^1} \leq \|r - \tilde{r}\|_{H^1} + \|\tilde{r} - \tilde{r}_{\Delta x_2}\|_{H^1}. \tag{5.12}$$

First we bound the term $\|r - \tilde{r}\|_{H^1}$. If we subtract (5.7) and (5.11), we obtain

$$-\nabla \cdot (q\nabla(r - \tilde{r})) = \nabla \cdot (h\nabla(z(q) - z_{\Delta x_1})).$$

Multiplying this by $(r - \tilde{r})$ and integrating by parts we obtain

$$\int q|\nabla(r - \tilde{r})|^2 = -\int h(\nabla(z(q) - z_{\Delta x_1})) \cdot \nabla(r - \tilde{r})) \leq h_{max}\|r - \tilde{r}\|_{H^1}\|z(q) - z_{\Delta x_1}\|_{H^1}.$$

Poincaré's inequality combined with the fact that $\alpha_0 \leq q$, from (5.5), we see that there exists a $C > 0$ such that

$$\|r - \tilde{r}\|_{H^1}^2 \leq C \int q|\nabla(r - \tilde{r})|^2.$$

Combining this with (5.12) we get:

$$\|r - \tilde{r}\|_{H^1}^2 \leq Ch_{max}\|r - \tilde{r}\|_{H^1}\|z(q) - z_{\Delta x_1}\|_{H^1}.$$

Thus

$$\|r - \tilde{r}\|_{H^1} \leq Ch_{max}\|z(q) - z_{\Delta x_1}\|_{H^1}. \tag{5.13}$$

We now bound the term $\|\tilde{r} - \tilde{r}_{\Delta x_2}\|_{H^1}$. By Theorem 9.1.10 from [15],

$$\|\tilde{r} - \tilde{r}_{\Delta x_2}\|_{H^1} \leq C_1 \Delta x_2 \|\Delta \tilde{r}\|_{L^2}. \tag{5.14}$$

This same theorem can be applied to the finite element approximation of z to find

$$\|z(q) - z_{\Delta x_1}\|_{H^1} \leq C_2 \Delta x_1 \|\nabla z\|_{L^2}.$$

Combining (5.13) and the above inequality with (5.12) we have

$$\|r - \tilde{r}_{\Delta x_2}\|_{H^1} \leq C_2 \Delta x_2 \|\Delta \tilde{r}\|_{L^2} + C_2 h_{max} \delta x_1 \|\Delta z(q)\|_{L_2}.$$

Applying the limit

$$\lim_{\Delta x_2 \to 0, \Delta x_1 \to 0} \|r - \tilde{r}_{\Delta x_2}\|_{H^1} \leq \lim_{\Delta x_2 \to 0, \Delta x_1 \to 0} C_2(\Delta x_2 \|\Delta \tilde{r}\|_{L^2} + \|z(q) - z_{\Delta x_1}\|_{H^1}) \tag{5.15}$$

$$= \lim_{\Delta x_2 \to 0} C_2(\Delta x_2 \|\Delta r\|_{L^2}) = 0, \tag{5.16}$$

since $r, z \in H^2(\Omega)$. □

Any linear operator defined on the finite dimensional space H, is uniquely defined by the image of a basis. Therefore if we project the admissible set Q^{ad} onto a finite dimensional space, typically a Galerkin finite element basis, we have a finite rank representation of the Fréchet derivative operator. The following theorem considers the convergence of the finite rank operator to the infinite dimensional counterpart.

Theorem 5.3.3 (Operator approximation)**.** *Let H^N be a finite dimensional space of $C^1(\Omega)$ with basis $\{\phi_i\}_{i=1}^{i=N}$. Let P^N be the projection of $C^1(\Omega)$ onto H^N. Furthermore, let D^N be the linear operator that is defined by $[D_q z(q)]_{q=q_0}(\phi_i) = v_{\phi_i}$, where v_{ϕ_i} is the solution of (5.7) when $\phi_i = h$. If P^N is the finite element projection operator satisfying*

$$\|P^n h - h\| \to 0 \quad as \quad N \to \infty, \qquad \forall h \in C^1(\Omega).$$

Then $D^N(P^N h) \to [D_q z(q)]_{q=q_0}$ as $N \to \infty$.

Proof. Let $z(q)$ be the solution of (5.4) and let $h \in C^1(\Omega)$ be given. Define $\tilde{h} = P^N h$ and $\sum_{i=1}^N h_i \phi_i$. Then $D^N(P^N h)$ is the solution to

$$-\nabla \cdot (q \nabla w) = \nabla \cdot (\tilde{h} \nabla z(q)), \qquad w \in H_0^1(\Omega). \tag{5.17}$$

Let v_h be the solution of (5.17) with $\tilde{h}$ replaced by h, then:

$$-\nabla \cdot (q \nabla (v_h - v_{\tilde{h}})) = \nabla \cdot ((h - \tilde{h}) \nabla z(q)).$$

Multiplying both sides by $v_h - v_{\tilde{h}}$ and integrating by parts we obtain

$$\int q |\nabla (v_h - v_{\tilde{h}})|^2 = \int (h - \tilde{h}) \nabla (v_h - v_{\tilde{h}}) \cdot \nabla z(q),$$

which leads to the following inequality

$$\|v_h - v_{\tilde{h}}\|_{H_0^1}^2 \leq C \|h - \tilde{h}\| \|v_h - v_{\tilde{h}}\|_{H_0^1} \|u\|_{H_0^1}$$

then $\|v_h - v_{\tilde{h}}\|_{H_0^1} \leq C \|h - \tilde{h}\| \|u\|_{H_0^1}$. By our assumption on P^N, $\|h - \tilde{h}\| \to 0$ as $N \to \infty$. Therefore $\|v_h - v_{\tilde{h}}\|_{H_0^1} \to 0$. □

Since we can find a finite dimensional representation of $[D_q z(q)]_{q=q_0}$, we would like to prove that the infinite dimensional operator is *Hilbert–Schmidt* or at least has a dominating eigenvalue. As we discuss below, the Hilbert–Schmidt triplets can be approximated by the singular value decomposition of our finite dimensional representation to $[D_q(q)]_{q=q_0}$. Furthermore, these triplets lead to interesting applications.

Definition 5.3.4 (Hilbert-Schmidt Operator)**.** *A linear operator $F : H_1 \to H_2$, where H_1 and H_2 are Hilbert spaces, is* Hilbert–Schmidt *if and only if there exists an H_1-orthornormal sequence in H_1, $\{v_k\}$, an H_2-orthornormal sequence in H_2, $\{u_k\}$, and positive decreasing sequence of numbers $\{\sigma_k\}$ satisfying $\sum_{k=1}^{\infty} \sigma_k^2 < \infty$, such that for any $h \in H_1$,*

$$F(h) = \sum_{k=1}^{\infty} \sigma_k \langle v_k, h \rangle_{H_1} u_k.$$

Theorem 5.3.5 (Sufficient Condition)**.** *If there exists an orthonormal basis $\{\phi\}_n$ of H_2 such that $\|F\phi_n\| \in l^2$, then F is a Hilbert-Schmidt operator [41].*

Lets consider the space $H^{2+\varepsilon}(\Omega)$ where $\varepsilon > 0$ and $\Omega \subset \mathbb{R}^2$ is bounded. It is well known that $X = H^{2+\varepsilon}(\Omega) \hookrightarrow C^1(\Omega)$. Thus convergence in $H^{2+\varepsilon}(\Omega)$ implies convergence in $C^1(\Omega)$. Therefore one can consider a Hilbert–Schmidt decomposition of the operator $D_q z(q)$ since both the domain $\mathcal{D} = H^{2+\varepsilon}(\Omega)$ and the range $\mathcal{R} \subset H_0^1$ are Hilbert spaces. Since $\mathcal{D}$ is a subset of $C^1(\Omega)$ the conditions of Theorem 5.3.1 are met. Thus the operator $[D_q z(q)]_{q=q_0}$ is well defined. Now we just have to prove the existence of the Hilbert-Schmidt decomposition.

Theorem 5.3.6 ($D_q(z(q))$ is a Hilbert–Schmidt Operator)**.** *If $\nabla z \in L^\infty(\Omega)$ then the Fréchet derivative of z(q) with respect to q, $[D_q z(q)]_{q=q_0}$, is a Hilbert-Schmidt operator in the domain $H^{2+\varepsilon}(\Omega)$ with $\varepsilon > 0$. As an example that the hypothesis, $\nabla z \in L^\infty(\Omega)$ can be achieved consider the case $f \in H^1(\Omega), q \in C^2(\Omega)$. This implies that $z(q) \in H^3(\Omega)$ and by the Sobolev imbedding theorem, $z(q) \in C^1(\Omega)$.*

Proof. As shown in Adams [1], there is an orthonormal basis, $\{\phi_n\}_{n\geq 1}$, of $H^{2+\varepsilon}(\Omega)$, where $\|\phi_n\|_{L^2(\Omega)}$ is an l^2 sequence. Then using the equation (5.7), we have that $v_{\phi_n} = [D_q z(q)]_{q=q_0}(\phi_n)$ is a solution of

$$-\nabla \cdot (q \nabla v_{\phi_n}) = \nabla \cdot (\phi_n \nabla z(q_0)), \qquad v_{\phi_n} \in H_0^1(\Omega).$$

Multipying both sides of the equation by v_{ϕ_n} and integrating by parts we obtain

$$\int q |\nabla v_{\phi_n}|^2 = -\int \phi_n \nabla z \cdot \nabla v_{\phi_n}.$$

Since q is bounded below by a constant, we apply the Hölder inequality to obtain $q_0 \|\nabla v_{\phi_n}\|_{L^2}^2 \leq \|\phi_n \nabla z\|_{L^2} \|\nabla v_{\phi_n}\|_{L^2}$. Using the Poincaré inequality, there is a constant c such that $\|v_{\phi_n}\|_{H_0^1}^2 \leq c \|\phi_n \nabla z\|_{L^2} \|v_{\phi_n}\|_{H_0^1}$. Thus $\|v_{\phi_n}\|_{H_0^1} \leq c \|\phi_n \nabla z\|_{L^2}$.

Note that $\phi_n = \frac{\sin(2\pi n x)}{1/2 + 2n^2\pi^2 + 8n^4\pi^4}$ is a orthonormal basis in $H^2(0,1) \cap H_0^1(0,1)$. However, the L^2 norm of the nth element is given by: $\frac{1}{1+4n^2\pi^2+16n^4\pi^4}$. Therefore the sequence $\|\phi_n\|_{L^2}$ is an l^2 sequence.

By hypothesis, $\nabla z \in L^\infty(\Omega)$. Thus we have $\|v_{\phi_n}\|_{H_0^1} \leq c\|\nabla z\|_{L^\infty}\|\phi_n\|_{L^2}$. This will lead us to $[D_q z(q)]_{q=q_0}(\phi_n) \in l^2$ since $\|\phi_n\|_{L^2} \in l^2$ using Theorem 5.3.5 we have that $[D_q z(q)]_{q=q_0}(\phi_n)$ is a Hilbert-Schmidt operator. □

Theorem 5.3.7. *The Fréchet derivative operator is compact, cf. [54].*

Theorem 5.3.8. *[Differentiability with a Parameterized Forcing Term] Let $p = (q, \theta)$ belong to the admissible set*

$$Q^{ad} = \{(q, \theta) : q, \theta \in C^1(\Omega) \text{ and } \exists \alpha_0 > 0 \text{ such that } q(x) \geq \alpha_0, \ \forall x \in \Omega\}.$$

Let $z(p)$ be the solution of

$$-\nabla \cdot (q \nabla z(p)) = f(\theta), \qquad f \in L^2(\Omega), \qquad \text{and} \qquad z \in H_0^1(\Omega). \tag{5.18}$$

where the norm on Q^{ad} is defined by $\|h\| \equiv \|q\| + \|\theta\|$. If f is Lipchitz continuous with respect to θ, then the Fréchet derivative of $z(q, \theta)$ with respect to (q, θ) at $p_0 = (q_0, \theta_0)$ applied to $h = (\delta q, \delta\theta)$, $v_h = [D_p z(p)]_{p=p_0}(\phi_n)$, is the weak solution of

$$-\nabla \cdot (q_0 \nabla v_h) = \nabla \cdot (\delta q \nabla z(p_0)) + [f_\theta]_{\theta=\theta_0}(\delta\theta), \qquad v_h \in H_0^1(\Omega). \tag{5.19}$$

where $[f_\theta]_{\theta=\theta_0}(\delta\theta)$ is the Fréchet derivative of f at θ_0 applied to $\delta\theta$.

Proof. Given

$$\begin{aligned} -\nabla \cdot (q_0 \nabla v_h) &= \nabla \cdot (\delta q \nabla z(p_0)) + [f_\theta]_{\theta=\theta_0}(\delta\theta), \\ -\nabla \cdot ((q_0 + \delta q)\nabla z_h) &= f(\theta_0 + \delta\theta), \end{aligned} \tag{5.20}$$

and

$$-\nabla \cdot (q_0 \nabla z(p_0)) = f(\theta_0). \tag{5.21}$$

We can combine the three equations above to obtain

$$-\nabla \cdot (q_0 \nabla (z_h - z(p_0) - v_h)) - \nabla \cdot (\delta q \nabla (z_h - z(p_0))) = f(\theta_0 + \delta\theta) - f(\theta_0) - [f_\theta]_{\theta=\theta_0}(\delta\theta).$$

Multiplying by $z_h - z(p_0) - v_h$ and integrating by parts,

$$\begin{aligned} \int q_0 |\nabla (z_h - z(p_0) - v_h)|^2 \leq &\int |\delta q| \nabla (z_h - z(p_0)) \cdot \nabla (z_h - z(p_0) - v_h) \\ &+ \int |f(\theta_0 + \delta\theta) - f(\theta_0) - [f_\theta]_{\theta=\theta_0}(\delta\theta)(z_h - z(p_0) - v_h). \end{aligned}$$

Applying applying Poincaré's inequality to the left hand side and Holder's inequality on the right hand side

$$\begin{aligned} \|(z_h - z(p_0) - v_h)\|_{H_0^1}^2 \leq &C\|\delta q\| \|(z_h - z(p_0))\|_{H_0^1} \|(z_h - z(p_0) - v_h)\|_{H_0^1} + \\ &\|f(\theta_0 + \delta\theta) - f(\theta_0) - [f_\theta]_{\theta=\theta_0}(\delta\theta)\|_{L^2} \|(z_h - z(p_0) - v_h)\|_{H_0^1}. \end{aligned}$$

This leads to

$$\begin{aligned}\|(z_h - z(p_0) - v_h)\|_{H_0^1} \leq & C\|\delta q\|\|(z_h - z(p_0))\|_{H_0^1} + \\ & C\|f(\theta_0 + \delta\theta) - f(\theta_0) - [f_\theta]_{\theta=\theta_0}(\delta\theta)\|_{L^2}.\end{aligned}$$

which, using (5.20) and (5.21) is

$$\|(z_h - z(p_0) - v_h)\|_{H_0^1} \leq C\|\delta q\|(\|\delta q\| + \|\delta\theta\|) + C\|f(\theta_0 + \delta\theta) - f(\theta_0) - [f_\theta]_{\theta=\theta_0}(\delta\theta)\|_{L^2}.$$

Dividing by $\|\delta h\| = \|\delta q\| + \|\delta\theta\|$,

$$\frac{\|(z_h - z(p_0) - v_h)\|_{H_0^1}}{\|\delta q\| + \|\delta\theta\|} \leq C\|\delta q\| + C\frac{\|f(\theta_0 + \delta\theta) - f(\theta_0) - [f_\theta]_{\theta=\theta_0}(\delta\theta)\|_{L^2}}{\|\delta q\| + \|\delta\theta\|}.$$

Now, using the fact that $[f_\theta]_\theta$ is the Fréchet derivative, the last terms vanishes as $\|\delta\theta\| \to 0$. Therefore,

$$\lim_{\|\delta q\|+\|\delta\theta\|\to 0} \frac{\|(z_h - z(p_0) - v_h)\|_{H_0^1}}{\|\delta q\| + \|\delta\theta\|} = 0.$$

□

Theorem 5.3.9. *Under the conditions of Theorem 5.3.8 and the added condition that $[f_\theta]_{\theta=\theta_0}(\delta\theta)$ be Lipchitz continuous, the Fréchet derivative, $v_\theta = [D_p z(p)]_{\theta=\theta_0}(\delta\theta)$ is the solution of*

$$-\nabla \cdot (q_0 \nabla v_\theta) = [f_\theta]_{\theta=\theta_0}(\delta\theta), \qquad v_\theta \in H_0^1(\Omega). \tag{5.22}$$

Proof. Follows from Theorem 5.3.8 with $\delta q = 0$. □

Theorem 5.3.10. *Let H^N be a finite dimensional subspace of $C^1(\Omega)$, let P^N be the projection of $C^1(\Omega)$ onto H^N, and $\{\phi_i\}_{i=1}^{i=N}$ be a basis for H^N. Furthermore, let D^N be the linear operator that is defined by $[D_\theta z(\theta)]_{\theta=\theta_0}(\phi_i) = v_{\phi_i}$ where v_{ϕ_i} is the solution of (5.22) when $\delta\theta = \phi_i$.*

If $\|P^n\theta - \theta\| \to 0$ as $N \to \infty$, for all $h \in C^1(\Omega)$, then $D^N(P^N h) \to [D_\theta z(\theta)]_{\theta=\theta_0}(\delta\theta)$ as $N \to \infty$.

