



Technisch-Naturwissenschaftliche
Fakultät

A Sobolev Space Smoother for Atmospheric Tomography in Multi Conjugate Adaptive Optics

MASTERARBEIT

zur Erlangung des akademischen Grades

Diplomingenieur

im Masterstudium

Mathematik in den Naturwissenschaften

Eingereicht von:

Markus Eslitzbichler

Angefertigt am:

Institut für Industriemathematik

Beurteilung:

Univ. Prof. Dr. Ronny Ramlau

Mitwirkung:

Dipl.-Ing. Dr. Clemens Pechstein

Linz, Januar, 2012

Abstract

Adaptive Optics are a system in optical and infrared telescopes to counter the effects of Earth's atmosphere on astronomical observations.

An important part of this is determining the distribution of variations in the refractive index in the atmosphere, which cause image distortions. This process is called "Atmospheric Tomography". Mathematically, this requires solving an ill-posed inverse problem.

With increasing telescope sizes and sensor resolutions, the dimensionality of this inverse problem increases rapidly, and solving this on next-generation telescopes such as the European Southern Observatory's *Extremely Large Telescope* requires novel mathematical techniques.

A candidate solver is the *Landweber-Kaczmarz* algorithm, an iterative method that is particularly well suited for ill-posed problems and problems in high dimensions. In the specific case of adaptive optics, this algorithm works on a domain consisting of overlapping ring-shapes. This causes the update step to introduce erroneous singularities in the solution.

In this thesis, we investigate whether adding a smoothing step to the update can improve the quality of the reconstruction. The smoothing is based on Sobolev space embeddings. This leads to an elliptic partial differential equation on such overlapping ringshape domains. This is a non-convex domain with singularities on the boundary. We will investigate and formulate detailed *regularity*, i.e. smoothness, results for solutions of this partial differential equation on these particular domains.

Finally we will present numerical results for a Landweber-Kaczmarz solver augmented by the smoother for realistic test scenarios.

Zusammenfassung

Adaptive Optik bezeichnet Systeme in optischen und Infrarot-Teleskopen, deren Aufgabe es ist die störenden Effekte der Erdatmosphäre auf astronomische Beobachtungen auszugleichen.

Ein wichtiger Bestandteil solcher Systeme ist die Bestimmung der Verteilung von Variationen im Brechungsindex in der Atmosphäre, die Verzerrungen in den Teleskopaufnahmen verursachen. Dieser Prozess wird als "Atmosphärische Tomographie" bezeichnet. Mathematisch gesehen erfordert sie die Lösung eines schlecht-gestellten inversen Problems.

Durch rasant ansteigende Teleskopgrößen und wachsende Sensorauflösungen wächst die Dimension dieses inversen Problems rapide an. Neue mathematische Ansätze sind erforderlich um dieses Problem für die nächste Generation moderner Teleskope, etwa das *Extremely Large Telescope* der Europäischen Südsternwarten, zu lösen.

Ein guter Kandidat ist die *Landweber-Kaczmarz* Methode, ein iteratives Verfahren, das sich besonders gut für schlecht gestellte und hochdimensionale Probleme eignet. Im speziellen Anwendungsgebiet der adaptiven Optik operiert dieses Verfahren auf einem Gebiet überlappender ebener Ringe. Update-Schritte in der Iteration führen dadurch zu inkorrekten Unstetigkeiten in der Lösung.

In dieser Arbeit untersuchen wir, ob das Hinzufügen einer Glättungsoperation zu diesen Update-Schritten die Qualität der Rekonstruktion verbessern kann. Die Glättung wird aus Sobolevraum-Einbettungen hergeleitet. Dies führt zu einer linearen elliptischen partiellen Differentialgleichung auf diesen Gebieten überlappender Ringe. Es handelt sich hierbei um eine nicht-konvexe Menge mit Singularitäten am Rand. Wir werden die *Regularität*, d.h. Glattheit, der Lösungen dieser Differentialgleichung auf solchen Gebieten detailliert untersuchen und formulieren.

Anschließend werden wir numerische Ergebnisse für den, um diesen Glätter erweiterten, Landweber-Kaczmarz Löser für realistische Testdaten präsentieren.

Acknowledgments

I would like to thank my advisor, Professor Ronny Ramlau for enabling me to work in the fascinating field of Adaptive Optics and for his valuable support and guidance throughout this project.

I am deeply grateful for all the advice and support I have received from Dipl.-Ing. Dr. Clemens Pechstein, my co-advisor.

Special thanks go to Dipl.-Ing. Dr. Andreas Obereder and Dipl.-Ing. Matthias Rosensteiner from the *Mathematical algorithms and software for ELT adaptive optics* project for their help in working with their Adaptive Optics simulation codes.

For interesting discussions and proofreading, I would like to thank my colleagues Dipl.Math. Tino Schulz, Daniela Saxenhuber and Peter Gangl.

Markus Eslitzbichler
Linz, Januar 2012

Contents

1	Introduction	1
2	Problem Formulation	3
2.1	Physical Background	3
2.1.1	Astronomy	3
2.1.2	Adaptive Optics	4
2.1.3	Multi Conjugate Adaptive Optics	6
2.1.4	A Model of the Atmosphere	8
2.2	Mathematical Background	10
2.2.1	Notation	10
2.2.2	Problem formulation	12
2.2.3	Solution Methods	12
2.3	An Inverse Problem	13
2.3.1	An Iterative Solution	13
2.3.2	Solving the adjoint embedding	16
3	Regularity in Non-Convex Domains	18
3.1	Regularity of Elliptic Partial Differential Equations	18
3.1.1	Structure of this Chapter	19
3.2	Prerequisites	19
3.2.1	Geometry and Notation	20
3.2.2	Sobolev Spaces	21
3.2.3	Variational Formulation	23
3.3	A Homogenization Based Result	24
3.3.1	Homogenization	24
3.3.2	A Limit Approach	25
3.3.3	Interpretation of the Smoothness Results	28
3.4	Modified Bessel Functions of the First and Second Kind	29
3.4.1	The Modified Bessel Function of the First Kind	31
3.4.2	The Modified Bessel Function of the Second Kind	32
3.5	A Closer Look at the Corners	33
3.5.1	Characterizing the Orthogonal Complement N	33
3.5.2	Constructing a Space for All Solutions	39

4 Numerical Realization	41
4.1 Implementation	41
4.1.1 Parameters	42
4.1.2 Evaluation	43
4.2 Tests	44
4.2.1 A first example	44
4.2.2 Smoothing Strength	44
4.2.3 Photon flux	45
4.2.4 Different Numbers of Reconstructed Layers and Deformable Mirrors	46
4.2.5 Large Iteration Numbers	51
4.2.6 Photon Flux 2	52
4.3 Interpretation of the Results	53
5 Conclusion	61
List of Figures	65

Chapter 1

Introduction

When making astronomical observations using earth-based telescopes, changes in the refractive index in the atmosphere due to temperature inhomogeneities cause strong distortions in the observation results. *Adaptive Optics* (AO) is a method that is used by modern telescopes to counter this effect. *Atmospheric Tomography* is an important part of AO systems. It attempts to reconstruct the structure of the atmosphere above the telescope, in particular the (turbulent) inhomogeneous variations in the refractive index that cause perturbations in the light wavefront.

This is not only a severely ill-posed problem, but also of high computational complexity. As telescope sizes and hardware capabilities increase, the systems of equations involved grow bigger and bigger, which requires new mathematical approaches to solve these systems in a setting that allows only a fraction of a second of computation time.

A newly proposed algorithm for tomography problems is the iterative *Landweber-Kaczmarz method*, which uses a component-wise approach to break down big systems of equations into more manageable chunks. In the specific case of atmospheric tomography, this algorithm however suffers from the fact that the single components are defined on two-dimensional domains of overlapping ring-shapes; an update step for a single one of those introduces erroneous discontinuities in the solution.

In this thesis we investigate incorporating an additional smoothing step into the Landweber-Kaczmarz iteration to improve both convergence speed and the quality of the solution.

The thesis is organized as follows:

- In Chapter 2, we will first give a short introduction to adaptive optics, its role in astronomy and its basic operating principles. We then move on to formulate a mathematical model of atmospheric tomography and analyse the spaces and operators involved. We will encounter the Sobolev embedding operator i and its adjoint i^* and derive from them an elliptic partial differential equation, the

solution of which is the result of the above-mentioned smoothing step.

- In Chapter 3, we will take a closer look at the partial differential equation derived in Chapter 2 and analyse its smoothness. It will turn out, that this depends on the exact shape of the domain, which the partial differential equation is defined on. This is part of *Regularity Theory*.
- In Chapter 4, numerical results of the smoothed Landweber-Kaczmarz method are presented for realistic datasets and parameters for two modern telescopes: the *European Southern Observatory's* already operating *Very Large Telescope* and their next-generation *European Extremely Large Telescope*, currently in development.
- Finally, in Chapter 5, we will review both theoretical and experimental results and discuss possible areas of improvement.

Chapter 2

Problem Formulation

Where we see how Adaptive Optics works, introduce mathematical notation and derive a system of equations that describe how atmospheric turbulences relate to perturbations in the incoming wavefronts and finally end up with a partial differential equation that acts as a smoothing operator for turbulence reconstruction.

Things then did not delay in turning curious.

The Crying of Lot 49
THOMAS PYNCHON

2.1 Physical Background

2.1.1 Astronomy

Astronomy is an observational science, built on measurements and data. The telescope, from its humble beginnings in the 17th century on to the age of NASA's *Great Observatories* space telescopes and earthbound *Very Large* telescopes, has been Astronomy's primary tool for investigating the heavens.

While modern telescopes are marvels of science, technology and engineering, it is a basic truth in Astronomy that *Size Matters*. Size in this case means aperture, the diameter of the telescope opening and its primary mirror. With bigger apertures a telescope can gather more light to make an observation, which enables it to observe a wider range of possible targets and increases angular resolution. The aperture places a strict physical limit, the *diffraction limit* on image quality. Informally speaking, this limit tells the minimum angular separation two light sources need to have in order to be distinguishable from one another. This value is inversely proportional to the aperture.

This fundamental limit on observation-performance has led to the construction of telescopes with ever bigger apertures.

At optical and near infrared (IR) wavelengths the current generation of cutting-edge telescopes range from a few to up to 12 meters of aperture, typical examples being the 10m *Keck* telescopes in Hawaii, the *European Southern Observatory*'(ESO) 8.2m *Very Large Telescopes* (VLT) in the Chilean desert (very high, very dry - ideal observing conditions) or the 11.9m *Large Binocular Telescope*. The next generation of earth-based optical telescopes, scheduled to be completed within the next ten years, features even larger apertures, e.g the 30 meters of the *Thirty Meter Telescope* or the 39.3 meters for ESO's *European Extremely Large Telescope* (E-ELT).

However, telescopes do not reach their theoretical diffraction limits due to imperfections in their optical systems and perturbations in the light wavefront caused by the atmosphere. Especially telescopes with big apertures suffer strongly from such atmospheric distortions. They are caused by temperature inhomogeneities in the atmosphere, where turbulently distributed pockets of warm and cold air lead to changes in the refractive index of air. This severely degrades telescope performance far below the diffraction limit. Indeed, it can be shown that for apertures over 20 cm, atmospheric conditions are the limiting factor for observational quality (Glindemann et al. [2000]).

In order to counter this effect, these big telescopes use *adaptive optics* systems to achieve high image quality.

2.1.2 Adaptive Optics

According to astronomers, Earth has one very annoying property: It has an atmosphere.

The vast majority of electromagnetic radiation incident on earth is either reflected or absorbed by the atmosphere. Only a few narrow ranges of wavelengths manage to penetrate it. These are the so-called *Optical*, *Infrared* and *Radio Windows*.

But even though the atmosphere is transparent to optical and near-IR wavelengths, it is an inhomogeneous body with turbulent temperature distributions causing refraction in incoming light wavefronts, i.e. variations in the phase of the incoming light.

One way to deal with these distortions are *Adaptive Optics* systems, where telescope designers attempt to counter the effects of a turbulent atmosphere with deformable mirrors that match those turbulences to even out the incoming wavefronts. See Figure 2.1 for the basic idea. Please note that this is a vast and complex subject and any attempt to do it justice would be outside the scope of this thesis. We will therefore

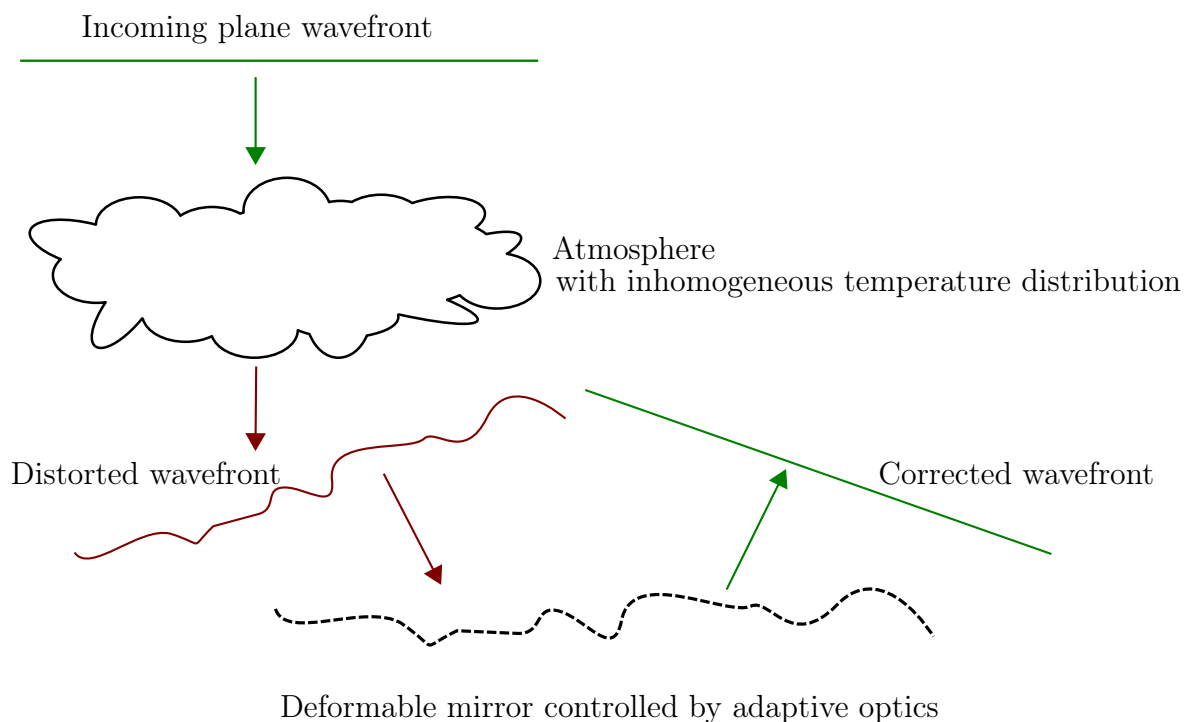


Figure 2.1: Concept of a deformable mirror.

focus on the bare minimum of background we need and direct the interested reader to e.g. Roddier [1999] and Hardy [1998].

These mirror deformations have to be made and adjusted on a millisecond basis, which is in itself a formidable technological challenge. In this thesis we will work on one of the prerequisite steps before the mirror deformation: *Atmospheric Tomography*, where one attempts to assemble an accurate picture of the turbulences in the atmosphere, which then allows one to calculate just how to deform one or more mirrors to flatten the wavefronts.

The basic idea is simple: If we look at a bright star, we know the light it emits should arrive with a virtually plane wavefront. If the wavefront is distorted however, we can measure these distortions and use that to attempt to reconstruct turbulence distributions in the atmosphere.

For this purpose, wavefront sensors that measure the incoming light wave are pointed at *guide stars*, which can be either *natural* ones (which have the problem of being somewhat sparsely distributed in the sky) or *laser guide stars* (which introduce additional errors as they are relatively close to the telescope). In this thesis we will only

consider natural guide stars.

Wavefront sensors measure the slopes of the incoming wavefront. The actual phase function can then be reconstructed from these measurements. In this thesis we assume that this has already been done, for example using the *Cumulative Reconstruction* method from Zhariy et al. [2011] and Rosensteiner [2011] .

2.1.3 Multi Conjugate Adaptive Optics

There are several variants of AO systems. The simplest variant is *Single Conjugate Adaptive Optics* (SCAO), where only a single guide star is used to correct for wavefront perturbations. This gives corrections only in a very narrow field of view. Atmospheric tomography is not required for SCAO, as perturbations coming from only one direction can be corrected for directly with the deformable mirror. An extension to SCAO is *Extreme Adaptive Optics*, which is an additional high-frequency high-resolution AO system on top of an SCAO loop. Using a coronagraph to cover the star, this can be used to look for planetary systems.

The next step in AO technology is *Multi Conjugate Adaptive Optics*, where multiple guide stars and multiple deformable mirrors, conjugated to different heights, are used to achieve phase corrections over a wider field of view. Figure 2.2 shows a simplified schematic of such a system.

A related approach to MCAO is *Multi Object Adaptive Objects*, which aims to enable observing several objects concurrently by using multiple deformable mirrors working in parallel, each of those focused on one specific object. MOAO, as well as MCAO, requires an atmospheric tomography process before the required mirror deformations can be calculated. In this thesis we will focus on MCAO systems.

The quality of the wavefront correction improves with the number of mirrors as well as with the number of guide stars and the resolution of the wavefront sensors. Such improvements in hardware capabilities however increase the size of the systems of equations that need to be solved in the atmospheric tomography computation. With the advent of the *Giant Telescopes*, with apertures in the tens of meters and up to 6 or more wavefront sensors, the question of how to efficiently solve the involved systems becomes a great concern.

It should be noted that while the systems in question are not necessarily “big”, the problem comes from the amount of time available for the reconstruction. These values vary from telescope to telescope, but are usually ≈ 1 ms, in some cases however as low as $200 \mu\text{s}$. These time limits stem from the *Frozen Flow* hypothesis, which assumes that turbulences in atmospheric layers don’t change over short time periods but are

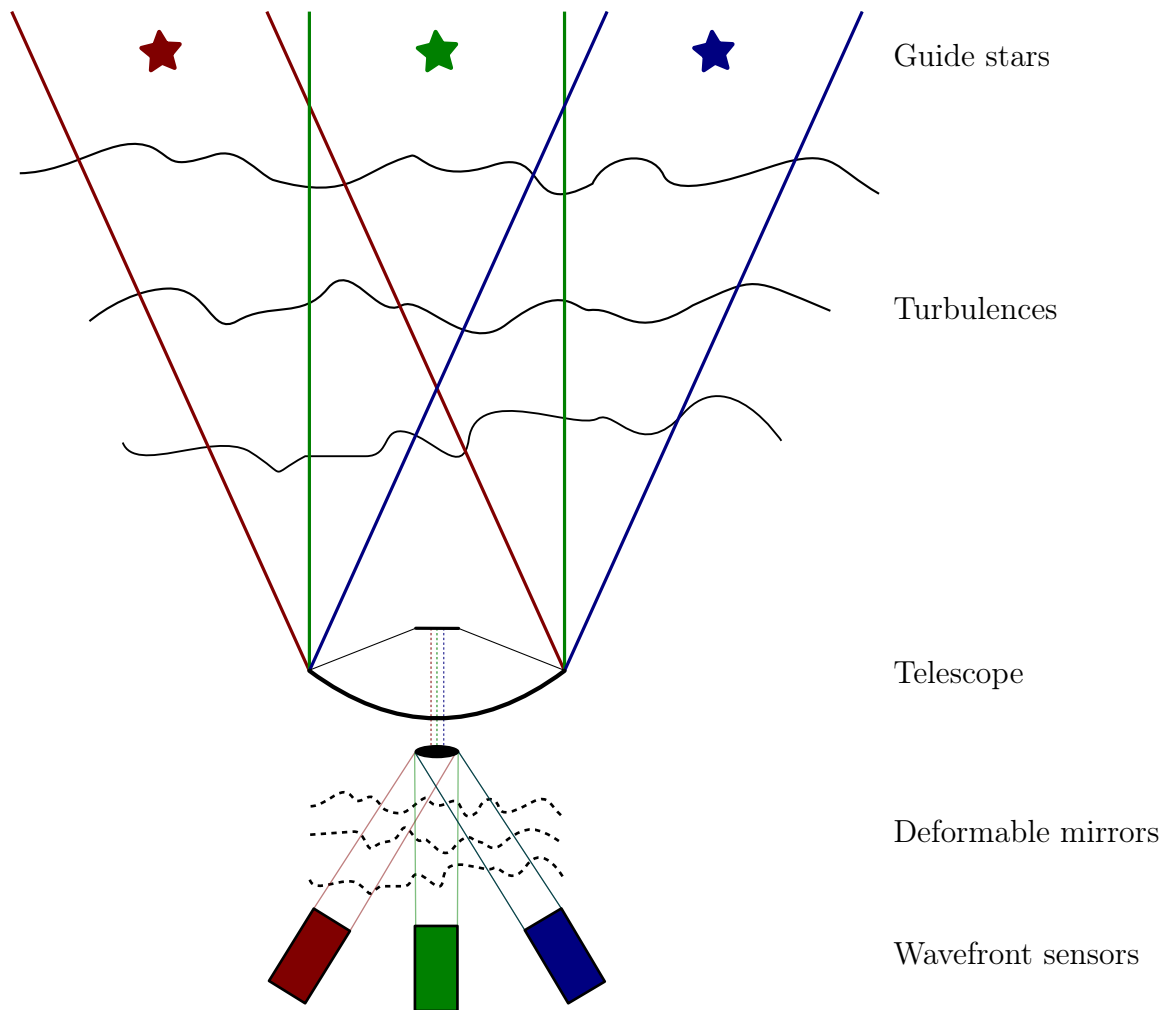


Figure 2.2: Schematics of an Multi Conjugate Adaptive Optics system. Note the overlapping volumes covered by the wavefront sensors.

instead simply transported by wind. Under this hypothesis, the *atmospheric time constant* τ_0 is defined as $\tau_0 := r_0/\bar{v}$, where \bar{v} is the average wind speed and r_0 denotes the *Fried parameter*, which roughly characterizes the size of a turbulence spot in the atmosphere. See Section 2.1.4 for more information on this parameter. The constant τ_0 expresses, how quickly the turbulence patterns in the atmosphere over the telescope change. This in turn dictates the time limits for adaptive optics systems.

For example, current plans for an MCAO system on the E-ELT call for 6 wavefront sensors with a resolution of 84x84 pixels with about 1 ms time available for the computation of the desired deformable mirror configuration which consists of the steps

wavefront reconstruction, atmospheric tomography and calculation of mirror actuator commands. In the atmospheric tomography step on the order of 3 to 9 atmospheric layers (more on this in Section 2.1.4) need to be reconstructed, resulting in system sizes ranging roughly from 100 000 to 400 000 unknowns. In Section 2.2.3 we discuss currently used mathematical methods for MCAO Systems.

2.1.4 A Model of the Atmosphere

As could already be glimpsed in Figure 2.2, we will assume that wavefront distortions caused by an inhomogeneous temperature distribution in the atmosphere can be accurately modelled as being caused by a discrete number of thin turbulent layers at different altitudes. We will further assume that the turbulences within each of these layers can be accurately described by the *Von Karman turbulence* model. Additionally, we will neglect dispersion effects, i.e. we will ignore that refraction depends on the wavelength.

Von Karman turbulence is a statistical model to describe the distribution of small scale turbulent motions. It can be seen as an extension of the *Kolmogorov* turbulence model, which we will discuss first.

The Kolmogorov model assumes that the distribution of turbulence in a layer depends only on spatial separation, i.e. that it is a stationary stochastic process. See Figure 2.3 for a few simulated turbulence fields based on Kolmogorov's model.

In order to characterize turbulences in the atmosphere, a few statistical functions prove very useful. Let u be a stationary stochastic process. Then the *structure function* is defined as

$$D_u(\Delta x) := \mathbb{E}[(u(x + \Delta x) - u(x))^2],$$

where Δx denotes spatial distance. Thus, the structure function gives the expectation of the squared differences of the process at points with varying distances.

In the Kolmogorov turbulence model, the structure function takes the following shape (Foy and Foy [2005]) for the refractive index function n :

$$D_n(\Delta x) := C_n^2(h)|\Delta x|^{2/3} \quad (2.1)$$

where C_n^2 describes the strength (amplitude of the phase function) of the turbulence. This *refractive index constant* varies with altitude and needs to be determined experimentally. Taking the Fourier transform of the structure function (2.1) yields the *power-spectrum*, which characterizes the distribution of energy across the range of frequencies:

$$\Phi_n(f) := 0.033(2\pi)^{-2/3}C_n^2(h)|f|^{-11/3}. \quad (2.2)$$

The von Karman model now uses the following energy spectrum (Vorontsov et al.

