From the Institute of Mathematics and Image Computing
of the University of Lübeck
Director: Prof. Dr. Jan Modersitzki

Dynamic Models and Coupling in Medical 4D Imaging

Dissertation
for Fulfillment of
Requirements
for the Doctoral Degree
of the University of Lübeck

from the Institute of Mathematics and Image Computing

Submitted by

Constantin Marius Sandmann
from Würzburg, Germany

Lübeck, 2016

Abstract

In this work we present different approaches to improve parameter estimation from voxel-wise medical 4D data using one-dimensional dynamic models. Whereas these models have been used successfully to describe generic regions of interest as e.g. renal parenchyma or the human brain, the transition to voxel-wise data can be challenging. Reasons for this include noise or patient motion during the image acquisition. In this work we describe methods to improve parameter estimation in such a setting. Results are demonstrated on phantom and on real data for three example problems (estimation of relaxation time T_1 in MRI, estimation of regional cerebral blood flow and estimation of renal function). We especially focus on the role of spatial coupling. To improve parameter estimation in the presence of noise, we extend parameter estimation by a class of spatial coupling terms, which is based on Schatten-p-norms of the Jacobian and was originally designed for RGB denoising. It is demonstrated that the novel methods can improve errors in T_1 estimation up to 8% as compared to established methods. We further present results which indicate that this coupling can also improve parameter estimation in a joint setting, where not only dynamic parameters but also motion or control parameters of the model are recovered. We conclude by demonstrating limits of established models for blood flow estimation in the case of highly developed capillary tissue and extend generic existing knowledge on limits of perfusion. Specifically, we simulate blood-flow through such a tissue patch and show that the estimated perfusion scales with the inverse voxel volume and thus leads to systematic overestimation.

Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit Strategien um die Schätzung voxelweiser Parameterkarten aus medizinischen 4D Daten mit Hilfe eindimensionaler dynamischer Modelle zu verbessern. Eindimensionale dynamische Modelle sind bereits etabliert um die dynamische Antwort größerer, generischer Regionen wie des Parenchyms der Niere oder des Gehirns zu beschreiben. Der Übergang auf Voxelgröße kommt jedoch mit Herausforderungen, da Bildrauschen und Patientenbewegung hier prominenter zu Tage treten. In dieser Arbeit präsentieren wir verschiedene Strategien, diese Herausforderungen zu bewältigen. Ergebnisse dieser Arbeit werden für drei verschiedene Modelle experimentell auf realen und auf Phantom Daten demonstriert: Schätzung der Filtrationsrate der Niere, Schätzung des regionalen cerebralen Blutflusses und Schätzung des MRT Relaxationsparameters T_1. Der Fokus dieser Arbeit liegt auf der Rolle räumlicher Kopplung. Um Robustheit gegen Rauschen zu verbessern, wird der etablierte Parameterschätzungsprozess um einen auf Schatten-p-Normen von Jacobimatrix basierenden Kopplungsterm aus der RGB-Bildentrauschung erweitert. Es wird demonstriert, dass mit diesem Term die T_1 Schätzung bei Rauschen um bis zu 8% verbessert werden kann. Weiterhin werden positive Effekte räumlicher Kopplung für Methoden demonstriert, bei denen neben dynamischen Parametern auch Bewegung oder Kontrollparameter der Modelle geschätzt werden. Abschließend werden Grenzen der Gültigkeit etablierter Modelle zur Blutflussschätzung aufgezeigt. Etabliertes Wissen um diese Grenzen wird erweitert, indem ein neuartiges Modell zur Kontrastmittelpropagation durch reines Kapillargewebe analysiert wird. Dort wird gezeigt, dass der rekonstruierte Blutfluss mit dem inversen Voxelvolumen skaliert und so zu einer systematischen Überschätzung führt.

Acknowledgments

My deepest gratitude goes to my supervisor, Jan Modersitzki, whose guidance and advice has shaped my thinking and thereby most of this work. He directed me both during hardship and success and I could always rely on his open and frank counsel. Without his innumerable suggestions and his perpetual patience, this thesis would not have been possible. Further thanks go to Erlend Hodneland and Erik Hanson from the University of Bergen. We enjoyed many fruitful discussions, gained further insights into the structure of perfusion and I was always greeted with warm hospitality on my visits to Norway. Also, I want to thank Benjamin Berkels, Martin Benning, Thomas Peter and Jan Lellmann, who broadened my horizon and improved this work with various suggestions. Further thanks go to the committee, Prof. Keller and Prof. Koch, who took the time to explain various technical details of MRI.

Also, I want to thank all members of the institute, many of whom grew close friends during the last years. Being part of the MIC and Fraunhofer MEVIS has been a special and unique experience. I apologize in advance for not contributing my weekly "DCE-MRI" to our morning rounds anymore.

Finally, I want to thank my wife Katja Sandmann and my daughter Henriette. Without them, without their patience and without their emotional support, this thesis would not have been possible. Last but not least I want to thank my parents Cornelius and Ulrike and my brother Julian, who always believed in me.

Contents

1 Introduction

In modern medicine there are various applications where quantitative information on velocities or rates are important indicators for diagnosis: In ischemic stroke, the amount of blood passing through brain tissue per unit time (perfusion) has been linked to a patient's potential for recovery [56, 158], in nephrology the total filtration rate of the kidneys (glomerular filtration rate, GFR) is a main indicator of kidney health and its decrease is linked to various malfunctions of the kidneys [78]. As the examples of GFR or perfusion of the brain show, such quantities are traditionally often referring to complete organs [52, 167, 78]. However, modern medical imaging methods are additionally enabling local estimates, yielding parameter maps which depict the local functionality of an organ in rich detail [56, 158, 139]. In this work we will present both potentials and pitfalls of a large class of techniques to obtain such local estimates with a focus on coupling methods for stabilization: We will show several applications where the inclusion of prior spatial information using coupling can improve local estimates if the data is degraded by noise or motion. However, we will also present coupled settings, where results of established methods are systematically biased and need to be interpreted with care. This work can be located at the boundary of mathematics and engineering: We will describe and analyze real-world problems using established mathematical theory and we will transfer techniques from applied mathematics to improve the solutions of our problems.

Let us start by outlining how imaging methods can be used to estimate temporal parameters from time series of medial images. Since the invention of X-Ray imaging by Wilhelm Conrad Röntgen in 1895, medical imaging has become one of the cornerstones of modern medical diagnosis [18]. From the early beginning, medical imaging was applied to display anatomical information: Historically, X-Ray and Computed Tomography (CT) images allowed to display bone structures, Magnetic Resonance Imaging (MRI) allowed to visualize soft tissue and Positron Emission Tomography (PET) allowed to highlight regions according to their metabolic activity. Such images are displaying a snapshot of the examined system at a given time point and are as such not including information on temporal processes. However, there are various applications where such temporal information is a desirable parameter to enable additional diagnosis [78, 56, 158]. In medical imaging, so-called dynamic imaging is a method which can be used to extract such time dependent parameters [111, 160, 140]. In dynamic imaging, multiple standard images are acquired, yielding a series of images which do not display the static system but changes of the system. A classic example of such imaging is so-called Dynamically Contrast Enhanced (DCE) imaging. Here, the

patient receives an intravenous injection of a contrast agent. After that, a time-series of images is acquired which shows the propagation of contrast agent through the blood stream and the tissue. The most common realizations of this technique which are used in routine clinical practice are Dynamic Contrast Enhanced MRI (DCE-MRI), Dynamic Susceptibility Contrast MRI (DSC-MRI) and Perfusion CT (CTP) [111, 161, 137]. If the changes of the system are additionally displayed in dependency of some known variable such as time, this allows not only to determine the absolute change but also rates of change. In this work we will focus on quantifying parameters which determine the change, such as velocities or flow rates, using DCE-MRI, DSC-MRI and CTP.

A way to extract these parameters from a series of dynamic images is by so-called inverse modeling: Key to this method is the observation that the measured dynamic images depend on the parameters of interest and will change if the parameters change. In some cases it is thus possible to set up a physical model which describes the expected data in dependency of the parameters. Mathematically this can be expressed as the assumption that the (measured) data d and (to be determined) parameters p are connected by the relationship $M(p) \approx d$, where $M(p)$ is a suitable physical model and $\approx$ denotes a carefully chosen measure of similarity which depends on the expected noise and will be introduced later. Given data d, in some cases the most likely parameters p associated with the data can then be found by solving $M(p) \approx d$, see Figure 1.1 for a simple example. In this work we will not focus on improving the physical models themselves, but on improving the estimation process by including additional information. Concerning the models, in dynamic imaging a large class of models used to describe the data are so-called *Dynamic Parametric Models*, or short *Dynamic Models* [98, 35, 111, 161, 137]. Such models have a finite number of parameters and are describing the response of a single system over time or a similar continuous variable [41, 71]. Before we outline the contributions and the content of this work, we conclude this section with specifying how we understand the concept of a dynamic model:

Definition 1. Let $P \subseteq \mathbb{R}^m$, $Q \subseteq \mathbb{R}^l$ be open domains and let $T > 0$. The mapping

$$M : P \times [0, T) \times Q \to \mathbb{R}$$
$$(p, t, q) \mapsto M(p; t, q)$$

is called a *Dynamic Parametric Model* or shorter *Dynamic Model*. For ease of notation we will omit the control parameters t, q if they are clear from the context and simply write $M(p) := M(p; \cdot, q)$, i.e. $M(p) : [0, T) \to \mathbb{R}$.

1.1 Content and Contribution

Traditionally, the physical models which are used to describe data from dynamic imaging are tailored to describe generic regions of interest, as e.g. the renal parenchyma or
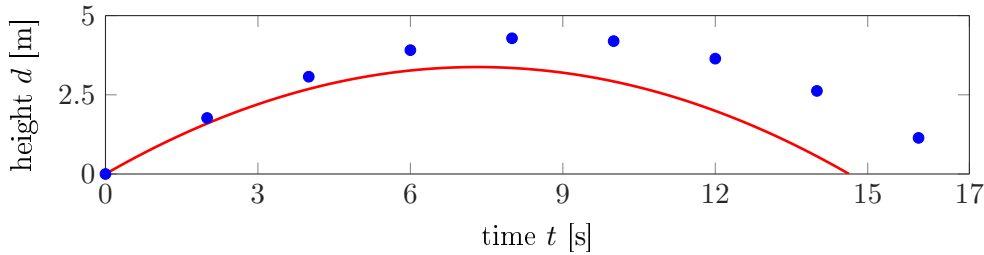
Figure 1.1: Illustrated is a simple example for an inverse problem: Given is the height $d = (d_1, \ldots, d_k)$ of a projectile at time points $(t_1, \ldots, t_k)$. The task is to determine the initial velocity v_0 and the launching angle α of the projectile. Measured data (d_i, t_i) is given in blue, a possible model simulation $M(p)$ in red where $p = (v_0, \alpha)$. The task to find parameters p such that $M(p) \approx d$ is an inverse problem.

the human brain [167, 139]. In current applications, the same models are increasingly used to quantify dynamic parameters on a voxel-by-voxel basis [139, 74, 12, 90, 79, 104]. As we will argue in Section 1.2.4, this transition in scale can be challenging: To avoid motion corruption and to achieve a high time-sampling, scanning times need to be low, which can introduce additional noise artifacts in the data [111, 161, 137]. Also, as the example of abdominal DCE imaging shows, if patient motion is inevitable, motion between images can heavily influence especially voxel-wise estimates [111, 161, 137]. A third problem is the validity of the model, which might not be given anymore if it is used on a finer scale, see Chapter 6.

We now give a general outline of this thesis: In Chapter 2 we start by describing the three dynamic models which we will use to demonstrate the merits of our contributions. A basic overview of the physical models and the expected data is given. In Chapter 3 we will present general analytical results to determine if and how stable parameters can be recovered from noisy data using the dynamic models described in the previous chapter. The Chapter 3 culminates in Section 3.5, where we will use these results to construct a novel technique for T_1 estimation from variable flip-angle data which has minimal sensitivity with respect to errors in the data: This novel technique thus combines advantages of established techniques and allows fast, simple and robust parameter estimation with errors reduced by up to 24% on software phantom data as compared to an established linear recovery approach and closely performing like the gold-standard nonlinear approach. In Chapter 4 we will present different ways to couple dynamic models in order to improve parameter estimation from spatially structured data. Based on methods which were originally developed for RGB image denoising, we will introduce tailored coupling terms for multiple parameters and high-resolution data with distinct spatial layout. As opposed to many existing approaches [114, 155], these terms are not only enforcing spatial regularity in the estimated parameters separately, but are indeed coupling the different parameter maps to additionally improve the re-

sults. We will demonstrate advantages of these terms for the example of T_1 estimation, where we show an improvement of up to 8% as compared to established techniques on software phantom data. In Chapter 5 we will further show how redundancy of 4D data can be exploited to add additional parameters to the recovery. We will both show how parameter estimation can be combined with motion correction and how additional control-parameters can be included in the recovery process. Specifically, we will show how coupling terms as introduced in Chapter 4 can improve joint motion correction and T_1 estimation and how the same terms can improve voxel-wise GFR estimation with missing or low-quality control parameters. We conclude this thesis with Chapter 6, where we demonstrate limits of dynamic models, which occur if they are used to describe only parts of a coupled system. Specifically we will show that established dynamic models will overestimate perfusion if they are only applied to parts of a highly coupled system, as in the case of pure capillary tissue. This effect is demonstrated on phantom data, which is constructed using a novel porous media model for indicator dilution. Additionally we present computations on real data which indicate that this effect can also be found on coarse scale in real life applications.