Proof. Let r be the solution of (5.22), and let $\delta\theta \in C^1(\Omega)$ and $\delta\tilde{\theta} = P^N\delta\theta = \sum_{i=1}^N \delta\theta_i\phi_i$. $D^N(P^N\delta\theta)$ is the solution of:

$$-\nabla \cdot (q_0 \nabla r) = [f_\theta]_{\theta=\theta_0}(\delta\tilde{\theta}), \qquad r \in H_0^1(\Omega). \tag{5.23}$$

Let $v_{\delta\theta}$ be the solution of (5.3.10) then

$$-\nabla \cdot (q\nabla(v_{\delta\theta} - v_{\tilde{\theta}})) = [f_\theta]_{\theta=\theta_0}(\delta\theta - \delta\tilde{\theta})$$

Multiplying both sides by $v_{\delta\theta} - v_{\delta\tilde{\theta}}$ and integrating by parts we obtain

$$\int q_0|\nabla(v_{\delta\theta} - v_{\delta\tilde{\theta}})|^2 = \int (v_{\delta\theta} - v_{\tilde{\theta}})[f_\theta]_{\theta=\theta_0}(\delta\theta - \tilde{\theta}).$$

Using similar arguments used in the proof of Theorem 5.3.8, the following inequality holds

$$\|v_{\delta\theta} - v_{\delta\tilde{\theta}}\|_{H_0^1}^2 \leq C\|\delta\theta - \delta\tilde{\theta}\|\|v_{\delta\theta} - v_{\delta\tilde{\theta}}\|_{H_0^1},$$

thus

$$\|v_{\delta\theta} - v_{\delta\tilde{\theta}}\|_{H_0^1} \leq C\|\delta\theta - \delta\tilde{\theta}\|.$$

Due to the approximation property of P^N, $\|\delta\theta - \delta\tilde{\theta}\| \to 0$ as $N \to \infty$. Therefore

$$\|v_{\delta\theta} - v_{\delta\tilde{\theta}}\|_{H_0^1} \to 0.$$

□

Combining Theorems 5.3.3 and 5.3.10, we have the following result

Theorem 5.3.11. *Under the hypotheses of Theorems 5.3.3 and 5.3.10,*

$$[D_p z(p)]_{p=(q_0,\theta_0)}(\delta h, \delta\theta) = [D_q z(q)]_{q=q_0}(\delta h) + [D_\theta z(\theta)]_{\theta=\theta_0}(\delta\theta). \tag{5.24}$$

Proof. Let H^N be a finite dimensional subspace of $C^1(\Omega)$ and let P^N be the projection of $C^1(\Omega) \times C^1(\Omega)$ into H^N with basis $\{(\phi_i, 0); (0, \phi_j)\}_{i,j=1}^{i,j=N}$. Let D^N be the linear operator that is defined by $[D_p z(p)]_{p=p_0}(\phi_i, \phi_j) = v_{(\phi_i,\phi_j)}$ where $v_{(\phi_i,\phi_j)}$ is the solution of (5.19) when $(\delta h, \delta\theta) = (\phi_i, \phi_j)$.

Since $\|P^N h - h\| \to 0$ as $N \to \infty$, $\forall h \in C^1(\Omega)$, we have

$$D^N(P^N(\delta q, \delta\theta)) \to D_p[z(p)]_{p=p_0}((\delta q, \delta\theta)),$$

as $N \to \infty$. □

Theorem 5.3.12. *The Fréchet derivative of $z(q)$ with respect to q, $[D_p z(p)]_{p=p_0}$, is a Hilbert-Schmidt operator.*

Proof. Since the space of Hilbert-Schmidt operators is a vector space one can infer that since $[D_q z(q)]_{q=q_0}(\delta h)$ and $[D_\theta z(\theta)]_{\theta=\theta_0}(\delta\theta)$ are Hilbert-Schmidt then the sum $[D_p z(p)]_{p=p_0}$ also is. □

5.3.3 Generalization to the Steady Advection-Diffusion Equation

Now we will discuss the following steady equation:

$$A(q)z(q) = f, \qquad f \in L^2(\Omega), \qquad \text{and} \qquad z(q) \in H_0^1(\Omega),$$

where $A(q)$ is defined on $H^2(\Omega)$ as $A(q)z = \nabla \cdot (a\nabla z) + b \cdot \nabla z + cz$ and its range is contained in $L^2(\Omega)$. We use the notation $q = (a, b, c)$ and $\delta q = (\delta a, \delta b, \delta c)$ in the following paragraphs. Note that b has two components, (b_x, b_y), thus q represents four parameters.

Theorem 5.3.13. *The solution operator $z(q)$ is weakly Fréchet differentiable as an operator between the spaces $L^\infty(\Omega)\times L^\infty(\Omega)\times L^\infty(\Omega)\times L^\infty(\Omega)$ and $\mathcal{H}= H^2(\Omega)\times H^2(\Omega)\times H^2(\Omega)\times H^2(\Omega)$ where $\Omega = (0,1)\times(0,1)$. Furthermore $[D_q z(q)]_{q=q_0} h$ is the $H_0^1(\Omega)$ weak solution of:*

$$A(q_0)v_h = -A(h)z(q_0), \qquad v_h \in H_0^1(\Omega).$$

Proof. The proof is equivalent to Theorem 5.3.1. □

Theorem 5.3.14. *Under the assumptions of the previous theorem, the operator $[D_q z(q)]_{q=q_0}(h)$ is Hilbert–Schmidt over $\mathcal{H}$.*

Proof. The proof is equivalent to the proof of Theorem 5.3.6. □

5.3.4 Differentiation of Boundary Conditions

Let $z(g)$ be the solution of the following steady equation with non-homogeneous boundary conditions,

$$\begin{aligned} &\nabla\cdot(a\nabla z(g)) + b\cdot\nabla z(g) + cz(g) = f \\ &z|_{\partial\Omega} = g \end{aligned} \qquad (\mathcal{P}_g)$$

The previous equation has a parameterized boundary condition, indeed we will consider the entire boundary condition itself as a parameter.

Theorem 5.3.15. *If $\Omega = (0,1)\times(0,1)$ and $f \in L^2(\Omega)$, then $z(g)$ is Fréchet Differentiable in $H^{\frac{1}{2}}(\partial\Omega)$. Moreover, for any $h \in H^{\frac{1}{2}}$, $[D_g z(g)]_{g=g_0} h = v_h$ is the solution of:*

$$\nabla\cdot(a\nabla v_h) + b\cdot\nabla v_h + cv_h = 0, \qquad v_h|_{\partial\Omega} = h.$$

Proof. Let z_h and v_h be the solution of:

$$\begin{aligned} &A(q)z_h = f, && f \in L^2(\Omega) \\ &z_h|_{\delta\Omega} = g + h \in h \in H^{\frac{1}{2}}(\Omega), \end{aligned}$$

and

$$\begin{aligned} &A(q)v_h = 0, && f \in L^2(\Omega) \\ &v_h|_{\delta\Omega} = h \in L^2(\Omega), \end{aligned}$$

respectively, then $z_h - z - v_h$ is the solution to

$$A(q)(z_h - z - v_h) = 0 \qquad z_h|_{\delta\Omega} = 0 \in h \in H^{\frac{1}{2}}. \tag{5.25}$$

Since solutions are unique and 0 is a solution, we have that:

$$\lim_{\|h\|\to 0} \frac{\|z_h - z - v_h\|}{\|h\|} = \lim_{\|h\|\to 0} \frac{0}{\|h\|} = 0 \tag{5.26}$$

thus $v_h = D_g z(g)$. □

Theorem 5.3.16. *If $\Omega = (0,1)\times(0,1)$ and $f \in L^2(\Omega)$, then the Fréchet Derivative $[D_g z(g)]_{g=g_0}$ is Hilbert–Schmidt over $H^1(\partial\Omega)$.*

Proof. The proof is identical to the time dependent case below in Theorem 5.4.14. □

5.4 Sensitivity of the Convection-Diffusion Equation

5.4.1 Background

In this section we consider the solution of the abstract convection-diffusion equation, $z(t,q)$, as a function of its distributed parameters and its boundary conditions. The Fréchet Derivative operator is well defined for the set of the parameters that leads to unique solutions of the PDE. As in the previous section when the Fréchet Derivative of the solution operator exists and when applied to suitable parameters variation,, it is the solution of the sensitivity equation. The existence of such a representation is discussed in several papers including Seubert and Wade [44] and Davis and Singler [17]. The convection-diffusion case is analysed in [44] and a special case is considered in [17]. Therefore the main focus of this chapter is restricted to the study of the spectral properties of the operator $[D_q z(t,q)]_{q=q_0}$, more specifically the existence of a Hilbert–Schmidt decomposition. Once we show the existence of such decomposition, we then derive a finite representation of the operator and then prove that its singular values and vectors converge to their infinite dimensional peers.

5.4.2 Fréchet Differentiability

The choice of the admissible set and the norm associate is crucial, these choices will determine if the solution operator $z(t,q)$, is well defined and moreover if it is F-differentiable. In the next paragraphs the parameter space Q, is the normed space: $Q = C^0(\Omega)\times C^0(\Omega)\times C^0(\Omega)\times C^0(\Omega)$ and q is the vector (a, b^x, b^y, c). The admissible set will be given by: $Q^{ad} = \{q \in Q$ such that $A(q)$ is strongly elliptic$\}$. The domain $\Omega \subset \mathbb{R}^2$ is closed, and $\Omega_f = \Omega \times [0,T]$ The spacial solution to the convection-diffusion equation $z(t,q)$, satisfies the equation:

$$(\mathcal{P})_q \begin{cases} \frac{d}{dt}z(t;q) = \nabla\cdot(a\nabla z(t;q)) + b\cdot\nabla z(t;q) + cz(t;q) + f(t) \quad z(t) \in H_0^1(\Omega)\bigcap H^2(\Omega) \\ z(0,\cdot) = z_0. \end{cases} \tag{5.27}$$

We will assume that $f, f' \in L^\infty(0,T;H_0^1(\Omega)) \cap L^\infty(0,T;H^2(\Omega))$ and $z_0 \in H_0^1(\Omega)\bigcap H^2(\Omega)$, this will guarantee that $z \in Z = L^\infty(0,T;H^2(\Omega))$. Defining the operator $A(q)$ on $H^2(\Omega)$ as $A(q)z = \nabla\cdot(a\nabla z) + b\cdot\nabla z + cz$ and rewriting the system $\mathcal{P}_q$ we have

$$(\mathcal{P})_q \begin{cases} \frac{d}{dt}z(t;q) = A(q)z(t;q) + f(t) \quad z(t) \in H_0^1(\Omega)\bigcap H^2(\Omega) \quad ; \\ z(0,x) = z_0 \end{cases} \tag{5.28}$$

Since all the spaces and the operator $A(q)$ satisfy the the hypothesis of the main result from Singler [9], we are in position to state the differentiability of $z(t,q)$.

Theorem 5.4.1 (Fréchet differentiability of z(t;q))**.** *The operator z(t;q), is weakly Fréchet differentiable over Q^{ad} and $v_h(t) := [D_q z(t,q)]_{q=q_0}h$ is the weak solution of the abstract Cauchy problem:*

$$(\mathcal{S})_q \begin{cases} \frac{d}{dt}v_h(t) = A(q_0)v_h(t) + A(h)z(t;q_0) & v_h(t) \in H_0^1(\Omega)\bigcap H^2(\Omega) \quad ; \\ z(0,x) = 0 & t \in [0,T]. \end{cases} \tag{5.29}$$

Proof. See [17]. □

Extension to the Groundwater Flow Equation:

We now extend this result to the groundwater flow equation $\mathcal{C}_q$. Let Ω be a closed bounded subset of $\mathbb{R}^2$ and
$Q^{ad} = \{q = (S, k^x, k^y)$ such that $S, k^x, k^y \in C^1(\Omega)$ and $S(x,y) \geq \alpha_0 > 0$, $k^x(x,y) \geq \alpha_1 > 0$, $k^y(x,y) \geq \alpha_2 > 0\}$ $\forall (x,y) \in \Omega$. Furthermore, let

Z=$L^1((0,T)$, $H_0^1(\Omega)\bigcap H^2(\Omega))$ and $\Omega_T = (0,T) \times \Omega$.

The Groundwater flow problem is given as the solution of $z \in Z$ of:

$$(\mathcal{P})_q \begin{cases} S\frac{d}{dt}z(t;q) = (k^x z_x(t;q))_x + (k^y z_y(t;q))_y + f(t,x) \\ z(0,x) = z_0 \end{cases} \tag{5.30}$$

Theorem 5.4.2 (Sensitivity Groundwater Flow equation)**.** *Let $z \in Z$ be the solution of 5.32. The weak Fréchet derivative of z, $[D_q z(t;q)]$, exists and $v_h = [D_q z(t;q)]h$ is the solution of:*

$$\begin{array}{l} S\frac{d}{dt}v_h(t;q) + (\delta S)\frac{d}{dt}z(t;q) = (k^x (v_h)_x(t;q))_x + (k^y (v_h)_y(t;q))_y + (\delta k^x z_x(t;q))_x + (\delta k^y z_y(t;q))_y; \\ z(0,x) = 0 \end{array} \tag{5.31}$$

Proof. Since $[D_q[z(t;q) - z_0]]_{q=q_0} = [D_q z(t;q)]_{q=q_0}$ one only needs to know the case for the zero boundary condition.
The zero initial condition is given by:

$$(\mathcal{P})_q \begin{cases} S\frac{d}{dt}z(t;q) = (k^x z_x(t;q))_x + (k^y z_y(t;q))_y + f(t); \\ z(0,x) = 0 \end{cases} \tag{5.32}$$

Changing variables to $w(t) = S^{\frac{1}{2}}z(t)$ and defining $f(t,x) = W(t,x) + (k^x z_{0x})_x + (k^y z_{0y})_y$, the previous equation is equivalent to:

$$(\mathcal{P})_q \begin{cases} \frac{d}{dt}w(t;q) = S^{-\frac{1}{2}}(k^x (S^{-\frac{1}{2}}w(t;q))_x)_x + S^{-\frac{1}{2}}(k^y (w(t;q)S^{-\frac{1}{2}})_y)_y + S^{-\frac{1}{2}}f(t) \\ w(0,x) = 0 \end{cases} \tag{5.33}$$

Now we are in condition to use Singler's result that proves that $w(t;q)$ is Fréchet differentiable and moreover $v_h(t)$ is the solution of the sensitivity equation:

$$(\mathcal{P})_q \begin{cases} \frac{d}{dt}w(t;q) = S^{-\frac{1}{2}}(k^x w_x(t;q))_x + S^{-\frac{1}{2}}(k^y w_y(t;q))_y + S^{-\frac{1}{2}}f(t); \\ w(0,x) = 0 \end{cases} \tag{5.34}$$

Let $B(S,k^x,k^y)=S^{-\frac{1}{2}}(k^x(S^{-\frac{1}{2}}z)_x)_x+S^{-\frac{1}{2}}(k^y(S^{-\frac{1}{2}}z)_y)_y$ and $A(q)z=(k^xz_x)_x+(k^yz_y)_y$ then :

$$B'_{q_0}(\delta S,\delta k^x,\delta k^y)z=S^{-\frac{1}{2}}(\delta k^x(S^{-\frac{1}{2}}[z])_x)_x+S^{-\frac{1}{2}}(\delta k^y(S^{-\frac{1}{2}}[z])_y)_y+ \tag{5.35}$$

$$-\frac{1}{2}(\delta S)S^{-\frac{3}{2}}[(k^x(S^{-\frac{1}{2}}[z])_x)_x+(k^y(S^{-\frac{1}{2}}[z])_y)_y]+ \tag{5.36}$$

$$S^{-\frac{1}{2}}(k^x(-\frac{1}{2}(\delta S)S^{-\frac{3}{2}}[z])_x)_x+S^{-\frac{1}{2}}(k^y(-\frac{1}{2}(\delta S)S^{-\frac{3}{2}}[z])_y)_y \tag{5.37}$$

So $[D_qw(t,q)]=v_h$ is then the solution of:

$$(\mathcal{P})_q\begin{cases}\frac{d}{dt}v_h(t;q)=B'(h)w(t,q)+B(S,k^x,k^y)v_h+[F'(q,t)]h\\ v_h(0,x)=0\end{cases} \tag{5.38}$$

Where $F(q_0,t)=S^{-\frac{1}{2}}f(t)$ and

$$[F'(q,t)]h=-\frac{1}{2}(\delta S)S^{-\frac{3}{2}}f(t) \tag{5.39}$$

$$=-\frac{1}{2}(\delta S)S^{-\frac{3}{2}}[S\frac{d}{dt}z(t;q)-(k^xz_x(t;q))_x-(k^yz_y(t;q))_y \tag{5.40}$$

So the forcing term, B'(h)w(t,q)+[F'(q,t)]h, from 5.32 can be decomposed as:

$$B'(h)w(t,q)+[F'(q,t)]h=B'(h)S^{\frac{1}{2}}z-\frac{1}{2}(\delta S)S^{-\frac{3}{2}}[S\frac{d}{dt}z(t;q)-A(q)z] \tag{5.41}$$

$$=S^{-\frac{1}{2}}A(\delta q)z-\frac{1}{2}\delta S^{-\frac{3}{2}}A(q)z-\frac{1}{2}A(q)((\delta S)S^{-1}z)-\frac{1}{2}(\delta S)S^{-\frac{3}{2}}[S\frac{d}{dt}z(t;q)-A(q)z] \tag{5.42}$$

$$=S^{-\frac{1}{2}}A(\delta q)z-\frac{1}{2}(\delta S)S^{-\frac{1}{2}}\frac{d}{dt}z(t;q)-\frac{1}{2}A(q)((\delta S)S^{-1}z) \tag{5.43}$$

Since $w=S^{\frac{1}{2}}z$ we have that $w_h=S^{\frac{1}{2}}v_h+\frac{1}{2}(\delta S)S^{-\frac{1}{2}}z$ so the equation 5.32 is equivalent to :

$$\frac{d}{dt}[S^{\frac{1}{2}}v_h+\frac{1}{2}(\delta S)S^{-\frac{1}{2}}z]=S^{-\frac{1}{2}}A(q)S^{-\frac{1}{2}}[S^{\frac{1}{2}}v_h+ \tag{5.44}$$

$$+\frac{1}{2}(\delta S)S^{-\frac{1}{2}}z]+S^{-\frac{1}{2}}A(\delta q)z-\frac{1}{2}(\delta S)S^{-\frac{1}{2}}\frac{d}{dt}z(t;q)-\frac{1}{2}A(q)((\delta S)S^{-1}z) \tag{5.45}$$

which is equivalent to

$$\frac{d}{dt}[S^{\frac{1}{2}}v_h+(\delta S)S^{-\frac{1}{2}}z]= \tag{5.46}$$

$$=S^{-\frac{1}{2}}A(q)[v_h+\frac{1}{2}(\delta S)S^{-1}z]+S^{-\frac{1}{2}}A(\delta q)z-\frac{1}{2}A(q)((\delta S)S^{-1}z) \tag{5.47}$$

which is equivalent to

$$\frac{d}{dt}[S^{\frac{1}{2}}v_h + (\delta S)S^{-\frac{1}{2}}z] = S^{-\frac{1}{2}}A(q)v_h + S^{-\frac{1}{2}}A(\delta q)z \tag{5.48}$$

$$\frac{d}{dt}[Sv_h + (\delta S)z] = A(q)v_h + A(\delta q)z \tag{5.49}$$

Which ends the proof. □

5.4.3 Hilbert–Schmidt Decomposition

In order to prove that the operator $[D_q z(t,q)]_{q=q_0}$ is Hilbert–Schmidt, it is necessary that both parameter and solution spaces be Hilbert spaces. Therefore one must change the spaces of differentiation, since the Q defined in the previoussection is a Banach space but not a Hilbert space.