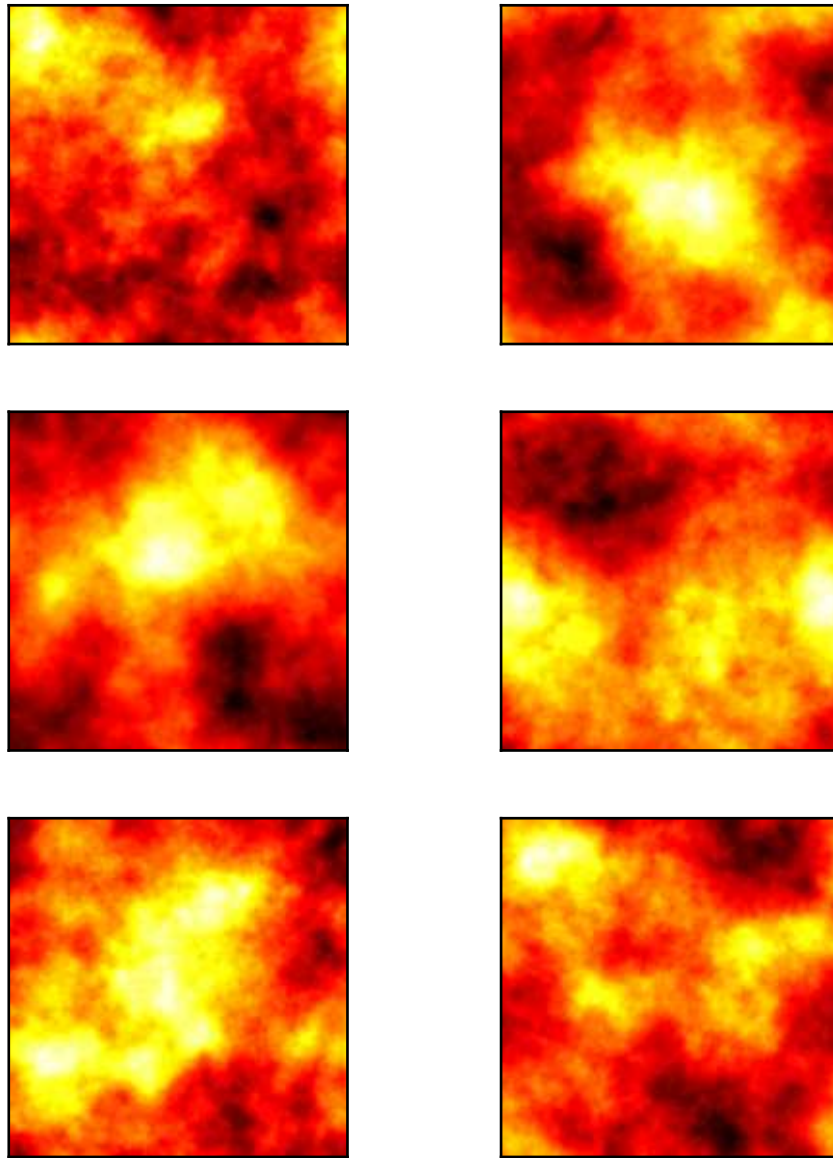


Figure 2.3: Simulations of Kolmogorov turbulence. Note the periodicity, which stems from the use of the Fourier-transformation in the simulation (based on (2.2)).

Layer	Height (m)	C_n^2 ($\text{m}^{-2/3}$)
1	47	0.5224
2	140	0.026
3	281	0.0444
4	562	0.116
5	1125	0.0989
6	2250	0.0295
7	4500	0.0596
8	9000	0.043
9	18000	0.06

Table 2.1: Layer parameters for an atmospheric model with 9 layers from the *European Southern Observatories*.

[2008]):

$$\Phi_n(f) := 0.033(2\pi)^{-2/3}C_n^2 \left(|f|^2 + \left(\frac{1}{L_0} \right)^2 \right)^{-11/6} \exp \left(-\frac{|f|^2}{|5.92/l_0|^2} \right), \quad (2.3)$$

where L_0 is the *outer length scale*, which is the scale of air motions that cause turbulences and l_0 is the *inner length scale*, at which turbulent motions experience vicious energy dissipation. The Kolmogorov model (2.2) is only valid in the range $1/L_0 \leq |f| \leq 1/l_0$, whereas the von Karman model (2.3) is defined for all $|f| \geq 0$.

See for example Roddier [1981] and Quirrenbach [2000] for more details on atmospheric turbulence and on the effects of an atmosphere to astronomical observations.

In this thesis we will work with a 9-layer atmospheric model developed by ESO. This model specifies altitudes and strengths, i.e. the amplitudes of the phase perturbations, for each layer. For closed loop simulations, which we neglect here, wind-parameters like speed and direction also need to be specified. Table 2.1 lists those settings for the 9-layer ESO model.

2.2 Mathematical Background

2.2.1 Notation

It is now time to formulate everything we have been talking about so far in a mathematical model. We will follow Ramlau and Rosensteiner [2011] step by step.

As explained in the previous section, we will assume that the atmospheric turbulences can be modelled by a discrete number of thin layers, the altitudes of which we denote

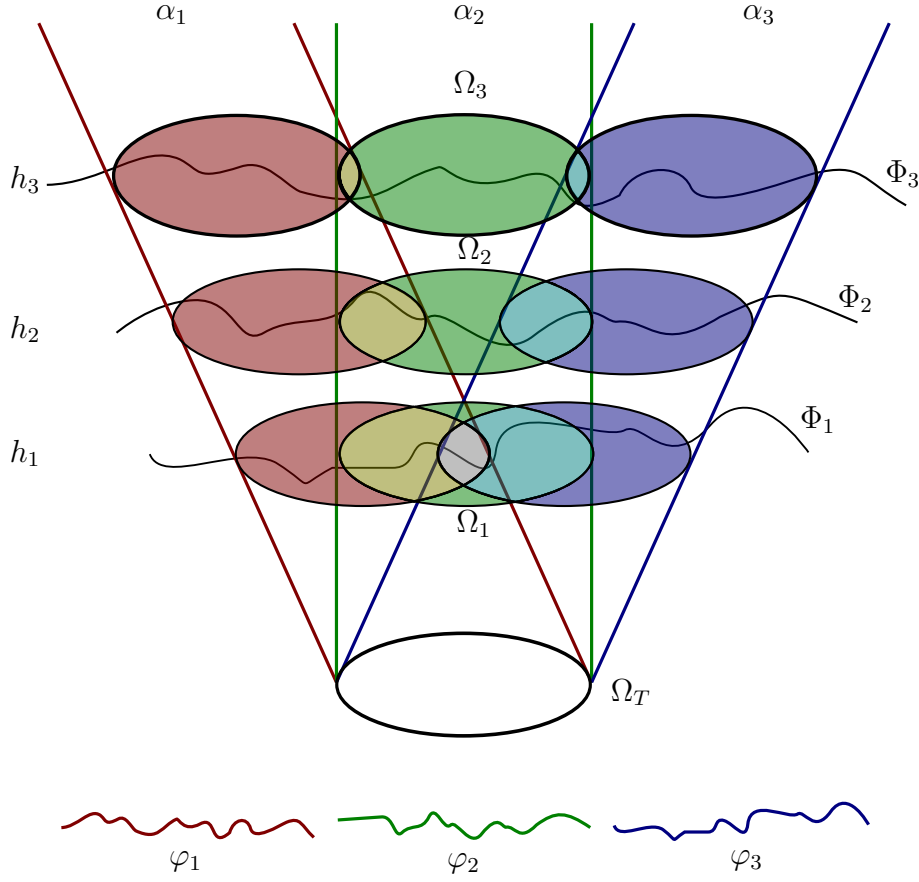


Figure 2.4: Mathematical notation for a model of atmospheric tomography.

by h_l , $l = 1, \dots, L$, for L layers. Furthermore, let G be the number of guide stars and denote their direction from the telescope's point of view by the unit vectors α_g , $g = 1, \dots, G$.

Defining Ω_T to be the telescope aperture with radius t :

$$\Omega_T := \{\mathbf{r} = (x, y) \in \mathbb{R}^2 : \|\mathbf{r}\| \leq t\},$$

we can specify the area seen by the g -th wavefront sensor in layer l as

$$\Omega_{l,g} := \Omega_T(h_l \alpha_g) := \{\mathbf{r} = (x, y) \in \mathbb{R}^2 : \mathbf{r} - h_l \alpha_g \in \Omega_T\}.$$

The total area of layer l covered by all wavefront sensors is then simply

$$\Omega_l := \bigcup_{g=1}^G \Omega_{l,g},$$

i.e. the union of shifted telescope apertures. Figure 2.5 shows these domains for the layers specified in the 9-layers atmospheric model from ESO.

We will use Φ_l , $l = 1, \dots, L$ to refer to the turbulence in the l -th layer. We will treat these turbulences as $L^2(\Omega_l)$ functions.

The (already reconstructed) wavefronts incident on the wavefront sensors are represented by φ_g for $g = 1, \dots, G$. We will assume that these φ_g are $L^2(\Omega_T)$ functions.

2.2.2 Problem formulation

Now that we have gathered enough notation, we can finally begin to formulate our problem: *Given measurements φ_g , $g = 1, \dots, G$, recover the turbulences Φ_l , $l = 1, \dots, L$.*

We start by observing that the connection between Φ_l and φ_g is given by (Roddier [1981])

$$\varphi_g(\mathbf{r}) = \sum_{l=1}^L \Phi_l(\mathbf{r} + h_l \boldsymbol{\alpha}_g), \quad \mathbf{r} \in \Omega_T, \quad (2.4)$$

which means that we sum up the perturbations of all layers in a given guide-star direction. Let us vectorize this as follows:

$$\begin{aligned} \boldsymbol{\Phi} &:= (\Phi_1, \dots, \Phi_L)^T \\ \langle \boldsymbol{\Phi}, \boldsymbol{\Psi} \rangle &:= \sum_{l=1}^L \langle \Phi_l, \Psi_l \rangle_{L^2(\Omega_l)} \\ (T^{\boldsymbol{\alpha}, h} \theta)(\mathbf{r}) &:= \theta(\mathbf{r} + h \boldsymbol{\alpha}) \text{ for } \mathbf{r} \in \Omega_T, \quad h \in \mathbb{R}, \quad \boldsymbol{\alpha} \in \mathbb{R}^2 \\ \mathbf{A}_g \boldsymbol{\Phi} &:= \sum_{l=1}^L T^{\boldsymbol{\alpha}_g, h_l} \Phi_l, \end{aligned} \quad (2.5)$$

so we have to solve the system of equations:

$$\mathbf{A}_g \boldsymbol{\Phi} = \varphi_g, \quad g = 1, \dots, G. \quad (2.6)$$

2.2.3 Solution Methods

There are a number of methods that could possibly be used to solve system (2.6), all with different strengths and advantages.

Matrix-Vector-Multiplication (MVM) is a direct method that works by pre-computing the generalized inverse of A . With respect to the ill-posedness, additional regularization terms can be added easily. However, even though in many cases the matrix A

will be sparse, its generalized inverse generally will not be, which leads to increased memory, bandwidth and computation power requirements. This method has for example been used in Tallon and Foy [1990] and Ragazzoni et al. [1999] to solve the tomography problem (2.6). Alternatively, the matrix resulting from (2.6) can be combined with the mirror-control matrix to skip tomography and generate deformable mirror commands directly (e.g. A. Tokovinin et al. [2001]). This approach has already been used in experimental AO systems (Marchetti et al. [2007]).

A popular iterative method is the (*preconditioned*) *conjugate gradient* (PCG) method. In adaptive optics, it has first been used for direct wavefront reconstruction in SCAO and MCAO, where good performance can be achieved using an appropriate preconditioning scheme (e.g. Thiébaud and Tallon [2010], Gilles et al. [2002] and Gilles et al. [2003]). PCG methods have also been extended to atmospheric tomography in MCAO (e.g. Yang et al. [2006], Gilles et al. [2007] and Gilles and Ellerbroek [2008]). However, the increased degree of ill-posedness in MCAO and atmospheric tomography can negatively affect the convergence speed of CG methods.

Another possible method for solving the atmospheric tomography problem is based on the Fourier-Transform, leading to a backprojection algorithm, see e.g. Tokovinin and Viard [2001] and Gavel [2004].

2.3 An Inverse Problem

2.3.1 An Iterative Solution

We will use a Kaczmarz iteration algorithm (Kaczmarz [1937]), which for a linear system $A\mathbf{x} = \mathbf{b}$ is defined as:

$$\mathbf{x}^{(k+1)} := \mathbf{x}^{(k)} + \frac{b_i - \langle a_i, \mathbf{x}^{(k)} \rangle}{\|a_i\|^2} a_i \quad i := 1 + (k \bmod m), \quad (2.7)$$

where a_i denotes the i -th row of the $m \times n$ matrix A and b_i is the i -th component of \mathbf{b} . The matrix A is assumed to be of full rank. The starting value $\mathbf{x}^{(0)}$ can be chosen arbitrarily.

As we can see from (2.7), this iteration uses only single “components” (in this case rows of the matrix) in each iteration, which makes this a computationally interesting approach, as it avoids large matrix-vector multiplications.

Kaczmarz iteration methods have a long history in image reconstruction and are particularly popular for tomography problems - see e.g. Eggermont et al. [1981], Natterer [2009], Haltmeier et al. [2007b], Haltmeier et al. [2007a] for different variants of the algorithm, convergence proofs, analyses of its regularization properties and many ap-

plications.

For the *projection* operators \mathbf{A}_g from (2.5), the algorithm takes the following shape (called *Landweber-Kaczmarz* iteration, see e.g. Haltmeier et al. [2007b] or Eggermont et al. [1981]):

$$\Phi^{(k+1)} := \Phi^{(k)} - \beta_g \mathbf{A}_g^* (\mathbf{A}_g \Phi^{(k)} - \varphi_g), \quad g := 1 + (k \bmod G), \quad (2.8)$$

where β_g is a relaxation parameter and \mathbf{A}_g^* denotes the adjoint operator.

A convergence result for the system (2.6) using the method (2.8) can be found in Ramlau and Rosensteiner [2011]:

Theorem 2.1. *Assume that (2.8) is consistent, i.e. admits a solution for exact data. If the scaling parameters β_g , $g = 1, \dots, G$, are chosen such that*

$$\beta_g \leq \frac{1}{\|\mathbf{A}_g\|}$$

holds, then the iterates of the Landweber-Kaczmarz method converge to a solution,

$$\Phi^{(k)} \rightarrow \Phi^\dagger \text{ for } i \rightarrow \infty,$$

with $\mathbf{A}_g \Phi^\dagger = \varphi_g$, $g = 1, \dots, G$, and Φ^\dagger is among all solutions the one that has a minimal distance to the initial iterate Φ^0 .

In order to apply the algorithm (2.8), we need the adjoint operator \mathbf{A}^* . This depends on the choice of function spaces that we let \mathbf{A} operate on. We will consider two cases:

$$\mathbf{A}_g : \bigotimes_{l=1}^L L^2(\Omega_l) \rightarrow L^2(\Omega_T), \quad g = 1 \dots G \quad (2.9)$$

$$\tilde{\mathbf{A}}_g : \bigotimes_{l=1}^L H^1(\Omega_l) \rightarrow L^2(\Omega_T), \quad g = 1 \dots G, \quad (2.10)$$

where H^1 is the Sobolev-Space of order one. (See Section 3.2.2 for more details.) The motivation for using the Sobolev space H^1 in (2.10) lies in the atmospheric turbulence model (2.3). The decay in this energy spectrum function indicates a smoothness in the turbulences that is ignored by working only on the space of L^2 functions as in (2.9).

In case of (2.9), Ramlau and Rosensteiner [2011] show that the adjoint operator \mathbf{A}_g^* is given by

$$\mathbf{A}_g^* \Psi = ((T^{g,1})^* \Psi, (T^{g,2})^* \Psi, \dots, (T^{g,L})^* \Psi), \quad (2.11)$$

$$\text{where } (T^{g,l})^* \Psi(\mathbf{r}) := \Psi(\mathbf{r} - \boldsymbol{\alpha}_g h_l) \chi_{\Omega_l, g}(\mathbf{r}). \quad (2.12)$$

This means that the adjoint is just a shift of the function defined on the telescope aperture towards the given guide star g at a specific layer's altitude.

Now denote by i_l the Sobolev embedding

$$i_l : H^1(\Omega_l) \rightarrow L^2(\Omega_l) \quad (2.13)$$

We then get $\tilde{\mathbf{A}}_g = \mathbf{A}_g \mathbf{i}$, and so $\tilde{\mathbf{A}}_g^* = \mathbf{i}^* \mathbf{A}_g^*$, where $\mathbf{i} := (i_1, \dots, i_L)$. This means that for the adjoint of the H^1 operator (2.10) we need to calculate the adjoints of the Sobolev embedding operators i_l , $l = 1, \dots, L$.

Ramlau and Teschke [2004a] and Ramlau and Teschke [2004b] give a thorough analysis of the adjoint Sobolev embedding operator, methods to implement it and example applications.

Using Green's formula, it is easy to show (Ramlau and Rosensteiner [2011]):

Theorem 2.2. *The adjoint of the embedding operator $i_l : H^1(\Omega_l) \rightarrow L^2(\Omega_l)$, $l = 1, \dots, L$, is given as $i_l^* v = u$ where*

$$\begin{aligned} u - \sigma \Delta u &= v \text{ in } \Omega_l \text{ (in the sense of distributions)} \\ \frac{\partial u}{\partial \mathbf{n}} &= 0 \text{ on } \partial\Omega_l \text{ (on an appropriate trace),} \end{aligned} \quad (2.14)$$

where $\sigma > 0$ and $\frac{\partial u}{\partial \mathbf{n}}$ is the conormal derivative. The constant σ can be used as a parameter to control the strength of the smoothing. (A high value of σ can intuitively be seen as ‘‘punishing’’ variance in the solution.)

Algorithm 1 now summarizes the findings of this chapter.

Algorithm 1 H^1 space Landweber-Kaczmarz iteration. (Extended from Ramlau and Rosensteiner [2011]).

Require: φ_g , $g = 1, \dots, G$	# Wavefront measurements
Require: Φ_0	# Starting value
for $i = 1, \dots$ do	
$\Phi_{i,0} \leftarrow \Phi_{i-1}$	
for $g = 1, \dots, G$ do	
$r \leftarrow \varphi_g - \mathbf{A}_g \Phi_{i,g-1}$	# Residual
$a \leftarrow \mathbf{A}_g^* r$	# Application of L^2 adjoint operator (2.11)
$s \leftarrow i^* a$	# Smoothing via equation (2.14)
$\Phi_{i,g} \leftarrow \Phi_{i,g-1} + \beta_g s$	
end for	
end for	

2.3.2 Solving the adjoint embedding

In Chapter 4, we will investigate the results of adding a smoothing step based on equation (2.14) to the Landweber-Kaczmarz iteration in an MCAO simulation based on real-world test data.

Before we get to that however, we will investigate an interesting theoretical aspect of (2.14):

We know from partial differential equation (PDE) theory, that solutions to elliptic PDEs are much smoother than their data (i.e. the right hand sides). This however is highly dependent on the shape and smoothness of the domain and particularly its boundary. In our specific case, we will see that the presence of non-convex corners (i.e. corners with interior angle $> \pi$) at the intersections between telescope apertures (i.e. the areas covered by the wavefront sensors) will play an important role in determining a solution's smoothness.

In the next chapter, we will take a closer look at this and derive some results for the smoothness of solutions of (2.14).

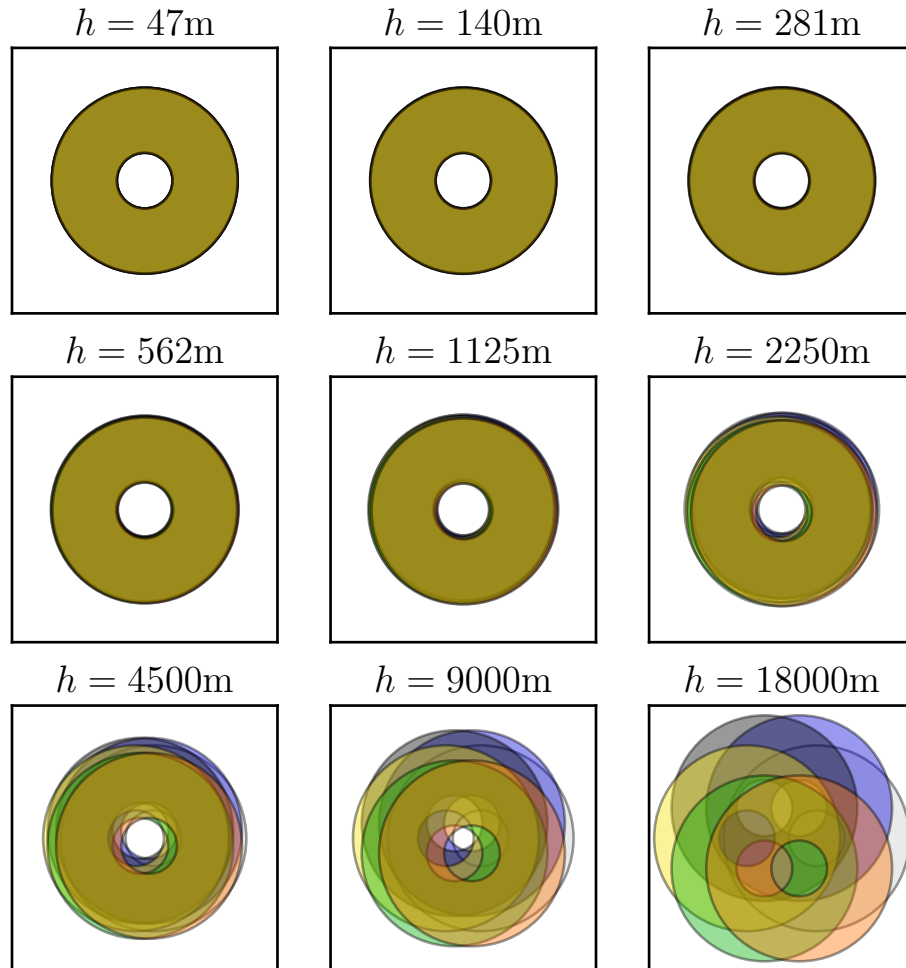


Figure 2.5: Overlapping apertures for 6 guide stars distributed in a 3 arc minutes field of view. The inner circular hole stems from a central obstruction in the telescope's optical pathway. Even on the lowest layer at $h = 47$ meters, the apertures are not precisely overlapping.

Chapter 3

Regularity in Non-Convex Domains

Where we investigate an elliptic partial differential equation on a non-convex domain with singularities on the boundary and find out that the solution's smoothness depends on the interior angles of the domain.

3.1 Regularity of Elliptic Partial Differential Equations

In this chapter, we will take a closer look at the elliptic PDE we have derived in Theorem 2.2:

$$\begin{aligned} u - \sigma \Delta u &= v \text{ in } \Omega \text{ (in the sense of distributions)} \\ \frac{\partial u}{\partial \mathbf{n}} &= 0 \text{ on } \partial\Omega \text{ (on an appropriate trace),} \end{aligned} \tag{3.1}$$

where the constant $\sigma > 0$ and Ω is the union of a finite number of overlapping rings from Chapter 2 (recall Figure 2.5).

This set Ω is non-convex, which affects the smoothness of solutions of (3.1). We will ignore cases where boundaries of the telescope apertures touch tangentially, i.e. in cusps, so we restrict ourselves to sets Ω that have a Lipschitz boundary.

A geometrically motivated approach to the problem of regularity with singularities on the boundary is presented in Grisvard [1992]. It works by analysing the space of solutions and characterizing the behaviour of functions near non-convex corners, which are precisely the cause of singularities in the solutions. This results in a constructive and insightful way to prove regularity for elliptic partial differential equations, which is why this chapter is devoted to applying it to the partial differential equation we have derived in the previous chapter. Grisvard [1992] notes such possible extensions of his proofs.

Note that throughout the entire chapter L will always refer to the operator

$$Lu := u - \sigma \Delta u, \quad (3.2)$$

for $\sigma > 0$.

3.1.1 Structure of this Chapter

This chapter is heavily based on Grisvard [1992], which derives regularity results for the Poisson problem on simply-connected bounded non-convex polygonal domains in \mathbb{R}^2 .

Simply connected means that each closed curve in Ω can be continuously deformed to a point. This implies that two-dimensional simply connected domains cannot have holes. This excludes the overlapping-rings domain. Luckily though, it is very easy to generalize the results for non-simply-connected domains as Grisvard also notes.

In Section 3.3.2, we investigate further generalizations of the regularity results to more general domains than polygons.

Another restriction is that the results which are relevant to us, are derived specifically for the Laplace operator Δ . In Sections 3.3.1 and 3.5 we will extend this to the operator L .

We will build on Grisvard's work in two ways. In Section 3.3 we will first use a homogenization ansatz to apply the original Laplace operator regularity results to the operator L . In a second step we will use a geometric limit approach to formulate a conjecture concerning the regularity of solutions of (3.1) for a domain of overlapping circles (no holes yet).

In Section 3.5 we shall derive a more precise result by adapting the proof-techniques from Grisvard [1992] to the operator L on a multiply connected domain. In some cases this will go hand in hand with some simplifications (as we do not have to deal with Dirichlet or mixed boundary conditions). A few lemmas apply without any modifications at all. However, eventually we will encounter more sophisticated structures when working with the operator L compared to the Laplace operator.

3.2 Prerequisites

We first need to introduce some notation as well as definitions and theorems we will use later.

3.2.1 Geometry and Notation

Throughout this thesis by *domain* we mean an open, connected and bounded subset of \mathbb{R}^2 . A *Lipschitz domain* is a domain with a Lipschitz boundary, which means that the boundary can locally be represented as the graph of a Lipschitz continuous function (in a suitable coordinate system). See McLean [2000] for a detailed discussion of such domains.

Definition 3.1.

- (a) Let $\Omega \subset \mathbb{R}^2$ be a Lipschitz polygon, i.e. a Lipschitz domain whose boundary is the union of finitely many straight line segments. Let $\{\Gamma_i\}_{i=1}^N$ denote the connected parts of $\partial\Omega$. An intersection point between two line segments is called a corner. Let $\{P_{(i,j)}\}_{j=1}^{M_i}$ denote the corners of Γ_i and let $w_{(i,j)}$ denote the interior angle at $P_{(i,j)}$, see Figure 3.1.
- (b) If $\Omega \subset \mathbb{R}^2$ is a Lipschitz domain with a piecewise C^1 boundary that splits into a finite number of connected parts (that are simple closed curves), the notion of corners and interior angles above can be extended straightforwardly (by using the two tangents).

Figure 3.1 shows an example of such domains with the specified notation.