Different pieces of this work are already published or are submitted and in the process of being reviewed. The following listing gives a brief overview of these results in chronological order:

[H1] **Heck, Ruthotto, Berkels and Modersitzki,** *Model-Based Parameter Estimation in DCE-MRI Without an Arterial Input Function*, **Proceedings of BVM 2014**: A major source of error for GFR estimation from DCE-MRI data is the accurate determination the so-called arterial input function, which is a main control parameter for the model fit and is related to the injection process of the contrast agent [99, 38]. For voxel-wise methods, the paper describes a strategy which exploits data redundancy to estimate the arterial input function as an additional unknown from the data. The main contribution of this paper is the additional introduction of spatial coupling to the estimation, which allows improvements of up to 45% as compared to unstabilized methods on a software phantom dataset for moderate noise.

[H2] **Heck, Benning and Modersitzki,** *Joint Registration and Parameter Estimation of T1 Relaxation Times Using Variable Flip Angles*, **Proceedings of BVM 2015**: A common problem for T1 estimation from abdominal data which was acquired with variable flip angles is patient motion between acquisitions [69]. This paper introduces a technique which is similar to the one proposed by [69] and combines model fit and motion correction. Conceptually this technique can be thought of as a sequential approach: First, estimated model parameters are used to simulate motion-free images, which are used as reference images to update the registration. Second, model parameters are updated using the updated registration and so on. The main contribution of this paper as compared to [69] is the introduction of spatial coupling and allows improved parameters of up to 30%.

[S3] **Sandmann, Hodneland and Modersitzki,** *A Practical Guideline for T1 Reconstruction from Various Flip Angles in MRI*, **to appear in Journal of Algorithms and Computational Technology, 2016**: Since in abdominal imaging short scanning times are crucial to avoid motion artifacts, data can come with limited quality due to noise effects, see Section 1.2.4. The paper presents strategies to improve the estimation of T_1 relaxation times from noisy datasets which were obtained with various flip angles. To do this, two strategies are presented: First, a novel estimation method is introduced based on a sensitivity analysis of the recovery. This novel method combines computational advantages of a linear parameter fit with the robustness of a nonlinear fit. Here improvements of up to 24% are achieved as compared to the linear method are achieved and results are close to the computationally more expensive nonlinear method. Second, the role of spatial coupling is investigated: In extension of the results published in [155], we find that for high resolution data with distinct spatial layout, results can be additionally improved if not only T_1 but also an additional variable M_0 is stabilized for a Total Variation coupling term. In this case improvements of up to 8% are achieved.

[S4] **Sandmann, Hanson, Malyshev, Lundervold, Modersitzki and Hodneland,** *Limitations of Traditional Models for Perfusion*, **submitted to IEEE Transactions on Biomedical Engineering**: Although traditional dynamic models to estimate perfusion from DSC-MRI data were originally developed with large regions of interest in mind [167], these models are increasingly used for voxel-wise analysis [56, 158]. In this paper we demonstrate limits of such models if they are used to describe parts of a system with heavy inter-voxel coupling, as found e.g. in pure capillary tissue. Specifically we will show that in this scenario established dynamic models will overestimate perfusion proportional to the voxel size. We demonstrated this effect on phantom data, which is constructed using a novel porous media model for indicator dilution. Additionally we present real data which indicates that this effect can also be found on coarse scale in real life applications.

[H5] **Amman, Derksen and Heck,** *A novel magnetometer-accelerometer calibration based on a least squares approach*, **Proceedings of ICUAS 2015**: Although the author has additionally contributed to a publication which describes a calibration method for magnetometers used for inertial navigation, this topic will not be covered in this thesis.

1.2 Three Dynamic Parameters and Clinical Relevance

All dynamic models which are under investigation in this work were carefully chosen and have significant clinical relevance. In this section we describe the clinical relevance of the three parameters which lie in the focus of this thesis: Cerebral Blood Flow

(CBF), Glomerular Filtration Rate (GFR) and T_1 relaxation time. Since the estimation of these parameters is following a similar protocol, in Section 1.2.4 we will additionally elaborate on common error sources which can lead to inaccurate parameter results.

1.2.1 CBF Estimation

In 2014 in Germany 15.298 people died of an ischemic stroke, making it the 14th cause for death in the country [141]. In order to describe the relevance of cerebral blood flow for treatment, we briefly outline common causes and therapy of ischemic stroke following [158]: In ischemic stroke, an arterial vessel of the brain is occluded, leading to hypo-perfusion of the fed area and possibly irreversible tissue damage. The affected area can roughly be divided into a hypo-perfused but possibly salvageable area, the so-called penumbra, and an infarct core where perfusion is drastically reduced. During the stroke, the penumbra is slowly replaced by the core. A well-known therapy for ischemic stroke is the usage of thrombolytic agents to clear the clogged vessel and to restore the perfusion. However, there are various contra-indications for this therapy: In the case of additional cerebral hemorrhage, thrombolysis might lead to additional tissue damage, in the case of little salvageable penumbra combined with a large core, studies indicate that the risk-benefit ratio of thrombolytic therapy is unfavorable. Accurate knowledge of the size of hypo-perfused areas is hence crucial to decide which therapy should be used.

CBF Estimation using Dynamic Imaging

To estimate regional cerebral blood flow (CBF) [ml/min/100ml], dynamic contrast enhanced imaging like CTP or DSC-MRI and inverse modeling can be used, see Figure 1.2. For a review of other techniques for CBF estimation see Section 2.1.6. In Section 2.1 we will introduce the so-called Meier-Zierler model, which can be used to describe the contrast agent concentration in the human brain in dependency of CBF and time.

1.2.2 GFR Estimation

The Glomerular Filtration Rate (GFR) [ml/min] describes how much blood-plasma is filtered by the kidneys per unit time [42, 78]. In healthy humans the GFR is typically varying with age and is decreasing from a mean of $116\,\text{ml/min}$ (20-29 years) to a mean of $75\,\text{ml/min}$ (≥ 70 years) [78]. In clinical practice the GFR is used as a main indicator of kidney health and its decrease is linked to a large variety of kidney diseases [78, 42]. This makes the GFR of large importance in clinical practice [42]. However, we will see in Section 2.2.3 that estimation of the GFR is delicate: Since it is difficult to measure directly, common methods aim to approximate it by estimating the speed at which an external marker substance is filtrated by the kidney [132, 42]. This value, which

Figure 1.2: Figure displaying 4D Perfusion CT (CTP) data of a human head. (a) The top-row visualizes a 2D slice at three from 12 time points, the bottom row the concentrations of contrast agent. (b) Displayed is the concentration timecurve from the location marked in blue. Note that motion artifacts at the boundary of the skull can be observed. Data was provided by courtesy of Rashindra Manniesing from the DIAG group at the Haukeland Unversity in Nijmegen, the Netherlands.

depends on the external substance, is called renal clearance and coincides only with the GFR if some additional assumptions are fulfilled. Specifically the marker needs to be filtrated freely and must not be metabolized, reabsorbed or secreted by the kidney [42]. In everyday clinical routine the GFR is typically estimated based on the blood concentration of a metabolic product, Creatinin, which fulfills all of the above criteria. Although this method is fast, well compatible with clinical routine and has been shown to be able to reliably detect kidney malfunctions, it comes with a variety of disadvantages: Most notably it only gives a GFR estimate for both kidneys and does not allow for separate, let alone local analysis of the kidney function [42]. For a more detailed discussion of methods for GFR estimation see Section 2.2.3.

GFR Estimation using DCE-MRI

In order to cope with this problem, it has been proposed to use DCE-MRI as an alternative method for GFR Estimation [68, 5, 139, 16]. As it will be described in Section 2.3.2, contrast agents in DCE-MRI are excreted nearly exclusively over the kidney and can be considered as markers whose clearance coincides with the GFR [42, 136]. Consequently, there are multiple approaches to determine the GFR from DCE-MRI data of the kidney [68, 5, 139, 16]. Estimation of the GFR from measured concentration time curves of the kidney follows a similar strategy as already described for CBF estimation, see Figure 1.3. It will be described in more detail in Section 2.2.
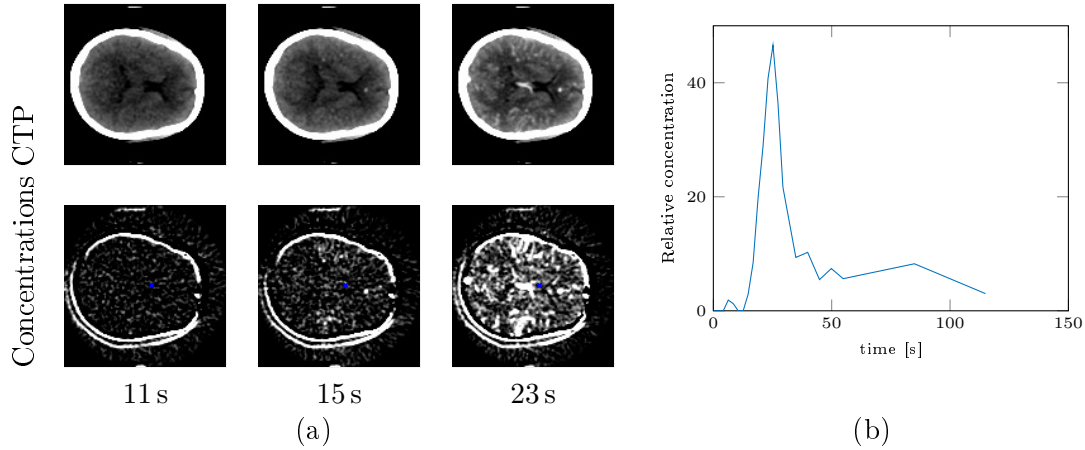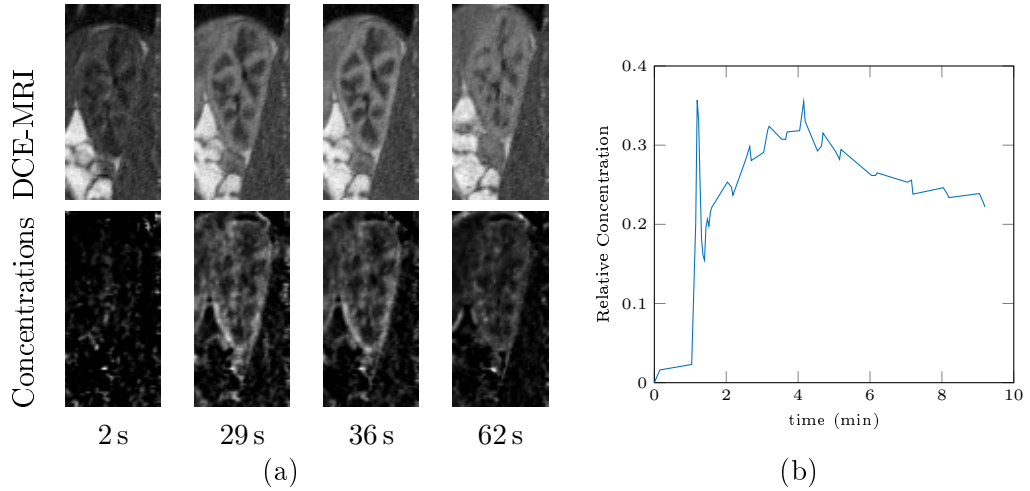
Figure 1.3: Figure displaying 4D DCE-MRI data of the human kidney. (a) The top-row visualizes a 2D slice at three from 49 time points, the bottom row the concentrations of contrast agent in the slice. (b) Displayed is the mean concentration of contrast agent in the kidney. Data was provided by courtesy of Jarle Rørvik of the Haukeland University Hospital in Bergen, Norway.

We will introduce the so-called Patlak-Rutland model for GFR estimation from DCE-MRI data in Section 2.2, which has been applied successfully in a variety of studies [68, 16, 166].

1.2.3 T_1 Estimation

All physical models which are used for GFR or CBF estimation from dynamical MRI or perfusion CT data are describing contrast agent concentrations in a volume of interest. However, in the case of MRI or CT, the measured data is typically not given in contrast agent concentration but is obtained in raw signal intensity. In order to relate the model simulation with the data, either the data has to be converted to contrast agent concentrations or the forward model simulation needs to be converted to signal intensities. In both cases a relationship between contrast agent concentration and measured signal needs to be established.