Existence Results

Theorem 5.4.3. *If $z(t) \in Z = W^{0,\infty}(0,T;W^{2,\infty}(\Omega))$ then for every $t \in [0,T]$ the operator $[D_q z(t,q)]_{q=q_0}h$ defined on the Hilbert space $\mathcal{H}$=$H^3(\Omega) \times H^2(\Omega) \times H^2(\Omega) \times H^2(\Omega)$ with range on $H_0^1(\Omega) \cap H^2(\Omega)$ is Hilbert–Schmidt.*

Proof. There are orthonormal basis $e_n \in H^3(\Omega)$ and $f_n \in H^2(\Omega)$ such as $\|e_n\|_{H^1(\Omega)}$ and $\|f_n\|_{H^0(\Omega)}$ are l^2 sequences by Adam's [1]. From (5.29) we know that

$$\|D_q z(t,q)]_{q=q_0}h_n\|_{H_0^1} \leq \|A(h_n)z(t;q_0)\|_{L^2(0,T,L^2(\Omega))} \leq [\|e_n\|_{H^1} + 3\|f_n\|_{H^0}]\|z\|_Z \tag{5.50}$$

Therefore by Theorem 5.3.5 the operator is Hilbert–Schmidt. □

Theorem 5.4.4. *For every $t \in [0,T]$ the operator $[D_q z(t,q)]_{q=q_0}h$ defined on the Hilbert space $\mathcal{H}$=$H^5(\Omega) \times H^4(\Omega) \times H^4(\Omega) \times H^4(\Omega)$ with range on $H_0^1(\Omega) \cap H^2(\Omega)$ is Hilbert–Schmidt.*

Proof. By Adams [2], there are orthonormal basis $e_n \in H^5(\Omega)$ and $f_n \in H^4(\Omega)$ such as $\|e_n\|_{H^3(\Omega)}$ and $\|f_n\|_{H^2(\Omega)}$ are l^2 sequences From $(\mathcal{S})_q$ we know that

$$\|D_q z(t,q)]_{q=q_0}h_n\|_{H_0^1} \leq \|A(h)z(t;q_0)\|_{L^2(0,T,L^2(\Omega))} \tag{5.51}$$

$$\leq [\|e_n\|_1 + 3\|f_n\|_0]\|z\|_{L^2(0,T,H^2(\Omega))} \tag{5.52}$$

$$\leq [\|e_n\|_{H^3} + 3\|f_n\|_{H^2}]\|z\|_{L^2(0,T,H^2(\Omega))} \tag{5.53}$$

Therefore by Theorem 5.3.5 the operator is Hilbert–Schmidt. □

Theorem 5.4.5. *For every $t \in [0,T]$ the operator $[D_q z(t,q)]_{q=q_0}h$ defined on the Hilbert space $\mathcal{H}$= $H^2((0,1)\times(0,1)) \times H^2((0,1)\times(0,1)) \times H^2((0,1)\times(0,1)) \times H^2((0,1)\times(0,1))$ with range on $\tilde{\mathcal{H}}$= $H_0^1((0,1)\times(0,1)) \cap H^2((0,1)\times(0,1))$ is Hilbert–Schmidt.*

Proof. The proof of this theorem is based on the fact that if the there is a orthonormal basis $\{\phi_n\}_{n\geq 1}$ of H such that $\|A\phi_n\| \in l^2$ then A is a Hilbert–Schmidt operator in $\mathcal{H}$, Reed and Simon [41]. It is know that $\phi^{n,m}(x,y) = a_{n,m}\sin(\pi nx)\sin(\pi ny)$ is a orthonormal basis over $H_0^1((0,1)\times(0,1))$ where:

$$a_{n,m} = \frac{2}{\sqrt{(1+n^2\pi^2+n^4\pi^4)(1+m^2\pi^2+m^4\pi^4)}}.$$

Then

$$\{(\phi_{n,m},0,0,0),(0,\phi_{n,m},0,0),(0,0,\phi_{n,m},0),(0,0,0,\phi_{n,m}),n,m\in N\}$$

form an orthonormal basis in $\mathcal{H}$. Next we will prove that image of each component is a l^2 sequence. From the last proof we known that:

$$\|D_q z(t,q)]_{q=q_0}(\phi_{n,m},0,0,0)\|_{H_0^1} \leq \|A((\phi_{n,m},0,0,0))z(t;q_0)\|_{L^2(0,T,L^2(\Omega))} \tag{5.54}$$

$$\leq [\|\phi_{n,m}\|_1\|z\|_{L^2(0,T,H^2(\Omega))} \leq (1+(n+m)\pi)a_{n,m}\|z\|_{L^2(0,T,H^2(\Omega))} \tag{5.55}$$

which obviously is a l^2 sequence. Also,

$$\|D_q z(t,q)]_{q=q_0}(0,\phi_{n,m},0,0)\|_{H_0^1} \leq \|A((0,\phi_{n,m},0,0))z(t;q_0)\|_{L^2(0,T,L^2(\Omega))} \tag{5.56}$$

$$\leq \|\phi_{n,m}\|_0\|z\|_{L^2(0,T,H^2(\Omega))} \leq a_{n,m}\|z\|_{L^2(0,T,H^2(\Omega))} \tag{5.57}$$

Which is also a l^2 sequence. With same type of argument one can prove for the cases $(0,0,\phi_{n,m},0)$ and $(0,0,0,\phi_{n,m})$. By adding these four inequalities together, it is easy to verify that the full basis is a l^2 sequence. Finally since $\tilde{H}$ is dense in $H^2((0,1)\times(0,1))$ this basis is a basis in $\mathcal{H}$ and therefore the operator is Hilbert–Schmidt in $\mathcal{H}$. □

5.4.4 Finite Dimensional Representation

Any linear operator defined on finite dimension space H, is uniquely defined by the image of a basis. Therefore, if we project the admissible set Q^{ad} onto the basis of a finite dimensional space, typically a finite element basis, and apply the Fréchet derivative, we have a finite rank representation of the Fréchet derivative operator. In this section, we prove that the finite dimensional operator converges in norm to infinite dimensional operator (defined on Q^{ad}). This results also holds for the singular vectors and singular values.

Operator definition: Let S^N be a finite element space of $H_0^1(\Omega)$, with polynomials of degree $s \geq 2$ such that $S^N \hookrightarrow S^{N+1}\ \forall N \in \mathbb{N}_2$, and let Q^M be a finite dimension space of H and let P^M be the projection of H onto Q^M. Let $D_q^{N,M}(t)$ be the linear operator that is defined in the following way:

Definition 5.4.6 (Approximate Operator)**.** $D_q^{N,M}(t) : H \to H_0^1$

$D_q^{N,M}(t)h = v^N_{P^M h}(t)$, *where* $v^N_{P^M h}(t)$ *is the finite element solution of* $v_{P^M h}(t)$.

Theorem 5.4.7. *Convergence of the Operator Estimation*
For all $t \in [0,T]$ the operator $D_q^{N,M}(t)$ converges uniformly to $[D_q z(t,q)]_{q=q_0}$.

Proof. For any norm the following inequality holds:

$$\|[D_q^{N,M}(t) - [D_q z(t,q)]_{q=q_0}]h\| =$$
$$= \|D_q^{N,M}(t)h - [D_q z(t,q)]_{q=q_0}P^M h + [D_q z(t,q)]_{q=q_0}P^M h - [D_q z(t,q)]_{q=q_0}h\| \quad (5.58)$$
$$\leq \|D_q^{N,M}(t)h - [D_q z(t,q)]_{q=q_0}P^M h\| + \|[D_q z(t,q)]_{q=q_0}P^M h - [D_q z(t,q)]_{q=q_0}h\| \quad (5.59)$$

Obviously the second term converges to zero since $[D_q z(t,q)]_{q=q_0}$ is continuous. The second term converges by the finite element theory:
From theorem 5.5 from [50] we have the following inequality:

$$\|D_q^{N,M}(t)h - [D_q z(t,q)]_{q=q_0}P^M h\| = \|v^N{}_{P^M h}(t) - v_{P^M h}(t)\|_{H^1} \quad (5.60)$$
$$\leq C\Delta^{r-1}[\|v_{P^M h}(t)\|_{H^1} + \|A(P^M h)z_0\|_{H^1} + \|D_t v_{P^M h}\|_{L^2(0,t;H^{r-1})}] \quad (5.61)$$

where Δ is maximum diameter of the triangles, using r=2.

$$\Delta[\|v_{P^M h}(t)\|_{H^1} + \|A(P^M h)z_0\|_{H^1} + \|D_t v_{P^M h}\|_{L^2(0,t;H^1)}] \quad (5.62)$$
$$\leq C\Delta[\|[D_q z(t,q)]_{q=q_0}\|\|P^M h\| + C\|P^M h\|\|z_0\| + \|A(P^M h)z\|]_{H^1(0,T,L^2)} \quad (5.63)$$
$$\leq C\Delta[\|[D_q z(t,q)]_{q=q_0}\|\|P^M h\| + C\|P^M h\|\|z_0\| + \|P^M h\|\|z\|]_{H^1(0,T,H^2)} \quad (5.64)$$
$$\leq C(t)\Delta\|P^M h\| \leq C(t)\Delta\|h\| \quad (5.65)$$

Therefore $D^{N.M}(t)$ converges in norm to $[D_q z(t,q)]_{q=q_0}$, since both parts converge to zero so does the left-hand side.

Now the focus will be to prove that the singular values of the finite dimension operator converges to the singular values of the infinitesimal operator, before we will need to prove some preliminary results.
In following paragraphs we will use the following notation $D(t) := [D_q z(t,q)]_{q=q_0}$. Firstly we will prove the the convergence of the singular decomposition of $D_1{}^M(t) = D(t)P^M$ for the one of $D(t)$, the second step is to proof the same convergence of decomposition of the $D_2{}^N(t) = v_h^N(t)$, finally by an triangular type of argument we will prove the decomposition convergence of $D^{N.M}(t)$. □

Corollary 5.4.8. *Let ${D_1}^M(t) = D(t)P^M$ then ${D_1}^M(t) \to D(t)$ and ${D_1}^{M^*}{D_1}^M \to D(t)^*D(t)$ uniformly.*

Proposition 5.4.9. *If Q^M are nested in such a way that $Q^M \subset Q^{M+1}$ and $P^M \to I$ uniformly then: The eigenvalues of $R_1^M(t) = {D_1}^M(t)^*{D_1}^M(t) = P^M R_1(t)P^M$ converge to the eigenvalues of $R_1(t) = D(t)^*D(t)$.*

Proof. The proof is a result from [19]-page 160 □

Corollary 5.4.10. *Let ${D_2}^N(t) = v_h^N(t)$ then ${D_2}^N(t) \to D(t)$ and ${D(t)_2}^N(t)^*D(t){_2}^N(t) \to D(t)^*D(t)$ uniformly.*

Proposition 5.4.11. *If H^N are nested in such a way that $H^N \subset H^{N+1}$ and $P^N \to I$ uniformly then: The eigenvalues of $R_2^N(t) = {D_2}^N(t){D_2}^N(t)^*$ converge to the eigenvalues of $R_2(t) = D(t)D(t)^*$.*

Proof. Let $\delta > 0$ and the following u's be in H then

$$\lambda_1^n \leq \max_{\|u\|=1}\langle R_2^n(t)u, u\rangle \leq \max_{\|u\|=1}\langle [R_2^n(t) - R_2(t)]u, u\rangle + \langle R_2(t)u, u\rangle \tag{5.66}$$

$$\leq \|R_2^n(t) - R_2(t)\| + \max_{\|u\|=1}\langle R_2(t)u, u\rangle \leq \delta + \lambda_1 \tag{5.67}$$

for N sufficiently lagre.
Then $\lambda_1^n \leq \lambda_1$. Using the same argument we can prove that $\lambda_1 \leq \lambda_1^n$ □

Theorem 5.4.12. *In the hypothesis from Propositions 4.3 and 4.5 for every $t > 0$ the singular values of $D^{M,N}(t)$ converge to the singular values of $D(t) := [D_q z(t,q)]_{q=q_0}$.*

Proof. The proof follows immediately from Propositions 5.4.9 and 5.4.11:

$$\lim_N \lim_M \sigma_1^{N,M}(t) = \lim_N \lim_M \sup_{\|u\|=1} \langle D^{N,M}(t)^* D^{N,M}(t)u, u\rangle \tag{5.68}$$

$$= \lim_N \sup_{\|u\|=1} \langle D^N(t)^* D(t)^N u, u\rangle = \lim_N \sup_{\|v\|=1} \langle D^N(t)D^N(t)^* v, v\rangle \tag{5.69}$$

$$= \sup_{\|v\|=1} \langle D(t)D(t)^* v, v\rangle = \sigma_1 \tag{5.70}$$

to prove the convergence of the greater singular values, comes from repeating the proof on the complement of the expansion of *span* $\{u_1(t), ...u_k(t)\}$. □

Theorem 5.4.13. *Under the same Hypothesis as the previous theorem , plus that $D(t)$ has no double singular vectors and values, then the singular vectors of $D^{N,M}(t)$ converge to their infinitive dimension pears.*

Proof. By definition of of $[D^N(t)]D^N(t)$ we are find the eigenvalues in H^N of $D^*(t)D(t)$ since the H^∞ is dense on H and their are nested, this means that we are solving the problem in a sequence of problem that close and close to H.

$$\sigma_N(t) = \min_{\|u\| \in H^n} \langle D(t)v, D(t)v \rangle \tag{5.71}$$

Then for $\varepsilon > 0$ exists p such $N > p$

$$|\min_{\|v\| \in H^n} \langle D(t)v, D(t)v \rangle - \min_{\|v\| \in H} \langle D(t)v, D(t)v \rangle| < \|\sigma_N(t) - \sigma(t)\| < \varepsilon \tag{5.72}$$

Since there is an unique solution v^* to

$$v^* := arg \min_{\|v\| \in H} \langle D(t)v, D(t)v \rangle \tag{5.73}$$

and the the spaces H^n are nested this means that v^{*N} does converge to v^*

□

5.4.5 Differentiation of Boundary Conditions

In this section we will Fréchet differentiate the solution operator, $z(t;g)$, of the convection equation, with respect to a boundary condition, $u(t)g(x)$. Once proven the differentiability, we present some spectral decomposition results. The convection diffusion equation written as follows:

$$(\mathcal{P})_q \begin{cases} \frac{d}{dt} z(t;g) = \nabla \cdot (a \nabla z(t;g)) + b \cdot \nabla z(t;g) + cz(t;g) + f(t) & z(t) \in H^1(\Omega) \\ z(t;g)|_{\partial\Omega} = z_0|_{\partial\Omega} + u(t)g; & \\ z(0;g) = z_0 & t \in [0,T]. \end{cases} \tag{5.74}$$

$Q = C^0(\Omega) \times C^0(\Omega) \times C^0(\Omega) \times C^0(\Omega)$ and q is the vector (a, b^x, b^y, c). The admissible set will be given by: $Q^{ad} = \{q \in Q$ such as $A(q)$ is strongly elliptic$\}$ $f \in L^2(0,T;L^2(\Omega))$ and $z_0 \in H^1(\Omega)$ then we will have the guarantee that $z \in Z = L^\infty(0,T;H^1(\Omega))$ and $u(0) = 0$.

Theorem 5.4.14. *For $q \in Q^{ad}$ and if $\Omega \subset \mathbb{R}^2$ is a bounded Lipschitz domain and $u, u' \in L^2([0,T])$, with u(0)=0 then $z(t;g)$ is weakly Fréchet differentiable with respect to g on $H^{\frac{3}{2}}(\partial\Omega)$. Moreover, $v_h := [D_g z(g)]_{g=g_0} h$ is the solution of the sensitivity equation:*

$$(\mathcal{S})_g \begin{cases} \frac{d}{dt} v_h(t) = \nabla \cdot (a \nabla v_h(t)) + b \cdot \nabla v_h(t) + c v_h(t) = 0 \\ v_h(t)|_{\partial\Omega} = u(t)h \\ v_h(0) = 0 \end{cases} \tag{5.75}$$

Proof. Let $D(t,h) = z(t, g+h) - z(t,g) - v_h(t)$then D(t,h) is the weak solution of:

$$(\mathcal{R})_h \begin{cases} \frac{d}{dt} r(t) = \nabla \cdot (a \nabla r(t)) + b \cdot \nabla r(t) + c r(t) = 0 \\ r(t)|_{\partial\Omega} = 0 \\ r(0) = 0 \end{cases} \tag{5.76}$$

Since for all h the only solution of $(\mathcal{R})_h$ is the zero solution, we have $D(t,h)=0$; therefore $\forall t>0$

$$\frac{\|D(t,h)\|_{H^1(\Omega)}}{\|h\|_{H^{\frac{3}{2}}(\partial\Omega)}} \to 0 \tag{5.77}$$

□

The following theorem will provide a sufficient condition to the Hilbert–Schmidt decomposition of the previous operator.

Theorem 5.4.15. *Under the same assumptions as the previous theorem $[D_g z(g)]_{g=g_0}$ is Hilbert–Schmidt over $H^2(\partial\Omega)$ and range in $H^1(\Omega)$, where $\Omega=(0,1)\times(0,1)$.*

Proof. As be before we will use the fact that if the there is a orthonormal basis on $H^2(\partial\Omega)$ such as the image of it is an l^2 basis then the operator is Hilbert–Schmidt.
The basis,
$\phi_1^n = a_n(sin(n\pi x),0)$, $\phi_2^n = a_n(sin(n\pi x),1)$, $\phi_3^n = a_n(0,sin(n\pi y))$, $\phi_4^n = a_n(1,sin(n\pi y))$
where

$$a_n = \frac{2}{\sqrt{1+n^2\pi^2+n^4\pi^4}}$$

obviously orthonormal over $H^2(\partial\Omega)$. Then for every t and i=1..4

$$\|D_q z(t,g)]_{q=q_0}(\phi_i^n)\|_{H^1(\Omega)} \leq 2T^2C\|z\|_{1,\infty}\|\phi_i^n\|_{H^{\frac{1}{2}}(\partial\Omega))} \tag{5.78}$$

wich is a l^2 sequence. □

Computing the singular values is very expensive therefore it is mandatory to develop an efficient way to do it. In the following chapter is dedicated to the numerical aspect of the computation of the singular values and vectors. Since by definition the singular values and vectors of $[D_q z(q)]_{q=q_0}$ are the eigenvalues and eigenvectors of $[D_q z(q)]^*_{q=q_0}[D_q z(q)]_{q=q_0}$ respectively, then one can use their favorite eigenvalue estimation technic.