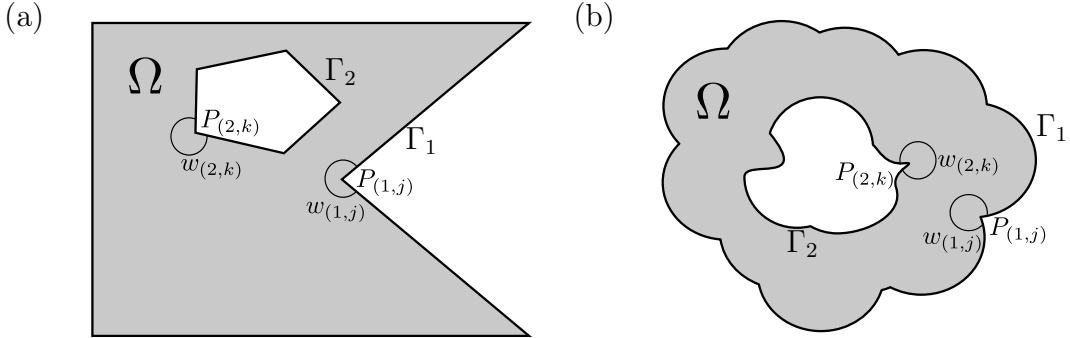


Figure 3.1: Corners and interior angles of (a) a Lipschitz Polygon and (b) a Lipschitz domain with a piecewise C^1 boundary.

In addition to that, we will need a truncation function near every corner. We define these as follows.

Definition 3.2. Let Ω be as in Definition 3.1. For every corner $P_{(i,j)}$ we fix a truncation function $\eta_{(i,j)} \in C^\infty(\overline{\Omega})$ such that

1. $\eta_{(i,j)}$ depends only on the distance $r_{(i,j)}$ to the corner $P_{(i,j)}$
2. $\eta_{(i,j)} \equiv 1$ in a neighbourhood of the corner $P_{(i,j)}$
3. $\eta_{(i,j)}$ vanishes near all other corners and near all edges of the domain that are not adjacent to the corner
4. All functions $\eta_{(i,j)}$ have disjoint support

Note that we do not require the functions $\eta_{(i,j)}$ to form a partition of unity of the domain.

3.2.2 Sobolev Spaces

Here we will quickly list and define all the spaces and theorems we will later need. For proofs or more general introductions into this subject, we refer to the excellent textbooks Adams and Fournier [2003] and McLean [2000].

Throughout this subsection, let $\Omega \subset \mathbb{R}^2$ be a Lipschitz domain. Let $L^2(\Omega)$ denote the Hilbert space of square integrable functions with the corresponding inner product and norm

$$\begin{aligned} \langle u, v \rangle_{L^2(\Omega)} &:= \int_{\Omega} u(x)v(x) \, dx & u, v \in L^2(\Omega) \\ \|u\|_{L^2(\Omega)} &:= \sqrt{\int_{\Omega} |u(x)|^2 \, dx} & u \in L^2(\Omega). \end{aligned}$$

For an arbitrary multi-index α and a function $u \in L^2(\Omega)$ the *weak partial derivative* $\partial^\alpha u$ is defined as a locally integrable function g_α such that:

$$\langle g_\alpha, \varphi \rangle_{L^2(\Omega)} = (-1)^{|\alpha|} \langle u, \partial^\alpha \varphi \rangle_{L^2(\Omega)} \quad \text{for } \varphi \in C_c^\infty(\Omega), \quad (3.3)$$

where $C_c^\infty(\Omega)$ is the space of smooth functions with compact support in Ω .

Definition 3.3. We denote by $H^n(\Omega)$ the Sobolev space of order $n \in \mathbb{N}$, defined as

$$H^n(\Omega) := \{u \in L^2(\Omega) : \partial^\alpha u \in L^2(\Omega) \text{ for } |\alpha| \leq n\}$$

where $\partial^\alpha u$ is the weak partial derivative as defined in equation (3.3).

It is a Hilbert space with the inner product:

$$\langle u, v \rangle_{H^n(\Omega)} := \sum_{|\alpha| \leq n} \int_{\Omega} \partial^\alpha u(x) \partial^\alpha v(x) \, dx,$$

and the corresponding norm

$$\|u\|_{H^n(\Omega)} := \sqrt{\sum_{|\alpha| \leq n} \int_{\Omega} |\partial^\alpha u(x)|^2 dx}.$$

Definition 3.4. For $s = n + \mu$ with $n \in \mathbb{N}$ and $\mu \in (0, 1)$, we denote by $H^s(\Omega)$ the (fractional) Sobolev space of order s , defined as

$$H^s(\Omega) := \{u \in H^n(\Omega) : |\partial^\alpha u|_{\mu, 2, \Omega} < \infty \text{ for } |\alpha| = n\}$$

where $\partial^\alpha u$ is the weak partial derivative as defined in equation (3.3) and $|u|_{\mu, p, \Omega}$ is the Slobodeckii semi-norm

$$|u|_{\mu, 2, \Omega} := \int_{\Omega} \int_{\Omega} \frac{|u(x) - u(y)|^2}{\|x - y\|^{2+2\mu}} dx dy.$$

(In case of space dimension n , the denominator needs to be replaced by $\|x - y\|^{n+2\mu}$.)

It is a Hilbert space with the inner product:

$$\langle u, v \rangle_{H^s(\Omega)} := \langle u, v \rangle_{H^n(\Omega)} + \sum_{|\alpha|=n} \int_{\Omega} \int_{\Omega} \frac{[\partial^\alpha u(x) - \partial^\alpha u(y)] [\partial^\alpha v(x) - \partial^\alpha v(y)]}{\|x - y\|^{2+2\mu}} dx dy,$$

and the corresponding norm

$$\|u\|_{H^s(\Omega)} := \sqrt{\|u\|_{H^n(\Omega)}^2 + \sum_{|\alpha|=n} |\partial^\alpha u(x)|_{\mu, 2, \Omega}^2 dx}.$$

See e.g. McLean [2000] for more details on Sobolev Spaces.

In order to work with boundary conditions, we need a precise definition of the restriction of Sobolev functions to the boundaries of domains. Without delving into too much detail, we will use the following definition:

Definition 3.5. Let Ω be a bounded open subset of \mathbb{R}^2 with a Lipschitz boundary $\partial\Omega$. Then the trace operator γ is defined as

$$\gamma : C^\infty(\bar{\Omega}) \rightarrow C^\infty(\partial\Omega) : \gamma u = u|_{\partial\Omega}. \quad (3.4)$$

It can be shown (see e.g. McLean [2000], Theorems 3.37 and 3.38) that there exists a unique extension of γ to a bounded linear operator

$$\gamma : H^s(\Omega) \rightarrow H^{s-\frac{1}{2}}(\partial\Omega)$$

for $s \in (1/2, 3/2)$.

Definition 3.6. Assume that Ω is a Lipschitz domain in \mathbb{R}^2 and $u \in H^2(\Omega)$. Then the conormal derivative $\frac{\partial}{\partial \mathbf{n}}u$ is defined as

$$\frac{\partial}{\partial \mathbf{n}}u = \gamma(\nabla u) \cdot \mathbf{n}, \quad (3.5)$$

where \mathbf{n} is the (unit, outward) normal vector of $\partial\Omega$.

This definition can be extended to functions $u \in H^1(\Omega)$ with $\Delta u \in L^2(\Omega)$ (see McLean [2000], Lemma 4.3).

3.2.3 Variational Formulation

We now have the necessary basics to properly formulate the partial differential equation (3.1).

Assume that Ω be a Lipschitz domain in \mathbb{R}^2 . Given $f \in L^2(\Omega)$ and $\sigma > 0$, find $u \in H^1(\Omega)$ such that

$$\begin{aligned} u - \sigma \Delta u &= f && \text{in } \Omega \\ \frac{\partial v}{\partial \mathbf{n}} &= 0 && \text{on } \partial\Omega. \end{aligned} \quad (3.6)$$

This corresponds to the following variational formulation:

$$a(u, v) := \int_{\Omega} uv + \sigma \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx, \quad \forall v \in H^1(\Omega). \quad (3.7)$$

It is easy to show boundedness and coerciveness of the bilinear form on the left side of (3.7):

$$\begin{aligned} |a(u, v)| &\leq \max\{1, \sigma\} \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} \\ a(u, v) &\geq \min\{1, \sigma\} \|u\|_{H^1(\Omega)}^2, \end{aligned}$$

so from the lemma of Lax-Milgram we know there exists a unique solution $u \in H^1(\Omega)$ of (3.7) such that

$$\|u\|_{H^1(\Omega)} \leq \frac{1}{\min\{1, \sigma\}} \|f\|_{L^2(\Omega)}. \quad (3.8)$$

In the interior and up to the boundary, excluding a neighbourhood of the corners, we know (see e.g. McLean [2000] and Grisvard [1992]) that the solution u is actually smoother than that, i.e. if φ is the restriction of a smooth function to Ω with no support near the corners, then $u\varphi \in H^2(\Omega)$, and for $\varphi \in C_c^\infty(\Omega)$ we have $u\varphi \in C_c^\infty(\Omega)$.

But: What do we know about the global smoothness of the solution, i.e. including the corners? *Is there an order $s > 1$ such that $u \in H^s(\Omega)$?*

3.3 A Homogenization Based Result

This section is based on the main regularity result in Grisvard [1992] (Corollary 2.4.4), which can be formulated for our purposes as follows:

Theorem 3.7. *Let Ω be a simply connected Lipschitz polygon in \mathbb{R}^2 , as in Definition 3.1 (a). Assume $u \in H^1(\Omega)$ such that*

$$\begin{aligned}\Delta u &= f \text{ in } \Omega \\ \frac{\partial u}{\partial \mathbf{n}} &= 0 \text{ on } \partial\Omega,\end{aligned}$$

for $f \in L^2(\Omega)$. Then $u \in H^s(\Omega)$ for $s \leq 2$ with $s < 1 + \pi / \max_{(i,j)} \omega_{(i,j)}$, where $\omega_{(i,j)}$ denotes the interior angles of corners.

This theorem tells us that the smoothness of the solution depends directly on the angles in corners. As non-convex angles get closer to π (i.e. convexity), the smoothness of the solutions of Theorem 3.7 approaches $H^2(\Omega)$. Since we do not consider cuts or cusps in the domain, we will always have $\pi / \max_{(i,j)} \omega_{(i,j)} > 1/2$.

Note that Theorem 3.7 is restricted to simply connected domains. In Section 3.5 we will generalize this to non-simply connected domains.

3.3.1 Homogenization

In order to switch from the Laplace operator Δ to the operator L , we will use a homogenization approach:

Given the assumptions in Theorem 3.7, we know from the variational formulation (3.7) that there exists a unique function $u \in H^1(\Omega)$ such that

$$\begin{aligned}u - \sigma \Delta u &= f && \text{in } \Omega, \\ \frac{\partial u}{\partial \mathbf{n}} &= 0 && \text{on } \partial\Omega\end{aligned}$$

for $\sigma > 0$. Defining $g := \sigma^{-1}(u - f)$, this function u also fulfils

$$\begin{aligned}\Delta u &= g && \text{in } \Omega \\ \frac{\partial u}{\partial \mathbf{n}} &= 0 && \text{on } \partial\Omega.\end{aligned}$$

Since $g \in L^2(\Omega)$, we can now apply Theorem 3.7 to find that $u \in H^s(\Omega)$, with s depending on the interior angles $\omega_{(i,j)}$.

3.3.2 A Limit Approach

We would like to generalize this result from a polygonal domain to a domain of overlapping circles. To do so, we work similarly to Grisvard [1985] (Theorem 3.2.1.3) and approximate a more general domain with piecewise C^1 boundaries by a polygonal domain from the outside. By choosing a series of special approximations of the general domain, we attempt to find a sequence of solutions of the PDE (3.6) defined on those approximations. We conjecture that this sequence of solutions is uniformly bounded in H^s , which implies that a weakly convergent subsequence exists. It is then relatively easy to show that this weak limit solves (3.6) on the general domain.

This approach leads to the following conjecture.

Conjecture 3.8. *Let Ω be as in Definition 3.1 (b) and additionally simply connected. Let f be an L^2 function on Ω and assume that $u \in H^1(\Omega)$ is the corresponding solution of (3.1).*

Then $u \in H^s(\Omega)$ with $1 < s \leq 2$ and $s < 1 + \pi / \max_{(i,j)} \omega_{(i,j)}$, where $\omega_{(i,j)}$ denotes the interior angles of corners of the domain (see Definition 3.1).

We give a sketch of a possible proof and will comment on the open issues.

We approximate Ω by polygonal domains from the outside: Let $(\Omega_n)_{n \in \mathbb{N}}$ be a series of bounded polygonal open subsets of \mathbb{R}^2 such that

$$\begin{aligned} \Omega &\subseteq \Omega_n \quad \forall n \in \mathbb{N} \\ \Omega_n &\subseteq \Omega_m \quad \forall m, n : m < n, \end{aligned}$$

and $\partial\Omega_n$ contains all non-convex corners of Ω . Since $\partial\Omega$ is Lipschitz, we can rule out cusps and choose our approximations Ω_n such that their interior angles in the corners are equal to the corresponding angles in $\partial\Omega$. In addition to that we require that the approximations are fine enough so that all non-convex interior angles in corners of Ω_n

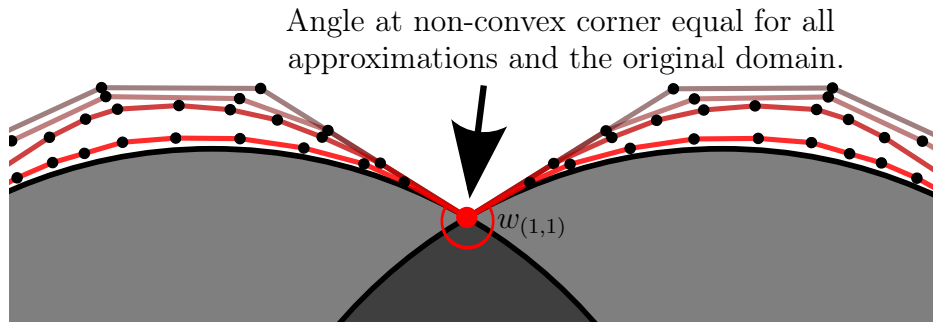


Figure 3.2: Example showing the construction of domain approximations as formulated in Conjecture 3.8.

are smaller than the largest non-convex corner in Ω . See Figure 3.2 for an example of such a construction.

Denote by \tilde{f} the continuation of f by zero in Ω^c , the complement of Ω in \mathbb{R}^2 . Then we know from Theorem 3.7 and Section 3.3.1 that for every n there exists a solution $u_n \in H^s(\Omega_n)$ to

$$\begin{aligned} Lu_n &= \tilde{f} \text{ in } \Omega_n \\ \frac{\partial u_n}{\partial \mathbf{n}} &= 0 \text{ on } \partial\Omega_n. \end{aligned}$$

We conjecture now, that these functions u_n form a bounded series in $H^s(\Omega)$. We sketch a possible line of proof as follows: We know from closer inspection of Grisvard [1992] (or Section 3.5) that the solutions u_n have the form

$$u_n = w_n + S_n, \quad (3.9)$$

where $w_n \in H^2(\Omega)$ can be considered the smooth solution and $S_n \in H^s(\Omega)$ a singular component with $LS_n \in L^2(\Omega)$. From (3.9) we find that w_n is the solution of

$$\int_{\Omega_n} w_n v + \sigma \nabla w_n \cdot \nabla v \, dx = \int_{\Omega_n} f v \, dx - \left\langle \frac{\partial S_n}{\partial \mathbf{n}}, v \right\rangle_{\partial\Omega_n} - \int_{\Omega_n} LS_n v \, dx, \quad \forall v \in H^1(\Omega), \quad (3.10)$$

where the term $\frac{\partial S_n}{\partial \mathbf{n}}$ vanishes as will be seen in Section 3.5.2.

We now formulate three subconjectures:

1. There exists a constant $C_1 > 0$ that is independent of n , such that

$$\|u_n\|_{H^s(\Omega_n)} \leq C_1 \|w_n\|_{H^2(\Omega_n)} + \|S_n\|_{H^s(\Omega_n)}. \quad (3.11)$$

2. There exists a constant $C_2 > 0$ that is independent of n , such that

$$\|w_n\|_{H^2(\Omega_n)} \leq C_2 (\|f - LS_n\|_{L^2(\Omega_n)} + \|w_n\|_{L^2(\Omega_n)}). \quad (3.12)$$

3. There exist constants $C_3 > 0$ and $C_4 > 0$ independent of n , such that

$$\|S_n\|_{H^s(\Omega_n)} \leq C_3 \|f\|_{L^2(\Omega_n)} \quad (3.13)$$

$$\|LS_n\|_{L^2(\Omega_n)} \leq C_4 \|f\|_{L^2(\Omega_n)}. \quad (3.14)$$

The first two conjectures are motivated by the simple and explicit constructions of the domains Ω_n as polygons with known and bounded angles. The third conjecture is motivated both by the geometry of the domains Ω_n as well as the precise knowledge of the form of the singular components S_n and the fact that these S_n depend only on the angles of Ω_n and on the source function f (see Section 3.5). In particular, in

(3.12) is motivated by Theorem 4.3.1.4 and Remark 4.3.1.5 in Grisvard [1985], where Grisvard indicates the possibility of finding an explicit expression for the constant C_2 as a function of the domain.

Using the conjectures (3.11) to (3.14) and the Lax-Milgram bound (3.8), one can now bound $\|u_n\|_{H^s(\Omega)}$ by

$$\begin{aligned} \|u_n\|_{H^s(\Omega)} &\leq C_1 \|w_n\|_{H^2(\Omega_n)} + \|S_n\|_{H^s(\Omega_n)} \\ &\leq C_1 \left(C_2 + \frac{1}{\min\{1, \sigma\}} \right) \|f - LS_n\|_{L^2(\Omega_n)} + \|S_n\|_{H^s(\Omega_n)} \\ &\leq C \|f\|_{L^2(\Omega_n)}, \end{aligned} \quad (3.15)$$

with C independent of n .

The bound (3.15) implies the existence of a function $u_\infty \in H^s(\Omega)$ that is the weak limit of a subsequence of (u_n) . From Section 3.2.2 we now recall that the embedding from $H^s(\Omega)$ into $H^1(\Omega)$ is compact, so that the weak convergence $u_n \rightharpoonup u_\infty$ in $H^s(\Omega)$ implies

$$\lim_{n \rightarrow \infty} u_n = u_\infty \quad \text{in } H^1(\Omega). \quad (3.16)$$

Using this, we will now prove that the limit function u_∞ is a solution of

$$\begin{aligned} Lu_\infty &= f \text{ in } \Omega \\ \frac{\partial u_\infty}{\partial \mathbf{n}} &= 0 \text{ on } \partial\Omega, \end{aligned}$$

by showing that for all $v \in H^1(\Omega_1)$

$$\int_{\Omega} f v \, dx = \int_{\Omega_n} u_n v + \sigma \nabla u_n \cdot \nabla v \, dx \xrightarrow{n \rightarrow \infty} \int_{\Omega} u_\infty v + \sigma \nabla u_\infty \cdot \nabla v \, dx. \quad (3.17)$$

The equality in (3.17) follows from the definition of \tilde{f} and from the variational formulation (3.7). The limit in (3.17) needs a closer inspection. Let $v \in H^1(\Omega_1)$ be an arbitrary but fixed test function. Firstly,

$$\begin{aligned} &\int_{\Omega} u_\infty v + \sigma \nabla u_\infty \cdot \nabla v \, dx - \int_{\Omega_n} u_n v + \sigma \nabla u_n \cdot \nabla v \, dx \\ &= \int_{\Omega} (u_\infty - u_n) v + \sigma \nabla (u_\infty - u_n) \cdot \nabla v \, dx - \int_{\Omega_n \setminus \Omega} u_n v + \sigma \nabla u_n \cdot \nabla v \, dx. \end{aligned} \quad (3.18)$$

For the first integral in (3.18), we easily see that

$$\begin{aligned} \int_{\Omega} (u_\infty - u_n) v + \sigma \nabla (u_\infty - u_n) \cdot \nabla v \, dx &\leq \max\{1, \sigma\} \langle u_\infty - u_n, v \rangle_{H^1(\Omega)} \\ &\leq \max\{1, \sigma\} \|u_n - u_\infty\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}, \end{aligned}$$

and using (3.16), we get

$$\lim_{n \rightarrow \infty} \int_{\Omega} (u_{\infty} - u_n)v + \sigma \nabla(u_{\infty} - u_n) \cdot \nabla v \, dx = 0. \quad (3.19)$$

For the second integral in (3.18), we proceed similarly:

$$\begin{aligned} \int_{\Omega_n \setminus \bar{\Omega}} u_n v + \sigma \nabla u_n \cdot \nabla v \, dx &\leq \max\{1, \sigma\} \langle u_n, v \rangle_{H^1(\Omega_n \setminus \bar{\Omega})} \\ &\leq \max\{1, \sigma\} \|u_n\|_{H^1(\Omega_n \setminus \bar{\Omega})} \|v\|_{H^1(\Omega_n \setminus \bar{\Omega})} \\ &\leq \max\{1, \sigma\} \|u_n\|_{H^1(\Omega_n)} \|v\|_{H^1(\Omega_n \setminus \bar{\Omega})} \\ &\leq \frac{\max\{1, \sigma\}}{\min\{1, \sigma\}} \|f\|_{L^2(\Omega)} \|v\|_{H^1(\Omega_n \setminus \bar{\Omega})}, \end{aligned}$$

where in the last step we used the inequality (3.8). From the construction of the approximations Ω_n , it follows that

$$\int_{\Omega_n \setminus \bar{\Omega}} u_n v + \sigma \nabla u_n \cdot \nabla v \, dx \xrightarrow{n \rightarrow \infty} 0. \quad (3.20)$$

Combining (3.18), (3.19) and (3.20) shows that

$$\int_{\Omega_n} u_n v + \sigma \nabla u_n \cdot \nabla v \, dx \xrightarrow{n \rightarrow \infty} \int_{\Omega} uv + \sigma \nabla u \cdot \nabla v \, dx,$$

which finally proves (3.17). This would complete the proof of Conjecture 3.8, if we had rigorous bounds (3.11), (3.12), (3.13) and (3.14).

Remark 3.9. *The proof strategy of Conjecture 3.8 can be applied to non-simply connected domains as well. The only difference is that a corresponding replacement for Theorem 3.7 is needed. We will derive such a replacement explicitly in Section 3.5, Theorem 3.19. This then leads to the following conjecture.*

Conjecture 3.10. *Let Ω be as in Definition 3.1 (b), but not necessarily simply connected. Let f be an L^2 function on Ω and assume that $u \in H^1(\Omega)$ is the corresponding solution of (3.1).*

Then $u \in H^s(\Omega)$ with $1 < s \leq 2$ and $s < 1 + \pi / \max_{(i,j)} \omega_{(i,j)}$, where $\omega_{(i,j)}$ denotes the interior angles of corners of the domain (see Definition 3.1).

3.3.3 Interpretation of the Smoothness Results

As we have seen, the regularity of the solution of the PDE (3.1) on non-convex domains depends on the largest interior angle. With regards to the domains in the adaptive optics scenario (ignoring the central obstruction for now), one can quickly derive equations for the non-convex angle at the intersection of two overlapping circles and then

use that to get an a-priori estimate of the regularity of (3.1) on that domain.

Figure 3.3 shows the smoothness for a domain of two equally sized circles intersecting at different distances. When the two circles exactly coincide (for $d = 0$), we have a smooth convex domain with the usual $H^2(\Omega)$ solution. As the circles move apart, the smoothness declines, approaching $H^{3/2}(\Omega)$ in the limit case where two circles touch in a single point.

We still need to generalise this result to domains with holes. Also note that the constant σ does not appear in the regularity result. However, as we will see in Section 3.5, a more detailed approach to proving the regularity of solutions of the partial differential equation (3.1) can shed some light on the role of this parameter.

3.4 Modified Bessel Functions of the First and Second Kind

In the next section, we will show that the behaviour of solutions of (3.1) in a neighbourhood of the corners can be expressed in terms of the *modified Bessel functions of the first and second kind*. Here we will formulate some of their properties, which we will need in the next section.

Consider the *modified Bessel equation*

$$r^2 u''(r) + ru'(r) - (\sigma^2 r^2 + \lambda^2)u(r) = 0, \text{ for } r > 0 \quad (3.21)$$

(we would get the *original* Bessel equation by writing $-\lambda^2$ instead of $+\lambda^2$ in (3.21)).

Two linearly independent solutions of (3.21) are I_λ and K_λ , the modified Bessel functions of the first and second kind, respectively. The general solution of (3.21) then has the form (see e.g. 17.6.1.1 in Jeffrey and Dai [2008]):

$$u(r) = \alpha I_\lambda(\sigma r) + \beta K_\lambda(\sigma r), \quad (3.22)$$

The behaviour of these smooth functions close to zero for different values of λ is depicted in Figures 3.4 and 3.5. We see that for $\lambda \geq 0$ and $r \rightarrow 0$ $I_\lambda(r)$ converges, whereas $K_\lambda(r)$ diverges for $r \rightarrow 0$.

Next, let us investigate some properties of these functions, starting with I_λ .