Whereas in CT the signal varies approximately linearly with the concentration of contrast agent [161], this relationship is more complicated in MRI. To relate MR signal to contrast agent concentration, two steps are necessary: First, the MR signal has to be modeled in dependency of tissue properties. Second, the effect of contrast agent on the tissue properties has to be described. Since the MR signal is determined by relaxation processes, the main tissue parameters used to describe the signal are T_1 and T_2, which are representing the speed of the relaxation in longitudinal and in transversal direction. We will introduce the relationship between T_1, T_2 and the measured MR signal, the
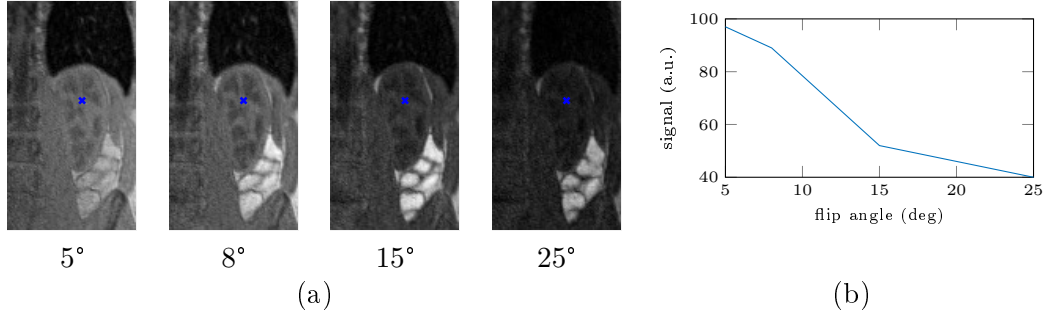
Figure 1.4: Figure displaying 4D human kidney data from a variable flip-angle experiment. (a) Shown are 2D slices of left kidney obtained with various flip angles. (b) Displayed are the signal intensities at the point highlighted in blue. Data was provided by courtesy of Jarle Rørvik of the Haukeland University Hospital in Bergen, Norway.

so-called signal equation, for a large class of MR sequences in Section 2.3.1. In order to describe the effect of contrast agent on T_1 and T_2, in DCE-MRI usually a linear relationship between contrast agent concentration and apparent relaxation $1/T_1$ is assumed, see Section 2.3.2. In order to quantify the amount contrast agent in the tissue based on strength of the MR signal, this means that accurate knowledge of the baseline T_1 of the tissue needs to be known. As described in [69], errors in baseline T_1 can have significant effects on the recovered parameters.

T_1 Estimation Using the Variable Flip Angle Technique

A method to estimate the baseline T_1 of the tissue is the so-called variable flip angle technique [35, 66, 55]. For a review of other techniques for T_1 estimation we refer to Section 2.3.4. The main idea of the variable flip angle technique is to recover the longitudinal relaxation time T_1 from a series of magnitude MR measurements which are obtained with different flip-angles. Since at each flip angle the longitudinal magnetization is tilted by a different amount, the resulting series of signals contains information on T_1. After having obtained such a series of measurements $(d_1, \ldots, d_k)$, the unknown T_1 can be determined such that $M(p) \approx d$, where M describes the MRI signal in dependency of parameters p, among which is T_1, see Figure 1.4.

1.2.4 Error Sources in Dynamic Imaging

All of the above methods come with distinct sources of error, which might affect the accuracy of the estimated parameters. A common challenge for CBF and GFR estimation is that the expected concentration time-curve of the tissue will depend not only on the CBF or GFR, but also on the injection protocol and on the individual physiology

of the patient. To determine CBF or GFR, these factors have to be taken into consideration. In the above modeling this is reflected by including the time-dependent contrast agent concentration in a feeding vessel into the estimation, the so-called arterial input function (AIF). Results in [47, 38] indicate, that the AIF has indeed a high impact on the estimated values of both CBF and GFR. Accurate measurement of the AIF is hence highly desirable. For the case of GFR estimation using the Patlak-Rutland model, we will introduce a novel method to cope with this problem in Section 5.3.

A second challenge for all three applications is patient motion: In abdominal imaging, especially breathing can lead to misalignment of the data [23, 74, 165]. If concentration time curves of the kidney or of other structures are to be established, it is necessary to account for this motion. This is typically challenging, since non-homogeneous variances in image contrast are introduced by contrast agent or variable flip angles, making it difficult to relate images which were acquired at different time-points or flip-angles. In Section 5.1 we will address this problem and introduce a general method for dynamic imaging which combines the parameter estimation with the motion correction.

A third challenge is the image quality. In DCE-MRI, DSC-MRI and in Perfusion CT fast sampling rates are necessary during the early uptake phase. For CTP sampling rates are proposed which are not slower than $4\,s$ [157], in DSC-MRI the rate is approximately $1.9\,s$ [144]. In CTP additionally the radiation exposure of the patient is to be limited. For all imaging modalities these temporal and dose constraints lead to data which comes often with limited quality due to noise effects, see [137, 161, 111] and the data presented in Figures 1.2,1.3 and 1.4. To cope with this problem, we will introduce two novel methods: For the case of T_1 estimation, we will introduce a new, efficient and robust way to estimate parameters from variable flip angle data. Additionally, we will add information on the spatial layout by introducing spatial coupling methods in Chapter 4.

2 Dynamic Models in 4D Imaging and Clinical Applications

We start by describing the three dynamic models, for which the theory of this work was implemented: The Meier-Zierler Model [98] to estimate Cerebral Blood Flow, the Patlak-Rutland Model [116, 129, 115] to estimate the Glomerular Filtration Rate of the kidneys and the signal equation for spoiled gradient echo sequences [93, 55] to estimate the MR parameter T_1. All of the presented models will be subject to further analysis in the following chapters. The selected models have been chosen with care: We will see in Chapter 3 that solving the respective parameter estimation problems requires a large variety of numerical methods (deconvolution, linear parameter fit, nonlinear parameter fit). Additionally we will see in Section 3.3.2 that parameter estimation using the Meier-Zierler model is indeed ill-posed. This makes these models representative for a variety of applications where the contributions of this work might be applied.

All these models will describe the dynamic behavior of a single black-box system. The *Meier-Zierler Model* describes the concentration of contrast agent which is flowing through a tissue volume under few prior assumptions. We will show in Lemma 2 that in this case the concentration $C(t)$ of contrast agent in the volume at time t can be described by

$$C(t) = P_\Omega \int_0^t c_{\mathrm{a}}(s) R(t-s) \, \mathrm{d}s = P_\Omega (c_{\mathrm{a}} * R)(t).$$

Here the so-called residue function $R : [0, \infty) \to \mathbb{R}$ describes the fraction of contrast agent which remains in the tissue after a certain time and P_Ω is the perfusion, the parameter of interest. The function $c_{\mathrm{a}} : [0, \infty) \to \mathbb{R}$ is the so-called arterial input function, which describes the concentration of contrast agent in a feeding vessel and is typically assumed to be known.

Whereas the Meier-Zierler Model is only valid for tissue where no extravasation of contrast agent is expected, the *Patlak-Rutland Model* is designed to model simple accumulation effects [137]. It was originally used to described blood brain barrier leakage [115], but since then has also been employed successfully for GFR estimation [68, 166]. It models the time-dependent contrast agent concentration $C(t)$ in a volume as

$$C(t) = v_a c_{\mathrm{a}}(t) + K^{\mathrm{trans}} \int_0^t c_{\mathrm{a}}(s) \, \mathrm{d}s.$$

Here $v_a \in [0,1]$ describes the fractional blood volume, $c_a : [0, \infty) \to \mathbb{R}$ is the arterial input function and $\text{Vol(Kidney)} \cdot K^{\text{trans}} = \text{GFR}$ describes the normalized glomerular filtration rate.

We conclude this section with describing the signal equation for spoiled gradient echo sequences [55, 168], which models the magnitude MR signal S in dependency of the flip-angle $\alpha \in [0, 2\pi]$ as

$$S = M_0 \sin \alpha \frac{1 - e^{-\text{TR}/T_1}}{1 - \cos \alpha e^{-\text{TR}/T_1}}.$$

Here $T_1 > 0$ is the longitudinal relaxation time, $M_0 > 0$ depends on proton density, echo time and transversal relaxation and $\text{TR} > 0$ is the repetition time. Although on the first glance the signal equation looks very different to the above models, it is indeed structurally very similar: In all cases a dynamic variable is described in dependency of a continuous control parameter: Whereas the Meier-Zierler and the Patlak-Rutland model are describing the concentration of contrast agent in dependency of time, the signal equation describes the MR signal in dependency of the flip-angle.

2.1 The Meier-Zierler Model for CBF Estimation

We will now introduce the so-called Meier-Zierler model [98] which can be used to estimate Cerebral Blood Flow (CBF) from a time series of contrast-enhanced images. This section will have three main results: The *Meier-Zierler Model*, the *Maximum Slope Model* and the *Well-Mixed Compartment Model*. In practical perfusion imaging, the *Meier-Zierler Model* and the *Maximum Slope Model* are the most important model to describe perfusion of the human brain and used in a large variety of applications [86, 76, 12]. These models will be essential to the considerations of Chapter 6, where we will show that these models can be used to equivalently describe a novel PDE model for perfusion.

All models can be derived from the continuity equation for tracer kinetic modeling, which will be derived in the following section. The continuity equation will depend on a large number of free variables, making it for all practical means an underdetermined system. In oder to apply it in clinical practice, the number of unknowns is reduced by adding additional assumptions. For the *Meier-Zierler Model* it will be assumed that the transit times of particles flowing through the tissue are distributed according to some probability density function. The unknown of the Meier-Zierler model will be a function, the so-called residue function. For the *Maximum Slope Model* it will be assumed that no outflow is taking place when the peak of the bolus is passing through the tissue. This means that during early uptake the concentration in the tissue patch will depend mainly on the blood flow CBF, leaving only one unknown parameter. For the *Well-Mixed Compartment Model* it will be assumed that the contrast agent is distributed homogeneously in the tissue patch. As we will see, in this case the number
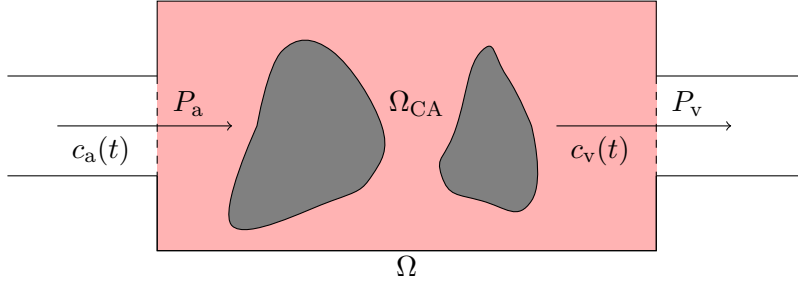
Figure 2.1: Displayed is a black-box domain Ω for the Meier-Zierler Model. Depicted are concentrations at the in- and outlet $c_\mathrm{a}, c_\mathrm{v}$ [mmol/mm^3], perfusion constants $P_\mathrm{a}, P_\mathrm{v}$ [mm^3/(s mm^3)] and the contrast agent distribution volume Ω_CA, which is the total volume of Ω without the gray cell-like structures. The perfusion P [mm^3/(s mm^3)] and absolute flow F [mm^3/s] are connected by $P \cdot \mathrm{Vol}(\Omega) = F$, see Lemma 1.

of free parameters can be reduced to two, the blood flow CBF and the Mean Transit Time (MTT). We will also introduce the *Central Volume Theorem*, which is another main result from perfusion imaging and establishes a connection between CBF, MTT and the Cerebral Blood Volume (CBV). However, in this work it is introduced mainly to simplify some proofs and will thus be stated without proof.

2.1.1 The Continuity Equation for Tracer Kinetic Modeling

We start by introducing the continuity equation for tracer kinetic modeling, which is basic to nearly all pharmacokinetic models [137]. Given some basic assumptions (see Assumption 1), the continuity equation describes the average contrast agent concentration in a black-box tissue which is fed with at one inlet and drained at one outlet. An example for such a black-box tissue is depicted in Figure 2.1.

Assumption 1 (Continuity equation for tracer kinetic modeling).

A1) The contrast agent in- and outflow is linear and stationary. This means that the absolute contrast agent in- and outflow $J_\mathrm{a}, J_\mathrm{v}$ [mmol/s] depends linearly on the contrast agent concentration at the respective locations with time-independent constants $F_\mathrm{a}, F_\mathrm{v}$ [mm^3/s], i.e.

$$J_\mathrm{a}(t) = F_\mathrm{a} c_\mathrm{a}(t) \text{ and } J_\mathrm{v}(t) = F_\mathrm{v} c_\mathrm{v}(t).$$

A2) No contrast-agent is created or destroyed in Ω (conservation of mass).

The following Lemma shows that contrast agent transport can be modeled under Assumption 1 using standard ordinary differential equations (ODEs).

Lemma 1 (See [137]). Let $c_a, c_v \in C^0([0, \infty))$ and let $C(t)$ denote the contrast agent concentration in Ω at time point t. Under Assumption 1 the concentration $C(t)$ is given by

$$\left| \begin{aligned} &C'(t) = P_a c_a(t) - P_v c_v(t), \\ &C(0) = 0. \end{aligned} \right. \tag{2.1}$$

where $P_a := F_a / \mathrm{Vol}(\Omega)$, $P_v := F_v / \mathrm{Vol}(\Omega)$ have units $\mathrm{mm}^3/(\mathrm{s}\,\mathrm{mm}^3)$.

PROOF. Let $m(t)$ [mmol] denote the absolute number of contrast agent molecules in Ω at time point t. Assumption 1 allows us to model the change of contrast agent in an infinitesimal small time interval as

$$\dot{m}(t) = J_a(t) - J_v(t) = F_a c_a(t) - F_v c_v(t).$$

Dividing the equation by the total volume $\mathrm{Vol}(\Omega)$ of Ω yields (2.1). $\qquad\square$

Note that P_a, P_v [$\mathrm{mm}^3/(\mathrm{s}\,\mathrm{mm}^3)$] are so-called perfusion constants, which are normalized by the distribution volume. Specifically the units of P_a, P_v are [volume of fluid] per [second] per [volume of tissue] and are describing how much fluid passes through the volume per unit time [119]. Note that in physical applications flow is usually normalized with respect to surface and not with respect to volume [37, 89]. Indeed we will see in Chapter 6 that this normalization will indeed lead to scaling problems for coupled systems.