5.5 Application to Parameter Estimation and Second Derivative

Even is highly costly one can use the Fréchet derivative for parameter estimations purposes, using the chain rule one can deduce:

$$[D_q J(q,z(q))]_{q=q_0} = D_q\left[\frac{1}{2}\|z(q)-z_{data}\|_X^2+\frac{\beta}{2}\|q-q_{data}\|_Q^2\right]_{q=q_0}(h) =$$

$$\langle z-z_{data}, v_h\rangle_H + \langle q-q_{data}, h\rangle_Q \tag{5.79}$$

Since v_h is the solution of the sensitive equation, then to a quadratic approximation the most indicated would be use BFGS or any other type of approximation to the second derivate. In some cases the second derivative can be evaluate explicitly as in example 5.4.

5.6 Second Derivative

Theorem 5.6.1. *If $z(q)$ is the solution of:*

$$\mathcal{P}_q \begin{cases} -\nabla \cdot (q \nabla z(q)) = f, \quad f \in L^2(\Omega); \\ z(q) \in H_0^1(\Omega). \end{cases} \tag{5.80}$$

where $\Omega \subset \mathbb{R}^n$ is bounded with a C^0 boundary and $q \in Q^{ad} = \{q:\ q \in C^1(\Omega)$ and $\exists \alpha_0 > 0$ such as $q(x) \geq \alpha_0, \forall x \in \Omega\}$. The norm on Q^{ad} is defined by $\|h\| = max_{x\in\Omega}\{|h(x)|\} + max_{x\in\Omega}\{|h'(x)|\}$.

Then the Fréchet derivative of $z(q)$ with respect to q applied to h, $[D_q z(q)]_{q=q_0}(h) = v_h$, is the weak solution of, the well-posed elliptic equation:

$$\mathcal{S}_q \begin{cases} -\nabla \cdot (q \nabla v_h) = \nabla \cdot (h \nabla z(q)); \\ z \in H_0^1(\Omega). \end{cases} \tag{5.81}$$

And the second derivative is the solution of:

$$\mathcal{SS}_q \begin{cases} -\nabla \cdot (q \nabla r_{\theta,h}) - \nabla \cdot (\theta \nabla v_h) = \nabla \cdot (h \nabla v_\theta); \\ r \in H_0^1(\Omega). \end{cases} \tag{5.82}$$

Proof. This is identical to the proof of Theorem 5.3.1 □

This can be the used to compute the inverse problem trough a Newton Method. Note that the finite representation is a 3 dimensional matrix, where the entry (i,j,k) is defined by $[D^{N,N,M}(\phi_i, \phi_j)]_k$.

Chapter 6

Numerical Implementation

6.1 Approximation of Singular Values and Vectors

Computing the singular values is computationally intensive. Therefore it is mandatory to develop efficient algorithms to approximate them. This section is devoted to numerical issues associated with the computation of the singular values and vectors.

6.1.1 Using the Operator to Evaluate the Singular Values

This method is the most intuitive and most expensive one. The idea is to construct the operator $D^{M,N}(t)$ from Definition 5.4.6. Let H^N be an N dimensional finite element subspace for approximating a model solution. Also, let Q^M be an M dimensional finite element subspace approximating the parameter q. We

$for\ i = 1 : N$
$for\ j = 1 : M$

- $\left(D_q^{N,M}(t)\right)_{i,j} = \left([D_q z(q,t)]|_{q=q_0} (\phi_i) \right)_j$ the coefficient of $[D_q z(q,t)]|_{q=q_0} (\phi_i)$ on the $j - element - basis$ on the space Q^M

$end\ for$
$end\ for$

$[u(t), s(t), v(t)] := svd(D^{M,N}(t))$

There are several issues with this approach, firstly it is very expensive since it requires to compute the solution of $(M+1)$ PDEs. Furthermore it requires to compute at each time step Δt the computation of the singular value decomposition of a $N \times M$ matrix, thus in a realistic point of view this is not feasible in the sense that a system might have millions of free parameters. The computation time decreases significantly by applying the previous

algorithm in parallel. Finally by computing the full SVD one might be doing unnecessary work since the original operator is Hilbert-Schmidt which implies that only the first singular directions are relevant where the last ones are just essentially noise. This motivated then to find different ways to compute those directions. One is instead of computing it in all time steps one can just take snapshots. The next sections are then dedicated to that purpose.

6.1.2 The Power Method for the Steady Case

Since by definition the singular values and vectors of $[D_q z(q)]_{q=q_0}$ are the eigenvalues and eigenvectors of $[D_q z(q)]^*_{q=q_0}[D_q z(q)]_{q=q_0}$ respectively, then one can use their favorite eigenvalue estimation technic. The first one is the power method discussed thoroughly at [10]. It's description is as follows:

1. Start with an h^0 in Q^M
2. For $n \geq 1$ repeat until convergence
   - Evaluate $\tilde{h}^{n-1} = [D_q z(q)]_{q=q_0} h^{n-1}$ by solving the sensitivity equation.
   - Evaluate $h^n = [D_q z(q)]^*_{q=q_0} \tilde{h}^{n-1}$ by solving the adjoint sensitivity equation.
   - Set $h^n = \frac{h^n}{\|h^n\|}$ and $\lambda^n = \|\tilde{h}^{n-1}\|^2$.

Example for the interior steady case:

Let $A(q)z = -\nabla \cdot (q\nabla z)$ be a strongly elliptic operator, $\Omega \subset \mathbb{R}^n$ be closed and z(q) the $H_0^1(\Omega)$ weak solution of:

$$\mathcal{P}_q \begin{cases} A(q)z(q) = f, \quad f \in L^2(\Omega); \\ z(q) \in H_0^1(\Omega). \end{cases} \tag{6.1}$$

Then $[D_q z(q)]_{q=q_0} h$ is the $H_0^1(\Omega)$ weak solution of:

$$\mathcal{S}_q \begin{cases} A(q_0)v_h = -A(h)z(q_0) \\ v_h \in H_0^1(\Omega). \end{cases} \tag{6.2}$$

And $v^*{}_\theta = [D_q z(q)]^*_{q=q_0}(\theta)$ is the $H_0^1(\Omega)$ weak solution of:

$$v^*{}_\theta = \nabla(A^{-1}(q_0)\theta)) \cdot \nabla z(q_0) \tag{6.3}$$

Example for the Boundary steady case:

Let $A(q)z = -\nabla \cdot (q\nabla z)$ be a strongly elliptic operator, $\Omega \subset \mathbb{R}^n$ be closed and $z(q)$ the $H_0^1(\Omega)$ weak solution of:

$$(\mathcal{P})_g \begin{cases} A(q)z(g) = f, \quad f \in L^2(\Omega); \\ \\ z(g)|_{\partial\Omega} = g \end{cases} \tag{6.4}$$

Then $[D_q z(g)]_{g=g_0} h$ is the $H^1(\Omega)$ weak solution of:

$$(\mathcal{S})_g \begin{cases} A(q)v_h = 0; \\ \\ v_h|_{\partial\Omega} = h \end{cases} \tag{6.5}$$

And ${v^*}_\theta = [D_q z(q)]^*_{q=q_0}(\theta)$ is the realization at the boundary of:

$${v^*}_\theta = q\nabla(A^{-1}(q)\theta)) \cdot \vec{n} \tag{6.6}$$

This method is extendable to the convection-ciffusion case, which is discussed in the Appendix $B.2$.

6.1.3 Evolving the Singular Values and Vectors in Time

The next paragraph will describe an effective way to calculate the singular values of the Hilbert-Schmidt decomposition of the Fréchet Derivative operator $s_h(t) := [D_q z(t,q)]_{q=q_0} h$. We know that $s_h(t)$ is the solution of:

$$(\mathcal{S})_q \begin{cases} \frac{d}{dt}s_h(t) = A(q_0)s_h(t) + A(h)z(t;q) & s_h(t) \in H_0^1(\Omega) \bigcap H^2(\Omega) \quad ; \\ s_h(0) = 0 & t \in [0,T]. \end{cases} \tag{6.7}$$

Where

$$(\mathcal{P})_q \begin{cases} \frac{d}{dt}z(t) = A(q)z + f(t) & z(t) \in H_0^1(\Omega) \bigcap H^2(\Omega) \quad ; \\ z(0) = z_0 & t \in [0,t_f]. \end{cases} \tag{6.8}$$

Since $[D_q z(t,q)]_{q=q_0}$ is Hilbert–Schmidt it can be decomposed in the following fashion:

$$s_h(t) = \sum_{k=1}^{\infty} \sigma_k \langle v_k(t), h \rangle_{H_1} u_k(t) \tag{6.9}$$

Then

$$s_{v_k(t)}(t) = \sigma_k(t) u_k(t) \tag{6.10}$$

So the pair $v_k(t), \tilde{u}_k(t) = u_k(t)\sigma_k(t)$ is the solution of the PDE:

$$(\mathcal{SVD})_q \begin{cases} \frac{d}{dt}\tilde{u}_k(t) = A(q_0)\tilde{u}_k(t) + A(v_k(t))z(t) \quad ; \\ \tilde{u}_k(0) = 0 \qquad\qquad\qquad\qquad\qquad\qquad t \in [0,T]. \end{cases} \tag{6.11}$$

Once we know a pair $v_k(t)$ and $\tilde{u}_k(t)$ at time t_1 one can solve $(SVD)_q$. The system S_q on the finite dimensional space can be written as

$$(\mathcal{S})_q^{\Delta x} \begin{cases} M^{\Delta x} s_h(t) = S^{\Delta x} s_h(t) + R^{\Delta x}(t)h \quad ; \\ s_h(0) = 0 \qquad\qquad\qquad\qquad\qquad t \in [t_1,T]. \end{cases} \tag{6.12}$$

using finite differences in time:

$$(\mathcal{S}^{\cdot\S})_q \begin{cases} (M^{\Delta x} - \frac{\Delta t}{2} S^{\Delta x}) s_h(t+\Delta t) = (M^{\Delta x} + \frac{\Delta t}{2} S^{\Delta x}) s_h(t) + \frac{R^{\Delta x}(t+\Delta t) + R^{\Delta x}(t)}{2} h \quad ; \\ s_h(0) = 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad t \in [0,T]. \end{cases} \tag{6.13}$$

for $t = 0$ we have:

$$(M^{\Delta x} - \frac{\Delta t}{2} S^{\Delta x}) s_h(\Delta t) = \frac{R^{\Delta x}(\Delta t) + R^{\Delta x}(0)}{2} h \tag{6.14}$$

$$s_h(\Delta t) = [(M^{\Delta x} - \frac{\Delta t}{2} S^{\Delta x})^{-1} \frac{R^{\Delta x}(\Delta t) + R^{\Delta x}(0)}{2}] h. \tag{6.15}$$

Defining

$$E = [(M^{\Delta x} - \frac{\Delta t}{2} S^{\Delta x})^{-1} \frac{R^{\Delta x}(\Delta t) + R^{\Delta x}(0)}{2}] \tag{6.16}$$

the problem is now reduced to a singular value problem or eigenvalue problem for $E'E$. There are several different ways to compute those eigenvalues, and a vast literature on it [42] Y. Saad, [31] R. B. Lehoucq and [47] G. L. Sleijpen among others. Once we have $v_k(\Delta t)$ and $\tilde{u}_k(\Delta t)$ we use them to evolve $(\mathcal{S}VD)_q$ and find their evolution in time. An intelligent construction of R reduces significantly the computation time, since $M^{\Delta x}$ and $S^{\Delta x}$ are inherited from the from forward problem. The method is then summarized by

- Use the Matrices M and R, inherited from the forward problem (6.8) (defined in (6.12)) to compute E defined in (6.16). Find the singular triplets, $[u(t), \sigma(t), v(t)]$ of E then evolve them trough time:

$$(\mathcal{SVD})_q \begin{cases} \frac{d}{dt}\tilde{u}_k(t) = A(q_0)\tilde{u}_k(t) + A(v_k(t))z(t) & ; \\ \tilde{u}_k(0) = 0 & t \in [0, T]. \end{cases} \tag{6.17}$$

- Compute $u_i(t) = [D_q z(q,t)]|_{q=q_0} (v^i(t))$
- Compute $\sigma_i(t) = \|u^i(t)\|_{l^2}$

6.2 Numerical results

6.2.1 Motivation

Recalling Example 5.1.

$$\mathcal{P}_q \begin{cases} (1 + sin(\pi x))\frac{d}{dt} z(t,x;q) = (e^{-4x} z(t,x;q)_x)_x + 1 + x + t & z(t,.;q) \in H^1([0,1]) \\ z(0,x) = x & x \in [0,1] \\ z(t,0) = 0 & t \in [0,1] \\ z(t,1) = e^{-t} & t \in [0,1] \end{cases} \tag{6.18}$$

The most sensitive directions from the system are shown in Figure 6.1.

This explains how such a small perturbation in the second half of the interval $[0,1]$ leads to such a great change in the solution while a small perturbation in the first half of the

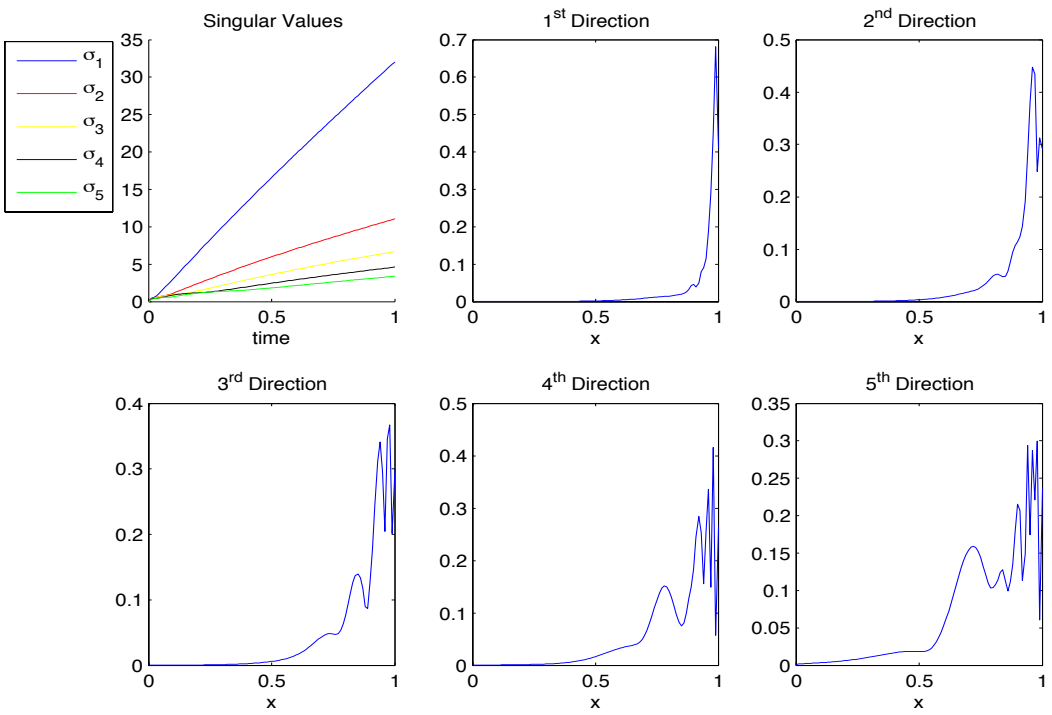

Figure 6.1: The most sensitive perturbations averaged in time

interval doesn't affect the solution significantly. This example shows how relevant the study of the most sensitive directions is. In this example two perturbations with the same L^2 norm $\|\delta_1(x)\|_{L^2([0,1])} = \|\delta_2(x)\|_{L^2([0,1])} = 2.4 \times 10^{-3}$ have two completely different responses. The relative change in z when is perturbed by $\delta_1(x)$ is $\frac{\|z(t;q(x))-z(t;q(x)+\delta_1(x))\|_{L^2([0,1])}}{\|z(t;q(x))\|_{L^2([0,1])}} = 9.9688 \times 10^{-4}$ where the perturbation $\delta_2(x)$ leads to a much greater relative change on z, $\frac{\|z(t;q(x))-z(t;q(x)+\delta_2(x))\|_{L^2([0,1])}}{\|z(t;q(x))\|_{L^2([0,1])}} = 4.37 \times 10^{-2}$. This can be easily explained by the fact that

$$z(t; q_0 + h) = z(t; q_0) + D_q[z(t; q)]_{q=q_0} h + o(\|h\|) \tag{6.19}$$

Since $D_q[z(q]_{q=q_0}$ is Hilbert-Schmidt the previous equation can be rewritten as:

$$z(t; q_0 + h) = z(t; q_0) + \sum_{i=1}^{\infty} \langle v_i(t), h\rangle \sigma_i(t) u_i(t) + o(\|h\|) \tag{6.20}$$

So those perturbations that have following property $|\langle v_i(t), h\rangle| s_i >> 0$ for small i and $t > 0$ will have a major impact on the solution. That is the main difference between $\delta_1(x)$ and $\delta_2(x)$, since $|\langle v_1(t), \delta_1(x)\rangle| s_1(t) = 0.0420$ and $|\langle v_1(t), \delta_2(x)\rangle| s_1(t) = 0.1005$ which explains why the second perturbation has a greater impact. One can argue that the differentiation is in a smooth space, since all parameters belong to $H^2(\Omega)$, and in GWF applications the parameters are $PWCT$. This can be solved by using the following elements $tanh(x)$

6.3 Convection-Diffusion Equation

In this section, we will exemplify the theoretical results. We have some empirical proof of the singular values convergence. We used an uniform triangular mesh, over the square $[0, 1] \times [0, 1]$ over the time interval [0,1]. The finite elements used were the cubic Hermite A. They were chosen because to compute the sensitivity equation it requires the partial derivatives of z. In our test case the parameters are the following: $q_1(x, y) = e^{-10xy} + e^{-10(1-x)y} + e^{-10x(1-y)} + e^{-10(1-x)(1-y)} +$; $q_{2,1}(x, y) = 1 + x + y$; $q^y_{2,2}(x, y) = 1.1 + \cos(3\pi(x - y))$; $cq_3x, y) = 1.1 + \cos(2\pi x)\sin(2\pi y)$; the forcing term is given by $f(x, y, t) = ty + x$ the initial condition is $z_0 = 0$, the time step that as used was $\Delta_t = 10^{-5}$ the time solver used was the mid point and backward Euler, as in (6.13). The next graph shows the different first directions for different space dimensions and time steps. In all examples M=N, meaning that the solution space as the same dimension as the parameter space. The following picture shows the average in time of the first four singular directions:

By close analysis of these graphs it becomes elucidative that there are 2 very sensitive regions, one centered on $(1, 0.5)$ and another on $(0.4, 0.5)$. This information can be interpreted as the finite element mesh should be finer than in another regions. Analyzing Figure 6.2 around t=0.017 there is a change in direction that is because the most sensitive direction changes form $(1, 0.5)$ to $(0.4, 0.5)$. This can be seen in the next picture:

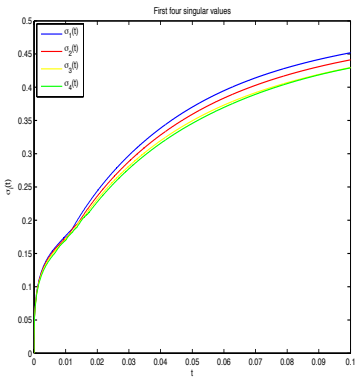

Figure 6.2: Singular values

The following table shows the comparison between the first singular value (SV) when the dimension is 3600 and the first SV for smaller dimensions. It also compares the first SV to consecutive dimensions.