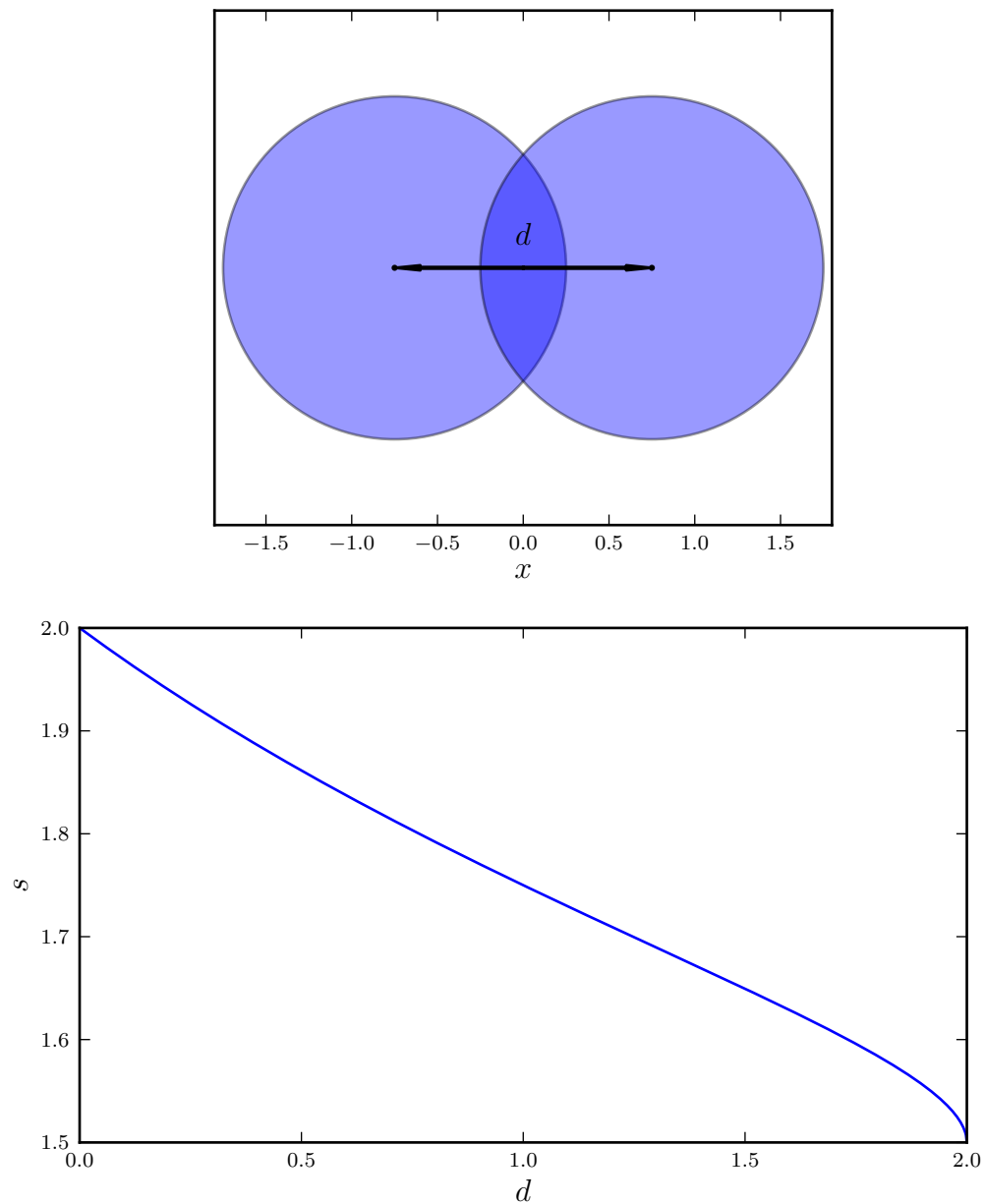
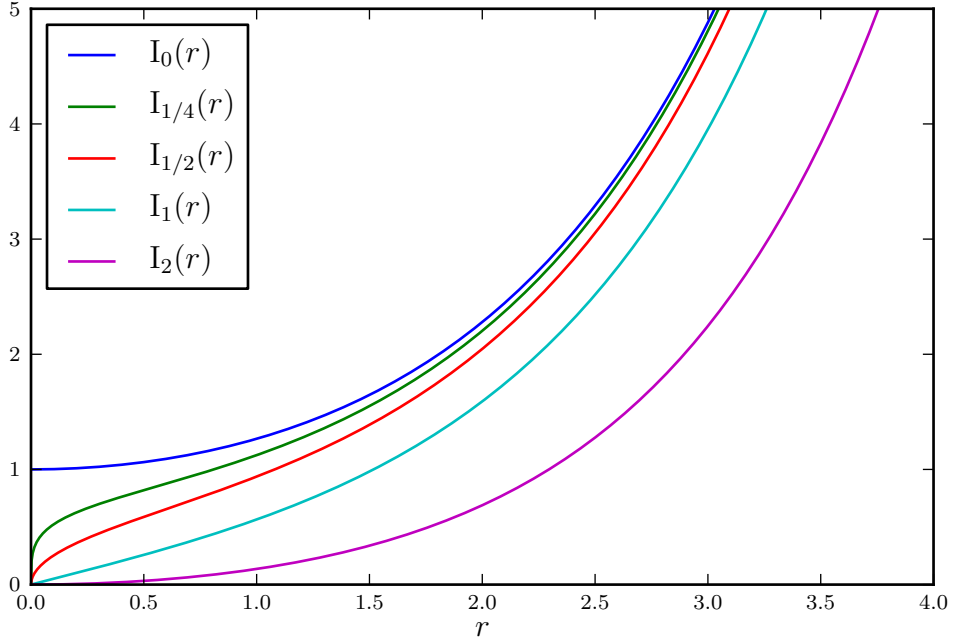


Figure 3.3: Regularity of the solution of (3.1) on the domain depicted on the upper plot. Here d is the distance between the centers of two circles with equal radii. The value s in the second plot indicates in which space $H^s(\Omega)$ the solution of the partial elliptic equation we investigate in this chapter lies, when considered on the corresponding overlapping circles domain.

Figure 3.4: Plots of I_λ for various values of λ .

3.4.1 The Modified Bessel Function of the First Kind

Using equation 17.10.2.1.4 from Jeffrey and Dai [2008] we find:

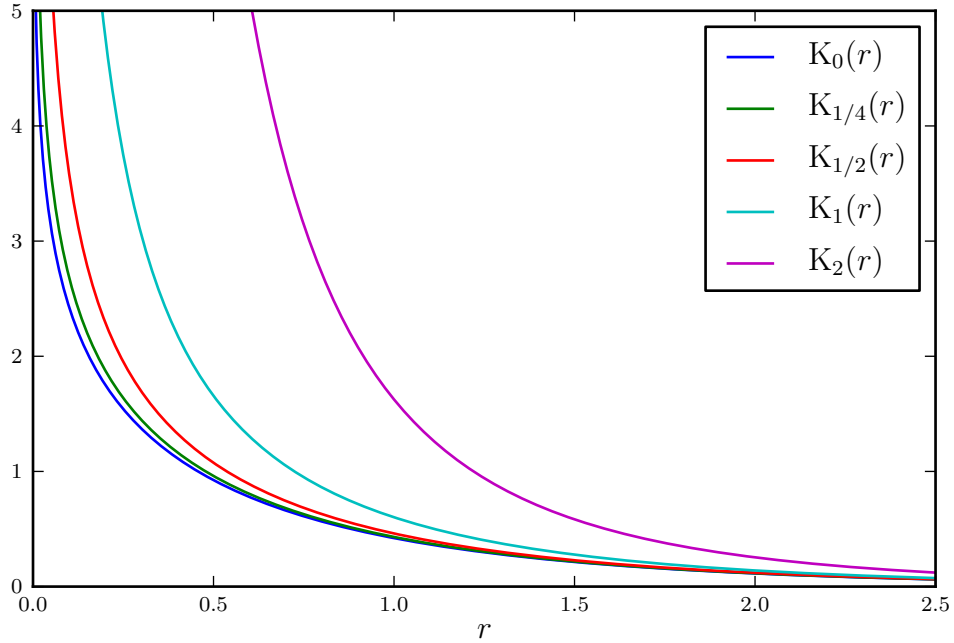
$$\frac{d}{dr}I_\lambda(r) = \frac{\lambda}{r}I_\lambda(r) + I_{\lambda+1}(r) \quad \text{for } r > 0 \quad (3.23)$$

$$\frac{d^2}{dr^2}I_\lambda(r) = \left(1 - \frac{\lambda}{r^2} + \frac{\lambda^2}{r^2}\right)I_\lambda(r) - \frac{1}{r}I_{\lambda+1}(r) \quad \text{for } r > 0. \quad (3.24)$$

From this and the asymptotic behaviour ($f(x) \sim g(x)$ iff $\lim_{x \rightarrow \infty}(f(x)/g(x)) = 1$)

$$I_\lambda(r) \sim \frac{1}{\Gamma(\lambda)} \left(\frac{1}{2}r\right)^\lambda \quad \text{for } \lambda \in \mathbb{R}^+, r \rightarrow 0 \quad (3.25)$$

from 9.6.7 in Abramowitz and Stegun [1964] (Γ is the Gamma function), we find that $\|\frac{d^2}{dr^2}I_\lambda\|_{L^2((0,\varepsilon))}$ is unbounded for $\lambda \in (0, 1)$ and $\varepsilon > 0$. However, the norm of the first derivative, $\|\frac{d}{dr}I_\lambda\|_{L^2((0,\varepsilon))}$, is bounded for $\lambda \geq \frac{1}{2}$ and for $\lambda = 0$.


 Figure 3.5: Plots of K_λ for various values of λ .

3.4.2 The Modified Bessel Function of the Second Kind

Turning now to the function K_λ , we again consider the asymptotic behaviour close to the origin. From 9.6.9 in Abramowitz and Stegun [1964] we get

$$K_\lambda(r) \sim \frac{1}{2}\Gamma(\lambda) \left(\frac{1}{2}r\right)^{-\lambda} \quad \text{for } \lambda \in \mathbb{R}^+, r \rightarrow 0 \quad (3.26)$$

where Γ is again the Gamma function. We then get that for $\varepsilon > 0$, the L^2 -norm $\|K_\lambda\sqrt{r}\|_{L^2((0,\varepsilon))}$ is bounded for $\lambda \in [0,1)$ but unbounded for $\lambda \geq 1$. Why we include the \sqrt{r} term in the norm will become clear in Lemma 3.15.

Summarizing all of this gives us:

Lemma 3.11. *Let I and K be the modified Bessel functions of the first and second kind respectively and let $\varepsilon > 0$. Then*

$$I_\lambda \in H^1((0, \varepsilon)) \quad \text{for } \lambda \geq \frac{1}{2} \text{ or } \lambda = 0 \quad (3.27)$$

$$I_\lambda \notin H^2((0, \varepsilon)) \quad \text{for } \lambda \in (0, 1) \quad (3.28)$$

$$\sqrt{r}K_\lambda \in L^2((0, \varepsilon)) \quad \text{for } \lambda \in [0, 1) \quad (3.29)$$

$$\sqrt{r}K_\lambda \notin L^2((0, \varepsilon)) \quad \text{for } \lambda \geq 1 \quad (3.30)$$

While the modified Bessel functions are very interesting mathematical objects, this is all we need to know about them for now, so we return to the topic of regularity of the PDE (3.1).

3.5 A Closer Look at the Corners

Grisvard [1992] notes the extensibility of his results to multiply connected domains, which we show here, using his original proof-techniques. In addition to the more general domains, we will get a precise description of the singularities that cause regularity problems near non-convex corners.

Our proof will follow Grisvard's closely. In some places we can use theorems without modification. Others require minor to major modifications, in which case we will give an entire proof. A certain simplification is achieved by using the fact that we are working on a pure Neumann problem.

The basic idea is to split up the function space $L^2(\Omega)$ into two orthogonal subspaces: One for which the solution of the partial differential equation (3.1) is an $H^2(\Omega)$ function, and the orthogonal complement N of that set (i.e. source functions for which the solution is less smooth). In Section 3.5.1 we will characterize that second subspace in detail and calculate its dimension. In Section 3.5.2 we will then use that to find a space which contains all solutions of (3.1) for all source functions $f \in L^2(\Omega)$.

3.5.1 Characterizing the Orthogonal Complement N

Again we consider the elliptic PDE:

$$\begin{aligned} Lu &= f && \text{in } \Omega \\ \frac{\partial u}{\partial \mathbf{n}} &= 0 && \text{on } \partial\Omega, \end{aligned} \tag{3.31}$$

with $L = u - \sigma\Delta u$, $f \in L^2(\Omega)$ and $\sigma > 0$. In the remainder of this chapter, Ω will be a Lipschitz polygon in \mathbb{R}^2 as defined in Definition 3.1 (a). Also recall the variational formulation (3.7).

We already know that restricted to the interior of Ω and even up to the boundary $\partial\Omega$ (excluding neighbourhoods of the corners) the solution u is an $H^2(\Omega)$ function. But we also know that this fails in a neighbourhood of non-convex corners. We will now investigate the exact behaviour of solutions in these corners.

In this section we will make use of the following spaces:

$$S := \{u \in H^2(\Omega) : \frac{\partial u}{\partial \mathbf{n}} = 0\} \tag{3.32}$$

$$N := L(S)^\perp := \{v \in L^2(\Omega) : \langle v, w \rangle_{L^2(\Omega)} = 0 \ \forall w \in L(S)\}. \tag{3.33}$$

The set S is the space of possible solutions of (3.31) if we had convex corners only and N denotes the orthogonal complement in $L^2(\Omega)$ of the image of S under L .

We will start by showing in Theorem 3.17 that the dimension of the space N is equal to the number of non-convex corners in the domain. This means that in the presence of such corners there exist functions $f \in L^2(\Omega)$ for which the solution of (3.31) is not in S . The next step is then to develop an expression for more general solutions of (3.31) for arbitrary $f \in L^2(\Omega)$, which allows us to precisely characterize the behaviour of such functions near the non-convex corners.

Lemma 3.12. *Let $v \in N$. Then $v \in \{u \in L^2(\Omega) : \Delta u \in L^2(\Omega)\}$ and v solves the adjoint boundary value problem:*

$$\int_{\Omega} (Lv)w \, dx = 0, \quad \forall w \in C_c^\infty(\Omega). \quad (3.34)$$

Proof. From (3.33), we have

$$\int_{\Omega} vLw \, dx = 0 \quad \forall w \in S,$$

so especially for all $w \in C_c^\infty(\Omega)$.

From Weyl's lemma (see e.g. Theorem 2 in Browder [1952] for a suitable formulation) we then know that $v \in C^\infty(\Omega)$ (but not necessarily $v \in C^\infty(\bar{\Omega})$) and $Lv = 0$ in Ω . The rest of the proof follows analogously to Grisvard [1992] Lemma 2.3.1 (when using an extension theorem suited for general Lipschitz domains, see e.g. McLean [2000] Appendix A). \square

Remark 3.13. *Note that the function v in Lemma 3.12 is not necessarily in $H^1(\Omega)$ but only in $L^2(\Omega)$. Therefore one can not conclude from the variational formulation that v vanishes entirely.*

Remark 3.14. *It can even be shown that v is C^∞ in the whole of $\bar{\Omega}$ excluding neighbourhoods of the corners, analogously to Lemma 2.3.4 in Grisvard [1992].*

Next, we analyse the behaviour of functions near the non-convex corners in more detail. For every corner $P_{(i,j)}$ with interior angle $\omega_{(i,j)}$, we define the space

$$\mathcal{A}_{(i,j)} := \{q \in H^2((0, \omega_{(i,j)})) : q'(0) = q'(\omega_{(i,j)}) = 0\} \quad (3.35)$$

and the operator

$$\Lambda_{(i,j)}\varphi = -\varphi'' \quad (3.36)$$

from $\mathcal{A}_{(i,j)}$ to $L^2((0, \omega_{(i,j)}))$. (See Section 2.3 in Grisvard [1992] for more details, especially on how to deal with Dirichlet and mixed boundary conditions.)

The operator $\Lambda_{(i,j)}$ has eigenvalues and normalized eigenfunctions given by

$$\begin{aligned}\varphi_{(i,j),1}(\theta) &= \omega_{(i,j)}^{-\frac{1}{2}} & \lambda_{(i,j),1} &= 0 \\ \varphi_{(i,j),m}(\theta) &= \sqrt{\frac{2}{\omega_{(i,j)}}} \cos(\lambda_{(i,j),m}\theta) & \lambda_{(i,j),m} &= \frac{(m-1)\pi}{\omega_{(i,j)}} \text{ for } m \geq 2.\end{aligned}\quad (3.37)$$

We now switch to polar coordinates $(r_{(i,j)}, \theta_{(i,j)})$ from the point of view of the (i, j) -th corner.

Define a neighbourhood of the corner in Ω :

$$D_\varepsilon^{(i,j)} := \Omega \cap B_\varepsilon(P_{(i,j)}),$$

where $B_\varepsilon(P_{(i,j)})$ is the open disk with radius ε and origin at $P_{(i,j)}$. It then follows from Lemma 3.12 by transformation to polar coordinates, that every $v \in \mathbb{N}$ is a solution of

$$v - \sigma \left(\frac{d^2}{dr^2} v + r^{-1} \frac{d}{dr} v + r^{-2} \frac{d^2}{d\theta^2} v \right) = 0 \quad \text{for } \theta \in (0, \omega_{(i,j)}), r \in (0, \varepsilon), \quad (3.38)$$

for a sufficiently small $\varepsilon > 0$.

Now comes a critical step: by considering v not as a function from a neighbourhood of the corner into \mathbb{R} , but rather as a function of r into $\mathcal{A}_{(i,j)}$ (it is easy to see that $v \in \mathbb{N}$ fulfils the conditions of (3.35)), we can replace the second derivative w.r.t. θ in equation (3.38) by the operator $\Lambda_{(i,j)}$:

$$v - \sigma \left(\frac{d^2}{dr^2} v + r^{-1} \frac{d}{dr} v + r^{-2} \Lambda_{(i,j)} v \right) = 0 \quad \text{for } r \in (0, \varepsilon). \quad (3.39)$$

(Keep in mind that we know from Lemma 3.12 and Remark 3.14, that v is smooth for $r > 0$.)

In the next lemma, we will expand $v \in \mathbb{N}$ (again seen as a function of r into $\mathcal{A}_{(i,j)}$) in the eigensystem (3.37). This is where we encounter the first deviation from Grisvard's original proof for the Laplace operator.

Lemma 3.15. *Let $v \in \mathbb{N} \cap C^\infty((0, \varepsilon); \mathcal{A}_{(i,j)})$ be a solution of (3.39) and assume that $v \in L^2(D_\varepsilon^{(i,j)})$. Then, in $D_\varepsilon^{(i,j)}$,*

$$v(re^{i\theta}) = \sum_{m \geq 1} \alpha_m I_{\lambda_{(i,j),m}}(\sigma^{-1/2}r) \varphi_{(i,j),m}(\theta) - \sum_{0 < \lambda_{(i,j),m} < 1} \beta_m K_{\lambda_{(i,j),m}}(\sigma^{-1/2}r) \varphi_{(i,j),m}(\theta), \quad (3.40)$$

with $\alpha_m, \beta_m \in \mathbb{R}$ for all $m \in \mathbb{N}$.

Proof. We start by developing v in the basis $(\varphi_{(i,j),m})_{m \in \mathbb{N}}$ from (3.37):

$$\begin{aligned} v(re^{i\theta}) &= \sum_{m \geq 1} v_m(r) \varphi_{(i,j),m}(\theta), \\ \text{with } v_m(r) &:= \int_0^{\omega_{(i,j)}} v(re^{i\theta}) \varphi_{(i,j),m}(\theta) d\theta. \end{aligned} \quad (3.41)$$

From (3.39) we now find after some reorganizing:

$$r^2 v_m'' + r v_m' - (\sigma^{-1} r^2 + \lambda_{(i,j),m}^2) v_m = 0 \quad \text{for } r \in (0, \varepsilon),$$

which has the solution

$$v_m(r) = \alpha_{(i,j),m} I_{\lambda_{(i,j),m}}(\sigma^{-1/2} r) + \beta_{(i,j),m} K_{\lambda_{(i,j),m}}(\sigma^{-1/2} r), \quad (3.42)$$

as we know from (3.22), with I_λ and K_λ being the modified Bessel functions of the first and second kind, respectively.

Now let us take a closer look at the function $K_{\lambda_{(i,j),m}}(\sigma^{-1/2} r)$. From (3.41) we find:

$$\int_0^\varepsilon v_m(r)^2 r dr \leq \|v\|_{L^2(D_\varepsilon)}^2 < \infty.$$

(The extra r in the integral stems from the use of polar coordinates.)

From Section 3.4 we know that in a neighbourhood of the corner the L^2 norm of $\sqrt{r} K_{\lambda_{(i,j),m}}$ is bounded only for $\lambda_{(i,j),m} \in [0, 1)$. ($I_{(i,j),m}$ converges as $r \rightarrow 0$.) So β_m must equal zero for $\lambda_{(i,j),m} \geq 1$.

It now remains to be shown that $\beta_1 = 0$. For this, first recall the truncation functions $\eta_{(i,j)}$, which we introduced in Definition 3.2. Note that the truncation functions $\eta_{(i,j)} \in \mathcal{S}$ since they are smooth and depend only on the radial distance to the corner, and therefore, with $v \in \mathcal{N}$, we find that

$$\int_{\Omega} v L \eta_{(i,j)} dx = 0.$$

But we also have

$$\int_0^{\omega_{(i,j)}} \varphi_{(i,j),m}(\theta) d\theta = 0 \quad \text{for } m \geq 2,$$

and after a lengthy calculation using (3.42) we obtain

$$\begin{aligned} 0 &= \int_{\Omega} v L \eta_{(i,j)} dx = \int_{\Omega} v_1 L \eta_{(i,j)} dx \\ &= \sqrt{\omega_{(i,j)}} \int_0^\infty (\alpha_1 I_0 + \beta_1 K_0) (\eta_{(i,j)} - \sigma(\eta_{(i,j)}'' + r^{-1} \eta_{(i,j)}')) r dr. \end{aligned} \quad (3.43)$$

With some of the properties of Bessel functions we have discussed in Section 3.4 and additionally the relations (Jeffrey and Dai [2008])

$$\begin{aligned} \frac{d}{dr} I_0(r) &= I_1(r) & \frac{d}{dr} K_0(r) &= -K_1(r) \\ \frac{d}{dr} (r I_1(r)) &= r I_0(r) & \frac{d}{dr} (r K_1(r)) &= -r K_0(r), \end{aligned}$$

one can show from (3.43) after a lengthy but simple calculation that $\beta_1 = 0$ as well, which gives equation (3.40). \square

We see that the major difference between the Laplace operator Δ and the operator L is that for the former we get local expressions in the form of

$$w_m(r) = \alpha_{(i,j),m} r^{\lambda_{(i,j),m}} + \beta_{(i,j),m} r^{-\lambda_{(i,j),m}},$$

whereas the operator L gives rise to (3.42). Interestingly though, both expressions have a very similar behaviour for the purpose of classifying regularity, as will become clear soon (and as could be expected, based on the asymptotics (3.25) and (3.26)).

Now we will try to combine the local functions derived in Lemma 3.15 into functions on all Ω .

Lemma 3.16. *For all corners (i, j) and $\lambda_{(i,j),m} \in (0, 1)$ there exists a function*

$$\rho_{(i,j),m} \in N,$$

such that

$$\rho_{(i,j),m} - \eta_{(i,j)} K_{\lambda_{(i,j),m}}(\sigma^{-1/2} r) \varphi_{(i,j),m}(\theta) \in H^1(\Omega).$$

This lemma can be proven with only minor modifications to Grisvard's original proof for the Laplace operator, so we skip it here.

The following theorem now wraps up the first part of the regularity proof. It shows that the dimension of the space N , the orthogonal complement of the image under L of $H^2(\Omega)$ functions satisfying homogeneous Neumann boundary conditions, is equal to the number of non-convex corners of the domain.

Theorem 3.17. *We have $\dim N = |\{(i, j, m) : \lambda_{(i,j),m} \in (0, 1)\}|$.*

Before proving this theorem, we recall the expressions for the eigenvalues $\lambda_{(i,j)}$ in (3.37). From this we find that

$$\lambda_{(i,j),m} \in (0, 1) \iff \frac{(m-1)\pi}{\omega_{(i,j)}} \in (0, 1) \iff (m-1)\pi < \omega_{(i,j)},$$

which is only possible if $m = 2$ and $\omega_{(i,j)} > \pi$, i.e. if the corner $P_{(i,j)}$ is non-convex. So Theorem 3.17 simplifies to

$$|\{(i, j, m) : \lambda_{(i,j),m} \in (0, 1)\}| = |\{(i, j) : \omega_{(i,j)} \in (\pi, 2\pi)\}|,$$

which counts the number of non-convex corners in the domain Ω .

Proof of Theorem 3.17. Let $v \in \mathbb{N}$. Due to Lemma 3.15, we have for every corner $P_{(i,j)}$:

$$\begin{aligned} v(re^{i\theta}) - \sum_{m \geq 1} \alpha_m \mathbf{I}_{\lambda_{(i,j),m}}(\sigma^{-1/2}r) \varphi_{(i,j),m}(\theta) \\ - \sum_{0 < \lambda_{(i,j),m} < 1} \beta_m \mathbf{K}_{\lambda_{(i,j),m}}(\sigma^{-1/2}r) \varphi_{(i,j),m}(\theta) = 0 \text{ in } D_\varepsilon^{(i,j)}, \end{aligned}$$

and so this difference is in $H^1(D_\varepsilon^{(i,j)})$.

Using Lemma 3.16, we furthermore find

$$v(re^{i\theta}) - \sum_{m \geq 1} \alpha_m \mathbf{I}_{\lambda_{(i,j),m}}(\sigma^{-1/2}r) \varphi_{(i,j),m}(\theta) - \sum_{0 < \lambda_{(i,j),m} < 1} \beta_m \rho_{(i,j),m} \in H^1(D_\varepsilon^{(i,j)}),$$

for some $\rho_{(i,j),m} \in \mathbb{N}$.