2.1.2 The Meier-Zierler Model and the Central Volume Theorem

We will now derive the Meier-Zierler Model from the continuity equation and additional assumptions presented in Assumption 2. The Meier-Zierler Model describes the contrast agent concentration $C(t)$ in a black-box tissue at time point t as

$$C(t) = P_a \int_0^t c_a(s) R(t-s) \, \mathrm{d}s.$$

Here P_a [$\mathrm{mm}^3/(\mathrm{s}\,\mathrm{mm}^3)$] is the flow at the inlet and $R : [0, \infty) \to \mathbb{R}$ is a tissue specific function, the so-called *residue function*.

In order to derive (2.2), we need to make additional assumptions, namely

Assumption 2 (Meier-Zierler Model).

1. The transit times t of particles passing through Ω are distributed according to a probability distribution $h \in C^0([0, \infty))$ with

   a) $h \geq 0$,

   b) $h(0) = 0$,

   c) $\int_0^\infty h(s)\, \mathrm{d}s = 1$.

2. The concentration at the inlet is given by a function $c_\mathrm{a} \in C^0([0, \infty))$.

The following lemma shows how the continuity equation can be simplified under Assumption 2:

Lemma 2 (See [137]). Let $C \in C^1([0, \infty))$ and Assumption 1, 2 hold. Then

$$C(t) = P_\mathrm{a} \int_0^t R(s) c_\mathrm{a}(t - s)\, \mathrm{d}s = P_\mathrm{a}(c_\mathrm{a} * R)(t) = (c_\mathrm{a} * I)(t) \qquad (2.2)$$

where

$$R(t) := 1 - \int_0^t h(s)\, \mathrm{d}s \quad \text{and} \quad I(t) := P_\mathrm{a} R(t) \qquad (2.3)$$

and the convolution of two functions $g, h \in C^1([0, \infty)$ is given by $g * h(t) := \int_0^t g(t - s) h(s)\, \mathrm{d}s$

PROOF. Assumption 2 allows us to describe the outflow $P_\mathrm{v} c_\mathrm{v}(t)$ of the tissue at time point t as

$$P_\mathrm{v} c_\mathrm{v}(t) = \int_0^t (P_\mathrm{a} c_\mathrm{a}(s)) h(t - s)\, \mathrm{d}s.$$

Combining this equation with the continuity equation (2.1) yields

$$C'(t) = P_\mathrm{a} \left(c_\mathrm{a}(t) - c_\mathrm{a} * h(t) \right) \text{ with } C(0) = 0.$$

Note that the solution of this ODE can be obtained by simple integration, yielding a solution which is differentiable due to the Fundamental Theorem of Calculus [83]. However, to simplify the calculations note that the integration can also expressed as convolution with the function $1 : [0, \infty) \to \mathbb{R}$, $t \mapsto 1$. Due to the algebraic properties

Table 2.1: Summary of some main vocabulary of perfusion imaging.

Symbol	Name
$P_\Omega := P_a$	Blood Flow or Perfusion
$T_\Omega := \int_0^\infty s h(s)\,\mathrm{d}s$	Mean Transit Time
$V_\Omega := \mathrm{Vol}(\Omega_{CA})/\mathrm{Vol}(\Omega)$	Blood Volume.
c_a	Arterial Input Function
R	Residue Function
I	Impulse Response Function

of the convolution (associativity, commutativity and distributivity) this allows us to write the solution as

$$C(t) = 1 * P_a(c_a - h * c_a) = P_a c_a * (1 - 1 * h)(t).$$

which concludes the proof. $\qquad\qquad\square$

Implications of the Meier-Zierler Model and the Central Volume Theorem

The following definition introduces some main vocabulary of perfusion imaging. The names can be explained as follows: The *Residue Function R* describes which fraction of contrast agent is still present after t seconds whereas the *Impuls Response Function I* describes how the system responds to a Dirac Input.

The following Lemma summarizes basic important properties of R:

Lemma 3 (See [137]). Let R, I be given by (2.3). Then

1. $R(0) = 1$,

2. R is monotonous decreasing

3. $\lim_{t\to\infty} R(t) = 0$.

4. $P_\Omega = \max_t I(t)$,

5. $T_\Omega = \int_0^\infty R(s)\,\mathrm{d}s$, if $\lim_{t\to\infty} t^2 h(t) = 0$.

6. $V_\Omega = \left(\int_0^\infty C(s)\,\mathrm{d}s\right) / \left(\int_0^\infty c_a(s)\,\mathrm{d}s\right)$.

PROOF. All steps follow by direct calculation.

1. $R(0) = 1 - \int_0^0 h(s)\, \mathrm{d}s = 1$,

2. Let $u < v$. Then $R(u) - R(v) = -\int_u^v h(s)\, \mathrm{d}s \le 0$ since $h \ge 0$.

3. Finally it holds that $\lim_{t \to \infty} R(t) = 1 - \int_0^\infty h(s)\, \mathrm{d}s = 0$ due to the properties of h.

4. Follows directly by $I(t) = P_\mathrm{a} R(t)$ and properties 1-3.

5. By definition it holds that $T_\Omega = \int_0^\infty sh(s)\, \mathrm{d}s$. However, it also holds that $h(t) = -\frac{d}{dt} R(t)$. Integrating by parts yields

$$\int_0^t sh(s)\, \mathrm{d}s = \int_0^t s(-\frac{d}{dt} R(s))\, \mathrm{d}s = -tR(t) + \int_0^t R(s)\, \mathrm{d}s$$

Using l'Hôpital's rule yields $\lim_{t \to \infty} tR(t) = \lim_{t \to \infty} t^2 h(t) = 0$ due to the additional assumption on the decay. This yields the claim

6. We first observe that for two functions $f, g \in C^0([0, \infty)$ the integral over $f * g$ equals the product of the integrals over f and g. This means that

$$\int_0^\infty C(s)\, \mathrm{d}s = \int_0^\infty (I * c_\mathrm{a})(s)\, \mathrm{d}s = \left(\int_0^\infty c_\mathrm{a}(s)\, \mathrm{d}s \right) \left(\int_0^\infty I(s)\, \mathrm{d}s \right) \qquad (2.4)$$

On the other hand it holds that

$$\int_0^\infty I(s)\, \mathrm{d}s = P_\mathrm{a} \int_0^\infty R(s)\, \mathrm{d}s = P_\mathrm{a} T_\Omega = V_\Omega \qquad (2.5)$$

where the last equality holds due to the central volume theorem (2.6) which will be introduced in the next theorem. Combining (2.4) and (2.5) yields the claim.

This concludes the proof. $\qquad\qquad\square$

Another important relationship used in perfusion estimation is the so-called *Central Volume Theorem* [142, 98, 118, 137]. The Central Volume Theorem relates three of the main indicators for perfusion, the bloodflow, the blood volume and the mean-transit time. Since we will need the theorem only occasionally, we refer to [98] for the proof.

Theorem 1 (Central Volume Theorem [142, 98]):
Let Assumption 1 and Assumption 2 hold. Then

$$V_\Omega = T_\Omega P_\Omega. \qquad (2.6)$$

2.1.3 The Well-Mixed Compartment Model

We have seen that assuming that transit times are distributed according to some probability distribution allowed to simplify the continuity equation (2.1) significantly. We now introduce another simplification, which will be of importance in Chapter 6. Specifically we will assume that the plasma compartment is well-mixed, i.e. that the contrast agent concentration at the outlet equals the average contrast agent concentration in the tissue. We will show that this assumption leads to exponential impulse-response functions $I(t) = P_\Omega \exp(-P_\Omega/V_\Omega t)$. In current literature this family of functions is often employed to simulate tissue curves, which are subsequently used as ground-truth values to evaluate novel recovery techniques [111, 160].

The Assumptions for the well-mixed compartment can be summarized as follows:

Assumption 3.

1. The contrast agent concentration at the outlet equals the average contrast agent concentration in the Ω, i.e. $V_\Omega c_\mathrm{v}(t) = C(t)$.

2. The inflow equals the outflow (incompressibility of flow), $P_\mathrm{a} = P_\mathrm{v}$.

The following lemma shows, that in this case the impulse response functions are exponentials:

Lemma 4. Under Assumption 1, 3 it holds that $C \in C^1[0, \infty)$ is given by

$$C(t) = (I * c_\mathrm{a})(t) \text{ with } I(t) = P_\Omega e^{-(P_\Omega/V_\Omega)t}. \tag{2.7}$$

PROOF. Combining Assumption 1 and Assumption 3 and denoting $C(t) = V_\Omega c(t)$ yields the ODE

$$V_\Omega c'(t) = P_\Omega(c_\mathrm{a}(t) - c(t)), \text{ with } C(0) = 0.$$

The above ODE is linear and inhomogenous. It can be solved by first solving the homogeneous system and then varying the constant, see [152, §2(II)] for an example.$\square$

2.1.4 The Maximum Slope Model

Another model to determine the perfusion from a tissue concentration curve is the so-called maximum slope model [80]. The maximum slope model relies on the assumption that no venous outflow is present, i.e. $P_\mathrm{v} = 0$. In this case the basic continuity equation for pharmakokinetic modeling (2.1) simplifies to

$$C'(t) = P_\mathrm{a} c_\mathrm{a}(t) - P_\mathrm{v} c_\mathrm{v}(t) = P_\mathrm{a} c_\mathrm{a}(t).$$

From this relationship one can see that if c_a has a maximum, C' must also have a maximum. It follows that

$$P_\mathrm{a} = \frac{\max_t C'(t)}{\max_t c_\mathrm{a}(t)}. \tag{2.8}$$

2.1.5 Discussion

We have presented a general equation for tracer kinetic modeling, the so-called continuity equation for tracer kinetic modeling (2.1). Since we have found that this equation had too many unknowns to be of practical use, we have described different assumptions under which the continuity equation could be simplified significantly. Specifically we presented three simplifications: The Meier-Zierler Model was derived under the assumptions that transit times are distributed according to some probability distribution. For the Well-Mixed Compartment Model it was assumed that contrast agent is distributed homogeneously in Ω_CA. For the Maximum-Slope Model it was supposed that not venous outflow was present during the peak of contrast agent.

In current literature, the most common model which is used to describe Cerebral Blood Flow is the Meier-Zierler Model [86, 76, 12]. Although an evaluation in [2] has shown that the maximum slope model performs comparably and yields equal therapy decisions in clinical practice, the assumption that no venous outflow is present when c_a peaks requires high injection speeds which are hard to fulfill [84]. However, we will see in Section 3.3.2 that parameter estimation using the Meier-Zierler Model is usually ill-posed and requires advanced numerical methods. This makes blood flow estimation using the Meier-Zierler Model highly complicated: In [86] it was demonstrated that different commercial software is generating different CBF results for the evaluation of perfusion data of the same patient. The Well-Mixed Compartment Model is often used for simulations [111, 160, 161], but its strong assumptions are usually not valid for current scanner resolutions. We will demonstrate other, more basic limits of perfusion in Chapter 6. Specifically we will study perfusion in highly coupled systems and demonstrate, that in this case perfusion is highly dependent on scale.

2.1.6 Alternative Methods for CBF Estimation

For sake of completeness, we conclude this section by giving a short introduction to CBF estimation using alternative estimation techniques. Specifically we will focus on a tailored MR sequence, so-called Arterial Spin Labeling (ASL) and an emerging but still highly experimental technique, so-called Magnetic Particle Imaging (MPI).

CBF estimation by Arterial Spin Labeling (ASL)

We describe ASL following closely [25]: ASL can be thought of dynamic contrast enhanced imaging without the usage of an external contrast agent. Instead, ASL uses the blood itself as contrast agent by so-called tagging: Tagging means that the longitudinal magnetization of arterial blood is inverted in a region upstream to the volume of interest. In a later measurement of downstream tissue, this magnetization can be detected. As compared to CBF estimation using external contrast agents, this has two main advantages: First, no administration of external contrast agent is necessary, second the arterial input will be known: If the tagging pulse is applied with some length, the arterial input function should have approximately the shape of a smoothed rectangle, which simplifies the pharmacokinetic modeling significantly. Since experiment times are approximately in the range of $1\,\mathrm{s}$, it is often additionally assumed that no venous outflow is present, yielding the following equation for the concentration $C(t)$ of tagged arterial blood in the volume of interest:

$$C(t) = P_\Omega \int_0^t c_\mathrm{a}(s)\,\mathrm{d}s$$

where c_a is the (rectangular) input function. If it is further assumed that all tagged blood has arrived in the tissue and no venous outflow is present, this yields theoretically the following expected ASL signal at time T after labeling:

$$S(T) = S_0 + P_\Omega(2f\tau M_{0a})d \tag{2.9}$$

here $(2f\tau M_{0a})$ represents the integral over the c_a, where $\tau > 0$ is the duration of the tagging pulse M_{0a} the steady-state magnetization of arterial blood and $f \in (0,1)$ models the efficiency of the tagging pulse. The factor $d := e^{-T/T_{1a}}$ accounts for the loss of magnetization in the time T between tagging and measurement.