The following table shows the comparison between the first singular value SV when the dimension is 3600 and the first sv for smaller dimensions, It also compares the first SV to consecutive dimensions.

Difference sequence	norm	Cauchy sequence	norm
$\|\sigma^{n_6} - \sigma^{n_1}\|_{L^2(0,0.1)}$	3×10^{-2}	$\|\sigma^{n_2} - \sigma^{n_1}\|_{L^2(0,0.1)}$	1.21×10^{-2}
$\|\sigma^{n_6} - \sigma^{n_2}\|_{L^2(0,0.1)}$	1.3×10^{-2}	$\|\sigma^{n_3} - \sigma^{n_2}\|_{L^2(0,0.1)}$	6×10^{-3}
$\|\sigma^{n_6} - \sigma^{n_3}\|_{L^2(0,0.1)}$	7.2×10^{-3}	$\|\sigma^{n_4} - \sigma^{n_3}\|_{L^2(0,0.1)}$	3.4×10^{-3}
$\|\sigma^{n_6} - \sigma^{n_4}\|_{L^2(0,0.1)}$	3.8×10^{-3}	$\|\sigma^{n_3} - \sigma^{n_2}\|_{L^2(0,0.1)}$	2.2×10^{-3}
$\|\sigma^{n_6} - \sigma^{n_5}\|_{L^2(0,0.1)}$	1.6×10^{-3}	$\|\sigma^{n_6} - \sigma^{n_5}\|_{L^2(0,0.1)}$	1.6×10^{-3}

$n_1 = 100, n_2 = 400, n_3 = 900, n_4 = 1600, n_5 = 2500, n_6 = 3600$. As one can see the difference converges to zero as $n \to \infty$ Even though we use implicit scheme in time, if $\Delta t < \Delta x \Delta y$ there are some problems with stabilization specially at the initial time steps. Which gives us the empirical proof of Theorem 5.4.12.
The next example includes differentiation on the boundary.

Boundary Numerical Example

Let $z(g)$ be the solution on $[0,1] \times [0,1]$ of:

$$(\mathcal{P})_g \begin{cases} ((1 + x + 13y)z(g)_x)_x + ((1 + e^{-x+5y}))z(g)_y)_y = x + y + 3 \\ z(g)|_{\partial\Omega} = g \end{cases}$$

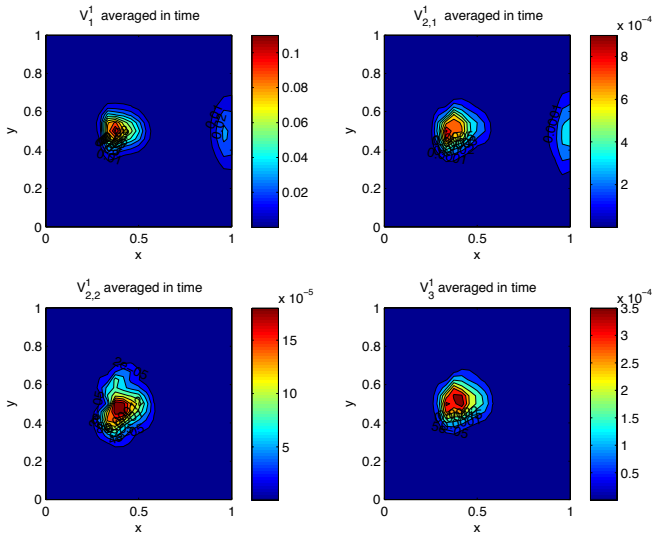

Figure 6.3: First singular vector

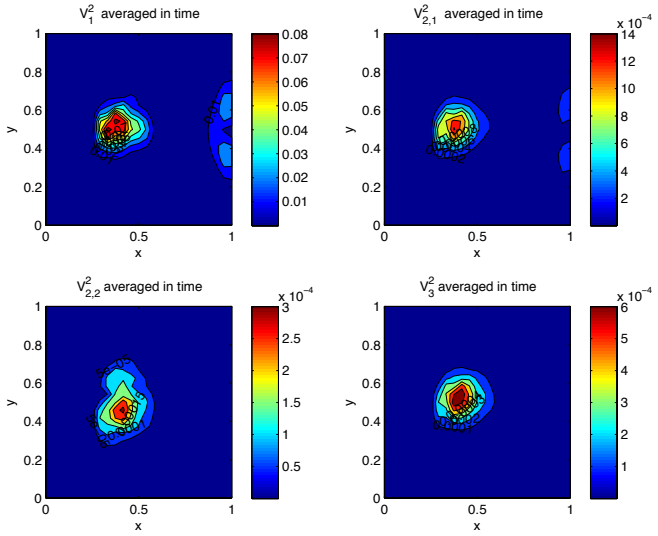

Figure 6.4: Second singular vector

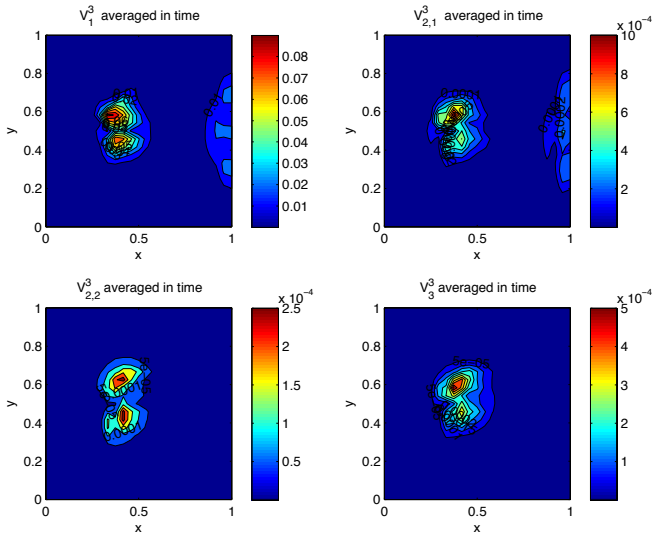

Figure 6.5: Third singular vector

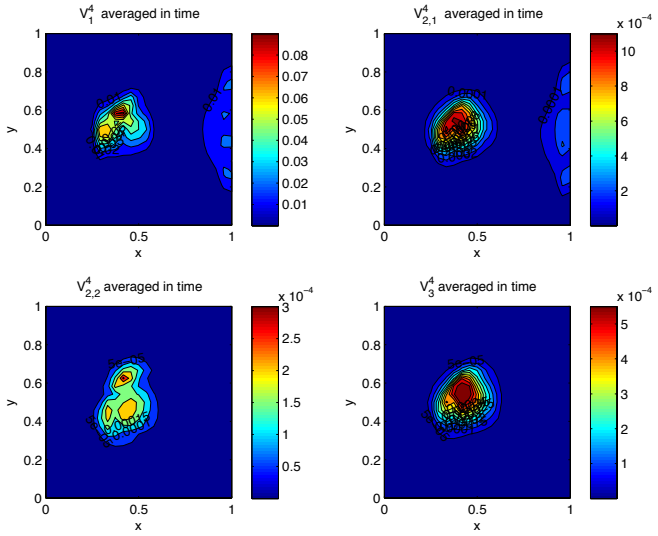

Figure 6.6: Fourth singular vector

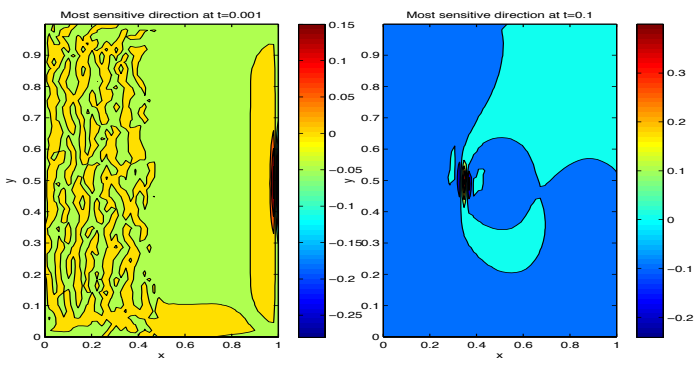

Figure 6.7: First singular value in time step t=0.001(left) and t=0.1 (right)

Then the Fréchet Derivative of $z(g)$, with respect to the boundary condition g, when applied to a direction h, $v_h = [D_g z(g)]_{g=g_0} h$, is the solution of the sensitivity equation:

$$(\mathcal{S})_h \begin{cases} ((1+x+13y)v_{hx})_x + ((1+e^{-x+5y}))v_{hy})_y = 0 \\ v_h|_{\partial\Omega} = h \end{cases}$$

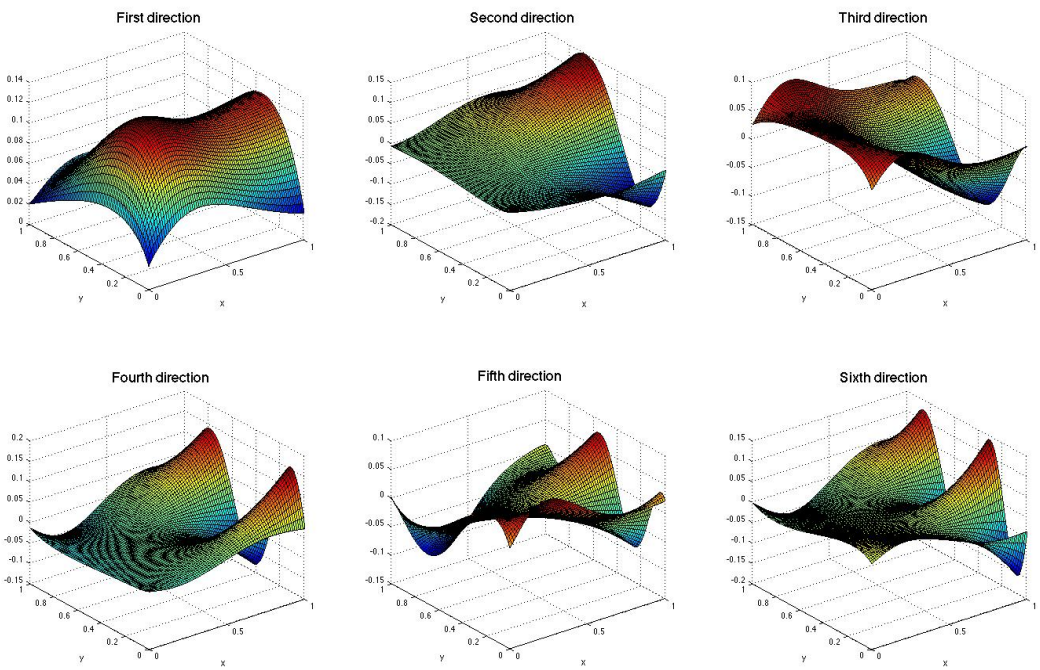

Figure 6.8: Response to the first most sensitive six directions

6.4 Applications

The applications of doing such analysis vary from, reduction on the dimension of the gradient, reduction on the parameter space dimension, mesh refinement and finally uncertainty quantification.

6.4.1 Parameter Space Reduction

The set if vectors $\{v_q^i\}_{i\leq\infty}$ form a basis of the parameter space Q^N, of dimension N, since their importance decay in the l^2 norm this means only the first ones are important to development of the system in the neighborhood of q_0. This means that any perturbation δq to q_0 that is represented by a linear combination in the latest singular values will have a small or irrelevant impact over the solution.

$$\delta q = \sum_{i\leq r}^{N} \alpha_i v_i \tag{6.21}$$

Then for $\delta q \in V_\epsilon^Q(q_0)$ for small ϵ such the linearization is so valid the following holds:

$$z(q_0+\delta q) = z(q_0) + \sum_{i\leq r}^{N} \alpha_i \sigma_i u_i + o(\|h\|_Q) \approx z(q_0) \tag{6.22}$$

This means that one can use the subspace of Q, $V_r := span\,\{v_i \text{ for } 1\leq i\leq r\}$ as reduced representation of the admissible set.

6.4.2 Uncertainty Quantification

In collaboration with Hans-Werner van Wyk and Jeff Borggaard, the results of Chapter 5 were developed further to an uncertainty quantification framework, this can be seen in more detail at [11]. One of the differences between the previous approach and the following is that one have multiple sets of data $\{(z_k, f_k)\}_{k=1}^{k=N}$ instead of an unique pair of (z, f) this will lead to N inverse problems and consequently N $q's$. Then the solution will be presented as a distribution and with its moments $\mathbb{E}(q(x,\cdot))$, $\mathbb{V}(q(x,\cdot))$ and so forth. This means that for instance the deterministic equation becomes the stochastic partial differential equation:

$$\mathcal{P}_q \begin{cases} -\nabla\cdot(q(\mathbf{x},\omega)\nabla z(q(\mathbf{x},\omega)) = f(\mathbf{x},\omega), \quad f\in L^2(\Omega); \\ z(q,\omega)\in H_0^1(\Omega). \end{cases} \tag{6.23}$$

The standard technique is the so called Monte Carlo method, which is explained in Appendix C. The sensitivity analysis and most sensitive directions come in place by reducing the Fréchet operator into and a collection of small rank derivatives that can be used in the inverse problem. It is known from previous chapters that r sufficiently large and ϵ sufficiently small $\forall h\in V_\epsilon^Q(q_0)$ one can write $\forall h\in V_\epsilon(q_0)$ we have the following estimation on the

deterministic level:

$$z(q_0 + h) \approx z(q_0) + \sum_{i=1}^{r} \langle v_i, h \rangle_Q \sigma_i u_i. \tag{6.24}$$

So q_0 can be written as:

$$q_0 + h := \sum_{i=1}^{r} \langle v_i, q_0 \rangle_Q v_i + \sum_{i=1}^{r} \frac{1}{\sigma_i} \langle z(q_0 + h) - z(q), u_i \rangle_H v_i. \tag{6.25}$$

This can be used then for parameter estimation where there is uncertainty in the parameters, in the following fashion:

$$q(\cdot, \omega) := \sum_{i=1}^{r} \langle v_i, q(\cdot, \omega_0) \rangle_Q v_i + \sum_{i=1}^{r} \frac{1}{\sigma_i} \langle z(q(\cdot, \omega_0)) - z(q(\cdot, \omega_0)), u_i \rangle_H v_i. \tag{6.26}$$

Which mean that all paths ω that lead to $z(\cdot, \omega)$ in a neighborhood of $z(\cdot, \omega_0)$ can be predicted, by (6.25) as the following algorithm shows:

Algorithm 1 Estimate the sample paths of the uncertain parameter q based on a random sample of measurements $\hat{z}$ of the model output z.

Input: tol, $\{\hat{z}(x, \omega_i)\}_{i=1}^{n_{\text{mc}}}$.
Let $k = 0$, $\mathcal{I} = \emptyset$.
Choose $i_k \in \{1, \ldots, n_{\text{mc}}\} \backslash \mathcal{I}$ and let $\mathcal{I} = \{\mathcal{I}, i_k\}$.
Compute the estimate $\hat{q}(\cdot, \omega_{i_k})$ of $q(\cdot, \omega_{i_k})$ from $\hat{z}(\cdot, \omega_{i_k})$.
Compute the operator $D_q[z(\hat{q}(\cdot, \omega_{i_k}))]$ and its singular value decomposition.
Use (6.26) to obtain estimates $\{\tilde{q}(\cdot, \omega_i)\}$ of $\{q(\cdot, \omega_i)\}$ for all $i \in \{1, \ldots, n_{\text{mc}}\} \backslash \mathcal{I}$.
while There are paths $\hat{z}(\cdot, \omega_i)$ so that

$$\| z(\tilde{q}(\cdot, \omega_i)) - \hat{z}(\cdot, \omega_{i_l}) \| \geq \text{tol} \tag{6.27}$$

for all i_l, $l = 0, ..., k$. **do**
  if (6.27) doesn't hold for $i \in \{1, \ldots, n_{\text{mc}}\} \backslash \mathcal{I}$ **then** let $\mathcal{I} = \{\mathcal{I}, i\}$.
  else
    Let $i_{k+1} = i$, $\mathcal{I} = \{\mathcal{I}, i_{k+1}\}$. $k \leftarrow k + 1$.
    Repeat steps ...
  end if
end while

Example 6.4.1 (Parameter Identification). *Here we apply Algorithm 1 to identify the diffusion coefficient $q_1(x,\omega)$ from stochastic measurements $\hat{z}$ of the output, using the method discussed above. We used a set of 65 sample paths $\{\hat{z}(\cdot,y_i)\}_{i=1}^{n_{\text{sc}}}$ corresponding to the quadrature points $y_i \in \Gamma$ of a sparse grid stochastic collocation approach, based on a Clenshaw-Curtis scheme (see [3, 35, 36, 51]). For the estimates $\tilde{q}$, we use a truncation level of $r = 99$ and $r = 20$ respectively. To approximate all sample paths $q(\cdot,y_i)$ so that the corresponding model output differs from the data to within an L^2-error tolerance of* $\text{tol} = 0.001$ *requires 9 linear models for both truncation levels $r = 99$ and $r = 20$, the linearization centers of which are depicted in Figure 6.9. $f(x) = \pi^2 \sin(\pi x)(1+x) - \pi\cos(\pi x)$,*
Random diffusion coefficients, q_1, q_2, and q_3, given by

- $q_1(x,\omega) := 1 + x + \frac{1}{2}\left(Y_1(\omega) - \frac{1}{2}\right)\cos(2\pi x) + \frac{1}{2}\left(Y_2(\omega) - \frac{1}{2}\right)\cos(3\pi x)$,
- $q_2(x,\omega) := q_1(x,\omega) + Y_3(\omega)e^{(\sqrt{1000}(x-\frac{1}{2}))^2}$, *and*
- $q_3(x,\omega) := q_1(x,\omega) + Y_3(\omega)e^{(\sqrt{1000}(x-\frac{1}{10}))^2}$.