Keeping in mind that $\lambda_{(i,j),1} = 0$ and $\lambda_{(i,j),m} > \frac{1}{2}$ for $m > 1$, Lemma 3.11 implies

$$\sum_{m \geq 1} \alpha_m \mathbf{I}_{\lambda_{(i,j),m}}(\sigma^{-1/2}r) \varphi_{(i,j),m}(\theta) \in H^1(D_\varepsilon^{(i,j)}),$$

which means that

$$v(re^{i\theta}) - \sum_{0 < \lambda_{(i,j),m} < 1} \beta_m \rho_{(i,j),m} \in H^1(D_\varepsilon^{(i,j)}).$$

Summing up over all corners j (so we now have coefficients $\beta_{(i,j),m}$ instead of β_m), we know that

$$w := v(re^{i\theta}) - \sum_{0 < \lambda_{(i,j),m} < 1} \beta_{(i,j),m} \rho_{(i,j),m} \in H^1(\Omega),$$

because $\rho_{(i,j),m} \in H^1(\Omega)$ due to Lemma 3.16 and v is smooth outside a neighborhood of the corners as we know from Remark 3.14.

Finally we show, that w actually vanishes, which means that v is a linear combination of the functions $\rho_{(i,j),m}$ that we have identified in Lemma 3.16. These functions are linearly independent, which yields the statement of this theorem.

To see that $w \equiv 0$, remember that $w \in \mathbb{N}$ (because $\rho_{(i,j),m} \in \mathbb{N}$ and $v \in \mathbb{N}$), and so Lemma 3.12 implies:

$$\int_{\Omega} (Lv)w \, dx = 0, \quad \forall w \in C_c^\infty(\Omega).$$

Furthermore, $w \in H^1(\Omega)$, so using inequality (3.8) we find that w vanishes, which completes the proof. \square

3.5.2 Constructing a Space for All Solutions

We now have a precise description of the space of $L^2(\Omega)$ functions f that admit no $H^2(\Omega)$ solutions to the partial differential equation (3.31). It is spanned by a space whose dimension equals the number of non-convex corners in our domain. Based on this insight, we now try to find a Sobolev space $H^s(\Omega)$ that contains all solutions of (3.31) for arbitrary $f \in L^2(\Omega)$.

Defining

$$S_{(i,j),m}(r_{(i,j)}, \theta_{(i,j)}) := \eta_{(i,j)} I_{\lambda_{(i,j),m}}(\sigma^{-1/2} r_{(i,j)}) \varphi_{(i,j),m}(\theta_{(i,j)}) \quad (3.44)$$

$$F_{(i,j),m}(r_{(i,j)}, \theta_{(i,j)}) := L S_{(i,j),m}(r_{(i,j)}, \theta_{(i,j)}), \quad (3.45)$$

we first find that

Lemma 3.18. $F_{(i,j),m}$ is not orthogonal to N for $\lambda_{(i,j),m} \in (0, 1)$.

This can easily be proven analogously to Lemma 2.4.2 in Grisvard [1992], using the expressions we have derived for the Bessel function I_λ in Section 3.4.

Theorem 3.19. Assume that Ω is a bounded polygonal open subset of \mathbb{R}^2 . For each $f \in L^2(\Omega)$ there exists a unique (variational) solution $u \in H^1(\Omega)$ of

$$\int_{\Omega} uv + \sigma \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx$$

for every $v \in H^1(\Omega)$ and in addition, there exist unique constants $c_{(i,j),m} \in \mathbb{R}$ such that

$$u - \sum_{(i,j)} \sum_{0 < \lambda_{(i,j),m} < 1} c_{(i,j),m} S_{(i,j),m} \in H^2(\Omega)$$

The proof is almost identical to the original one and requires no special care for the Bessel functions; we therefore skip it.

Theorem 3.19 tells the precise behaviour of solutions of (3.1) near non-convex corners: The part of the solution which is not in $H^2(\Omega)$ is proportional to $I_{\lambda_{(i,j),m}}(\sigma^{-1/2} r_{(i,j)})$ near the corner $P_{(i,j)}$. Hence the parameter σ does not influence the *degree of smoothness* of the solution, but it does contribute to the *magnitude* of the term which is less regular than $H^2(\Omega)$.

With Conjecture 3.8 and Remark 3.9 we now arrive at our final regularity result.

Conjecture 3.10. Let Ω be as in Definition 3.1 (b). Assume we have $u \in H^1(\Omega)$ and $f \in L^2(\Omega)$ such that

$$\begin{aligned} Lu &= f \text{ in } \Omega \\ \frac{\partial u}{\partial \mathbf{n}} &= 0 \text{ on } \partial\Omega. \end{aligned}$$

Then $u \in H^s(\Omega)$ with $1 < s \leq 2$ and $s < 1 + \pi / \max_{(i,j)} \omega_{(i,j)}$ (recall Definition 3.1).

Note that the H^s regularity follows from the asymptotic of the modified Bessel function, see Section 3.4.

So we see that again the smoothness of a solution depends on the largest interior angle. The larger the non-convex angles, the less smoothness we have.

Figure 3.6 shows the expected smoothness of solutions to the PDE (3.1) for the geometry from Chapter 2 (recall Figure 2.5). Note that even on the lowest layer, which is located at 47 meters, the apertures do not exactly overlap.

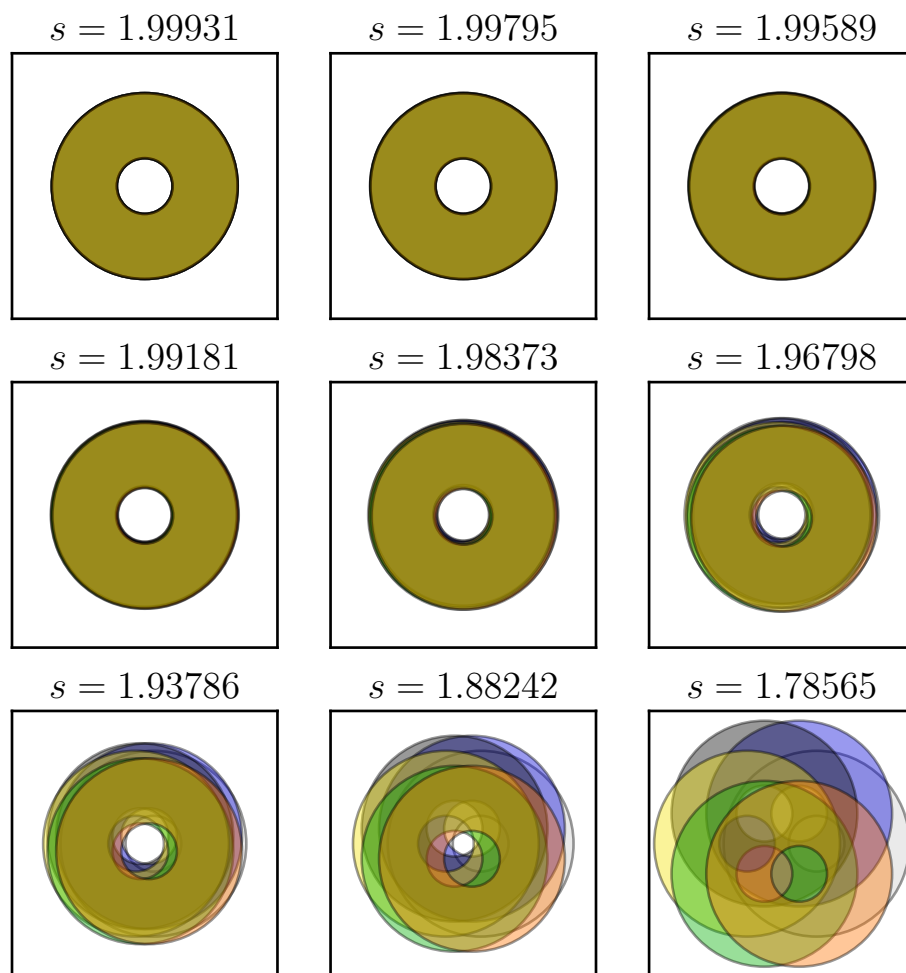


Figure 3.6: Smoothness estimates for solutions of (3.1) on different domains. The heights correspond to the ones in Figure 2.5. Compare these indices s with the decay power $11/6 \approx 1.83$ in the von Karman turbulence model (2.3).

Chapter 4

Numerical Realization

Where we test the H^1 space Landweber-Kaczmarz in a wide variety of testcases and interpret the results.

4.1 Implementation

To test the H^1 Sobolev space Kaczmarz method that was derived in Chapter 2, this functionality was implemented in an MCAO simulation framework developed by the *Mathematical algorithms and software for ELT adaptive optics* project, a joint work by the *Institute for Industrial Mathematics* at the *Johannes Kepler University*, the *Johann Radon Institute for Computational and Applied Mathematics* (RICAM) and the *Industrial Mathematics Competence Center* (IMCC). The tests were performed in a realistic setting using data from the *European Southern Observatories* (ESO) and designs for the *European Extremely Large Telescope* (E-ELT).

Since all (wavefront) data was given on regular grids (as specified by the geometry of the wavefront sensors), a finite difference method was used to model the partial differential equation (2.14). As possible solvers for the resulting sparse systems of linear equations, both PCG and direct methods based on sparse LU-factorization were considered. For the problem sizes at hand, both methods performed very well.

Von Karman turbulence was used to model turbulent layers in the atmosphere. There are different ways to generate random turbulences following this statistical model. One popular method is to multiply a complex noise vector elementwise with the square root of the Kolmogorov spectrum vector and then take the inverse Fourier-Transform of the resulting vector. The periodicity that is caused by using the Fourier transform (see again Figure 2.3) is particularly useful when shifting the layers to simulate winds for closed-loop operation. See e.g. Lane et al. [1992] and Vorontsov et al. [2008] for more information on simulating atmospheric turbulence.

4.1.1 Parameters

The MCAO simulation depends on a large number of parameters, some of which will be assumed fixed throughout this chapter due to the focus on the E-ELT. Other parameters will be tested with various settings to investigate their role in determining the reconstruction's operating performance.

Fixed Parameters

The fixed parameters include:

Telescope The telescope aperture was assumed to be 42m with a circular opening and a central obstruction.

Wavefront sensors The number of wavefront sensors was set to six, according to current plans for the E-ELT. They are Shack-Hartmann sensors with a resolution of 84x84 subapertures.

Guidestars Correspondingly, six guide stars were used, arranged in a circle surrounding the central region (see Figure 4.1). *All guide stars were assumed to be natural ones.*

Variable Parameters

For the following parameters, we tested various values and settings to evaluate their reconstruction performance :

Field of View As we have seen, the field of view directly influences the smoothness of the solution of the partial differential equation (3.1). A typical field of view for the E-ELT is 3 arc minutes, which was used in all tests unless otherwise noted.

Atmospheric layers In order to test AO systems, atmospheric turbulences in multiple layers need to be simulated. Unless noted otherwise, a 9 layers atmospheric model with real-world characteristics from ESO was used, which was introduced in Section 2.1.4. In a few cases we have restricted ourselves to a simple 3 layers model in order to simplify interpretation of results.

Layers to reconstruct MCAO systems typically use a-priori knowledge about how many layers the atmosphere consists of (or more precisely, how many layers at which heights match the actual atmospheric composition best). A typical choice is a 3 layers atmospheric model, considerably more simple than the full 9 layers model. In Section 4.2.4, we will take a closer look at this use of a-priori information.

Flux level The *photon flux* is an important characteristic in AO systems. This value expresses, how many photos from a guide star are received by a wavefront sensor, i.e. how bright the guide-star is. The brighter a guide star, the higher the

accuracy of the wavefront measurement. As guide stars become dimmer, the level of noise in the system increases quickly (The lack of sufficiently bright natural guide stars led to the introduction of laser guide stars into AO systems.). We measure the photon flux p in photons per subaperture of the wavefront sensor per image (i.e. exposure time). A typical value in MCAO is $p = 500$, which we use for most of our tests.

In Section 4.2.3 we will investigate further photon flux levels.

Strength of smoothing The strength of the smoothing can be adjusted using the parameter σ in (3.1). Low values of σ correspond to sharp solutions of (3.1), whereas high values of σ lead to very blurry results, as we will see in Section 4.2.2.

Deformable mirrors We will generally assume that the telescope uses three deformable mirrors conjugated to different altitudes. We will however in one test case investigate varying the number of deformable mirrors to see what potential AO performance improvements can be achieved by using a higher number of deformable mirrors.

4.1.2 Evaluation

There are several possible choices to evaluate the performance of an AO system and of atmospheric tomography, and the *Strehl* ratio has proven to be the most useful measure. The Strehl ratio measures how much light from a single source is spread out, compared with the best possible focus (which is dictated by the diffraction limit, which depends on the wavelength and the telescope aperture). The maximum value is 1.0 and a lower value indicates a decrease in image resolution due to optical imperfections in the telescope optics and image distortions due to atmospheric turbulences.

In order to measure the quality of a reconstruction of the atmosphere, we define a number of view-directions that are uniformly distributed in the sky. Figure 4.1 shows such an arrangement.

For every one of those view directions, the perturbations caused by all atmospheric layers when making an observation in that given direction are summed up which gives the total wavefront distortion (compare equation (2.4)). The same process is repeated with the reconstructed layer-functions replacing the real ones. The difference in the resulting data is then measured using the Maréchal approximation for the Strehl ratio (see e.g. Ross [2009]). By computing these measures for multiple positions on the sky, we can evaluate and plot the quality of the perturbation correction over the telescopes entire field of view.

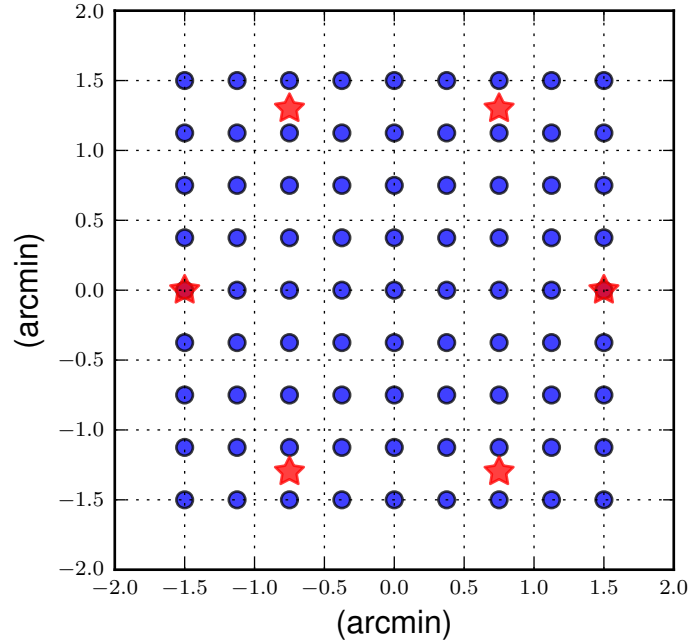


Figure 4.1: Angular placement of views (blue circles) and guide stars (red stars) for a 3 arcminutes field of view.

4.2 Tests

4.2.1 A first example

In Figure 4.2 the original layers of a 3 layers atmosphere were plotted, together with a reconstruction of those layers. The layers were located at altitudes 0m (directly above the telescope), 4000m and 12700m. Note the different shapes of the reconstructions at different heights (particularly the hole caused by the central obstruction), matching the expectations from Figure 2.5.

4.2.2 Smoothing Strength

In a very simple test, we will investigate the role of the smoothing parameter σ . For this we will use a simple 3 layers atmosphere. In Figure 4.3, we can see the original atmospheric layers and the reconstructions with different values of σ . Note how small values of σ lead to sharp pictures, with higher values leading to successively more blurry results.

Tests have shown that the best smoothing performance was achieved with a factor $\sigma \approx 0.2$, which we will use in the remaining tests.

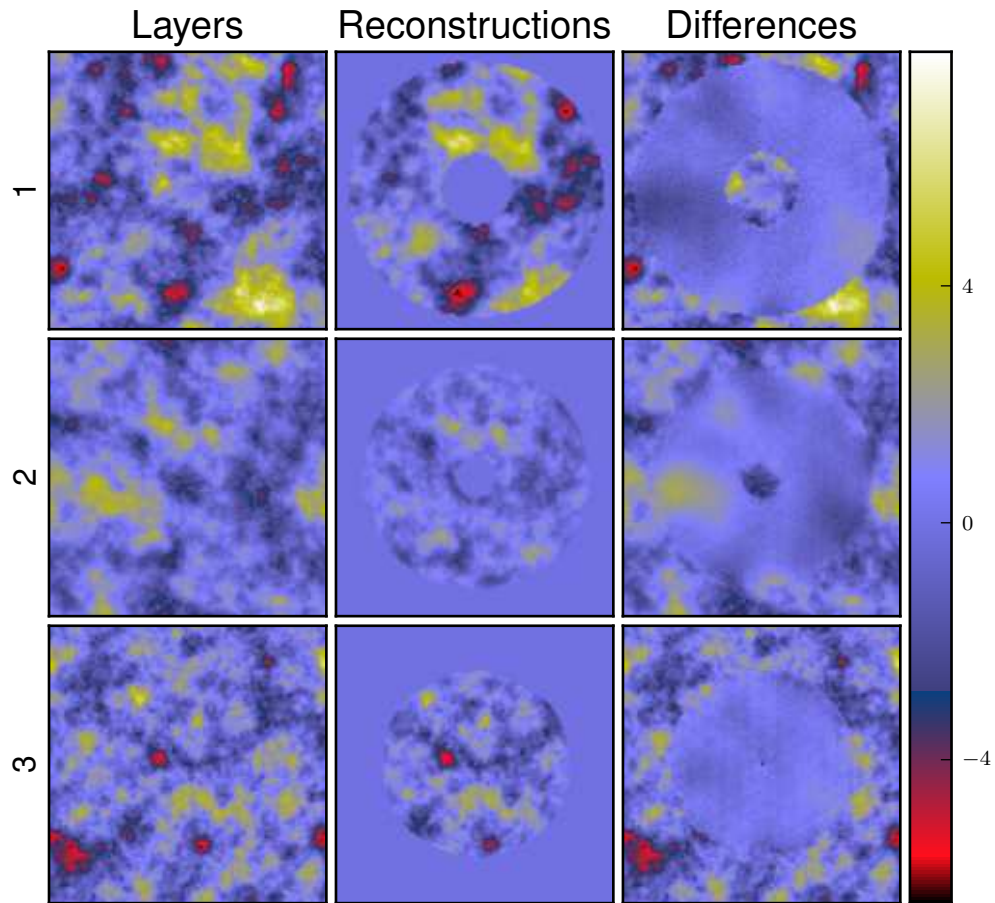


Figure 4.2: An atmosphere consisting of three layers and their reconstructions using 10 iterations of the Landweber-Kaczmarz iteration.

4.2.3 Photon flux

The brighter the guide stars are, the better the reconstruction, as the wavefronts can be measured more precisely. In Figures 4.5 and 4.6 we see reconstruction results for a 9 layers atmosphere for different iteration numbers with a high photon flux of $p = 10^4$ and a low flux of $p = 3$, respectively. Figures 4.7 and 4.8 show the average Strehl-ratios plotted against the distance from the centre of the telescope's field of view.

As can be seen, the smoothing generally hurts the reconstruction quality, however, with very dim guide stars the smoothing can actually help spread out the correction over a wider area, towards the centre of the field of view. Results for noise strengths between those two extremes show the same behaviour.

In Section 4.2.6 we will take another look at the effect of photon noise on atmospheric tomography.

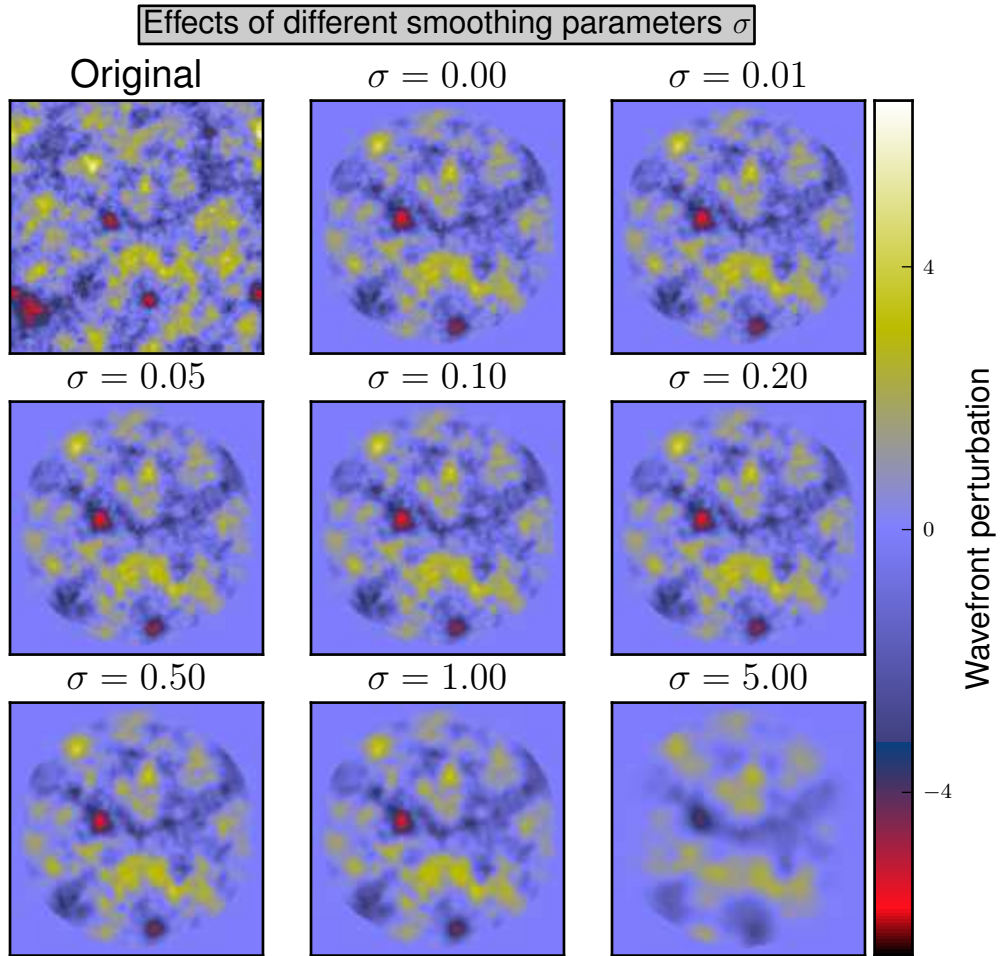


Figure 4.3: Highest layer of a 3 layers atmosphere and reconstructions using 10 iterations of the Landweber-Kaczmarz iteration with different smoothing strengths.

4.2.4 Different Numbers of Reconstructed Layers and Deformable Mirrors

MCAO Systems depend on a-priori information on the shape of the atmosphere. Particularly information about how many layers to reconstruct and at which altitudes these layers are located can have an enormous impact on reconstruction performance.

Computational concerns for example, can lead to the decision to reconstruct only a small number of layers. Similarly, the number of deformable mirrors is generally much smaller than the number of atmospheric layers and sometimes even smaller than the number of reconstructed layers. In such cases, a mirror's deformation is determined by projecting or "fitting" a subset of all reconstructed layers onto the mirrors conjugated

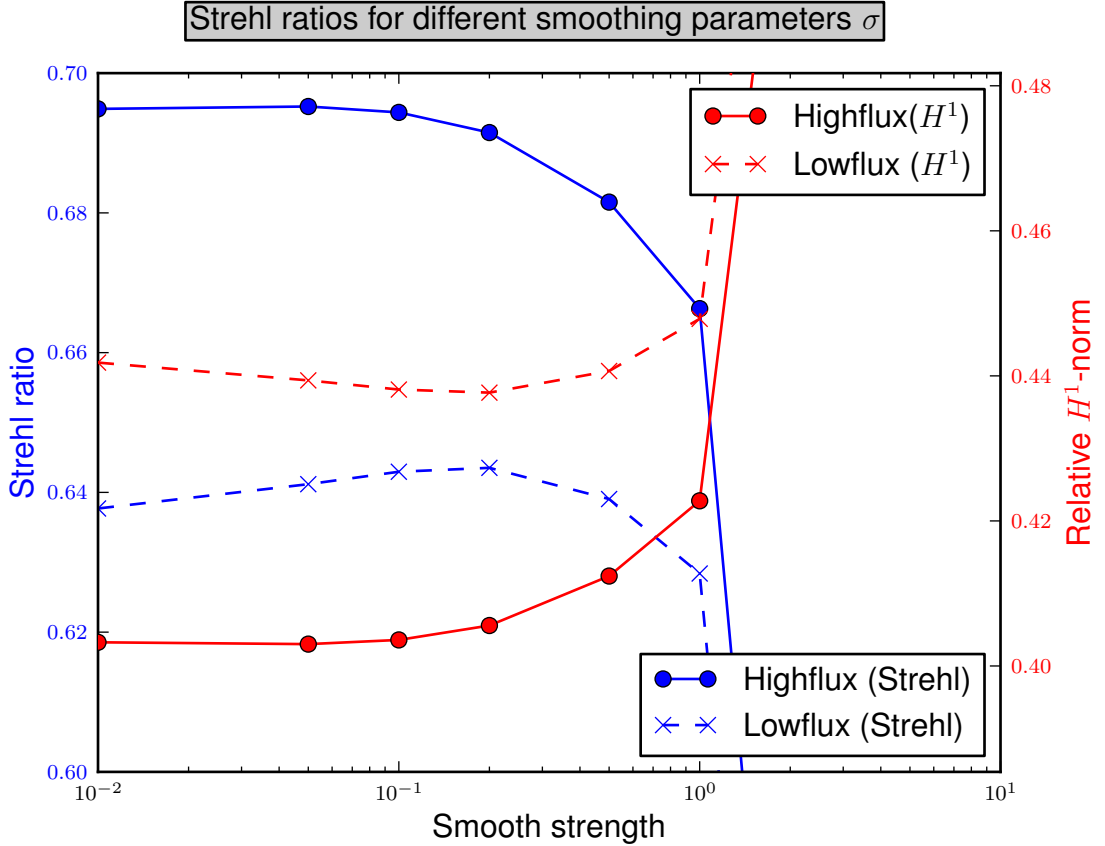


Figure 4.4: Strehl ratios and H^1 norms for the different reconstructions seen in Figure 4.3. Two different noise cases (see Section 4.2.3) are covered: *Highflux*, i.e. bright guide-stars and *Lowflux*, i.e. dim guide stars. For the Lowflux case the ratio improves for small smoothing strengths and then quickly degenerates after a value of $\sigma = 1$. In the Highflux case the smoothing generally impairs reconstruction quality.

height. See e.g. Yang and Vogel [2006].