This modeling already shows some of the challenges which have to be overcome if ASL is to be employed in routine clinical practice: Equation (2.9) is only valid if indeed all tagged blood has arrived at the volume of interest. This makes ASL sensitive to arterial delay [164, 4]. Also, in the above modeling it is assumed that the tagged spins are relaxing with the T_1 of arterial blood: However, since it is expected that there is water exchange with the tissue, these effects have to be clarified as well [25, 4]. Nevertheless ASL is an alternative, very promising technique for CBF measurement and expected to be employed in clinical practice soon [4].

CBF estimation by Magnetic Particle Imaging (MPI)

Another emerging method in this field is so-called Magnetic Particle Imaging (MPI), where the propagation of a ferromagnetic contrast agent is displayed using MRI-like techniques [82]. As compared to the previously mentioned DCE protocols, by design MPI has several advantages, as the lack of tissue background, radiation and increased resolution [82]. However, since MPI is not yet used routinely in clinical practice, this work will focus on the previously mentioned DCE protocols.

2.2 The Patlak-Rutland Model for GFR Estimation

In this section we describe the so-called Patlak-Rutland Model to estimate the Glomerular Filtration Rate (GFR) of the kidney from a time series of contrast-enhanced images [68, 166]. Since the Meier-Zierler model does not describe extravasation of contrast agent explicitly, its usage to to extract filtration rate is difficult. The Patlak-Rutland model on the other hand is tailored to model accumulation effects in highly vascularized tissue: As compared to the Meier-Zierler Model it is assumed that contrast agent can be in either of two areas of the black-box tissue: A vascular space and extra-vascular extra-cellular space (EES), where it is assumed to accumulate. These two spaces are assumed to be connected by a directed flow, which represents the filtration. This concept makes applications of the Patlak-Rutland broader than only GFR estimation: In hemorrhagic stroke, it also used to model blood brain barrier leakage effects [115]. Note that since the GFR refers to the amount of blood plasma, which is filtered by the kidney per unit time, in this section we will not describe blood volume as in the section on CBF, but plasma volume. These two quantities are connected by the hematocrit value, which measures the volume fraction of solid particles of the blood. The hematocrit value will be introduced in more detail at the end of this section. The Patlak-Rutland model will be basic to the considerations of Section 5.3, where we will use it to simultaneously estimate the arterial input function.

2.2.1 The Patlak-Rutland Model

We will now show how the marker concentration can described in dependency of the plasma concentration c_a, the plasma volume v_a and the filtration K^trans, see Figure 2.2 for details.

Lemma 5. Let $\Omega \subseteq \mathbb{R}^3$ be a bounded domain and let $E \subseteq \Omega$ and $A \subseteq \Omega$ be disjunct. Let us further assume that there is some linear and stationary flow K^trans [$\mathrm{mm}^3/(\mathrm{s\,mm}^3)$] from A to E (cf. Section 2.1) and that contrast agent can only be in A or E. Let $v_a := |A|/|\Omega|$ and let $c_\mathrm{a} \in C^0([0,\infty))$ denote the contrast agent
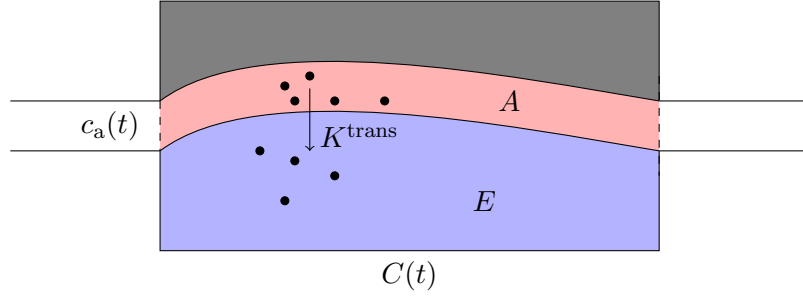
Figure 2.2: Schematic of tissue for which the Patlak-Rutland Model is valid. The tissue is supposed to consist of a plasma compartment (A, red) and an extra-vascular-extra-cellular-space (E, blue) and a compartment which is not accessible to contrast agent (gray). The perfusion constant K^{trans} [mm^3/(s mm^3)] models flow from plasma to EES (filtration) and $C(t)$ models the concentration in the whole volume.

concentration in A. Then it holds for the contrast agent concentration in Ω at time point t that

$$C(t) = v_a c_{\text{a}}(t) + K^{\text{trans}} \int_0^t c_{\text{a}}(s)\,\mathrm{d}s. \tag{2.10}$$

PROOF. We can derive the Patlak Model as follows: Let v_a, v_e [mm^3/mm^3] denote the fractional plasma and EES volume respectively and $c_{\text{a}}(t), c_e(t)$ [mmol/mm^3] the contrast agent concentrations in plasma and EES. We can express the contrast agent concentration in Ω as

$$C(t) = v_a c_{\text{a}}(t) + v_e c_e(t).$$

However, if we assume that there is some flow K^{trans} [mm^3/(s mm^3)] from Plasma to EES, then the same derivation as in Lemma 1 shows that

$$\left| \begin{aligned} (v_E c_E)'(t) &= K^{\text{trans}} c_{\text{a}}(t), \\ c_E(0) &= 0. \end{aligned} \right| \tag{2.11}$$

Straight-forward integration yields the claim. $\qquad\square$

2.2.2 Discussion

In this section we have introduced a model to describe the contrast agent concentration in the kidney. Note that the Patlak-Rutland model assumes that once contrast agent

is brought into the EES, it stays there and is not excreted. Since the Patlak-Rutland model also explicitly models the contribution of the plasma compartment, it is valid for highly perfused tissue where contrast agent extravasation but no outflow is expected [137]. In the kidney the Patlak Rutland Model is typically only valid in the early uptake phase since no outflow is modeled [68, 137]. It acute stroke the Patlak model is also used to model blood-brain-barrier leakage in the early uptake phase [75].

Apart from the Patlak-Rutland-Model there are manifold different pharmacokinetic models which are used for GFR estimation. To cope with the disadvantage that the Patlak Model is only valid during early uptake, in [5] the following ODE was assumed to govern the contrast agent concentration in the EEC:

$$(v_E c_E)'(t) = K^{\text{trans}} c_{\text{a}}(t) - K^{\text{trans}}(1 - f)c_E(t) \text{ and } c_E(0) = 0.$$

Note that as compared to (2.11) tubular outflow is modeled with a rate of $(1-f)K^{\text{trans}}$ for the extraction fraction $f \in (0, 1)$. In [139] this model is even extended by not using the arterial plasma concentration c_{a}, but by using a possibly diffused version c_P. However, note that significantly more parameters are introduced, which are scaling exponentially in the model. This makes the numerical model-fitting difficult and sensitive with respect to noise [22, 16, 151, 46].

Practical Remarks and the Role of Hematocrit

If the pharmacokinetic models like the Patlak model are used for GFR estimation, it is important to note that the GFR refers to the amount of blood plasma which is filtered by the kidneys. If the AIF is selected from measurements in a major vessel, these measurements typically only reflect the contrast agent concentration in the blood. This means that the measured AIF needs to be normalized by the plasma fraction, which is given by $(1 - \text{Hct})$, where Hct is the hematocrit-value of the patient For large vessels hematocrit is often assumed to be in the range $0.43 - 0.45$ [68, 139]. Note, however, that the hematocrit value varies both by patient and by vessel size [161]. Also note that K^{trans} is given in normalized units $[\text{mm}^3/(\text{s}\,\text{mm}^3)]$ whereas the GFR is given in absolute units [ml/min]. In order to convert K^{trans} to the GFR, a typical recalculation is given by $\text{GFR} = K^{\text{trans}} \cdot \text{Vol(Kidney)}$.

2.2.3 Alternative Methods for GFR Estimation

In this section we will introduce alternative methods to estimate the GFR. Specifically we will introduce GFR estimation based on Creatinin, which is the most important method in daily clinical routine. This section is mainly based on [42].

Most methods aim to approximate the GFR by estimating not the GFR itself, but the *renal clearance* of a marker substance. The main idea to determine renal clearance is

based on Fick's principle [52] which relates the filtered excreted amount of a substance with the excreted amount, i.e.

$$\frac{n_F}{\Delta t} = \frac{n_E}{\Delta t} \tag{2.12}$$

where n_F, n_E [mmol] are the number of particles which are filtered and which are excreted respectively and Δt is some time interval. The total numbers of particles can be expressed in terms of concentration and volume, i.e.

$$n_F = V_F c_P \text{ and } n_E = V_U c_U. \tag{2.13}$$

where in the case of the kidney V_F [ml] is the total amount of filtered plasma, c_P [mmol/ml] is the plasma concentration, V_U [ml] is the total amount of urine and c_U [mmol/ml] is the urine concentration of the substance. Inserting (2.13) into (2.12) and rearranging yields

$$\text{clear} := \frac{V_F}{\Delta t} = \frac{V_U c_U}{c_P \Delta t}, \tag{2.14}$$

the renal clearance [ml/min] of a substance. As described above, for substances which are fulfilling certain criteria the renal clearance and the GFR are coinciding. The most important examples for such substances are *Creatinin*, which is a product of the muscle metabolism and is hence created by the body itself, and *Inulin*, which is a polyfructosid and needs to be injected. For these substances equation (2.14) shows that the GFR can be estimated from four parameters: The concentration in the plasma, the concentration in the urine, the amount of urine and time. However, especially the measurement of the amount of urine is difficult in clinical practice [42].

The most common methods in everyday clinical practice are hence based exclusively on the concentration of Creatinin in the plasma and neglect the measurement of urine volume and urine concentration. Instead default values are used which depend on parameters ranging from sex, weight and age (Cockcroft-Gault) to sex, weight, age, ethnicity and c_P (CKD-EPI). Advantages of this methods are their speed, good conformity with clinical routine and proven reliability to detect kidney failure [42]. However, since (2.14) shows that the GFR is proportional to $1/c_P$, a reliable diagnosis of a significantly reduced GFR based on these methods can only be performed for a GFR which is smaller than 50 ml/min [42]. Also, the kidneys are able to compensate for failing areas by increasing filtration in healthy areas, which means that heavy losses in overall glomerular filtration might stay undetected. Finally traditional methods only allow to estimate the GFR for both kidneys, the failing kidney can hence not be detected by these methods.

2.3 The SPGR Signal Equation for T_1 Estimation

In this section we introduce the so-called Signal Equation for Spoiled Gradient Echo (SPGR) Sequences. Although this might seem like a significant break, we will see that

structurally the models from the previous section are similar, since all are describing the one dimensional response of a system in dependency of a continuous variable: Whereas pharmacokinetic models describe contrast agent concentration in a reference volume over time, the signal equation describes the magnitude MR signal of a reference volume over the flip-angle. We will use the Signal Equation to recover the tuple (M_0, T_1), where M_0 is a constant which depends on the steady state magnetization and the transversal relaxation time T_2. Additionally, we will describe the influence of contrast-agent on the apparent T_1 of the tissue. We conclude this section by introducing alternative methods for T_1 estimation.

2.3.1 The Signal Equation for SPGR Sequences

In this section we model the signal equation for spoiled-gradient echo (SPGR) sequences, closely following [93]. However, note that this section follows argumentation designed to describe an NMR signal and is thus applicable only for one voxel: For a complete analysis of such sequences for MRI see [96, 168].

We start by giving a brief overview of SPGR sequences: In SPGR sequences, magnetization is not flipped by full 90°, thus avoiding long recovery times before the next measurement [54]. Instead a fast series of pulses with considerably smaller flip angles of e.g. 15° is employed until eventually a steady-state magnetization is established. After each pulse, so-called spoiling is used to destroy the transversal magnetization. Note that this use of spoiling considerably simplifies the description of the MR signal, since the signal is mainly governed by processes affecting the longitudinal magnetization and transversal magnetization can be neglected. We start this section by introducing a description of strictly longitudinal recovery processes. After that, we will introduce a recursive formula to describe the steady-state magnetization which is established after a finite number of pulses. We conclude this section with the main result: The signal-equation (2.19) for SPGR sequences.