Where the random variables $(Y_1, Y_2) \sim \text{unif}([0,1]^2)$ and $Y_3 = 1 - \widetilde{Y}_3$, with $\widetilde{Y}_3 \sim \text{beta}(2,5)$

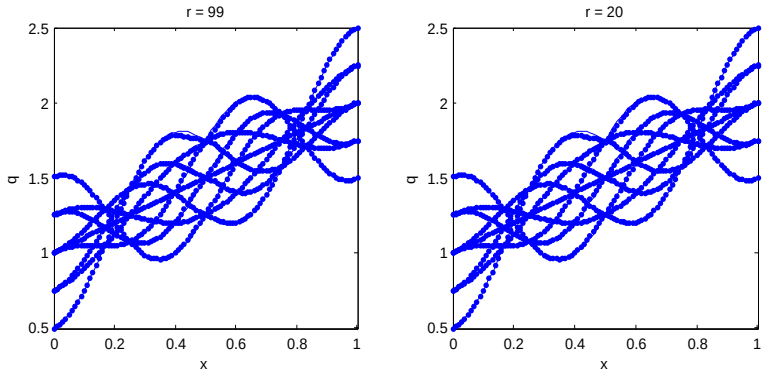

Figure 6.9: Estimates $\hat{q}(\cdot,y_{i_k})$ of the linearization centers $q(\cdot,y_{i_k})$, $k = 1,...,9$ (dotted lines) together with their true values (solid lines) for both truncation levels.

Although a lower truncation level r results in a less accurate reconstruction $\tilde{q}$ of q (see Figure 6.10), this doesn't seem to affect the validity of the linear model, attested to by the fact that 9 linear models with similar linearization centers (see Figure 6.9) are sufficient to explain the input-output map for both $r = 99$ and $r = 20$ to within the required relative error tolerance (Figure 6.11).

The model reduction was successfully implanted given that the errors are form the same order as the full rank operator. Another important fact is that the family of linear models successfully model the solution oprator. This methodology is being developed for the time dependent case.

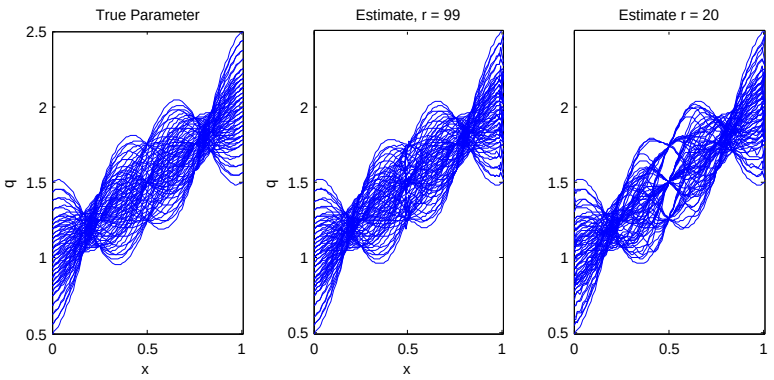

Figure 6.10: Sample paths of the true parameter q as well as its estimate $\tilde{q}$.

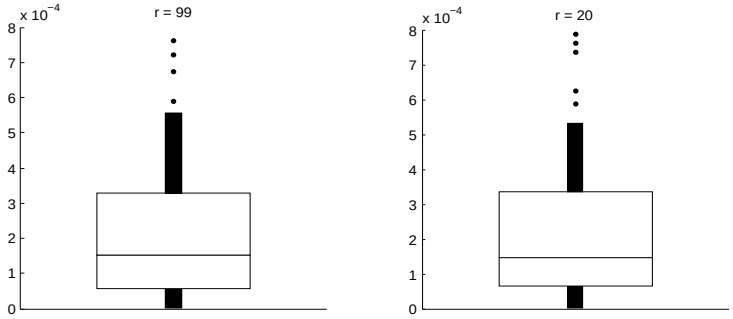

Figure 6.11: Boxplot of the relative L^2-error in the model output for the parameter estimate based on truncated Hilbert-Schmidt decompositions with $r = 99$ and $r = 20$ expansion terms respectively.

6.4.3 Paramater ranking

The Figure 6.8 is very elucidative how one can rank the parameters from the equation, since the first direction is a vector $v_1(t) = [v_1^{q_1}(t), v_1^{q_{2,1}}(t), v_1^{q_{2,2}}(t), v_1^{q_3}(t)]$ and the $X = L^1((0,1); L^2(0,1))$ of $\|v_1^{q_1}(t)\|_X \gg \|v_1^{q^{2,1}}(t)\|_X$, $\|v_1^{q_1}(t)\|_X \gg \|v_1^{q^{2,2}}(t)\|_X$ and $\|v_1^{q_1}(t)\|_X \gg \|v_1^{q^3}(t)\|_X$ this means that on the parameter dimension basis, the last three components have a much smaller impact on the system. The ordering in a decreasing order of importance we have $v_1^{q_1}(t), v_1^{q^{2,2}}(t), v_1^{q^{2,1}}(t)$ and $v_1^{q^3}(t)$ which means that for error purposes one should make sure that the measurements regarding q_1 need to be accurate. This is discussed with more detail in B.1.

Chapter 7

Parameter Estimations Results

7.1 Data

In this chapter will present the results of the parameter estimation techniques developed on Chapter 4, thus the objective is given subsidence, water heads and pumping cycles data, to estimate the values and spatial zonation of the aquifer's transmissivity and Interbed's skeletal storage. In the following synthetic example we use the forward problem from MODFLOW as data. The initial value for the interbed and aquifer is 800 m^3. We consider that there is no flow coming out of the boundaries (zero Neumann boundary condition). There is interchange of water between the interbed and the aquifer. The initial value of the preconditioned head is 750 m^3 throughout the domain. The pumping cycles are divided essentially into two types a pumping season and a non pumping season, each of them with a period of 182.5 days, over 15 years. The aquifer's storativity is assumed to be known and it is constant throughout the domain $S(x, y) = 0.002$. The vertical hydraulic conductivity of the interbed is known as well and is constant with values $K_v((x, y), z) = 6 \times 10{-}5\ m/d$. This is modeled by the following function:

$$W(t,x) := \begin{cases} \tilde{W}(x) & \text{odd stress period} \\ 0(m^3/d) & \text{even stress period} \end{cases} \tag{7.1}$$

Where x is in thousand of meters (km) and

$$\begin{cases} \tilde{W}(4,4) = -3000(m^3/d) \\ \tilde{W}(11,3) = -3000(m^3/d) \\ \tilde{W}(12,7) = -3000(m^3/d) \\ \tilde{W}(14,8) = -6000(m^3/d) \\ \tilde{W}(24,13) = -1000(m^3/d) \\ 0(m^3/d) \qquad otherwise \end{cases} \tag{7.2}$$

The variables of interest are the transmissivity of the aquifer and the specific storage of the interbed. The data is not collected continuously in time therefore one needs to interpolate

it in the time domain. The interbed thickness is given and it can be seen in the following figure.

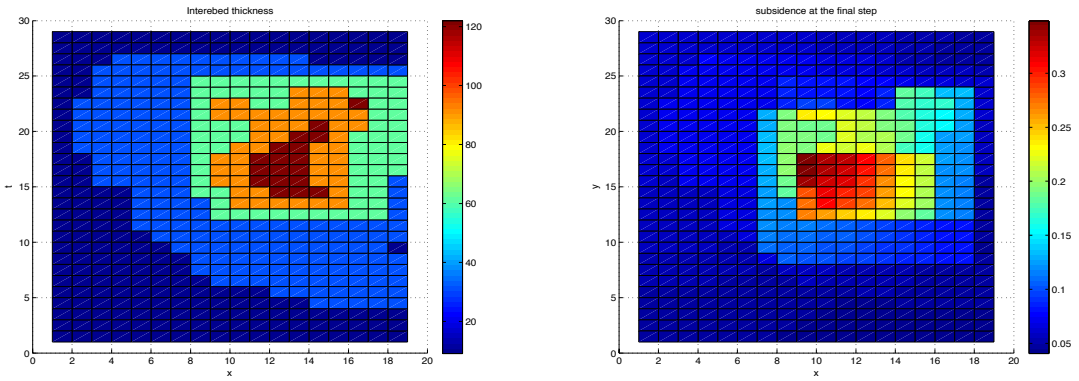

Figure 7.1: Left: Interbed thickness. Right: Subsidence over the last time step

The distribution of the subsidence can be used to be as initial zonation for the specific storage.

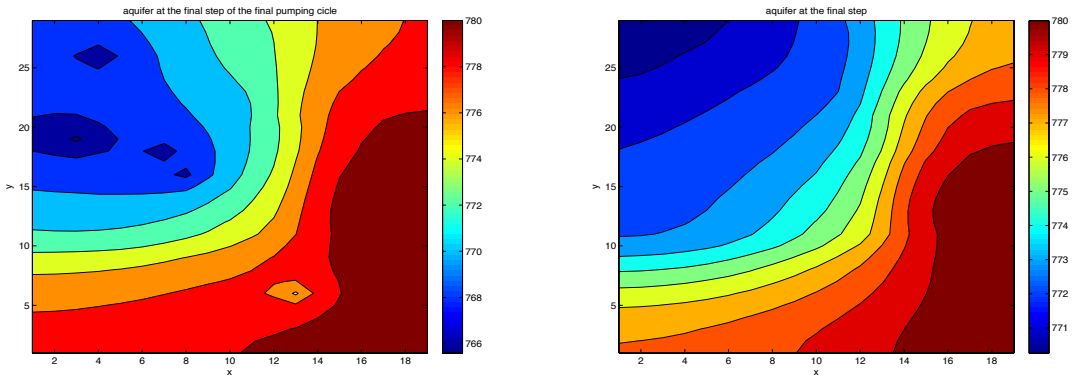

Figure 7.2: Left: last time step at a pumping cycle horizontal scale in km and vertical units in m, Right: Last time step on the last non pumping cycle. Horizontal scale in km and vertical in m^3.

As one can see the wells have a major impact in the water flow. They contribute to a better identifiably of the transmissivity by doing two parameter estimations one when there is pumping and other when there is no pumping, this technic is called optimal pumping test design, this is discussed by [53], where advantages and disadvantages of such an approach are discussed.

7.2 Agglutination Algorithm

Using Algorithm 4.3.3 with an initial zonation where all of the cells are a zone, and $T(x, y) := 800\ m^2/d$ across the whole domain with the penalty term:

$$\frac{1}{2}\left\{10^{-6}\|T - 800\|_2^2 + 10^{-6}\|S_{ske}\|_2^2 + 10^{-9}\|S_{skv}\|_2^2 + 10^{-9}\int_\Omega \sqrt{\|\nabla T\|_2^2 + 10^{-3}}d\mathbf{x}\right\} \quad (7.3)$$

The choice of the a penalty term for S_{ske} greater than the one on S_{skv} because $S_{skv} \gg S_{ske}$. The results for the zonation Algorithm 4.3.3 is as follows:

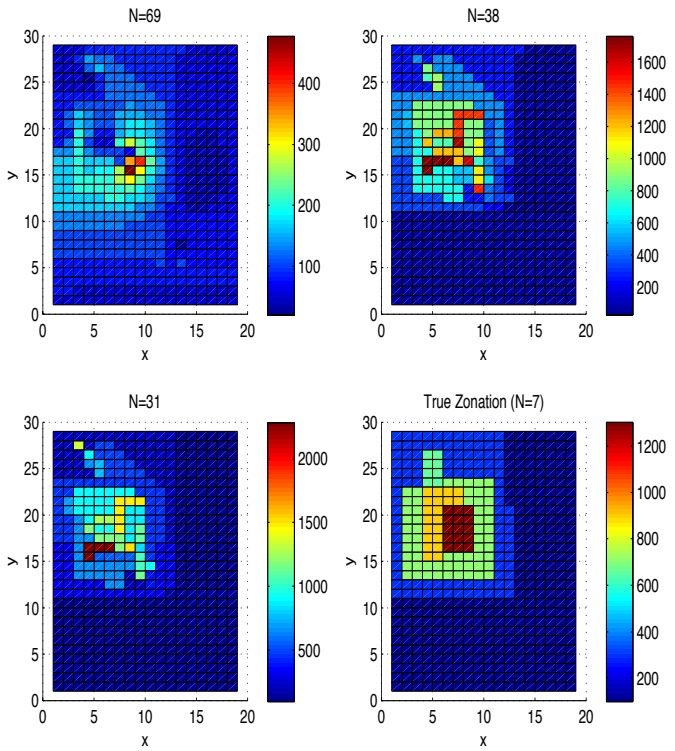

Figure 7.3: From left to right top to bottom the zonations of T, Horizontal scale in km and volume in m^3

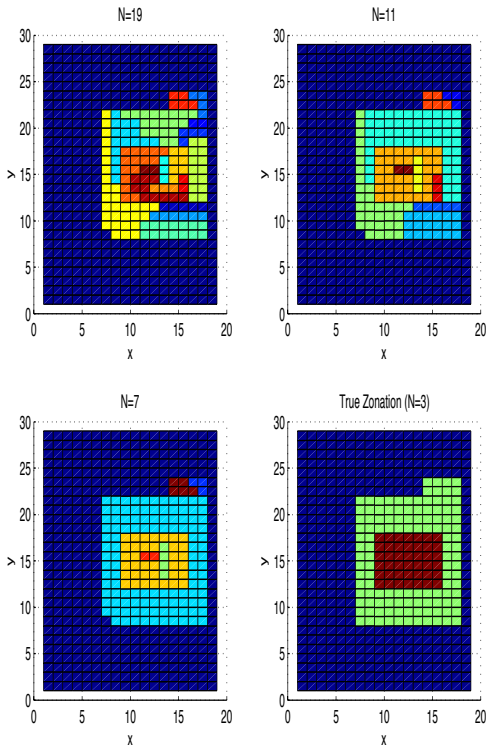

Figure 7.4: From left to right top to bottom the zonations of S_{sk}

Number of Zones (K)	Relative Error
$N = 69$	0.8011
$N = 38$	0.3672
$N = 31$	0.4365

Number of Zones (S_{sk})	Relative Error (S_{ske})	Relative Error (S_{skv})
$N = 19$	0.7526	0.5897
$N = 11$	199.6965	0.5942
$N = 7$	0.5787	0.2250

This was not a sufficient accurate estimation, therefore we used its information to start an multilevel Algorithm 4.3.3. Since the number of zones on S are stable enough we froze those and did a multi level on T.

7.3 Multi-Level

The last iteration of the agglutination one can see two very distinct areas, one in blue very homogeneous and another very unstructured, that will be a good start for the multi level

method. The initial guess was $T(x, y) = 500\ m^2/d$, and the values of S_{ske} and S_{skv} those inherited from the the last iteration of the adjoint method. The sequence of meshes are as follows:

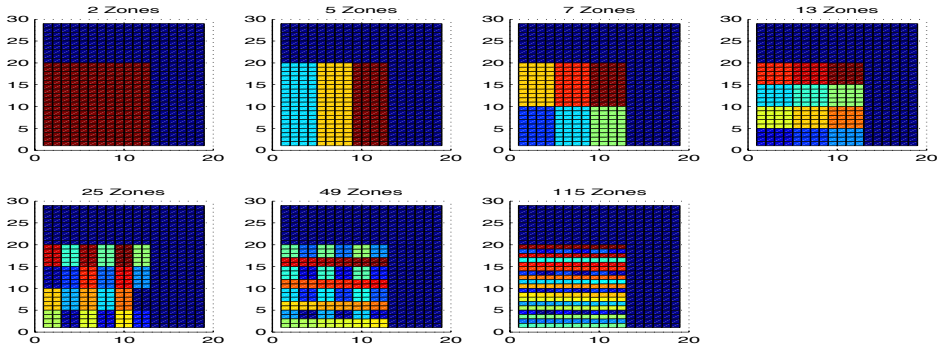

Figure 7.5: Values of T for different zones

The sequence of approximations are as follows:

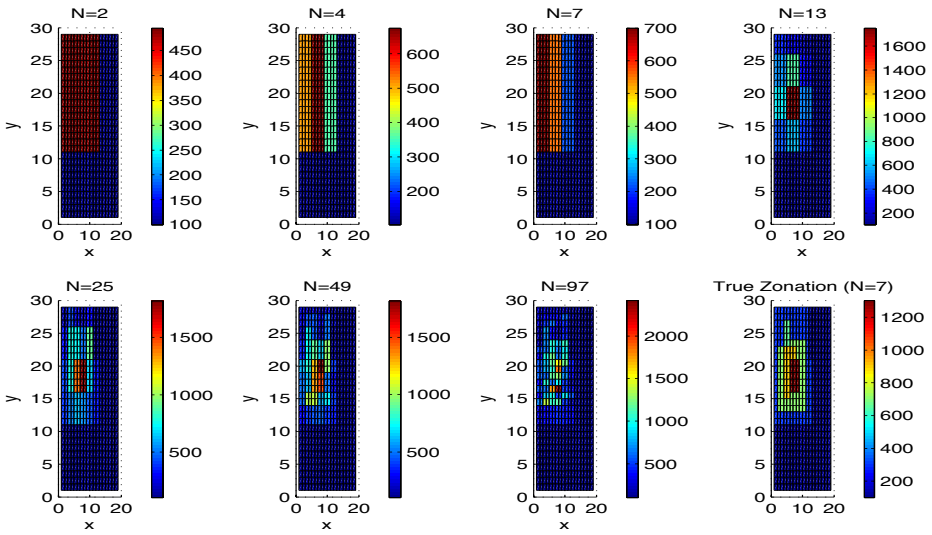

Figure 7.6: Values of T for different zones

Since this is synthetic data, we have access to the inverse problem solution, the errors are as follows.

7.3.1 Error Analysis

Number of Zones	Relative Error	Cost functional (UCODE)
$N = 2$	0.5087	3.91×10^7
$N = 4$	0.4610	1.86×10^7
$N = 7$	0.5162	3.01×10^7
$N = 13$	0.4079	8.65×10^6
$N = 25$	0.3550	5.26×10^6
$N = 49$	0.2878	3.76×10^6
$N = 97$	0.3885	3.48×10^6

Looking carefully to the table above the value of the cost functional is always decreasing and the same happens to the relative error except for two iterations. This means that the optimization problem might need to be adjusted in order to improve identifiably. The other important fact is that $N = 49$ has a well defined structure, which gives the indication that one should use the agglutination algorithm from that point on.