In the next scenario, we will investigate the effects of such limitations. Again, we will simulate the E-ELT with the 9 layers atmosphere. We will however now differentiate between two cases: In the first we will provide the reconstruction with exact information on how many layers to reconstruct and at which altitudes to locate them. This information will match the 9 layers atmosphere. In a second test-run, we will, due to presumed lack of a-priori information, only provide the reconstruction with a 3 layers approximation of the full 9 layers atmosphere. In both cases we will use a setup using only 3 deformable mirrors.

For the 9 layers reconstruction, we will then additionally investigate both using a full

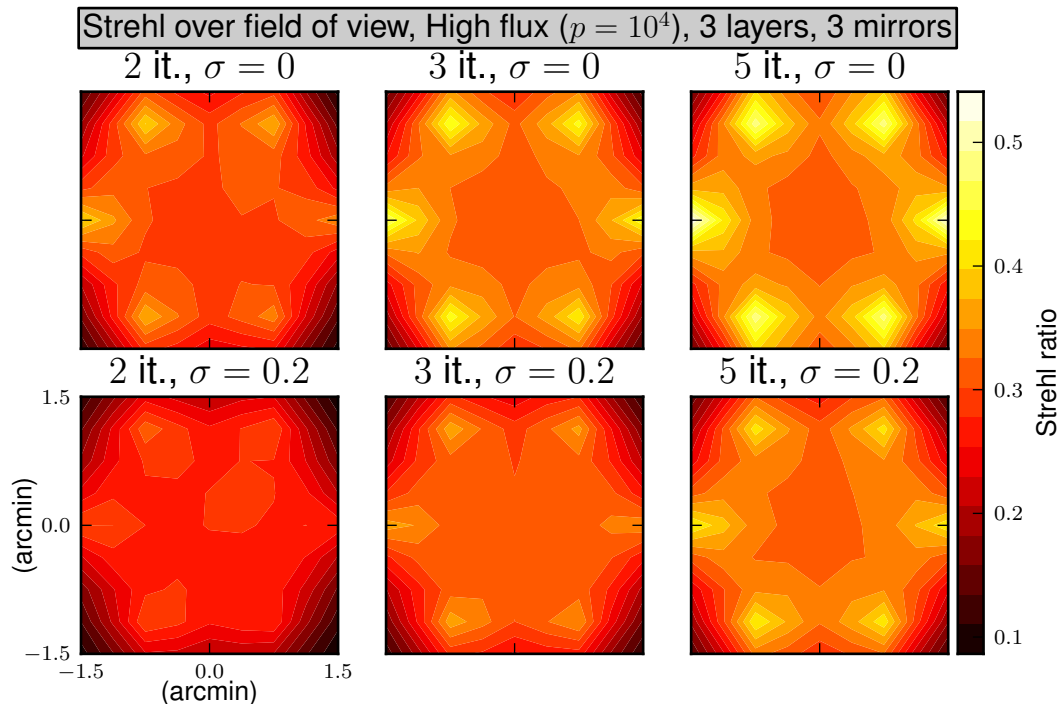


Figure 4.5: Strehl maps for a high flux ($p = 10^4$) observation with 3 reconstructed layers.

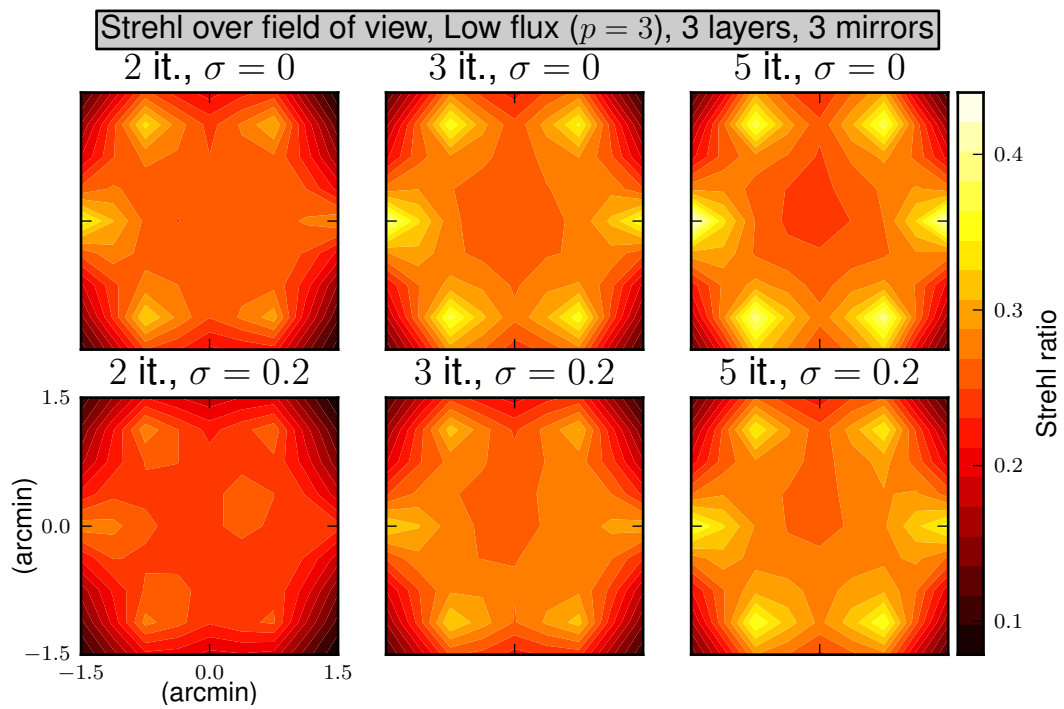


Figure 4.6: Strehl maps for a low flux ($p = 3$) observation with 3 reconstructed layers.

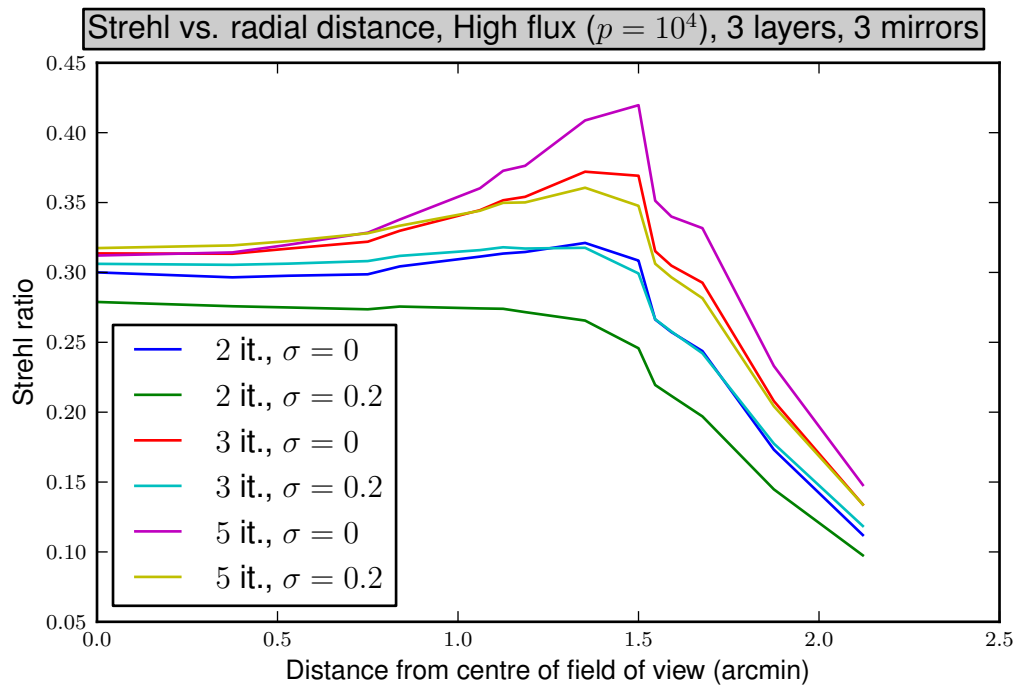


Figure 4.7: Strehl ratios plotted versus distance from the centre of the field of view for a high flux ($p = 10^4$) observation. Here, a 3 layers atmosphere was reconstructed.

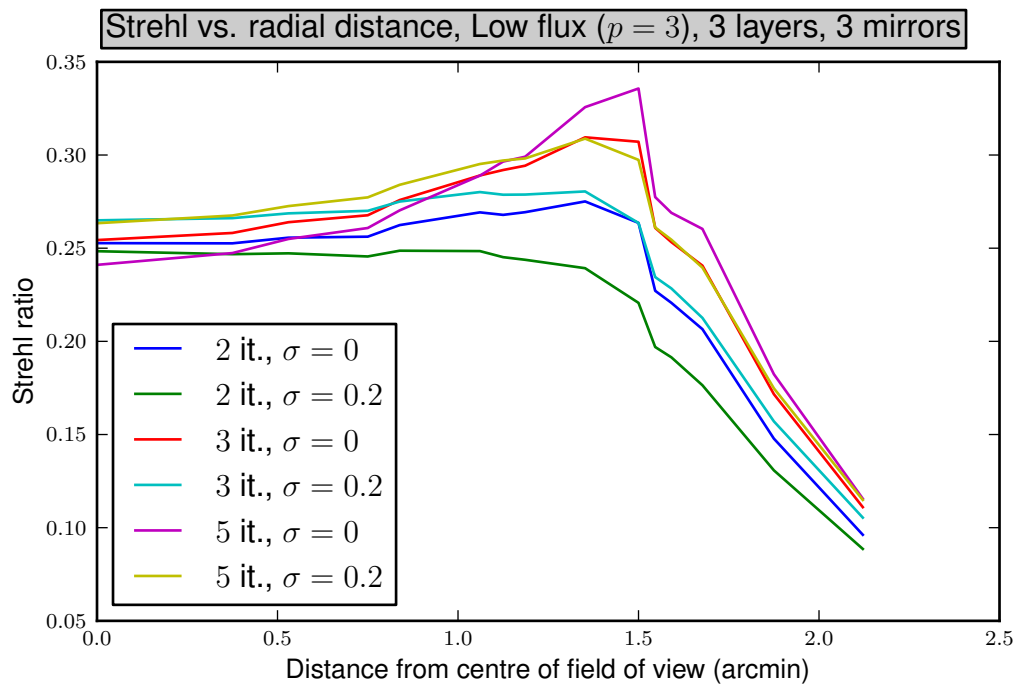


Figure 4.8: Strehl ratios plotted versus distance from the centre of the field of view for a low flux ($p = 3$) observation. Here, a 3 layers atmosphere was reconstructed.

9 mirrors setup with heights conjugated to the corresponding atmospheric layers.

As we will see, all these factors have a strong effect on the tomography results.

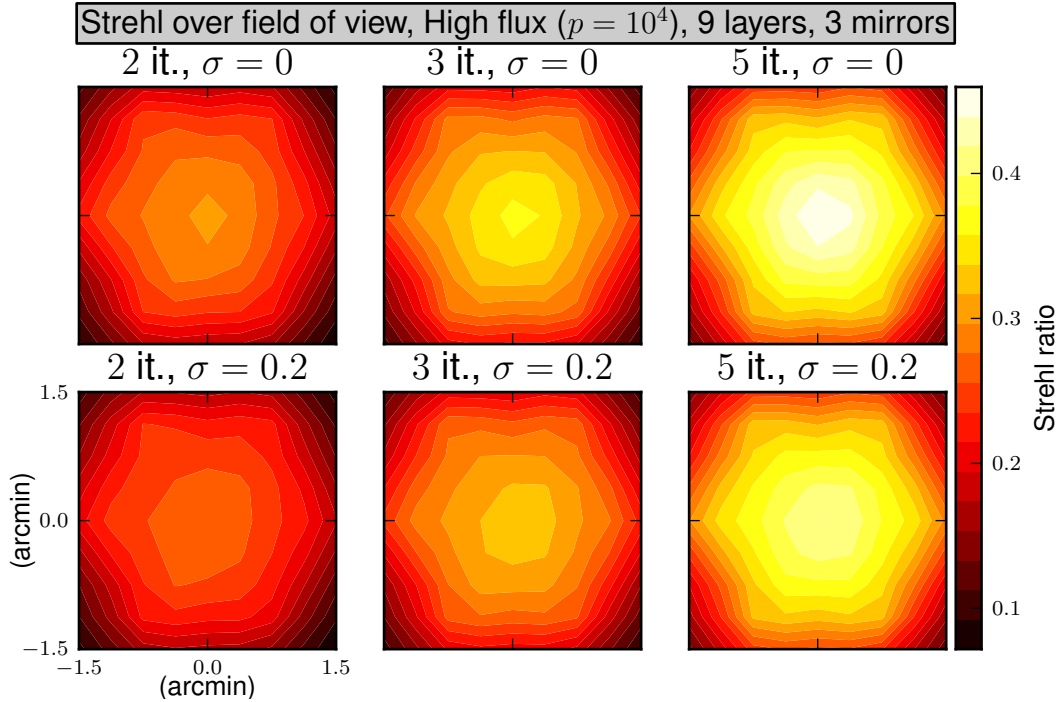


Figure 4.9: Strehl maps for a high flux ($p = 10^4$) observation with 9 reconstructed layers and 3 deformable mirrors. Note the high correction in the centre of the field of view.

In Figures 4.5 to 4.8 we already saw the results for a 9 layers atmosphere reconstructed in 3 layers. Figures 4.9 and 4.10 now show the Strehl maps for the same settings but with the atmosphere reconstructed in 9 layers. Figures 4.11 and 4.12 show the corresponding Strehl over distance plots. These are tests using 3 deformable mirrors. In Figures 4.13, 4.14, 4.15 and 4.16 we see the results of reconstructing 9 turbulent layers and using 9 matching deformable mirrors to correct for them.

As could be expected, the more information the reconstruction has at its disposal, the better the overall result. However, even the rough 3 layers approximation can yield good corrections over a large field of view. Note especially how in the 3 layers case the focus on the guide stars is much stronger, whereas in the 9 layers reconstruction the central region of the field of view is improved more uniformly.

Comparing in particular Figures 4.7 and 4.15, we see how well atmospheric tomography, without mirror projects, works. If we could correct every turbulent layer

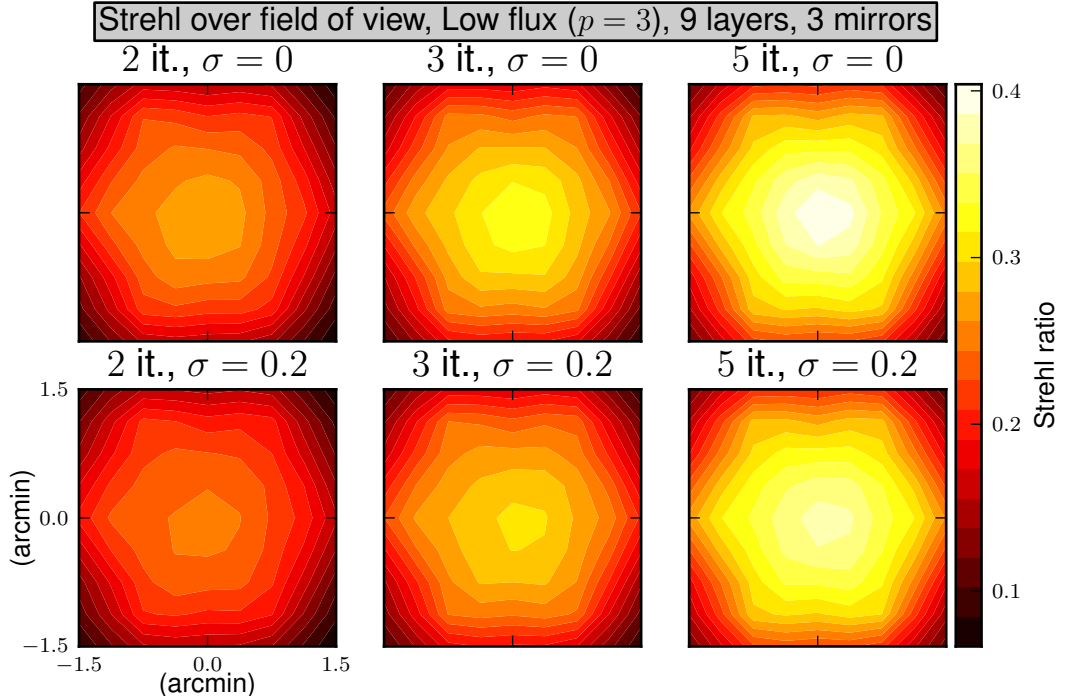


Figure 4.10: Strehl maps for a low flux ($p = 3$) observation with 9 reconstructed layers

directly, the MCAO system could achieve an almost uniform correction on the entire inner field of view. Even in the low flux case of Figure 4.16, the corrections are generally better than even in the high flux scenario of Figure 4.7.

In the average flux case $p = 500$, the algorithm behaves analogously with results in between the high flux ($p = 10^4$) and low flux ($p = 3$) cases.

Together with the previous section's result, we get a first indication of the performance characteristics for the H^1 -space Landweber-Kaczmarz method. Roughly speaking it seems to work best for low flux levels and with little available a-priori information.

4.2.5 Large Iteration Numbers

Due to performance concerns, during actual operation (especially in closed-loop mode), the Landweber-Kaczmarz method is executed for only a handful of iterations. Nevertheless, studying the behaviour of the algorithm and the smoothing for larger iteration numbers can yield important insights into the algorithm.

In Figures 4.17, 4.18, 4.19 and 4.20, we can see the effect of smoothing for high iteration numbers. In the 3 layers reconstruction case, we again see how the algorithm focuses on the guide star directions and how the smoothing helps to spread that concentration.

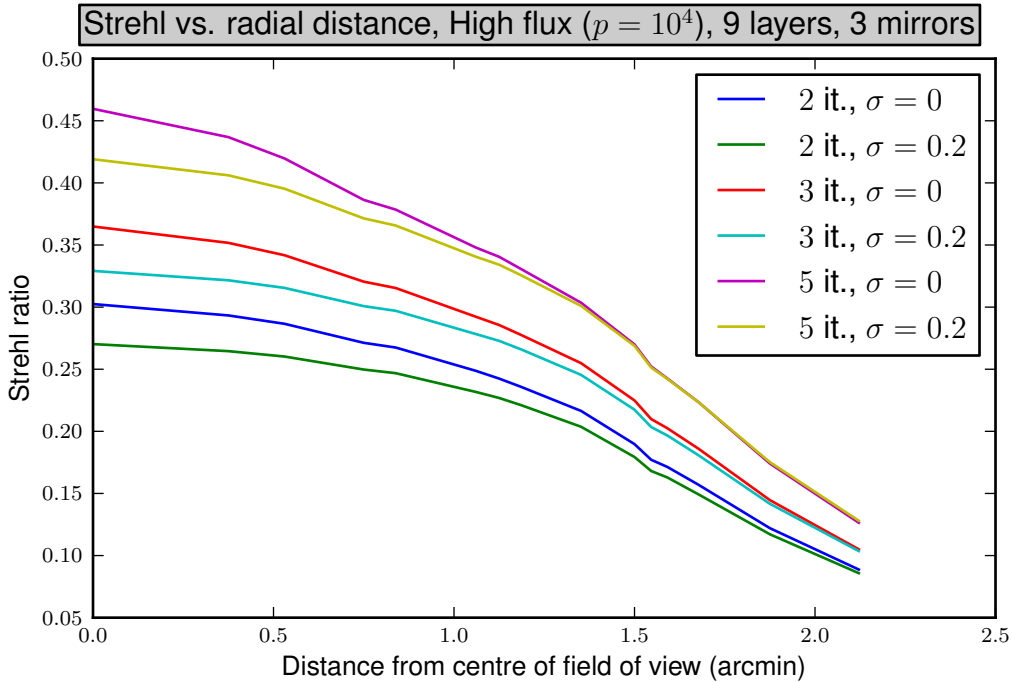


Figure 4.11: Strehl ratios plotted versus distance from the centre of the field of view for a high flux ($p = 10^4$) observation. Here, a 9 layers atmosphere was reconstructed.

Since atmospheric tomography is an ill-posed problem, successive iterations do not improve the result after a certain point. This can clearly be seen with the 50 and 100 iteration cases in Figure 4.19. Here, smoothing has countered the worsening of the reconstruction. Related to the discussion in the previous section about the importance of accurate a-priori information on the structure of the atmosphere, notice however how in the more accurate 9 layers reconstruction of Figure 4.20 even high iterates still improve the reconstruction quality.

4.2.6 Photon Flux 2

We now take one more look at photon flux induced noise, combining elements of the test cases from Sections 4.2.3, 4.2.4 and 4.2.5.

In Figures 4.22, 4.23, 4.24 and 4.25, we have plotted the average Strehl ratios for four specific reconstructors for multiple noise levels. The reconstructors are 5 iterations smoothed and not-smoothed and 20 iterations smoothed and not-smoothed. In Figures 4.22 and 4.24 the average Strehl ratio over the entire field of view was measured, whereas in Figures 4.23 and 4.25 only the more interesting inner field of view of approximately 0.53 arcminutes radius around the centre (w.r.t. Figure 4.1) was taken into account.

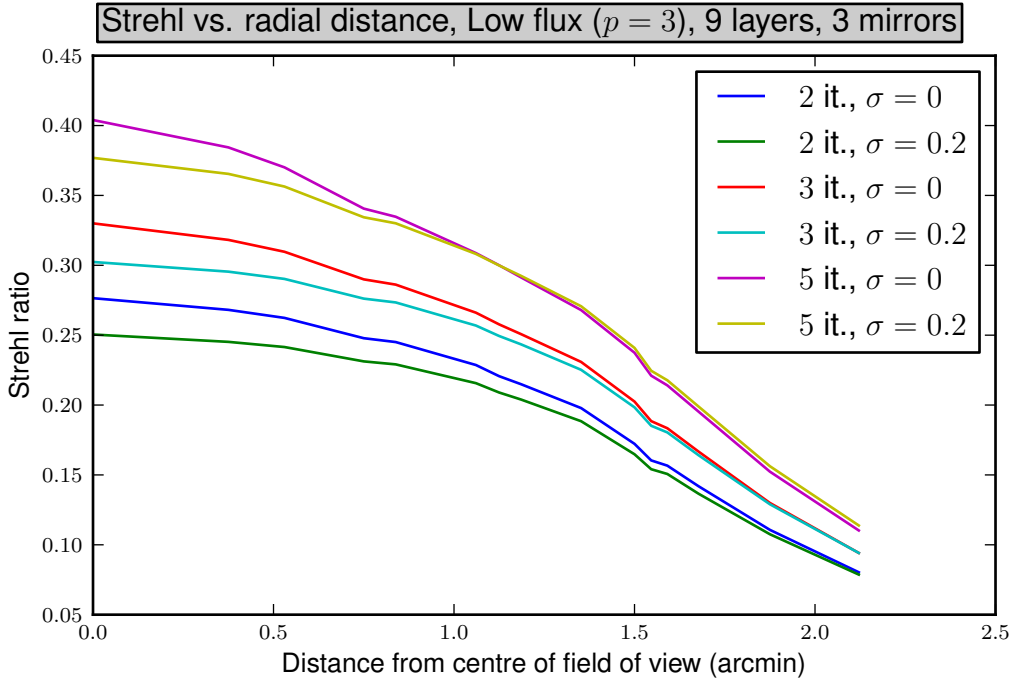


Figure 4.12: Strehl ratios plotted versus distance from the centre of the field of view for a low flux ($p = 3$) observation. Here, a 9 layers atmosphere was reconstructed.

These tests were run for two different reconstruction cases: in the first a 3 layers atmosphere was reconstructed and 3 matching deformable mirrors were used. In the second case the full 9 layers atmospheric model was reconstructed and 9 matching deformable mirrors were employed (i.e. we are measuring the quality of the tomographic reconstruction, ignoring mirror-fitting).

We again see that the smoothing helps only in low flux cases (which are the relevant ones in MCAO anyway). We also again see how due to the ill-posed nature of the tomographic problem, additional iterates can actually degrade reconstruction quality, as for example in Figure 4.22.

4.3 Interpretation of the Results

In this chapter, we have investigated a number of different test scenarios for an MCAO reconstructor. While it did not generally improve reconstruction results, it has however shown to be a handy addition to the Landweber-Kaczmarz iteration in high flux ($p = 10^4$) scenarios. In actual operation, many more factors contribute to noise and inaccuracies in the MCAO system, which warrants further investigation of the role of smoothing in MCAO.