We start by introducing the Bloch equations, which describe the magnetization of a system exposed to a magnetic field [14]. However, note that these will be significantly simplified due to additional assumptions to be made later. Given an external magnetic field $\mathbf{B} : [0, T] \rightarrow \mathbb{R}^3$, the Bloch equations are describing the magnetization $\mathbf{M} : [0, T] \rightarrow \mathbb{R}^3$ of a system by

$$\dot{\mathbf{M}}(t) = \gamma \mathbf{M}(t) \times \mathbf{B}(t) - \mathbf{R}(\mathbf{M}(t) - \mathbf{M_{00}}) \qquad \text{for } \mathbf{R} := \text{diag}(\tfrac{1}{T_2}, \tfrac{1}{T_2}, \tfrac{1}{T_1}).$$

Here $\mathbf{M_{00}} \in \mathbb{R}^3$ is the equilibrium magnetization of the system, $\gamma \in \mathbb{R}$ the gyromagnetic ratio and $\times$ is the cross-product. Since $\mathbf{B}(t) \times \mathbf{M}(t)$ is orthogonal to the plane spanned by $\mathbf{B}(t)$ and $\mathbf{M}(t)$, the first term can be interpreted as a torque which causes $\mathbf{M}$ to rotate if a magnetic field $\mathbf{B}$ and $\mathbf{M}$ are linear independent. The matrix $\mathbf{R}$ introduces so-called relaxation effects in x, y and in z direction, causing the magnetization to return to the original state with rates $1/T_2$ and $1/T_1$ respectively. In MRI, usually

the magnetic field $\mathbf{B_0}$ is applied to point only in one direction, i.e $\mathbf{B_0} = (0, 0, B_0)^\top$. In this case the equilibrium magnetization $\mathbf{M_{00}} = (0, 0, M_{00})^\top$ will point in the same direction as B_0. In full form, the Bloch equations are then given by

$$
\dot{\mathbf{M}}_x(t) = \gamma \left(\mathbf{M}_y(t)\mathbf{B}_z(t) - \mathbf{M}_z(t)\mathbf{B}_y(t) \right) - \frac{\mathbf{M}_x(t)}{T_2},
$$

$$
\dot{\mathbf{M}}_y(t) = \gamma \left(\mathbf{M}_z(t)\mathbf{B}_x(t) - \mathbf{M}_x(t)\mathbf{B}_z(t) \right) - \frac{\mathbf{M}_y(t)}{T_2},
$$

$$
\dot{\mathbf{M}}_z(t) = \gamma \left(\mathbf{M}_x(t)\mathbf{B}_y(t) - \mathbf{M}_y(t)\mathbf{B}_x(t) \right) - \frac{\mathbf{M}_z(t) - M_0}{T_1}.
$$

In MRI, so-called radio-frequency (RF) pulses can be used to create an additional external field $\mathbf{B_1}(t) = (B_1(t), 0, 0)^\top$ for a very short time. It can be shown that such a magnetic field tilts the magnetization by a flip-angle angle $\alpha \in [0, 2\pi)$. Additionally, after such a pulse the magnetization relaxes to the equilibrium magnetization while rotating around the z axis. Due to induction this rotation induces a voltage in receiving coils, the so-called free induction decay (FID) signal.

Following [93], we make the following assumptions:

Assumption 4. Let us assume that

- The external magnetic field $B(t)$ is constant, i.e. $B(t) = (0, 0, B_0)^\top$.

- After an RF pulse and subsequent spoiling the magnetization of the system is given by $M(0) = (0, 0, M^+)^\top$.

In this case the above equations simplify to

$$
\left|
\begin{aligned}
\dot{M}_z(t) &= -\frac{M_z - M_{00}}{T_1}, \\
M_z(0) &= M^+.
\end{aligned}
\right|
\iff M_z(t) = M_{00} + \left(M^+ - M_{00} \right) e^{-\frac{t}{T_1}}. \tag{2.15}
$$

Assume that a sequence of repeated pulses with constant flip-angle α is applied, see Figure 2.3 for an illustration of the situation. Further assume that directly after each pulse spoiling sequences are employed which are destroying all transversal magnetization. After that we wait a period of time TR and then the procedure starts again.

In the case of repeated pulses it holds that the magnetization M_n^+ directly after the n-th pulse and subsequent spoiling, which destroys the transversal magnetization, can be written as

$$
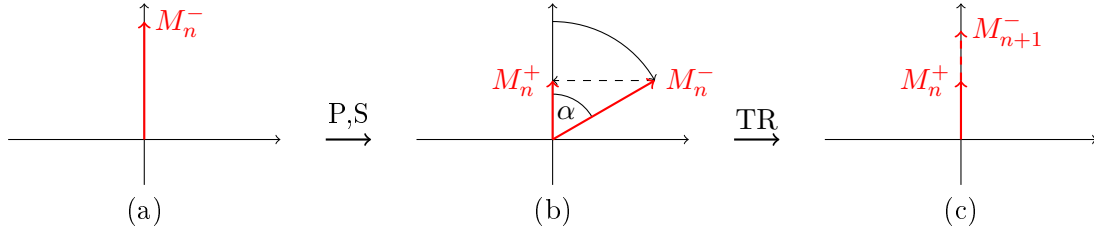M_n^+ = \cos(\alpha) M_n^- \tag{2.16}
$$

Figure 2.3: Figure displaying the magnetization during a SPGR sequence in a 2D co-ordinate system. (a) Displayed is the magnetization M_n^- before the n-th pulse. (b) Displayed is the magnetization M_n^+ after the n-th pulse and spoiling (P,S). The magnetization is tilted by the flip-angle α and imme-diately afterwards the transversal component is destroyed, see also (2.16). (c) Displayed is the magnetization M_{n+1}^- at time TR after the n-th RF pulse as determined by (2.17). An expression for M_{n+1}^- in dependency of M_n^- is the basis for the signal equation, cf. (2.18).

where M_n^- is the magnetization directly before the pulse, see Figure 2.3. Assuming that pulses are following a regular period TR, we can use (2.15) to model M_{n+1}^-, the magnetization right before the $(n+1)$-th pulse as

$$M_{n+1}^- = M_{00} + \left(M_n^+ - M_{00}\right) e^{-\mathrm{TR}/T_1}. \tag{2.17}$$

Combining (2.17) with (2.16) yields the following, recursive formula for the magneti-zation before the $n+1$-th pulse:

$$M_{n+1}^- = M_{00} + \left(\cos(\alpha)M_n^- - M_{00}\right) e^{-\mathrm{TR}/T_1}. \tag{2.18}$$

Letting $E_1 =: e^{-\mathrm{TR}/T_1}$ and $c := \cos\alpha$, we obtain by straight-forward recursion

$$M_n^- = M_{00}(1 - E_1) \sum_{j=0}^{n-1} (cE_1)^j + M_{00}(cE_1)^n$$

$$= M_{00}(1 - E_1)\frac{1 - (cE_1)^n}{1 - cE_1} + M_{00}(cE_1)^n.$$

Letting $n \to \infty$ and using $cE_1 < 1$ yields the steady-state solution

$$M_\infty^- = M_{00}\frac{1 - E_1}{1 - cE_1}.$$

To complete the modeling, two additional factors have to be included: First the mea-sured signal is proportional to the transverse component of the magnetization and second, not the free induction decay is measured by instead an echo of the signal after

33

time TE. Combining this yields the expression $S = \sin(\alpha)M_\infty^- e^{-\text{TE}/T_2^*}$ for the magnitude steady state signal S in dependency of the observed T_2 which is denoted by T_2^*. We thus obtain the signal equation for spoiled gradient echo sequences:

$$S(M_0, T_1) = M_0 \sin(\alpha) \frac{1 - e^{-\text{TR}/T_1}}{1 - \cos(\alpha)e^{-\text{TR}/T_1}}, \tag{2.19}$$

where $M_0 := M_{00}e^{-\text{TE}/T_2^*}$. See also [96, Section 3.4.7] and [168] for a different derivation of the signal equation.

2.3.2 Influence of Contrast Agent on T_1

In this section we will relate the contrast-agent concentration with the apparent T_1 of the tissue. The main sources for this section are [145, 117]. A large class of contrast agents in DCE-MRI is based on gadolinium (Gd^{3+}) ions. Since free metal ions are highly toxic to human tissue, they need to be embedded in larger molecules (chelates). In bound form, gadolinium mainly increases T_1-relaxivity, although it also affects T_2^* to some degree. Following [145], the effect of contrast agent on the effective T_1 can be modeled as

$$\frac{1}{T_1} = \frac{1}{T_{1,0}} + r_1 c \tag{2.20}$$

where $T_{1,0}$ [ms] is the T_1 of the tissue without contrast agent, r_1 [l/(mmol s)] is the (known) relaxivity constant of the contrast agent and c [mmol/l] is its concentration. Contrast agents like Magnevist or Omniscan are biologically almost inert and are distributed from the bloodstream to the extra-cellular-extra-vascular space rapidly [117]. They lose their effect within 1.3 h–1.6 h after injection and are excreted within approximately 24 h, mainly by the kidney [117]. Combined, these properties make this class of contrast agents especially interesting for measurement of the filtration capability of the kidney, since the excretion process can be observed in a time sequence of MR images.

2.3.3 Discussion

In this section we have introduced a model for the magnitude signal of spoiled gradient echo sequences. In various cases the actual signal might be described only imperfectly by the signal equation, the most important in clinical practice being imperfect flip angles. Since the flip-angle is determined by $\alpha = \gamma B_1 \tau$, where $\tau > 0$ is the duration of the pulse [96], local inhomogeneities in B_1 can cause the flip-angle to differ from the intended one [33]. This effect has been observed as a cause of error in various applications [20, 33]. Common methods to identify B_1 errors aim to measure B_1 by identifying the true signal zero around 180° [150, 44] or relate signals with constant flip-angle but varying repetition time [112, 163]. However, in this work we will assume that errors in B_1 are negligible. If additional measurements of B_1 are available, these can be incorporated into the reconstruction at will.

2.3.4 Alternative Methods for T_1 Estimation

As outlined in Section 1.2.3, the signal equation can be used to recover T_1 from multiple measurements of a volume which were obtained with different flip-angles but otherwise constant scan parameters. In this part we briefly describe other methods to estimate T_1. Most T_1 recovery methods make the assumption that the transversal magnetization M_{xy} is zero and that hence the signal intensity is mainly governed by the relaxing longitudinal magnetization. In this case one can see that in (2.19) apart from T_1 and M_0 there are only two free experimental parameters which describe this process: the flip-angle α and the repetition time TR. Where in the variable flip-angle method it is assumed that TR is constant and the flip-angles α are varied, so-called *Inversion or saturation recovery sequences* make a complementary assumption. Here, the flip-angle is assumed to be constant whereas the recovery time is varied. This can be interpreted as follows: Initially the magnetization is flipped by a fixed, known angle. After that the strength of the rebuilding magnetization is measured at different time points. The resulting curve has implicit information on the relaxation time T_1, which again can then be recovered by a model fit.

Saturation Recovery Sequences refer to the special case that the flip angle $\alpha = \pi/2$. In this case it follows for (2.15) that $M^+ = 0$ and hence

$$M_z(\text{TR}) = M_0(1 - e^{-\frac{\text{TR}}{T_1}}).$$

Using at least two different repetition times TR now allows to determine T_1 by a model fit. Note that the above equation only holds in the ideal scenario, that $\alpha = \pi/2$. Due to B_1 field inhomogeneities the actual α often differs from the one intended by the user. In this case we obtain

$$\begin{aligned}
M_z(\text{TR}) &= M_0(1 - (1 - \cos\alpha)e^{-\frac{\text{TR}}{T_1}}) \\
&= M_0(1 - (1 - (\pi/2 - \alpha))e^{-\frac{\text{TR}}{T_1}}) + \mathcal{O}(\alpha^2) \\
&= M_0((\pi/2 - \alpha)e^{-\frac{\text{TR}}{T_1}}) + \mathcal{O}(\alpha^2).
\end{aligned}$$

Note that now for $\text{TR} << T_1$ the measured signal will be dominated by α and hence will yield unreliable reconstruction results for slightly wrong flip-angles.

Inversion Recovery Sequences refer to a similar class of sequences, where the magnetization is inverted. Then the signal is allowed to relax for a variable time TI, the inversion time. In this case it follows from (2.15) that the remaining magnetization is given by

$$M_z(\text{TI}) = M_0(1 - 2e^{\frac{\text{TI}}{T_1}}).$$

In order to measure this magnetization, a 90° pulse is applied which rotates the remaining magnetization to the plane where $M_z = 0$. For an error-analysis for imperfect flip-angles see [77]. However, since these sequences have a very long repetition time, Look and Locker [93] introduced a modified type of these sequences. They proposed

to use the signal equation (2.19) with variable TR and low, but constant flip angles. This approach has been recently combined with ECG gating to estimate T_1 for cardiac MRI and termed Modified Look-Locker (MOLLI, [100]).

Note that for inversion or saturation recovery sequences one has to wait after each experiment for the equilibrium magnetization M_0 to be fully reestablished. Since it takes approximately a duration of $4.6 T_1$ for M_0 to return to 99% of its original value, this often yields long experiment times. Due to motion, it is difficult to use these sequences in abdominal imaging.

3 Uncoupled Dynamic Models: State of the Art Solution Strategies, Analysis and Approximation

Having described the models of interest in the previous chapter, we will now introduce state of the art methods to use these dynamic models to estimate tissue parameters. We will assume that parameters for only one dynamic curve are to be determined and neglect any spatial dependency. This issue will be addressed in Chapter 4 explicitly. Next to setting up a general framework for parameter estimation from noisy data, the main contribution of this section is a novel method to recover T_1 from variable flip-angle data. We will show in Section 3.6 and Section 7.1 that this novel method combines speed and efficiency of linear estimation with the accuracy and robustness of a nonlinear recovery.

We start this section by describing a general framework which establishes a connection between dynamic parameters and measured data. Since key to this connection is a suitable noise model, we start by arguing that for magnitude MRI images the expected noise is approximately Gaussian [73]. Based on this modeling, we introduce state of the art solution strategies to estimate parameters from noisy data using the models described in Chapter 2. We then describe analytical theory which can be used to determine if and how stable parameters can be estimated from noisy data with a given numerical method (so-called identifiability and sensitivity). These results are used to improve T_1 estimation: In Section 3.5 we construct a novel parameter estimation method for T_1 estimation, which has minimal sensitivity for a range of expected T_1. The novel method will be a linear estimation technique, which makes it computationally very efficient. However, as we will show its results are in accuracy comparable to a computationally more expensive nonlinear parameterfit. The benefits of this method will be demonstrated in Section 3.6 on software phantom data and in Section 7.1 on real data.