7.4 Inverting from our forward solver

The previous section was the inversion of MODFLOW's forward data. Using our forward model from (2.2.3) as data, with the same imputes as the previous sections we obtain better results, specially for local convertion. In this experiment we perturb the real $T's$ (shown below) and used Algorithm 4.3.3.

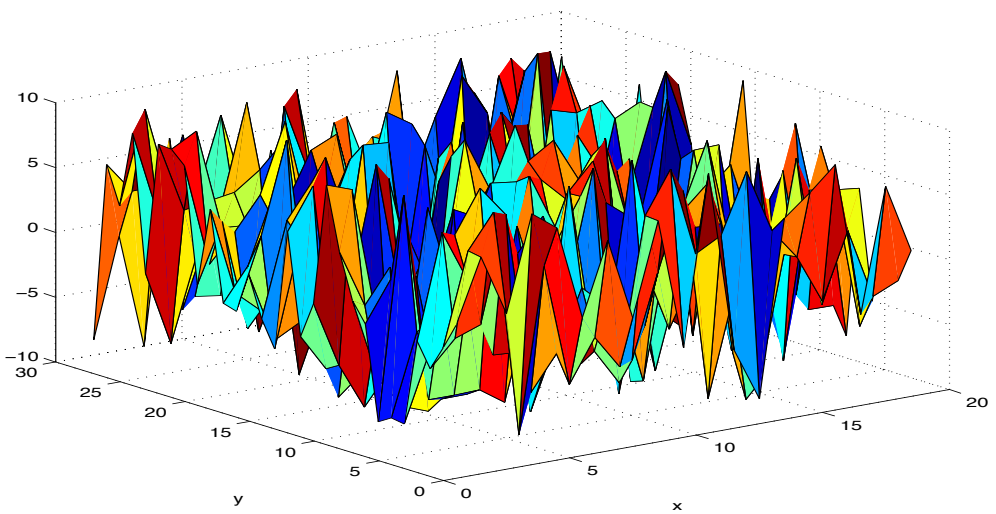

Figure 7.7: Real $T's$ perturbation, axis in km

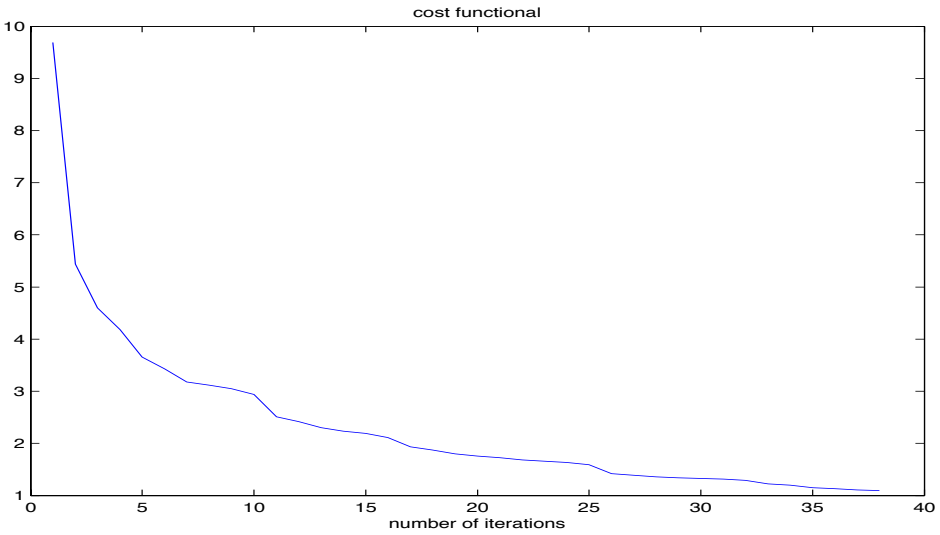

Figure 7.8: Cost functional's value

As one can see, the cost functional decreases faster then in UCODE and their values are considerably small when compared to the ones on Table 7.3.1. Specially the error shown below.

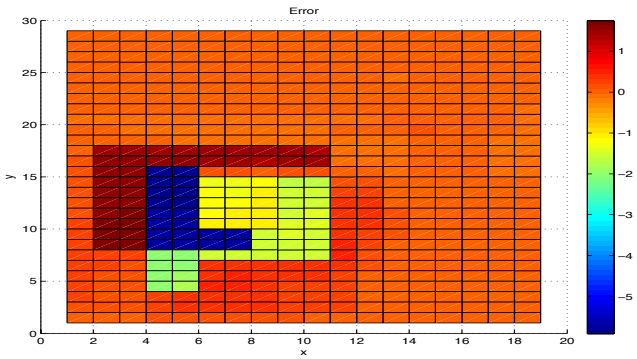

Figure 7.9: Cost functional's value

Which is much better than the UCODE inversion. This ends this section on parameter estimation.

Chapter 8

Conclusion

8.1 Groundwater Parameter Estimation

The main goal of this section was the to identify the parameters, transmissivity (aquifer) and skeletal storage, elastic and inelastic (interbed) by using two distinct strategies, an agglutination algorithm, Algorithm 4.3.3 and a multilevel algorithm, Algorithm 4.3.4, even though they both have the same objective they start from opposite principles. The first initiates with the maximum number of zones (fine mesh), whereas the multilevel begins with a reduced amount of zones (coarse mesh). Both have two distinct results and applicabilities, the multilevel algorithm captures small zones very well, unlike the agglutination algorithm that identifies the bigger zones. Therefore the combination of them is very valuable, since one can start guess the big zone with the agglutination and then just focus on the smaller ones with the multilevel. In a penalty term analysis, the bounded variation term introduced on (4.5) improved the rate of convergence specially when allied with the agglutination algorithm. Another conclusion is that from the subsidence data one can recover a good initial guess for the zonation of the skeleton storage, S_{sk}, over the interbed. This is done by applying the zonation algorithm directly to the last time steps of the subsidence data. This leads to a partition that has the real zonation as subpartition. In a more theoretical framework, we prove that the problem is parameter estimation convergent, (3.4.5), in a finite dimensional admissible set, meaning that the sequence of the solution of the finite dimensional optimization problems converges to the real parameter that is being identified. We also proved the existence of the Lagrangian and the fact that it the solution of the adjoint equation, that was used to compute the cost functional gradient for the optimization. Finally our experiments shown that by analyzing in non-pumping cycles one can estimate the quotient between T and S, that is theoretically justified by (3.4.2).

8.2 Sensitivity Analysis

In this parallel parameter estimation study, we developed results on the Fréchet differentiability of the solution of a PDE with respect to their parameters, such as transmissivity, specific storage and further more to the boundary conditions, diffusion coefficient, convection. Additionally to the results on the existence of such derivatives, we developed results on the spectral decomposition of the derivative operator, and its approximation. More specifically the Fréchet derivative operator is Hilbert-Schmidt, (5.4.4). This allowed us to reduced the parameter space, find the most sensitive directions that made possible to find a reduced form of the operator, compute an explicitly inverse of the linearization of the parameter estimation operator (steady case), which combine with a sampling based framework, enabled the clustering of parameters of interest into a small set of basis. The use of cubic Hermit polynomials on solving the forward problem made possible to improve the rate of convergence of the finite element solution of the sensitive equation from linear to quadratic. By differentiating on the boundary one can determine where the flux will coming from outside of the domain will impact the system. For the computation of such directions and decomposition we introduced two algorithms, a power method and an evolution in time, which saves computational time and memory, and other by evolving the the triplet $[u(t), s(t), v(t)]$ through time.

Chapter 9

Future and Current Work

I would like to proceed by extending these sensitivity analysis results to a wider class of model equations. The singular value decomposition (Hilbert-Schmidt decomposition) has been applied to sensitivity analysis from a statistical viewpoint [49]. My approach differs in that it is derived from a partial differential equation that models the physical behavior of a problem. Additionally, since the decomposition is performed over Hilbert spaces, this allows me to efficiently rank the importance of the parameters, describe their effects, and identify the best- and worse-case scenarios. In the near-term, I intend to seek other applications of the sensitivity analysis for poro-elastic media. In addition to groundwater modeling, with different material laws, there are direct applications to modeling biological tissues, soil contamination, and the distribution of oil fields. However, this research is not restricted to poro-elastic media. There are natural extensions to radar imaging, calibration of machinery, and tomography. I am interested in modeling and analysis of new problems arising in physical or biological systems. Motivated by applications, I have a strong desire to continue collaborating with interdisciplinary teams since this is a source of interesting and important mathematical problems. My theoretical research efforts would be to build a bridge between the statistical and deterministic sensitivity analysis. Specifically, I would like to establish a correlation between the statistical and deterministic singular value decomposition of the parameter space. Another objective is to develop results in the projection of the Fréchet derivative operator over a reduced-order basis (such as those generated by the proper orthogonal decomposition also known as the Karhunen-Loeve expansion). This would be an important result in the combination of model reduction and optimization algorithms. I am also interested in do adaptive sampling. Using the most sensitive directions one can use information of where the data must be the most accurate and by that reducing significantly the number of measurements. Finally, as current work I am incorporating the algorithms develop in Chapter 4 into UCODE in oder to be available to other users.

Appendix A

Hermite Cubic Finite Element Methods

A.1 Implementation and Approximation Results

As one can read in more detail in, e.g. [12], Hermite cubic finite elements are piecewise polynomial functions defined over a triangle such that directional derivatives and function values match along triangle edges. For each triangular elements, the degrees of freedom are

- function values at the 3 vertex nodes
- directional derivatives at each vertex (2 values at each vertex)
- 1 interior node.

The ten basis functions that satisfy the Lagrange interpolation conditions on the reference triangle (with (r, s) coordinates and vertices $(0, 0)$, $(0, 1)$, and $(1, 1)$) are

$$\begin{aligned}
\phi_1(r,s) &:= 1 - 3r^2 - 3s^2 - 13rs + 13r^2s + 13rs^2 + 2r^3 + 2s^3 \\
\phi_2(r,s) &:= 3s^2 - 7rs + 7r^2s + 7rs^2 - 2s^3 \\
\phi_3(r,s) &:= 3r^2 - 7rs + 7r^2s + 7rs^2 - 2r^3 \\
\phi_4(r,s) &:= r - 2r^2 - 3rs + 3r^2s + 2rs^2 + r^2 \\
\phi_5(r,s) &:= -r^2 + 2rs - 2r^2s - 2rs^2 + r^3 \\
\phi_6(r,s) &:= -rs + r^2s + 2rs^2 \\
\phi_7(r,s) &:= s - 2s^2 - 3rs + 2r^2s + 3rs^2 + s^3 \\
\phi_8(r,s) &:= -rs + 2r^2s + rs^2 \\
\phi_9(r,s) &:= -s^2 + 2rs - 2r^2s - 2rs^2 + s^3 \\
\phi_{10}(r,s) &:= 27rs - 27r^2s - 27rs^2.
\end{aligned}$$

Figure A.1 shows these basis functions on the reference triangle.

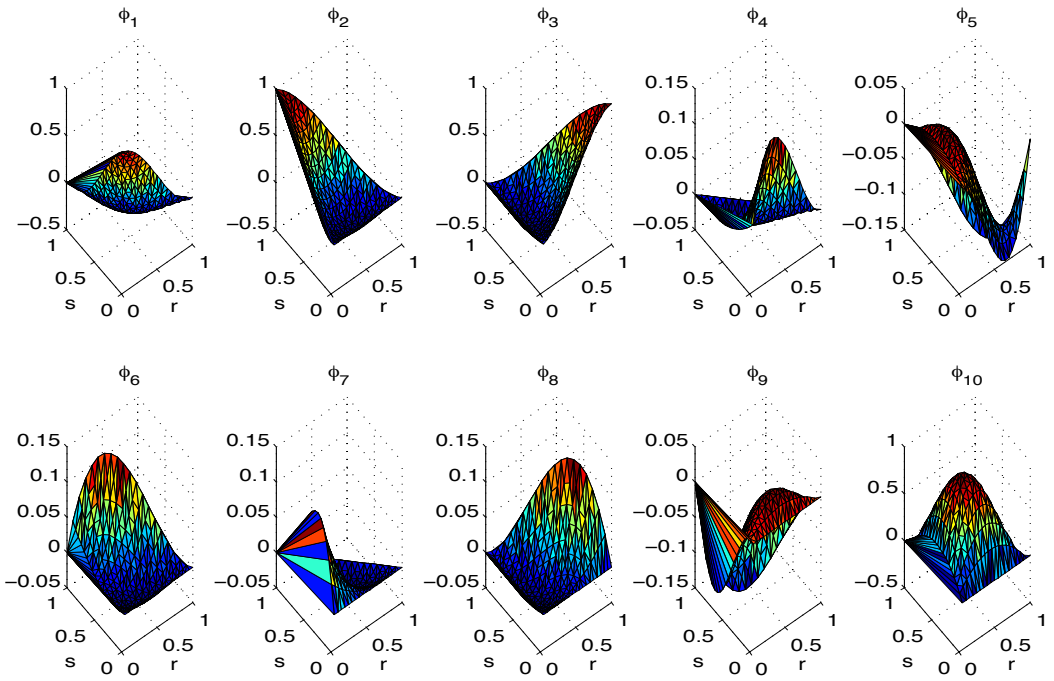

Figure A.1: Hermite cubic basis functions

In order to implement the finite element method with these basis functions, one builds a transformation from the reference element to each element in the domain. It is crucial that the mapping takes into consideration the fact that element rotation will change the values and the direction of the normal derivative. This is demonstrated in Figure A.2.

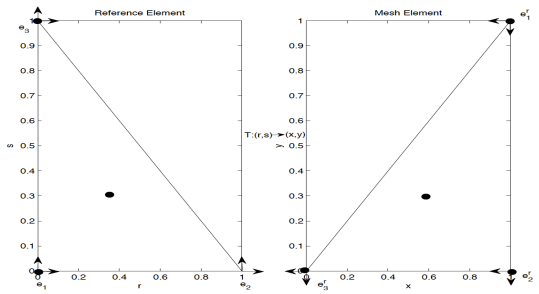

Figure A.2: Mapping from the reference element (left) to a general element (right)

While some simpler implementations match the derivatives along edges, the output is no longer the gradient but a directional derivative. As one can see in in the weak formulation

of the sensitivity equation,

$$(\mathcal{S})_q \begin{cases} \frac{d}{dt}v_h(t) = A(q_0)v_h(t) + A(h)z(t;q_0) & v_h(t) \in H_0^1(\Omega) \bigcap H^2(\Omega), \\ z(0,x) = 0 & t \in [0,T], \end{cases} \tag{A.1}$$

the presence of $A(h)z$ requires computation of the gradient of z, ∇z. To have a better rate of convergence and accuracy in the computation of the sensitivities, we use Hermite cubic polynomials to solve the forward problem

$$(\mathcal{P})_q \begin{cases} \frac{d}{dt}z(t;q) = A(q)z(t;q) + f(t) & z(t) \in H_0^1(\Omega) \bigcap H^2(\Omega), \\ z(0,x) = z_0 & t \in [0,T]. \end{cases} \tag{A.2}$$

These details are discussed in Banks and Kunish [4]. Note that one does not need to use the Hermite cubic elements in the approximation of the sensitivity equation. However, while Hermite cubic elements do not form a subspace of $H^2(\Omega)$, the do provide adequate (nonconforming) approximations to many finite element formulations naturally posed in $H^2(\Omega)$.

Theorem A.1.1 (Ciarlet [15]). *The finite element space X_k generated from Hermite cubic elements is a subset of $C^0(\overline{\Omega}) \cap H^1(\Omega)$.*

Proof. The proof is given as Theorem 2.2.10 in [15]. □

We provide error estimates for our finite element approximations below.

A.1.1 Error Estimates for Sensitivity Equations Solutions

Theorem A.1.2 (Interpolation, Ciarlet [15]). *If $u \in H^3(\Omega)$, and Π_k is the interpolation operator mapping to X_k. Then the interpolated function satisfies the following error estimate*

$$\|u - \Pi_k u\|_{m,\Omega_k} \leq C\Delta_k^{3-m} \quad \text{if} \quad 0 \leq m \leq 3,$$

where Δ_k is the diameter of element Ω_k.

Theorem A.1.3. *Let v_h and be the solution of (A.1) and $\tilde{v}_h$ be the solution of (A.1) where the forward problem is the finite element solution of (A.2). Then if $z \in H^3(\Omega)$, there exists a constant C, independent of z, such that:*

- $\|\tilde{v}_h - v_h\|_{H^1(\Omega)} \leq C\|h\|_0\Delta^2$, *if the forward problem was evaluated with Hermite cubic elements, and*
- $\|\tilde{v}_h - v_h\|_{H^1(\Omega)} \leq C\|h\|_0\Delta$, *if evaluated with linear elements,*

where Δ is the maximum diameter of the elements.

Proof. Let $\tilde{z}$ be the finite element solution of (A.2) then

$$\int_{\Omega} \nabla \cdot (q_0 \nabla (v_h - \tilde{v}_h))(v_h - \tilde{v}_h) d\mathbf{x} = - \int \nabla \cdot (h \nabla (z - \tilde{z}))(v_h - \tilde{v}_h) d\mathbf{x}. \tag{A.3}$$

Using the Poincaré inequality, Cauchy-Schwarz inequality, and Theorem A.1.2,

$$\|v_h - \tilde{v}_h\|_{H^1(\Omega)} \leq \frac{\|h\|_0}{\inf\limits_{\mathbf{x} \in \Omega} q_0(\mathbf{x})} \|z - \tilde{z}\|_{H^1(\Omega)} \leq \frac{\|h\|_0}{[\inf\limits_{\mathbf{x} \in \Omega} q_0(\mathbf{x})]^2} \|f\|_{L^2(\Omega)} C \Delta^2. \tag{A.4}$$

The same arguments can be used to prove the linear convergence rate with linear finite elements. □

Appendix B

Extension Results on Sensitivity Analysis

This appendix is a quick review of sensitivity analysis for discretized problems (or problems set in finite dimensional spaces). For instance, the consider the discretization of the 1D Poisson equation with varying conductivity parameter q,

$$-\nabla\left(qz(q)\right) = f(x), \qquad f \in L^2(\Omega), \qquad \text{and} \qquad z \in H_0^1(\Omega). \tag{B.1}$$

If q is discretized as $q(x) \approx q^N(x) := \sum_{i=1}^{M} q_i \phi_i(x)$ and the solution z is discretized as $z(x;q) \approx z^{N,M}(x;q^M) := \sum_{i=1}^{N} z_i \psi_i(x)$ then one can define the sensitivity as follows:

$$\left[D_q z^{N,M}(q^M)\right]_{q^M=q_0^M}(h^M) := \lim_{\epsilon\to 0} \frac{z^{N,M}(x;q_0^M+\epsilon h^M) - z^{N,M}(x;q^M)}{\epsilon}$$

To find the partial derivative q_i one just need to set $h_i = e_i$ where e_i is the base i on $\mathbb{R}^M$. Another way is to differentiate directly the solution equation:

$$A^{N,N}(q^M)z^N(q^M) = F^N,$$

as follows:

$$D_{q^M}(A^{N,N}(q^M)z^N(q^M)) = D_{q^M}(F^N)$$

Under some assumptions this implies

$$D_{q^M}\left[A^{N,N}(q^M))z^N(q^M)\right]_{q^M=q_0^M}(h) = A^{N,N}(q^M)D_{q^M}\left[z^N(q^M)\right]_{q^M=q_0^M}$$

Which is very similar to its infinite dimensional peer (5.6).