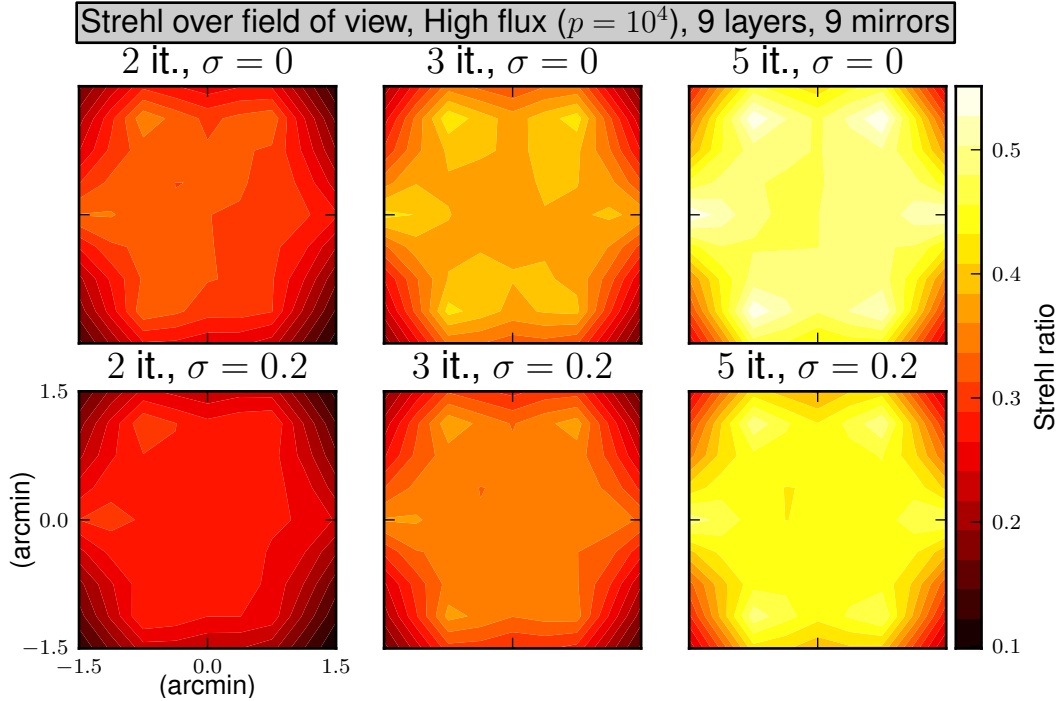


Figure 4.13: Strehl maps for a high flux ($p = 10^4$) observation with 9 reconstructed layers and 9 deformable mirrors.

The results so far indicate that smoothing is generally harmful to low-iteration reconstructions. For ill-posed problems, such as MCAO, the smoothing seems to just throw away too much information early on in the algorithm.

However, it did have a positive effect on the reconstruction for large iteration numbers, where it prevents the algorithm from focusing too narrowly on the guide-stars. In some cases, we have seen the same effect even for low iteration numbers, where the smoothing could be used to spread the correction effort over a larger field of view.

There are a number of extensions and refinements to possibly improve the performance of the H^1 Landweber-Kaczmarz method, which we will discuss in the next chapter.

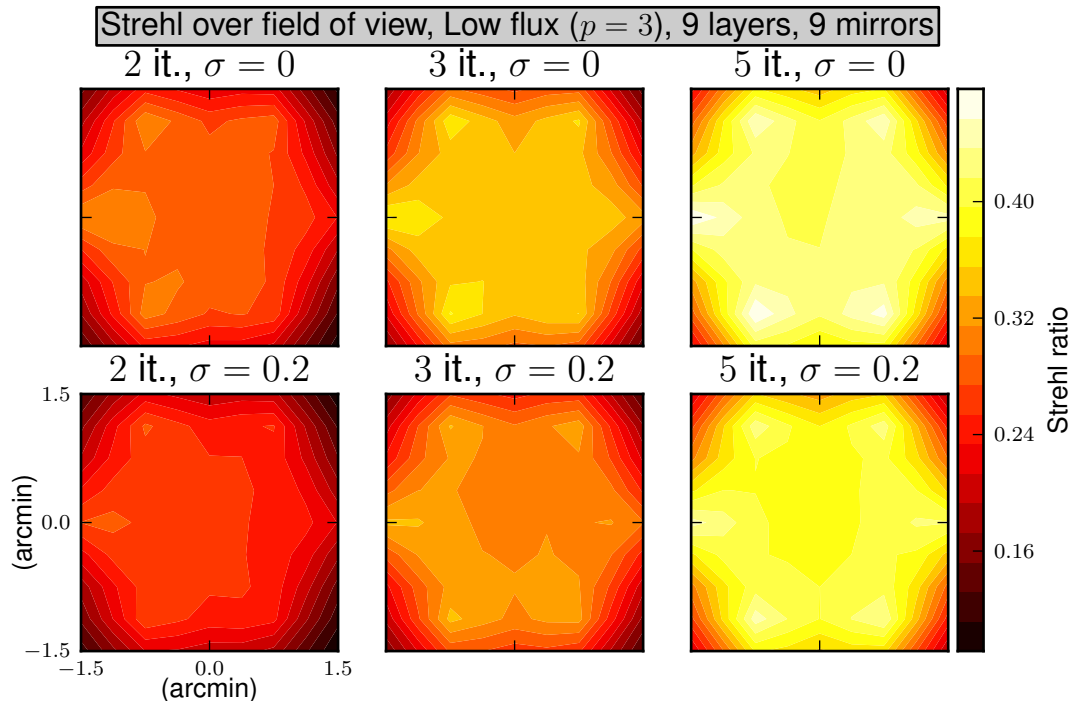


Figure 4.14: Strehl maps for a low flux ($p = 3$) observation with 9 reconstructed layers and 9 deformable mirrors

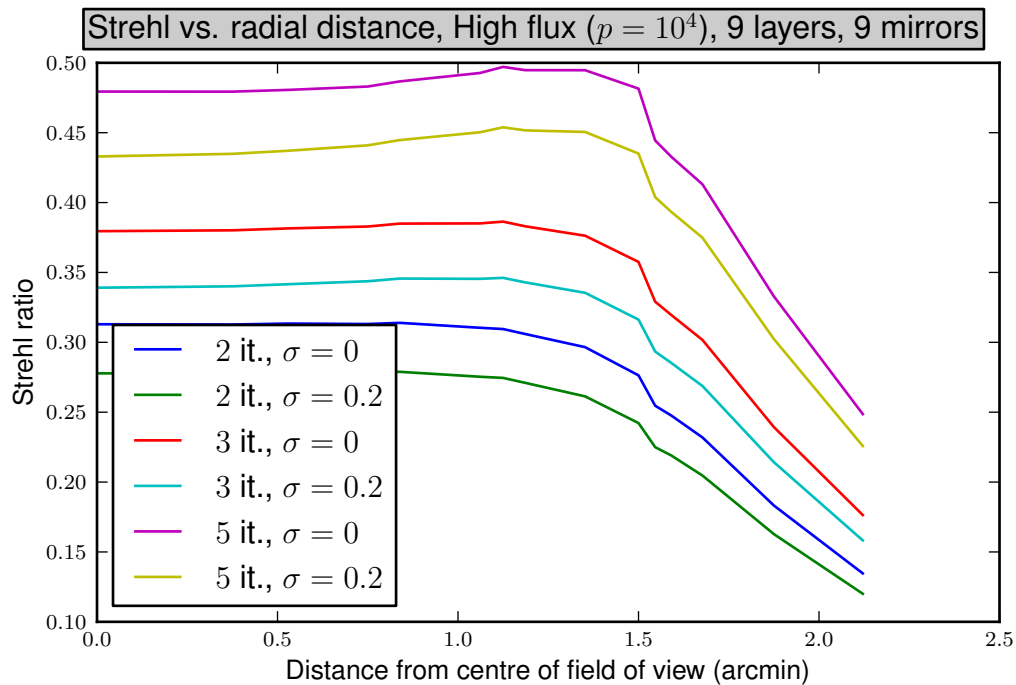


Figure 4.15: Strehl ratios plotted versus distance from the centre of the field of view for a high flux ($p = 10^4$) observation with 9 reconstructed layers and 9 mirrors.

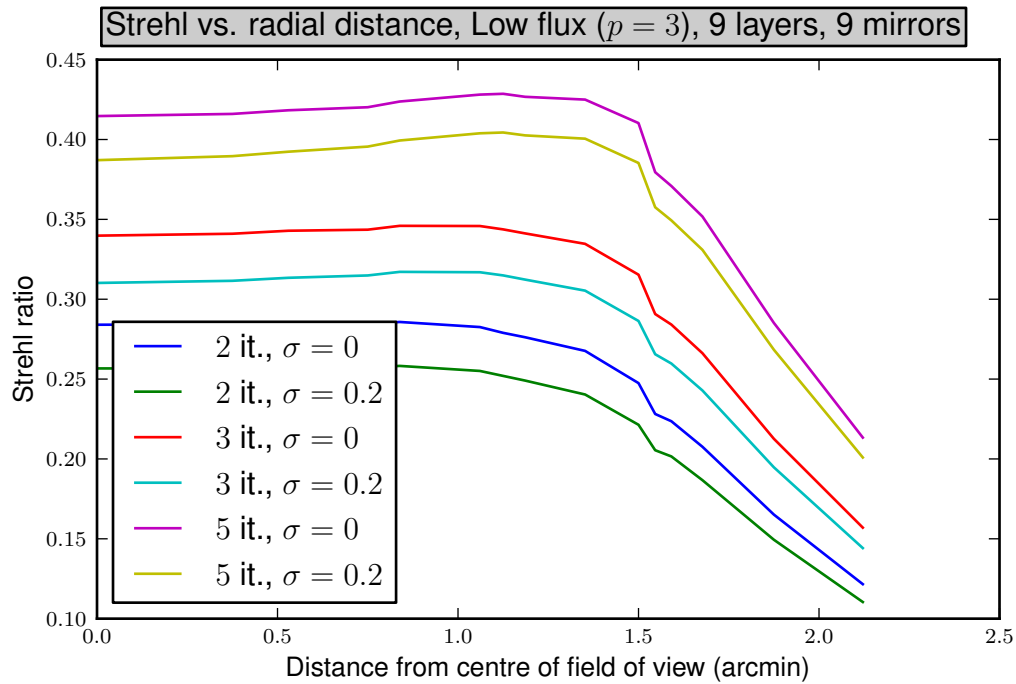


Figure 4.16: Strehl ratios plotted versus distance from the centre of the field of view for a low flux ($p = 3$) observation with 9 reconstructed layers and 9 mirrors.

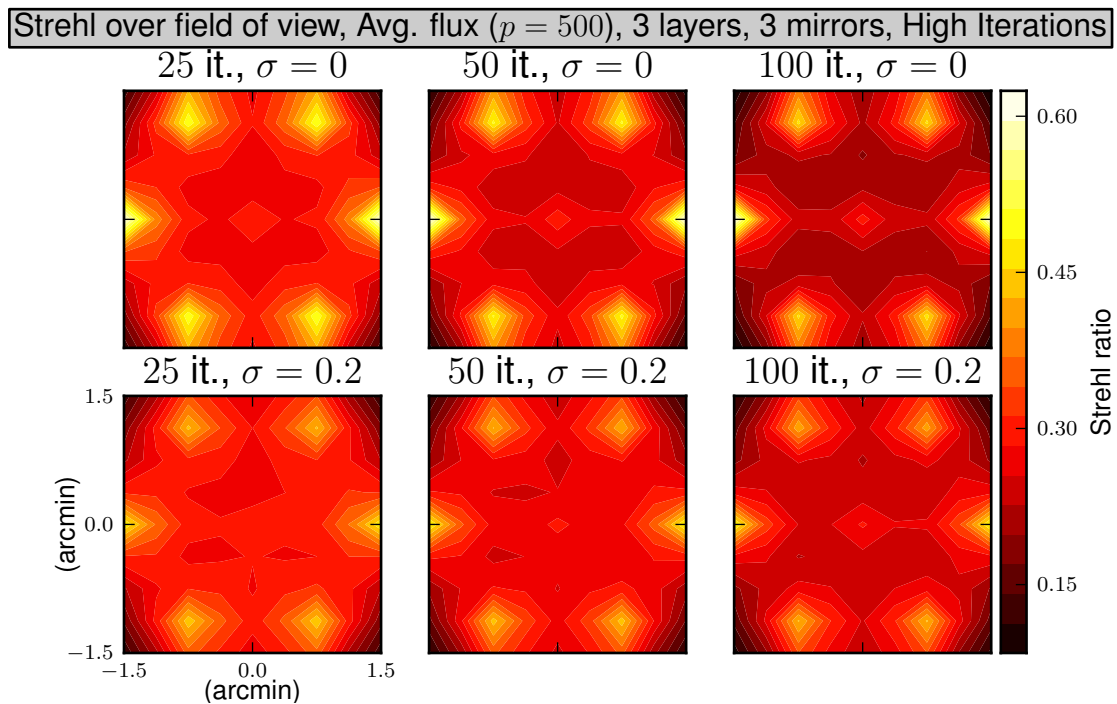


Figure 4.17: Strehl maps for an observation with 3 reconstructed layers, average flux ($p = 500$) and high iteration numbers.

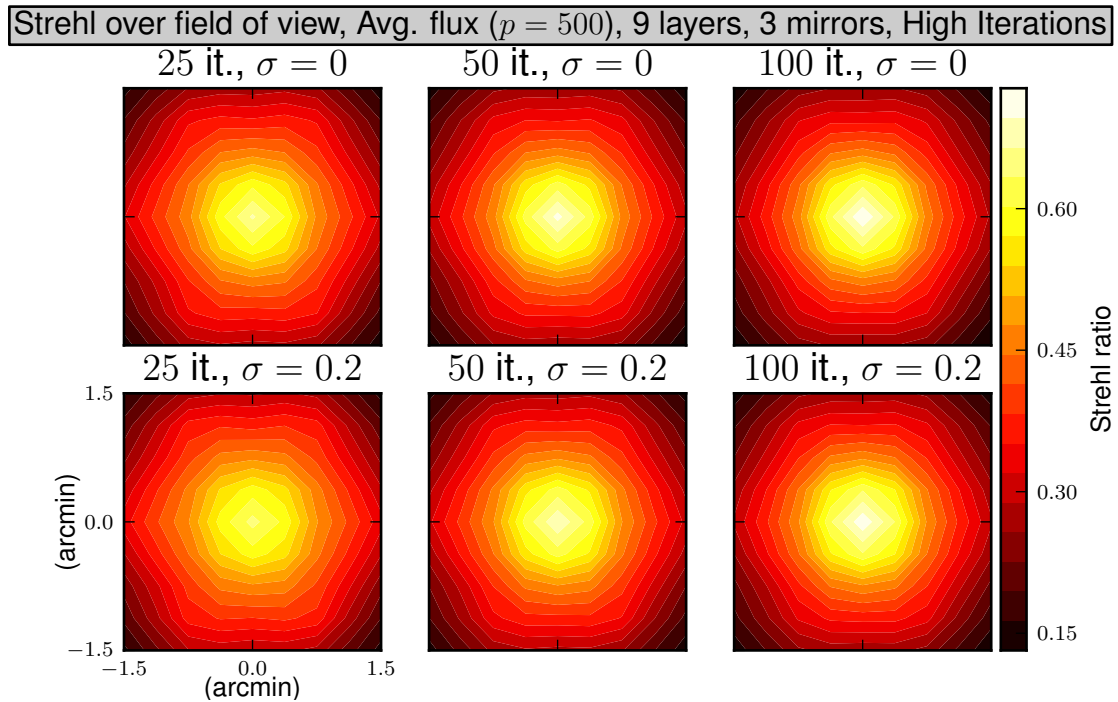


Figure 4.18: Strehl maps for an observation with 9 reconstructed layers, average flux ($p = 500$) and high iteration numbers.

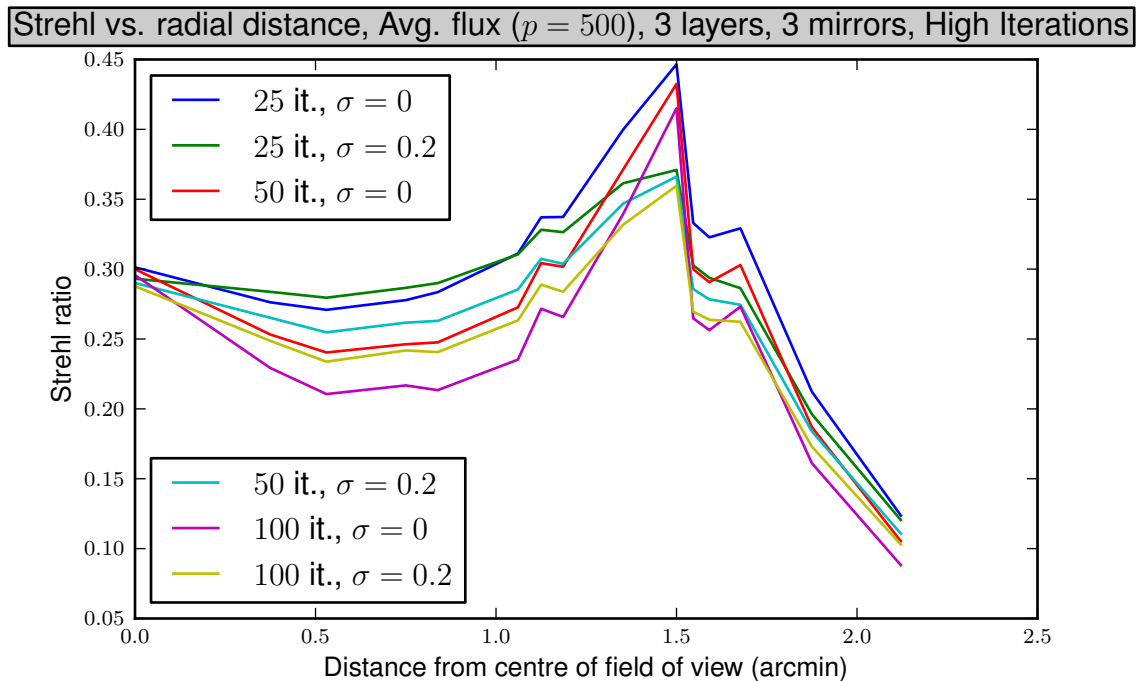


Figure 4.19: Strehl ratios plotted versus distance from the centre of the field of view for 3 reconstructed layers, average flux ($p = 500$) and high iteration numbers.

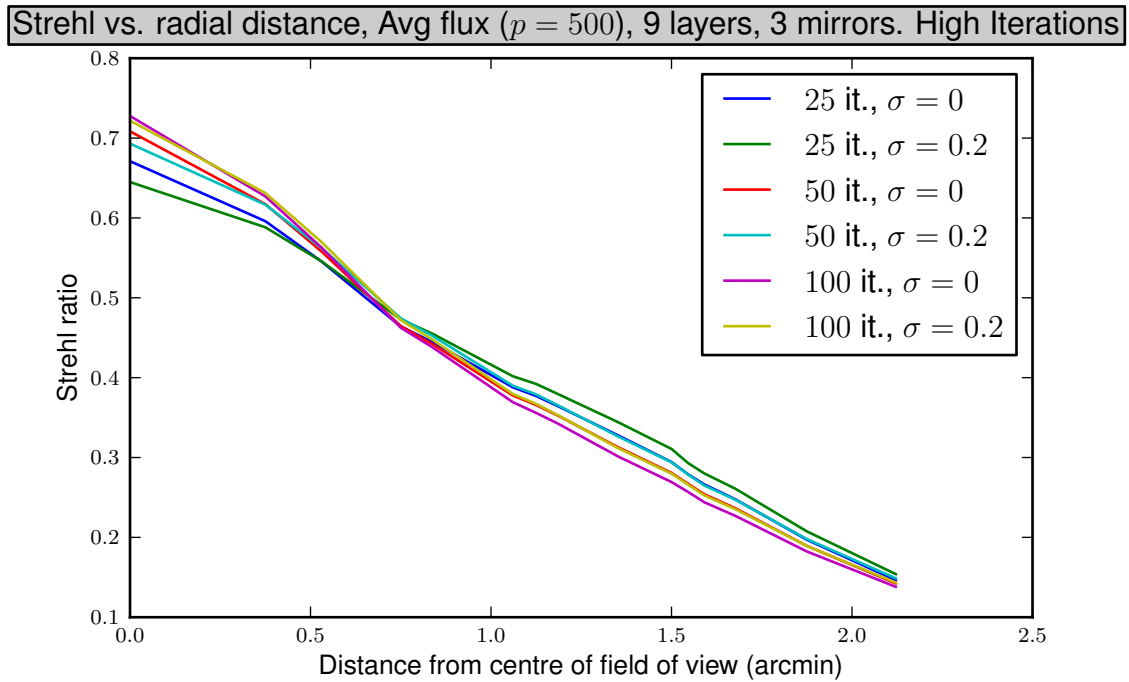


Figure 4.20: Strehl ratios plotted versus distance from the centre of the field of view for 9 reconstructed layers, average flux ($p = 500$) and high iteration numbers.

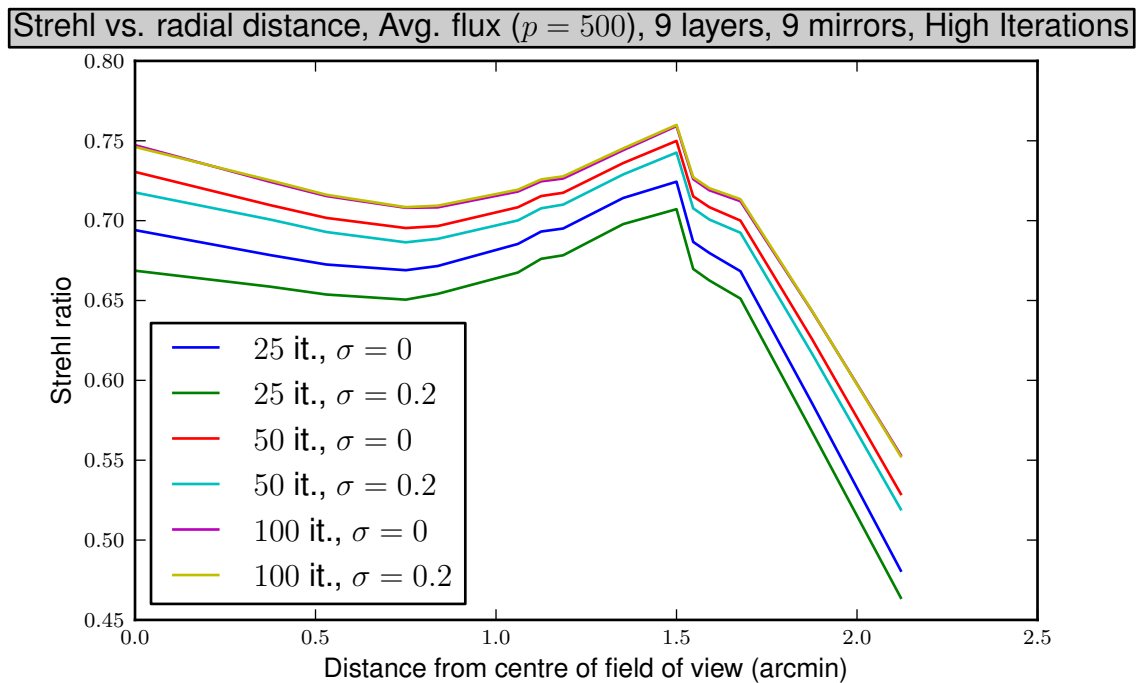


Figure 4.21: Strehl ratios plotted versus distance from the centre of the field of view for 9 reconstructed layers, 9 mirrors, average flux ($p = 500$) and high iteration numbers.

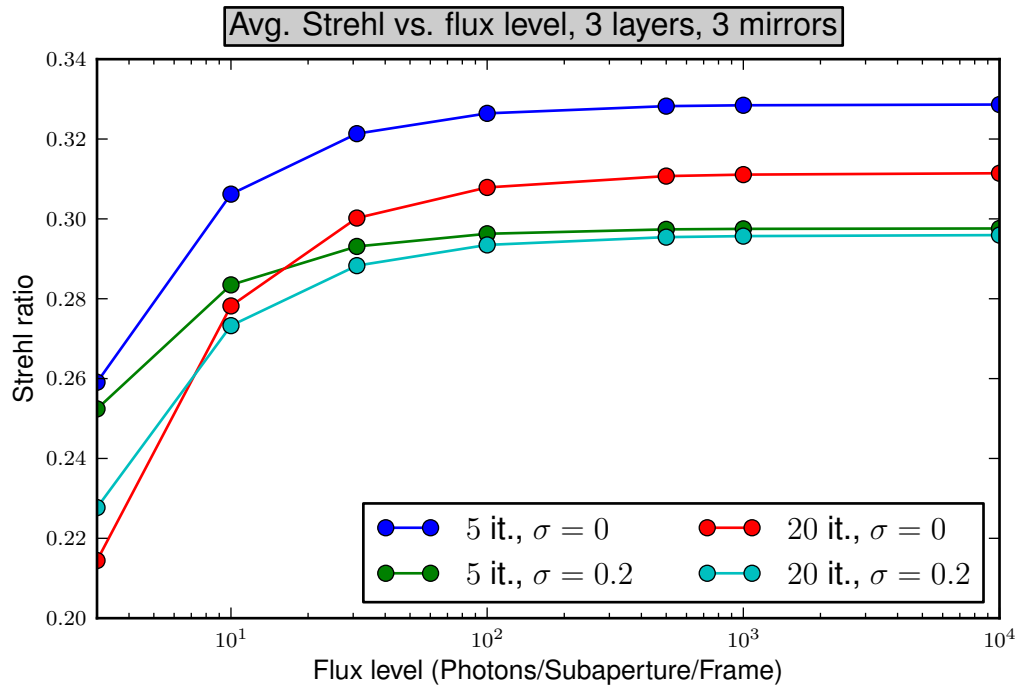


Figure 4.22: Strehl ratios (Averaged over the entire field of view) plotted versus photon flux level for a 3 layers reconstruction.