We start this chapter with a general remark: Whereas the theory which was introduced in Chapter 2 has been formulated in the continuous setting, in real-life applications usually only discrete measurements are available. In order to separate elements from the respective spaces, we will characterize elements from the continuous setting only in this chapter using bold fonts. Given a continuous dynamic curve $\mathbf{M}(p) \in C^0[0, \infty)$, we denote the discrete counterpart by $M(p) \in \mathbb{R}^k$, which we assume to be given by

$$M(p) := (\mathbf{M}(p; t_1, q), \dots, \mathbf{M}(p; t_k, q))^\top \in \mathbb{R}^k.$$

The continuous variable t might denote time in the case of pharmacokinetic models (2.2), (2.10) but might also denote the flip angle as in the case of the signal equation (2.19). If t denotes time, we will assume that the acquisition time points $(t_1, \ldots, t_k)$ are given by $t_i = \Delta t \cdot (i-1)$ for some interval $\Delta t > 0$. Note however, that there are various real-life situations in which the assumption that the timeline is evenly spaced is not fulfilled. Reasons for this might be breath-hold protocols in MRI [74] or dose reduction in CT [111, 161, 136]. In this case we will follow common practice and interpolate the data on a regular time grid [161, 136].

3.1 Noise Modeling in MRI

In the previous chapter we have introduced relationships between dynamic parameters and theoretical, perfect data. Since in imaging data is often degraded by noise we will now establish a relationship between theoretical, perfect data and noisy measurements. Since the focus of this work is MRI, we will follow [73] and show that for magnitude MRI, noise is approximately Rician distributed [122]. However, as another result we will see that the Rician Distribution can be approximated with a Gaussian distribution for good SNR. In all following work we will hence assume a Gaussian noise model for parameter recovery from MRI signal data. In software phantom data we will nevertheless use Rician noise.

As a short outline, we will follow [73] and assume that the Fourier transformed image data (so-called k-space data) is degraded by white Gaussian noise with uniform variance. However, as typical images used in clinical practice are magnitude images, we will have to describe the distribution of absolute values $|z| = |w + n| \in \mathbb{C}$ for white Gaussian noise $n \in \mathbb{C}$. From the example $w = 0$ one can see, that the expected distribution will not be Gaussian since no negative values can be attained. If on the other hand the variance of the noise is much smaller than the absolute value of w, one would expect an approximate Gaussian Distribution.

More specifically, let us assume that the measured (complex) k-space value $z \in \mathbb{C}$ can be decomposed as $z = w + n$ where $w = w_x + iw_y$ is the mean and $n = n_x + in_y$ is white Gaussian noise with zero mean and uniform variance σ^2. As for clinical applications the MR signal is usually given as a magnitude image without the phase information, we aim to model the distribution of $|z|$. To do this, we first observe that the probability density function for the cartesian coordinates of the measured signal $z = z_x + iz_y$ is given by

$$p_{\mathrm{C}}(z_x, z_y) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\left((z_x - w_x)^2 + (z_y - w_y)^2\right)}.$$

Since rotating the frame of reference will not change the distribution of absolute values, we can assume without loss of generality that $w_y = 0$ and $w_x = |w|$. In this case the above equation simplifies to the following expression for the rotated cartesian

coordinates:

$$p_{\mathrm{RC}}(z_x, z_y) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\left((z_x - |w|)^2 + z_y^2\right)}$$

Changing to polar coordinates $z_x = |z|\cos\phi$, $z_y = |z|\sin\phi$ yields the transformed density

$$p_{\mathrm{P}}(|z|, \phi) = \frac{|z|}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\left((|z|\cos\phi - |w|)^2 + (\sin\phi|z|)^2\right)}.$$

To establish the notational connection to the previous and following chapters, let us denote $d := |z|$ and $\bar{d} := |w|$. Integrating over the phase yields the following distribution of d given $\bar{d}$:

$$\begin{aligned} p_{\mathrm{A}}(d) &= \frac{d}{\sigma^2} e^{-\frac{1}{2\sigma^2}(d^2 + \bar{d}^2)} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\frac{\bar{d}d}{\sigma^2}\cos\phi} \, d\phi \\ &= \frac{d}{\sigma^2} e^{-\frac{d^2 + \bar{d}^2}{2\sigma^2}} I_0\left(\frac{d\bar{d}}{\sigma^2}\right) \end{aligned} \tag{3.1}$$

where I_0 is the (even) modified Bessel function of first kind of 0th order, which is equivalent to [3]

$$I_0(x) = \frac{1}{\pi} \int_0^\pi e^{x\cos\phi} \, d\phi.$$

The distribution described in (3.1) is the so-called Rice distribution [122].

Two limits of the Rician Distribution are of interest: For large SNR, i.e. if $|w|/\sigma >> 0$, the Rician Distribution converges to a Gaussian-Distribution with mean $\sqrt{\bar{d}^2 + \sigma^2} \approx |w|$ and variance σ^2 [65]. This means that for good SNR a Gauss-Distribution is a suitable approximation. However, it is worth noting that the magnitude signal will generally overestimate the true mean signal. The other limit of interest is if $\bar{d} = 0$. In this case a similar calculation as for (3.1) shows that

$$p(d) = \frac{d}{\sigma^2} e^{-\frac{d^2}{2\sigma^2}}.$$

This distribution is the so-called Rayleigh-Distribution and is significantly different to the Gaussian-Distribution, see Figure 3.1.

3.2 State of the Art Solution Strategies

In the last section we have completed the task to relate dynamic parameters with the measured data. In this section we will describe state of the art solution strategies to recover parameters from noisy measurements. Given (noisy) data d and a model $M(p)$, we will describe ways to determine $\bar{p}$ such that $M(\bar{p}) \approx d$.

In Section 3.2.1 we start by clarifying the exact meaning of $M(p) \approx d$: Specifically, we will model the distribution $f(d|p)$ of data given parameters by using the results of Chapter 2 and Section 3.1. Using a standard Log-Likelihood method we will then derive
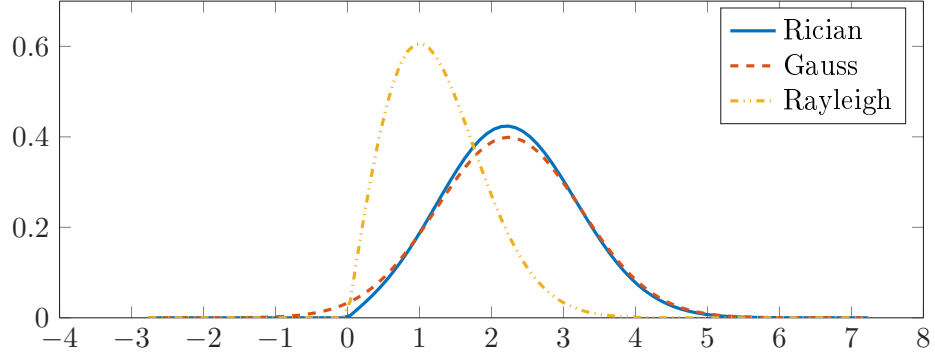
Figure 3.1: Comparison of different distributions for $\sigma = 1$ and $\bar{d} = 2$. Shown is the Rician distribution, Gauss normal distribution with mean $\sqrt{\bar{d}^2 + \sigma^2}$ and the Rayleigh distribution. One can see that even at SNR=2 the Rice distribution is well approximated by the Gauss distribution.

an L^2 distance measure for the parameter fit and obtain the well-known formulation for parameter estimation:

$$\text{Minimize } J(p) := \frac{1}{2\sigma^2} |M(p) - d|^2,$$

where σ^2 is the variance of the (uniform) noise. However, in practice this formulation is not a magic bullet for all parameter estimation problems, as the following examples show: In the case of CBF estimation using the Meier-Zierler Model, we will show in Section 3.3.2 that parameter estimation is ill-posed and data might be explained by multiple parameters. Suitable regularization is hence necessary to cope with this problem and will be introduced in Section 3.2.2. In the case of T_1 estimation, many authors use a linear reformulation of the quadratic model since it is easier to implement [66, 55, 154, 33], although results from [154, 33] indicate that it is less robust.

We will hence present state of the art solution strategies for the three considered models. In Section 3.2.2 we will show how the Meier-Zierler Model can be used to estimate blood flow, blood volume and mean transit time. To cope with the problem of ill-posedness, we will present a deconvolution approach which is regularized using a truncation method [111, 161]. In Section 3.2.3 we will show how (v_a, K^{trans}) can be estimated from a dynamic concentration curve C by using a straight-forward least-squares formulation of the Patlak-Rutland model. We conclude with Section 3.2.4, where we describe two state of the art solution strategies to determine (M_0, T_1) from variable flip angle data. Specifically we will introduce two methods for T_1 estimation: A linear method and a nonlinear method, which are both widely used in literature [66, 55, 154, 33].

3.2.1 Data Distribution and Log-Likelihood

We start by deriving the L^2 distance measure to recover parameters from measured data and Gaussian noise. In Chapter 2 we have established a relationship between dynamic parameters $\bar{p}$ and (perfect) data $\bar{d}$ and in the previous section we have established a relationship between (perfect) data $\bar{d}$ and measured data d. More thoroughly, for we have modeled $d = M(\bar{p}) + n$ for additive noise n. In the case of Gaussian noise this allows us to describe the expected distribution of measurements given parameters $\bar{p}$ as:

$$f(d|\bar{p}) = \frac{1}{\sqrt{2\pi \det \Sigma}} e^{-\frac{1}{2}(M(\bar{p})-d)^\top \Sigma^{-1}(M(\bar{p})-d)}.$$

where Σ is the (symmetric, positive definite) covariance matrix which models the noise.

In parameter estimation one is not interested in describing the probability to measure data d given some parameter p, but instead of finding parameters p which are explaining given data d. Given d, a straight-forward approach is to determine p such that the probability $f(d|p)$ is maximal [6]. To simplify the calculations, a well-known approach is to not maximize f but to use a monotonous logarithmic transformation $\log f$, the so-called log-likelihood function, yielding:

$$\text{Maximize } g(p) := \log f(d|p).$$

In the case of Gaussian noise, straight-forward calculations show that this maximization problem is equivalent to the minimization problem:

$$\text{Minimize } J(p) := |M(p) - d|_\Sigma^2.$$

In the rest of this work, we will hence use a standard L^2 distance to recover parameters. Note that this means that we are assuming an approximate Gaussian noise model for the data. In case of CBF or GFR estimation from DCE MRI data this is in line with current literature [161, 136]. However, since in this case the fitting is performed with respect to CA concentrations and not the raw signal intensities (see Section 2.3.2), future work could either establish a more suitable relationship or make use of concatenated models. Also, in the case of CT usually Poisson noise is expected. In this case usage of the Kullback-Leibler distance could improve parameter estimation [64], but has not been implemented by the author.

3.2.2 The Deconvolution Method for the Meier-Zierler Model

We will now describe a state-of the art method to estimate perfusion from DCE data using the Meier-Zierler Model, the so-called block-circular SVD (bSVD) method [160, 86, 143, 95, 79]. Indeed we will show in Lemma 8 that CBF estimation using the Meier-Zierler model is an ill-posed problem. Standard recovery methods which were

introduced in the previous section hence cannot be used to recover blood flow from measured data. Instead, suitable regularization has to be introduced. In the bSVD method [160], the deconvolution problem is reformulated as a matrix equation $LI = C$. For the convolution matrix L a block-circular structure is assumed, which makes the deconvolution robust against delay in contrast agent arrival. The equation is then solved using a standard truncated singular-value decomposition.

A Matrix Formulation for Convolution

Let us assume that a discrete tissue curve $C \in \mathbb{R}^k$ and an arterial input function $c_a \in \mathbb{R}^k$ are given. Our goal is to determine the impulse response function $I \in \mathbb{R}^k$. In the continuous setting, Lemma 2 shows that the three functions are connected by the relationship

$$\mathbf{C}(t) = (\mathbf{I} * \mathbf{c}_a)(t) = \int_0^t \mathbf{c}_a(t-s)\mathbf{I}(s) \, \mathrm{d}s.$$

This in connection with the results of the previous section motivates to formulate the discrete problem as follows:

$$\text{Minimize: } J(I) := |c_a *_{\mathrm{D}} I - C|^2. \tag{3.2}$$

Here $I, C \in \mathbb{R}^k$ and $*_{\mathrm{D}}$ is a suitable discretization of the convolution operator which will be described in the next paragraph.