This is the type of approach used by *UCODE* [40] where they repeatedly evaluate the cost functional to approximate the gradient. Let $J(q)$ be the cost functional then the gradient is approximated by the forward difference quotient

$$\left[\nabla J(q)\right]_i \approx \frac{J(q+\epsilon e_i) - J(q)}{\epsilon},$$

with ϵ "significantly small." This approximates the Gâteaux derivative, as discussed in 5.2.1. This type of sensitivity approximation is discussed in, e.g. [43].

Other approach for computing sensitivities are considered in, e.g. Borggaard and Burns [9], Griewank [20], and Turgeon, et al. [48].

B.1 Evaluating Partial Derivatives

Theorem B.1.1. *Let $F(x,y) : X \times Y \to Z$ be functional then if the Frechét derivatives $\left[D_{(x,y)}F\right]_{(x,y)=(x_0,y_0)}$, $\left[D_xF\right]_{x=x_0}$ and $\left[D_yF\right]_{y=y_0}$ exists then the following equality holds:*

$$\left[D_{(x,y)}F\right]_{(x,y)=(x_0,y_0)}(\theta,\nu) = \left[D_xF\right]_{x=x_0}(\theta) + \left[D_yF\right]_{y=y_0}(\nu) \tag{B.2}$$

This fact it is very important one needs to evaluate the total derivative of the solution convection diffusion equation. By defining $A(q)(.) = \nabla\cdot(q_1\nabla z) + (q_{2,1}, q_{2,2})\cdot\nabla z + q_3 z$ and observing that the partial and vector derivative of the convection-diffusion exists then one can easily see that:

$$\left[D_q z(t,q)\right]_{q=q_0}(\delta q) =$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \left[D_a z(t,q)\right]_{a=a_0} & 0 & 0 & 0 \\ 0 & \left[D_{b^1} z(t,q)\right]_{b^1=b_0^1} & 0 & 0 \\ 0 & 0 & \left[D_{b^2} z(t,q)\right]_{b^2=b_0^2} & 0 \\ 0 & 0 & 0 & \left[D_c z(t,q)\right]_{c=c_0} \end{bmatrix} \begin{bmatrix} \delta a \\ \delta b^1 \\ \delta b^2 \\ \delta c \end{bmatrix}.$$

Therefore one can compute all these derivative separately and taking full advantage of parallel computation. It is relevant to mention that the SVD the whole matrix can be written as SVD of the individual matrices. That same sum will provide which direction of which parameter has the largest value. Whereas the SVD of the whole matrix gives the perturbation on the whole system. This can be seen in the following set of equations:

$$z(q_0+h) \approx z(q_0) + \sum_{i=1}^{N} \langle v_i, h\rangle_Q \sigma_i u_i \tag{B.3}$$

and

$$z(q_0+h) \approx z(q_0) +$$

$$\sum_{i=1}^{N} \langle v_i^a, h^a\rangle_{Q^a} \sigma_i^a u_i^a + \sum_{i=1}^{N} \langle v_i^{b^1}, h^{b^1}\rangle_{Q^{b^1}} \sigma_i^{b^1} u_i^{b^1} + \sum_{i=1}^{N} \langle v_i^{b^2}, h^{b^2}\rangle_{Q^{b^2}} \sigma_i^{b^2} u_i^{b^2} + \sum_{i=1}^{N} \langle v_i^c, h^c\rangle_{Q^c} \sigma_i^c u_i^c \tag{B.4}$$

This means that if you perturb the whole system in the direction v_1 you will have a major impact, whereas if you just want to perturb the diffusion term then is v_1^a.

B.2 Power Method for Partial Derivatives of Order Greater Than Two

In the following paragraph we will discuss the power method to evaluate the most sensitive directions of the steady advection-diffusion equation.

Lets consider the following PDE:

$$P_q \begin{cases} A(q)z(q) = f, & f \in L^2(\Omega); \\ u \in H_0^1(\Omega). \end{cases}$$

Where: $A(q)z = \sum_{i,j}^{n} (a_{i,j} z_{x_i})_{x_j} + bz$

$$S_q \begin{cases} -A(q)v_h = A(h)z(q), & f \in L^2(\Omega); \\ v_h \in H_0^1(\Omega). \end{cases}$$

Then one given h^{m-1} one must find $h^m = D_q z^*(D_q z(h^{m-1}))$ which is equivalent to:

$$(h^m, \phi)_{L^2(\Omega)^2} = (D_q z^*(D_q z(h^{m-1})), \phi))_{L^2(\Omega)^4} =$$
$$= (D_q z(h^{m-1}), D_q z(\phi))_{L^2(\Omega)} = (D_q z(h^{m-1}), -A^{-1}(q)A(\phi)z(q))_{L^2(\Omega)} =$$
$$= (-A^{-1}(q)^* D_q z(h^{m-1}), A(\phi)z(q))_{L^2(\Omega)} = (-A^{-1}(q) D_q z(h^{m-1}), A(\phi)z(q))_{L^2(\Omega)} =$$
$$= (-A^{-1}(q) D_q z(h^{m-1}), \sum_{i,j}^{n} (\phi_{i,j} z_{x_i})_{x_j} + \phi_2 z)_{L^2(\Omega)} =$$
$$= -\sum_{i,j}^{n} ((A^{-1}(q) D_q z(h^{m-1})_{x_{i,j}}) z_{x_{i,j}}, \phi_{i,j})_{L^2(\Omega)} + ((A^{-1}(q) D_q z(h^{m-1}))z, \phi_2)_{L^2(\Omega)}$$

This leads to :

$$h^m = \begin{bmatrix} -((A^{-1}(q) D_q z(h^{m-1})_{x_{1,1}}) z_{x_{i,j}} \\ \vdots \\ -((A^{-1}(q) D_q z(h^{m-1})_{x_{n,n}}) z_{x_{n,n}} \\ [(A^{-1}(q) D_q z(h^{m-1})]\, z(q) \end{bmatrix} \tag{B.5}$$

Which enables a fast power method discussed in the next section.

Theorem B.2.1 (Power method convergence)**.** *Under the same assumptions of Theorem 5.4.13, the power method converges. At each iteration k of a discretization on H^N and Q^M the error is bounded by*

$$c \left[\|D_q(z) - D_q^{N,M}(z)\| + \left(\frac{\sigma_2^{N,M}}{\sigma_1^{N,M}} \right)^k \right].$$

Furthermore, there is a subsequence $(M, N, \alpha(M,N))$ such that

$$\lim_{N,M} \|\sigma_1 - \sigma_1^{N,M,\alpha(M,N)}\| \to 0. \tag{B.6}$$

Where $\alpha(M,N)$ is the iteration number for the power method.

Proof.

$$\|\sigma_1 - \sigma_1^{N,M,n}\| \leq \|\sigma_1 - \sigma_1^{N,M}\| + \|\sigma_1^N - \sigma_1^{N,M,n}\| \tag{B.7}$$

Let $\varepsilon > 0$ then by Theorem 5.4.13 we know that $\exists\ p$ such that for $N, M > p$ we have $\|\sigma_1 - \sigma_1^{N,M,n}\| \leq C\|D_q(z) - D_q^{N,M}(z)\| < \frac{\varepsilon}{2}$. By the same arguments of the power method for a matrix we have that for fixed N, M there $\exists \rho_{M,N}$ such that for $k > \rho_{M,N}$ the following inequality $C\|\sigma_1^N - \sigma_1^{N,M,n}\| \leq \left(\frac{\sigma_2^{N,M}}{\sigma_1^{N,M}}\right)^k < \frac{\varepsilon}{2}$ holds, since by Hypothesis $\sigma_1 > \sigma_2$ This then proves that for $N, M > p$ and $k > \rho_{M,N}$ this implies that :

$$\|\sigma_1 - \sigma_1^{N,M,n}\| < \varepsilon \tag{B.8}$$

Finally by choosing $\alpha(M, N) = 1 + \rho_{M,N}$, we have that there is a subsequence that converges to the infinitesimal singular value which completes the proof. □

Appendix C

Monte Carlo Method

The Monte Carlo method is a very common in several types of inverse problems and in geosciences [28] and [34]. The idea is to given some data compute the parameter estimation individually and then take the moments of those estimated parameters, such as mean, variance and so forth. The main problem with this method is that the rate of convergence is of the order of $o(\frac{1}{\sqrt{N}})$. Which is relatively slow, thus the method needs a lot of data to have a reasonable accuracy, therefore a lot of inverse problems must be computed in order to have a good accuracy. So for the simple case explained in Section 6.4.2 would have been a simply inversion for each data z_i, q_i^* and then find $\mathbb{E}(q)$ by its natural approximation, the sample mean $\hat{q} := \frac{1}{N_{sample}} \sum_{i=1}^{N_{sample}} q_i^*$ and variance $\mathbb{V}(q)$ by it sample counterpart S. This is the classical method and it is known to be very inefficient. There are several extensions and improvements of these ideas and applications to a range of application areas, cf. [14], [16], and [52].

Bibliography

[1] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces*, volume 140. Academic Press, 2003.

[2] K. E. Atkinson. *An Introduction to Numerical Analysis*. John Wiley & Sons, 2008.

[3] I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 45(3):1005–1034, 2007.

[4] H. T. Banks and K. Kunisch. *Estimation Techniques for Distributed Parameter Systems*. Birkhauser, Boston, 1989.

[5] J. Bear and A. H.-D. Cheng. *Modeling Groundwater Flow and Contaminant Transport*. Springer, 2010.

[6] P. B. Bedient, H. S. Rifai, and C. J. Newell. *Groundwater Contamination: Transport and Remediation*. Prentice-Hall International, Inc., 1994.

[7] M. A. Biot. General theory of three-dimensional consolidation. *Journal of Applied Physics*, 12:155–164, 1941.

[8] J.-F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. A. Sagastizábal. *Numerical Optimization: Theoretical and Practical Aspects*. Springer, 2006.

[9] J. Borggaard and J. Burns. A PDE sensitivity equation method for optimal aerodynamic design. *Journal of Computational Physics*, 136(2):366–384, 1997.

[10] J. Borggaard and V. L. Nunes. Fréchet sensitivity analysis for partial differential equations with distributed parameters. In *Proc. 2011 American Control Conference*, 2011.

[11] J. Borggaard, V. L. Nunes, and H.-W. van Wyk. Sensitivity and uncertainty quantification. *Mathematics in Engineering, Science and Aerospace*, 4(2):115–127, 2013.

[12] S. C. Brenner and R. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer, 2007.

[13] J. A. Burns, P. Morin, and R. D. Spies. Parameter differentiability of the solution of a nonlinear abstract Cauchy problem. *Journal of Mathematical Analysis and Applications*, 252(1):18–31, 2000.

[14] Y. Cao, H. Chi, C. Milton, and W. Zhao. Exploitation of sensitivity derivatives via randomized quasi-Monte Carlo methods. *Monte Carlo Methods and Applications*, 14(3):269–279, 2008.

[15] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North Holland, fourth edition, 1978.

[16] K. Cliffe, M. Giles, R. Scheichl, and A. L. Teckentrup. Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. *Computing and Visualization in Science*, 14(1):3–15, 2011.

[17] L. G. Davis and J. R. Singler. Computational issues in sensitivity analysis for 1D interface problems. *Journal of Computational Mathematics*, 29(1):111–130, 2011.

[18] L. C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 1998.

[19] S. H. Gould. *Variational Methods for Eigenvalue Problems: An Introduction to the Methods of Rayleigh, Ritz, Weinstein, and Aronszajn*. Dover Publications, 1995.

[20] A. Griewank. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*, volume 19 of *Frontiers in Applied Mathematics*. SIAM, 2000.

[21] T. Herdman and R. Spies. Fréchet differentiability of the solutions of a semilinear abstract cauchy problem. *Journal of Mathematical Analysis and Applications*, 307(2):656–676, 2005.

[22] M. C. Hill. *MODFLOW-2000: The US Geological Survey Modular Groundwater Model–User Guide to the Observation, Sensitivity, and Parameter-estimation Processes, and Three Post-processing Programs*. US Department of the Interior, US Geological Survey, 2000.

[23] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints*, volume 23 of *Mathematical Modelling: Theory & Applications*. Springer, 2008.

[24] J. Hoffmann, S. A. Leake, D. L. Galloway, and A. M. Wilson. MODFLOW-2000 groundwater model–user guide to the subsidence and aquifer-system compaction (SUB) package. Technical Report Open-File Report 03–233, U. S. Geological Survey, Tucson, Arizona, 2003.

[25] K. Ito and K. Kunisch. *Lagrange Multiplier Approach to Variational Problems and Applications*. Society for Industrial and Applied Mathematics, 2008.

[26] C. T. Kelley and E. W. Sachs. Mesh independence of Newton-like methods for infinite dimensional problems. *Journal of Integral Equations and Applications*, 3(4):549–573, 1991.

[27] C. Kravaris and J. H. Seinfeld. Identification of parameters in distributed parameter systems by regularization. *SIAM Journal on Control and Optimization*, 23(2):217–241, 1985.

[28] G. Kuczera and E. Parent. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. *Journal of Hydrology*, 211(1):69–85, 1998.

[29] K. Kunisch and L. White. The parameter estimation problem for parabolic equations and discontinuous observation operators. *SIAM Journal on Control and Optimization*, 23(6):900–927, 1985.

[30] S. A. Leake. Interbed storage changes and compaction in models of regional groundwater flow. *Water Resources Research*, 26(9):1939–1950, September 1990.

[31] R. B. Lehoucq, D. C. Sorensen, and C. Yang. *ARPACK Users' Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*, volume 6. Siam, 1998.

[32] M. W. Mendonca. *Multilevel Optimization: Convergence Theory, Algorithms and Application to Derivative-Free Optimization*. PhD thesis, Facultés Universitaires Notre-Dame de la Paix, Namur, Belgium, 2009.

[33] J. Modersitzki. *Numerical Methods for Image Registration (Numerical Mathematics and Scientific Computation)*. Oxford University Press, 2004.

[34] K. Mosegaard and A. Tarantola. Monte Carlo sampling of solutions to inverse problems. *Journal of Geophysical Research*, 100(B7):12431–12, 1995.

[35] F. Nobile, R. Tempone, and C. G. Webster. An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2411–2442, 2008.

[36] F. Nobile, R. Tempone, and C. G. Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2309–2345, 2008.

[37] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer, New York, 1983.

[38] P. J. Phillips and M. F. Wheeler. A coupling of mixed and continuous Galerkin finite element methods for poroelasticity I: the continuous-in-time case. *Computational Geosciences*, 11(2):131–144, 2007.

[39] P. J. Phillips and M. F. Wheeler. A coupling of mixed and continuous Galerkin finite element methods for poroelasticity II: the discrete-in-time case. *Computational Geosciences*, 11(2):145–158, 2007.

[40] E. P. Poeter and M. C. Hill. UCODE, a computer code for universal inverse modeling. *Computers & Geosciences*, 25(4):457–462, 1999.

[41] M. Reed and B. Simon. *Methods of Modern Mathematical Physics, I: Functional Analysis*. Gulf Professional Publishing, 1980.

[42] Y. Saad. *Numerical Methods for Large Eigenvalue Problems*. SIAM, 1992.

[43] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global Sensitivity Analysis: The Primer*. Wiley-Interscience, 2008.

[44] S. Seubert and J. Wade. Fréchet differentiability of parameter-dependent analytic semigroups. *Journal of Mathematical Analysis and Applications*, 232(1):119–137, 1999.

[45] R. E. Showalter. *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*, volume 49. American Mathematical Society, 1997.

[46] R. E. Showalter. Diffusion in poro-elastic media. *Journal of Mathematical Analysis and Applications*, 251(1):310–340, 2000.

[47] G. L. Sleijpen and H. A. Van der Vorst. A Jacobi–Davidson iteration method for linear eigenvalue problems. *SIAM Review*, 42(2):267–293, 2000.

[48] E. Turgeon, D. Pelletier, and J. Borggaard. A continuous sensitivity equation approach to optimal design in mixed convection. *Numerical Heat Transfer: Part A: Applications*, 38(8):869–885, 2000.

[49] J. F. Van Doren, S. G. Douma, P. M. Van den Hof, J. D. Jansen, and O. H. Bosgra. Identifiability: From qualitative analysis to model structure approximation. In *Proceedings of the 15th IFAC Symposium on System Identification*, Saint-Malo, France, 2009.

[50] T. Vidar. *Galerkin Finite Element Methods for Parabolic Problems*, volume 25. Springer, Berlin, 2006.

[51] C. G. Webster. *Sparse grid stochastic collocation techniques for the numerical solution of partial differential equations with random input data*. PhD thesis, The Florida State University, 2007.

[52] R. M. Wilcox. Exponential operators and parameter differentiation in quantum physics. *Journal of Mathematical Physics*, 8:962, 1967.

[53] W. W.-G. Yeh and N.-Z. Sun. An extended identifiability in aquifer parameter identification and optimal pumping test design. *Water Resources Research*, 20(12):1837–1847, 1984.

[54] K. Yosida. *Functional Analysis*, volume 35. Springer, Berlin, 1980.

[55] W. Yu, F. Zhang, and W. Xie. Differentiability of C^0-semigroups with respect to parameters and its application. *Journal of Mathematical Analysis and Applications*, 279(1):78–96, 2003.

[56] J. Zowe and S. Kurcyusz. Regularity and stability for the mathematical programming problem in Banach spaces. *Applied Mathematics & Optimization*, 5(1):49–62, 1979.