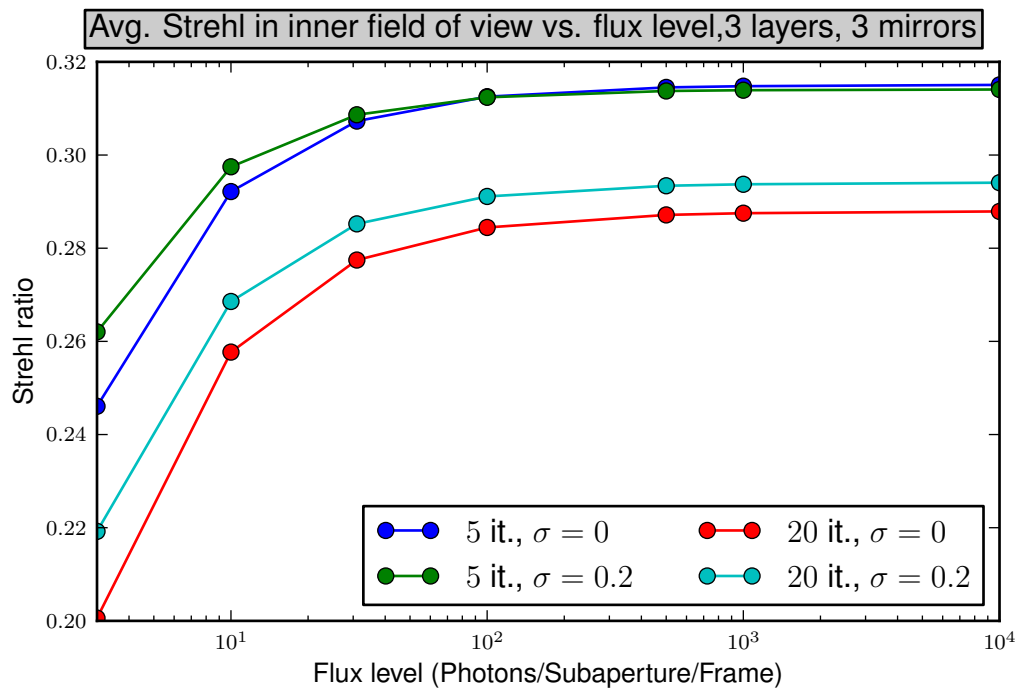


Figure 4.23: Strehl ratios (Averaged over the entire field of view) plotted versus photon flux level for a 3 layers reconstruction.

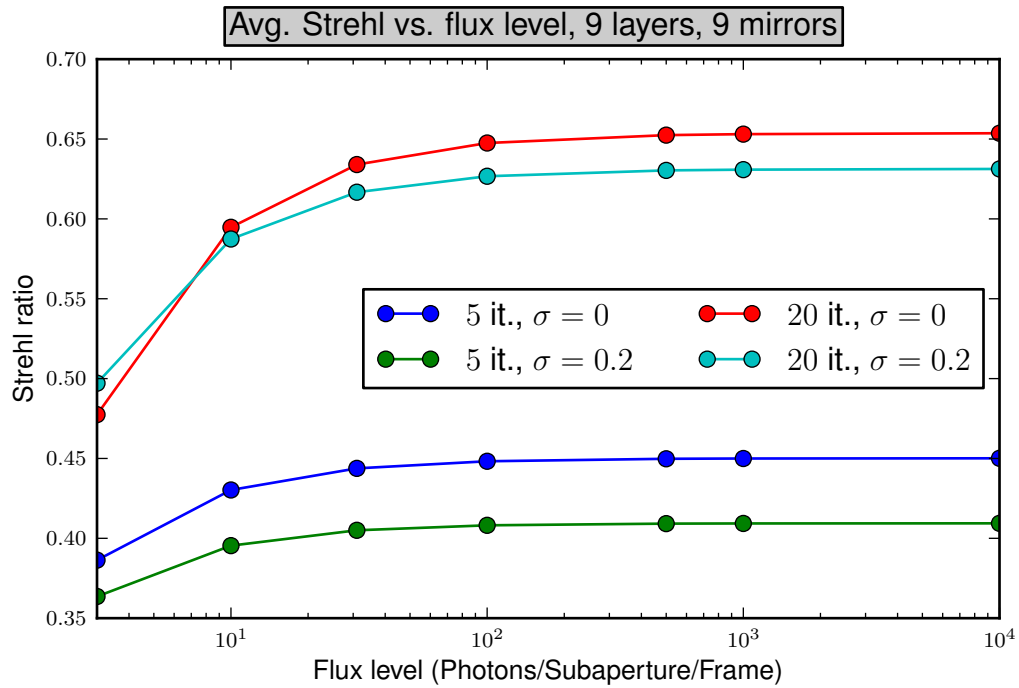


Figure 4.24: Strehl ratios (Averaged over the entire field of view) plotted versus photon flux level for a 9 layers reconstruction.

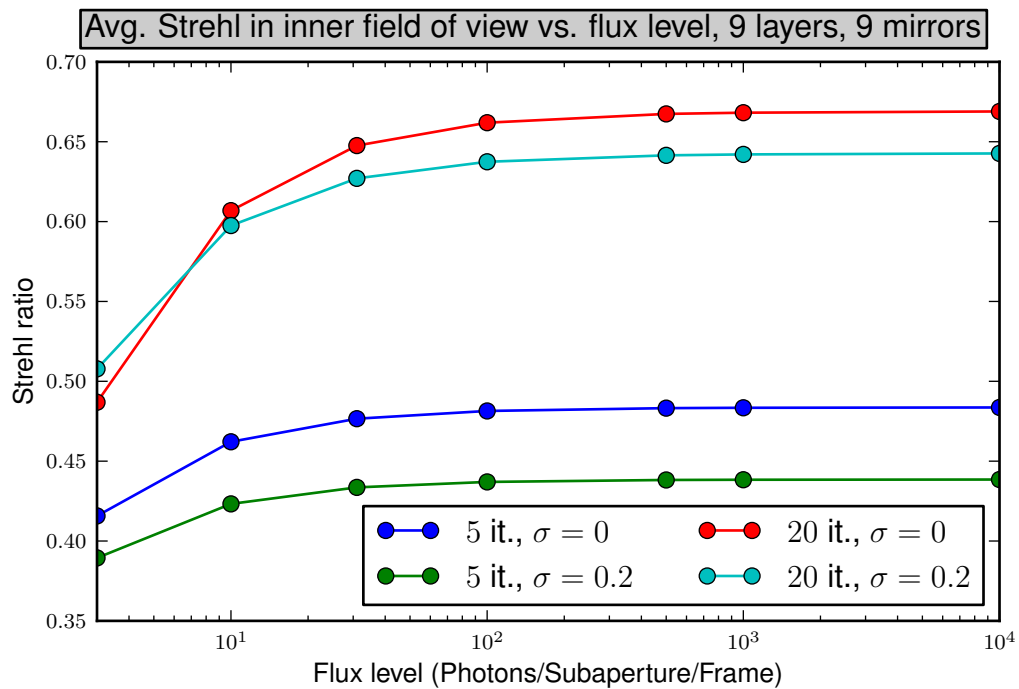


Figure 4.25: Strehl ratios (Averaged over the entire field of view) plotted versus photon flux level for a 9 layers reconstruction.

Chapter 5

Conclusion

We have investigated a Landweber-Kaczmarz method in an H^1 setting. This was accomplished by incorporating a smoothing operation into the iteration.

Numerical results

While the smoothing does help for large iteration numbers, it is in most cases detrimental to early iterates, presumably due to loss of too much information. Since MCAO systems are typically run in closed loop operation with only a few iterations, this is an undesirable behaviour.

However, for high noise levels and little or inaccurate a-priori information, the smoothing augmented Landweber-Kaczmarz iteration did improve reconstruction results. Due to a large number of additional error-introducing factors such as artificial guide-stars and varying guide-star layouts, smoothing could end up being an important part of atmospheric tomography for Multi Conjugate Adaptive Optics.

There is ample space for further study of such a smoothing approach. Some possible topics are:

Smoothing factors Could the smoothing factor σ be chosen more optimally, particularly different values for different atmospheric layers? Analogously: Should the smoothing be done only for specific (more/less turbulent) layers?

When to smooth In Chapter 4 we have, in strict accordance to the theoretical derivation of the smoothing, executed the smoothing for every single update step in the Landweber-Kaczmarz iteration. Applying the smoothing only after a number of iterations or possibly only at the end of the non-smoothed L^2 iteration might improve quality of the solution.

Closed-Loop Operation Chapter 4 focused on open-loop Adaptive Optics. A possible role of smoothing in closed-loop operations should still be investigated.

Laser Guide Stars In this thesis we have only used Natural Guide Stars. Laser Guide Stars (LGS) introduce additional errors into the reconstruction. Considering the good performance of smoothing in the photon-noise test-cases in the previous Chapter, the H^1 Landweber-Kaczmarz method might be particularly well suited for LGSs, which warrants further investigations.

Regularity results

From a theoretical viewpoint, we have adapted an approach developed in Grisvard [1985] and Grisvard [1992] to proving smoothness results for solutions of partial differential equations to equations of the form

$$\begin{aligned} u - \sigma \Delta u &= f \text{ in } \Omega \\ \frac{\partial u}{\partial \mathbf{n}} &= 0 \text{ on } \partial\Omega, \end{aligned}$$

for bounded Lipschitz domains in \mathbb{R}^2 with piecewise C^1 boundaries. This has resulted in three closely related conjectures concerning norm-bounds on Sobolev spaces on Lipschitz polygons.

Possible further work in this area could study the applicability of these proof-techniques to more general elliptic operators, as indicated by Grisvard [1992].

List of Figures

2.1	Concept of a deformable mirror.	5
2.2	Schematics of an Multi Conjugate Adaptive Optics system. Note the overlapping volumes covered by the wavefront sensors.	7
2.3	Simulations of Kolmogorov turbulence. Note the periodicity, which stems from the use of the Fourier-transformation in the simulation (based on (2.2)).	9
2.4	Mathematical notation for a model of atmospheric tomography.	11
2.5	Overlapping apertures for 6 guide stars distributed in a 3 arc minutes field of view. The inner circular hole stems from a central obstruction in the telescope's optical pathway. Even on the lowest layer at $h = 47$ meters, the apertures are not precisely overlapping.	17
3.1	Corners and interior angles of (a) a Lipschitz Polygon and (b) a Lipschitz domain with a piecewise C^1 boundary.	20
3.2	Example showing the construction of domain approximations as formulated in Conjecture 3.8.	25
3.3	Regularity of the solution of (3.1) on the domain depicted on the upper plot. Here d is the distance between the centers of two circles with equal radii. The value s in the second plot indicates in which space $H^s(\Omega)$ the solution of the partial elliptic equation we investigate in this chapter lies, when considered on the corresponding overlapping circles domain.	30
3.4	Plots of I_λ for various values of λ	31
3.5	Plots of K_λ for various values of λ	32
3.6	Smoothness estimates for solutions of (3.1) on different domains. The heights correspond to the ones in Figure 2.5. Compare these indices s with the decay power $11/6 \approx 1.83$ in the von Karman turbulence model (2.3).	40
4.1	Angular placement of views (blue circles) and guide stars (red stars) for a 3 arcminutes field of view.	44
4.2	An atmosphere consisting of three layers and their reconstructions using 10 iterations of the Landweber-Kaczmarz iteration.	45

4.3 Highest layer of a 3 layers atmosphere and reconstructions using 10 iterations of the Landweber-Kaczmarz iteration with different smoothing strengths. 46

4.4 Strehl ratios and H^1 norms for the different reconstructions seen in Figure 4.3. Two different noise cases (see Section 4.2.3) are covered: *Highflux*, i.e. bright guide-stars and *Lowflux*, i.e. dim guide stars. For the Lowflux case the ratio improves for small smoothing strengths and then quickly degenerates after a value of $\sigma = 1.$. In the Highflux case the smoothing generally impairs reconstruction quality. 47

4.5 Strehl maps for a high flux ($p = 10^4$) observation with 3 reconstructed layers. 48

4.6 Strehl maps for a low flux ($p = 3$) observation with 3 reconstructed layers. 48

4.7 Strehl ratios plotted versus distance from the centre of the field of view for a high flux ($p = 10^4$) observation. Here, a 3 layers atmosphere was reconstructed. 49

4.8 Strehl ratios plotted versus distance from the centre of the field of view for a low flux ($p = 3$) observation. Here, a 3 layers atmosphere was reconstructed. 49

4.9 Strehl maps for a high flux ($p = 10^4$) observation with 9 reconstructed layers and 3 deformable mirrors. Note the high correction in the centre of the field of view. 50

4.10 Strehl maps for a low flux ($p = 3$) observation with 9 reconstructed layers 51

4.11 Strehl ratios plotted versus distance from the centre of the field of view for a high flux ($p = 10^4$) observation. Here, a 9 layers atmosphere was reconstructed. 52

4.12 Strehl ratios plotted versus distance from the centre of the field of view for a low flux ($p = 3$) observation. Here, a 9 layers atmosphere was reconstructed. 53

4.13 Strehl maps for a high flux ($p = 10^4$) observation with 9 reconstructed layers and 9 deformable mirrors. 54

4.14 Strehl maps for a low flux ($p = 3$) observation with 9 reconstructed layers and 9 deformable mirrors 55

4.15 Strehl ratios plotted versus distance from the centre of the field of view for a high flux ($p = 10^4$) observation with 9 reconstructed layers and 9 mirrors. 55

4.16 Strehl ratios plotted versus distance from the centre of the field of view for a low flux ($p = 3$) observation with 9 reconstructed layers and 9 mirrors. 56

4.17 Strehl maps for an observation with 3 reconstructed layers, average flux ($p = 500$) and high iteration numbers. 56

4.18 Strehl maps for an observation with 9 reconstructed layers, average flux ($p = 500$) and high iteration numbers. 57

4.19	Strehl ratios plotted versus distance from the centre of the field of view for 3 reconstructed layers, average flux ($p = 500$) and high iteration numbers.	57
4.20	Strehl ratios plotted versus distance from the centre of the field of view for 9 reconstructed layers, average flux ($p = 500$) and high iteration numbers.	58
4.21	Strehl ratios plotted versus distance from the centre of the field of view for 9 reconstructed layers, 9 mirrors, average flux ($p = 500$) and high iteration numbers.	58
4.22	Strehl ratios (Averaged over the entire field of view) plotted versus photon flux level for a 3 layers reconstruction.	59
4.23	Strehl ratios (Averaged over the entire field of view) plotted versus photon flux level for a 3 layers reconstruction.	59
4.24	Strehl ratios (Averaged over the entire field of view) plotted versus photon flux level for a 9 layers reconstruction.	60
4.25	Strehl ratios (Averaged over the entire field of view) plotted versus photon flux level for a 9 layers reconstruction.	60

Bibliography

- A. Tokovinin, M. Le Louarn, E. Viard, N. Hubin, and R. Conan. Optimized modal tomography in adaptive optics. *A&A*, 378(2):710–721, 2001.
- M. Abramowitz and I.A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Number v. 55, no. 1972 in Applied mathematics series. U.S. Govt. Print. Off., 1964. ISBN 9780486612720.
- R.A. Adams and J.J.F. Fournier. *Sobolev spaces*. Number v. 140 in Pure and applied mathematics. Academic Press, 2003. ISBN 9780120441433.
- F. E. Browder. The Dirichlet Problem for Linear Elliptic Equations of Arbitrary Even Order with Variable Coefficients. *Proceedings of the National Academy of Science*, 38:230–235, March 1952.
- P.P.B. Eggermont, G.T. Herman, and A. Lent. Iterative algorithms for large partitioned linear systems, with applications to image reconstruction. *Linear Algebra and its Applications*, 40(0):37 – 67, 1981. ISSN 0024-3795.
- R. Foy and F.C. Foy. *Optics in astrophysics*. NATO science series: Mathematics, physics, and chemistry. Springer, 2005. ISBN 9781402034367.
- D. T. Gavel. Tomography for multiconjugate adaptive optics systems using laser guide stars. In D. Bonaccini Calia, B. L. Ellerbroek, & R. Ragazzoni, editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 5490 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 1356–1373, October 2004.
- L. Gilles, B. L. Ellerbroek, and C. R. Vogel. Layer-oriented multigrid wavefront reconstruction algorithms for multi-conjugate adaptive optics. In P. L. Wizinowich & D. Bonaccini, editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 4839 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 1011–1022, February 2003.
- Luc Gilles and Brent L. Ellerbroek. Split atmospheric tomography using laser and natural guide stars. *J. Opt. Soc. Am. A*, 25(10):2427–2435, Oct 2008.

- Luc Gilles, Curtis R. Vogel, and Brent L. Ellerbroek. Multigrid preconditioned conjugate-gradient method for large-scale wave-front reconstruction. *J. Opt. Soc. Am. A*, 19(9):1817–1822, Sep 2002.
- Luc Gilles, Brent Ellerbroek, and Curtis Vogel. A comparison of multigrid v-cycle versus fourier domain preconditioning for laser guide star atmospheric tomography. In *Adaptive Optics: Analysis and Methods/Computational Optical Sensing and Imaging/Information Photonics/Signal Recovery and Synthesis Topical Meetings on CD-ROM*. Optical Society of America, 2007.
- A. Glindemann, S. Hippler, T. Berkefeld, and W. Hackenberg. Adaptive optics on large telescopes. *Experimental Astronomy*, 10(1):5–47, 2000.
- P. Grisvard. *Elliptic problems in nonsmooth domains*. Monographs and studies in mathematics. Pitman Advanced Pub. Program, 1985. ISBN 9780273086475.
- P. Grisvard. *Singularities in boundary value problems*. Recherches en mathématiques appliquées. Masson, 1992. ISBN 9782225827709.
- M. Haltmeier, R. Kowar, A. Leitão, and O. Scherzer. Kaczmarz methods for regularizing nonlinear ill-posed equations ii: Applications. *Inverse Problems and Imaging*, 1(3):507–523, 2007a.
- M. Haltmeier, A. Leitão, and O. Scherzer. Kaczmarz methods for regularizing nonlinear ill-posed equations i: Convergence analysis. *Inverse Problems and Imaging*, 1(2):289–298, 2007b.
- J.W. Hardy. *Adaptive optics for astronomical telescopes*. Oxford series in optical and imaging sciences. Oxford University Press, 1998. ISBN 9780195090192.
- A. Jeffrey and H.H. Dai. *Handbook of mathematical formulas and integrals*. Elsevier Academic Press, 2008. ISBN 9780123742889.
- S. Kaczmarz. Angenäherte Auflösung von Systemen linearer Gleichungen. *Bulletin International de l'Académie Polonaise des Sciences et des Lettres*, 35:355–357, 1937.
- R Lane, A Glindemann, and J Dainty. Simulation of a kolmogorov phase screen. *Waves in Random Media*, 2(3):209–224, 1992.
- Enrico Marchetti, Roland Brast, Bernard Delabre, Robert Donaldson, Enrico Fedrigo, Christoph Frank, Norbert Hubin, Johann Kolb, Jean-Louis Lizon, Massimiliano Marchesi, Sylvain Oberti, Roland Reiss, Joana Santos, Christian Soenke, Sebastien Tordo, Andrea Baruffolo, Paolo Bagnara, and The CAMCAO consortium. On-sky testing of the multi-conjugate adaptive optics demonstrator. *The Messenger*, 189, 2007.
- W.C.H. McLean. *Strongly elliptic systems and boundary integral equations*. Cambridge University Press, 2000. ISBN 9780521663755.

- Frank Natterer. Algorithms in tomography. 2009.
- A. Quirrenbach. Observing Through the Turbulent Atmosphere. In P. R. Lawson, editor, *Principles of Long Baseline Stellar Interferometry*, page 71, 2000.
- R. Ragazzoni, E. Marchetti, and F. Rigaut. Modal tomography for adaptive optics. *Astronomy and Astrophysics*, 342:L53–L56, February 1999.
- Ronny Ramlau and Matthias Rosensteiner. A kaczmarz type reconstructor for mcao data. 2011. Submitted for publication.
- Ronny Ramlau and Gerd Teschke. Regularization of Sobolev embedding operators and applications. I. Fourier and wavelet based methods. *Sampl. Theory Signal Image Process.*, 3(2):175–196, 2004a. ISSN 1530-6429.
- Ronny Ramlau and Gerd Teschke. Regularization of Sobolev embedding operators and applications. II. Data driven regularization and applications. *Sampl. Theory Signal Image Process.*, 3(3):225–246, 2004b.
- F. Roddier. The effects of atmospheric turbulence in optical astronomy. *Progress in optics. Volume 19. Amsterdam, North-Holland Publishing Co., 1981, p. 281-376.*, 19:281–376, 1981.
- F. Roddier. *Adaptive optics in astronomy*. Cambridge University Press, 1999. ISBN 9780521553759.
- Matthias Rosensteiner. Cumulative reconstructor: fast wavefront reconstruction algorithm for extremely large telescopes. *J. Opt. Soc. Am. A*, 28(10):2132–2138, Oct 2011.
- T. Sean Ross. Limitations and applicability of the maréchal approximation. *Appl. Opt.*, 48(10):1812–1818, Apr 2009.
- M. Tallon and R. Foy. Adaptive telescope with laser probe - Isoplanatism and cone effect. *Astronomy and Astrophysics*, 235:549–557, August 1990.
- Eric Thiébaud and Michel Tallon. Fast minimum variance wavefront reconstruction for extremely large telescopes. *J. Opt. Soc. Am. A*, 27(5):1046–1059, May 2010.
- Andrei Tokovinin and Elise Viard. Limiting precision of tomographic phase estimation. *J. Opt. Soc. Am. A*, 18(4):873–882, Apr 2001.
- A M Vorontsov, P V Paramonov, M T Valley, and Mikhail A Vorontsov. Generation of infinitely long phase screens for modeling of optical wave propagation in atmospheric turbulence. *Waves In Random And Complex Media*, 18(1):91–108, 2008.

- Q. Yang and C. R. Vogel. Fourier domain algorithm for the fitting step in multi-conjugate adaptive optics. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 6272 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, July 2006.
- Qiang Yang, Curtis R. Vogel, and Brent L. Ellerbroek. Fourier domain preconditioned conjugate gradient algorithm for atmospheric tomography. *Appl. Opt.*, 45(21):5281–5293, Jul 2006.
- Maria Zhariy, Andreas Neubauer, Matthias Rosensteiner, and Ronny Ramlau. Cumulative wavefront reconstructor for the shack-hartmann sensor. *Inverse Problems and Imaging*, 5(4):893–913, 2011.

Eidesstattliche Erklärung

Ich, Markus Eslitzbichler, erkläre an Eides statt, dass ich die vorliegende Masterarbeit selbständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Linz, Januar 2012

Markus Eslitzbichler

Markus Eslitzbichler

Contact m.eslitzbichler@gmail.com, +43 650 9213142
Address Freistädterstraße 205, A-4040 Linz, Austria

Nationality Austrian
Date of birth February 1st, 1987

Education

October 2010 to January 2012 *Mathematics for Natural Sciences (Master)*
Thesis: “A Sobolev Space Smoother
for Atmospheric Tomography in Multi Conjugate Adaptive Optics”,
supervised by Univ. Prof. Dr. Ronny Ramlau and Dipl.-Ing. Dr. Clemens Pechstein.
Courses focused on functional analysis, integral equations and inverse problems,
geometry and stochastic processes.
Johannes Kepler Universität Linz, Austria

January to June 2011 *Erasmus Exchange semester*
Courses in applied mathematics and astrophysics.
Lunds Tekniska Högskola, Lund, Sweden

September 2010 *ECMI Modelling Week*
Project: “ER surveys of burial mounds”,
supervised by Andreas Asheim (Norges teknisk-naturvitenskapelige universitet, NTNU).
Università degli Studi di Milano, Italy

October 2007 to October 2010 *Technical Mathematics (Bachelor)*
Graduated with distinction.
Thesis: “Das Lemma von Pták in der Funktionalanalysis”,
supervised by Univ. Doz. Dr. Eva Kopecká.
Received merit-based scholarships 2008, 2009 and 2010.
Courses focused on numerical methods and mathematical modelling.
Johannes Kepler Universität Linz, Austria

September 2001 to Juni 2006 *Informatics and Software Engineering*
Graduated with distinction.
Final project: “Tracking of Hand Movements”
Höhere Technische Bundeslehranstalt Leonding, Austria
(Technical college, entitles to enter university)

Work Experience

Convergent Information Technologies GmbH *June 2007 to Present*
Software Developer, Robotics
Im Stadtgut A1, A-4407 Steyr-Gleink, Austria
Planned and developed user-friendly applications for programming industrial robots. Designed and implemented mission-critical software components for production environments in the automobile industry. Managed a team of programmers. Represented company management in meetings with customers, partners and investors, at industry fairs and at conferences. Interviewed and evaluated job applicants.

Echelon Software *February to May 2007*
Software Developer, Video games
34-01 38th Ave. Astoria NY 11101, USA
Developer in a 90 day prototype project. In charge of the implementation of gameplay features. Instrumental in analyzing and evaluating features and capacities of licensed technology. Coordinated design, graphics and programming departments. The project was successful and led to further funding negotiations with publishers.

AeroSpy Sense & Avoid Technology GmbH *November 2005 to June 2007*
Freelance Software Developer, Robotics
Hafenstr. 47-51, A-4020 Linz, Austria
Development of simulation, navigation and image processing software for Unmanned Aerial Vehicles (UAVs). This was the basis for a software package that is today sold to UAV producers for sense and avoid tasks.

Skills

Software Engineering Leadership experience in all stages of software project development, from requirements analysis to deployment and support.

Programming Experience in multiple programming languages and programming paradigms. Worked both on software architecture as well as on algorithm development.

Languages English (fluent), German (native)