In order to describe a suitable discretization for $*_{\mathrm{D}}$, let us start by elaborating on $*$ in the continuous setting: The Meier-Zierler Model describes the contrast agent concentration in the tissue as $\mathbf{C}(t) = \mathbf{I} * \mathbf{c}_a(t)$. Following Section 2.1, the impulse response function is given by $\mathbf{I}(t) = P_\Omega \mathbf{R}(t)$ and $\mathbf{R}(t)$ describes the fractional residues of contrast agent in the tissue at time t. The convolution can hence be interpreted as an integration over all contrast agent which has been introduced into the system up to time t, weighted with the respective probabilities that it resides in the tissue. For an evenly spaced timeline, a straight-forward discretization of $C = c_a *_{\mathrm{D}} I$ is hence given by

$$C_j = \Delta t \sum_{i=1}^j c_{a,j-i+1} I_i \iff C = \tilde{L}I \text{ for } \tilde{L} := \Delta t \begin{pmatrix} c_{a,1} & & & \\ c_{a,2} & c_{a,1} & & \\ \vdots & \ddots & \ddots & \\ c_{a,k} & \cdots & c_{a,2} & c_{a,1} \end{pmatrix}. \tag{3.3}$$

In practical perfusion imaging, this discretization comes with a significant problem: Specifically, contrast agent might be present in the tissue although no contrast agent has been measured in the arteries yet, i.e. $C(0) >> 0$ and $c_a(0) = 0$ [160]. This might be due to arteriovenous malformations or simply measurement errors [160]. Since it holds in this case by construction that $c_{a,1} = 0$, this formulation cannot account for
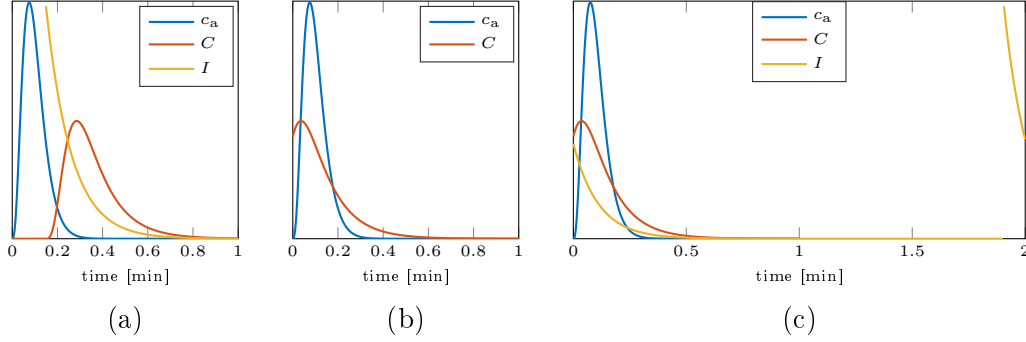
Figure 3.2: Effects of delay on the Impulse-Response function I obtained from a concentration curve C using deconvolution. Positive delay as in (a) will result in a shifted I. Negative delay as in (b) cannot be reconstructed with the standard deconvolution since by construction $(c_a * I)(0) = 0$ for all I. This behavior can be avoided by using a deconvolution with a block-circular matrix (3.4) as depicted in (c) where it is assumed that c_a is circularly shifted.

such data and will yield unreliable results, see Figure 3.2 for an illustration of this problem. To cope with this problem, it was proposed in [160] to use a block-circular matrix L with the following structure:

$$L := \Delta t \begin{pmatrix} l_1 & 0 & \dots & \dots & 0 & l_k & \cdots & \cdots & l_2 \\ l_2 & l_1 & \ddots & & & \ddots & l_k & \cdots & l_3 \\ \vdots & \ddots & \ddots & \ddots & & & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & & \ddots & l_k \\ l_k & \cdots & \cdots & l_2 & l_1 & 0 & \cdots & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{k, 2k-1}. \qquad (3.4)$$

where l_i depends on c_a and the chosen integration rule, in our case $l_i = c_{a,i}$. This structure allows to reconstruct I also from data C which shows earlier enhancement than the AIF, see Figure 3.2. Results from [160] have shown that this approach is indeed more robust to shifts than the standard approach for positive as well as negative shifts. Since its introduction, this so-called *block-circular matrix* has been used in a wide range of applications [86, 143, 95, 79] and will hence be used in the this work as the main method to recover perfusion using the Meier-Zierler Model.

Solution of the Linear System Using a Truncated SVD

We will show in Lemma 8 that CBF estimation using the discretization given in (3.4) is ill-posed and multiple parameters might explain the same data. To solve (3.2), it was proposed in [160] to use a truncated singular value decomposition of the block-circular convolution matrix L. Since the singular value decomposition describes a linear map in

rotated coordinate systems, where the axes are rearranged according to directions of maximal and minimal scaling, it is thus especially suited to isolate modes responsible for the ill-posedness of the problem. The following theorem shows that a singular value decomposition does indeed exists for any matrix and its proof gives further insight on how its results are to be interpreted:

Theorem 2 (Singular Value Decomposition, cf. [63]):
Let $A \in \mathbb{R}^{n,m}$ and $l := \min\{m, n\}$. Then there are orthonormal matrices $U \in \mathbb{R}^{n,n}$, $V \in \mathbb{R}^{m,m}$ and a diagonal matrix $\Sigma := \mathrm{diag}(\sigma_1, \dots, \sigma_l) \in \mathbb{R}^{n,m}$ for real numbers $\sigma_1 \geq \cdots \geq \sigma_l \geq 0$ such that
$$A = U\Sigma V^\top.$$

PROOF (CF. [63]). Let $A \in \mathbb{R}^{n,m}$ be arbitrary. We define
$$v_1 := \underset{|v|=1}{\mathrm{argmax}}\, |Av| \text{ and } \sigma_1 := |Av_1|.$$

Note that the maximum is attained as we are considering a continuous map over a compact set and $\sigma_1 \geq 0$. If $\sigma_1 = 0$ clearly $A = 0$. In this case it holds for $\Sigma := 0 \in \mathbb{R}^{n,m}$ and arbitrary orthonormal bases $U \in \mathbb{R}^{n,n}, V \in \mathbb{R}^{m,m}$ that $A = U\Sigma V^\top$.

Let us hence consider the case $\sigma_1 > 0$. We define $u_1 := Av_1/|Av_1| \in \mathbb{R}^n$. Let $V_1 \in \mathbb{R}^{m,m}$ and $U_1 \in \mathbb{R}^{n,n}$ be any extension of v_1 and u_1 to orthonormal bases of $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. Then it holds that
$$U_1^\top A V_1 = \begin{pmatrix} \sigma_1 & \tilde{v} \\ 0 & B \end{pmatrix} := \Sigma_1,$$

for $B \in \mathbb{R}^{n-1,m-1}$ and $\tilde{v} \in \mathbb{R}^{1,m-1}$.

We will now show that $\tilde{v} = 0$. As U_1, V_1 are orthonormal bases under which length is preserved, it holds that $\mathrm{argmax}_{|v|=1} |\Sigma_1 v| = \mathrm{argmax}_{|v|=1} |Av| = \sigma_1$. However, for the test-vector $w := (\sigma_1, \tilde{v})^\top$ it holds that
$$|\Sigma_1 w| \geq \sigma_1^2 + |\tilde{v}|^2 = \sqrt{\sigma_1^2 + |\tilde{v}|} \cdot |w| > \sigma_1 |w| \text{ if } |\tilde{v}| > 0$$

Hence $|\Sigma_1 w|/|w| > \sigma_1$ and it follows $\tilde{v} = 0$.

Repeating the same procedure for matrix $B \in \mathbb{R}^{n-1,m-1}$ again yields orthonormal matrices $V_1 \in \mathbb{R}^{n-1,n-1}, U_2 \in \mathbb{R}^{m-2,m-2}$, a positive real number $\sigma_2 \in \mathbb{R}$ and a matrix $C \in \mathbb{R}^{n-2,m-2}$. Now it holds that
$$\left(U_1 \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix} \right)^\top A \left(V_1 \begin{pmatrix} 1 & 0 \\ 0 & V_2 \end{pmatrix} \right) = \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & C \end{pmatrix}$$

and the claim follows by repeating the procedure until

$$U^\top AV = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_l \end{pmatrix}.$$

$\square$

Common thresholds for truncation are at approximately 10%, see [111, 161]. For deconvolution of perfusion data it has been shown that this procedure is very robust [111].

The bSVD Method for CBF Estimation and other approaches

For the rest of this work, the method to recover perfusion from CA concentration data by using the block-circular matrix and a truncated SVD will be referred to as *bSVD*. As compared to deconvolution with the matrix presented in (3.3), this formulation allows to account for delay in CA arrival. Note that the use of truncation acts as a regularizer, since small singular values are discarded in this formulation.

Naturally, there are many other approaches to solve the deconvolution problem: A straight-forward approach is to solve it in the Fourier domain [59, 110]. Since convolution in time-space corresponds to multiplication in the Fourier space and vice-versa (cf. [110]), a solution I can be found by

$$I = \mathcal{F}^{-1}(\mathcal{F}(C) / \mathcal{F}(c_\mathrm{a})).$$

Here $\mathcal{F}$ denotes some orthogonal transformation, as e.g. Fourier or Laplace Transform. Monte-Carlo simulations show that such approaches tend to underestimate the true flow even if additional filters are used [59, 159, 161]. Consequently, Fourier methods are rarely used in practice. Another class of deconvolution approaches are using Bayesian Modeling to additionally incorporate monotonicity constraints on R [17]. A recent evaluation [131] has shown that this method yields slightly improved estimates as compared to bSVD. However, since the method relies heavily on stochastic techniques like marginalization, expensive integrations need to be performed, yielding a slower performance than bSVD.

3.2.3 Linear Fit for Patlak-Rutland Model

The discretization of the Patlak-Rutland Model is straight-forward and leads to a standard linear system. Given discrete contrast agent concentration measurements in the kidney $C \in \mathbb{R}^k$ and arterial input function $c_\mathrm{a} \in \mathbb{R}^k$, a standard discretization of the Patlak-Rutland Model is given by

$$C = v_a c_\mathrm{a} + K^{\mathrm{trans}}(\tilde{L} c_\mathrm{a}) = Bp$$

where $\tilde{L} \in \mathbb{R}^{k,k}$ is the lower triangular matrix given by (3.3), $B := [c_{\mathrm{a}}, Lc_{\mathrm{a}}] \in \mathbb{R}^{k,2}$ and $p := (v_a, K^{\mathrm{trans}})^{\top} \in \mathbb{R}^2$. This means that given $C, c_{\mathrm{a}} \in \mathbb{R}^k$ it is possible to reconstruct p using a standard linear parameterfit, i.e. $p = VS^{\dagger}U^{\top}$ for the SVD $B := USV^{\top}$ and the pseudo-inverse $S^{\dagger}$ of S, see Theorem 2.

3.2.4 Two Solution Strategies for the Signal Equation

In this section we will introduce the most common formulations of T_1 estimation using the variable flip-angle technique, the so-called nonlinear and the linear approach. Historically, the nonlinear approach was the first method used for T_1 estimation [35]. Here T_1 is reconstructed with a so-called nonlinear parameter fit. It is known to yield more accurate results [154, 33] but can come with significant computational overhead, see [155] for a GPU-implementation. The so-called linear approach is based on a linear reformulation of the model [66, 55]. It is easier to implement but yields less accurate result if more than two flip angles are used for the reconstruction [154, 33].

We start by repeating the signal equation for spoiled gradient echo sequences (2.19):

$$M(M_0, T_1) = M_0 \sin \alpha \frac{1 - e^{-\mathrm{TR}/T_1}}{1 - \cos \alpha \, e^{-\mathrm{TR}/T_1}}$$

In the *nonlinear method*, the unknowns (T_1, M_0) are estimated as the solution of

$$\text{Minimize } J(M_0, T_1) = |M(M_0, T_1) - d|^2 \tag{3.5}$$

However, note that since there is no explicit solution formula, the minimization has to be carried out using iterative algorithm, cf. [107, 9]. In this work we used Gauss-Newton algorithms with hard constraints $T_1 \geq 0$ and $M_0 \geq 0$. These can be shown to converge Newton-like for low-residue least squares problems [9].

The *linear method* was first introduced by [66] for NMR and then used in [55] to estimate T_1 using the FLASH sequence. It relies on the following reformulation of the signal equation:

$$M(E_1, N) := \frac{N \sin \alpha}{1 - \cos \alpha E_1} \qquad \text{where } N := M_0(1 - e^{-\mathrm{TR}/T_1}), \ E_1 := e^{-\mathrm{TR}/T_1}.$$

Denoting $p := (E_1, N)^{\top}$ this allows to reformulate the relationship $M(T_1, M_0) = d$ as the following linear system:

$$Ap = b, \text{ for } A_i := [d_i/\tan \alpha_i \ , \ 1] \text{ and } b_i := d_i/\sin \alpha_i. \tag{3.6}$$

with $A \in \mathbb{R}^{k,2}$ and $b \in \mathbb{R}^k$. As for the Patlak-Rutland Model, the linear method comes with a closed form solution formula $p = A^{\dagger}b$ [87]. From a practical point of view it is preferable and has been used extensively in T_1 estimation [55, 154, 33]. However, note that since noisy data d will influence both A and b, the linear method is known to lack robustness [154, 33]. We will introduce an improved version of the linear method in Section 3.5.

3.3 Identifiability Analysis for Parametric Models

In the previous section we have presented solution strategies to determine dynamic parameters from measured data. However, it is not clear yet the recovery problem is indeed well-posed. Specifically, we could encounter the problem that different parameters might be linked to same measured data, yielding an ill-posed recovery problem.

In this section we will introduce a basic criterium to determine if parameter recovery using inverse modeling is possible, so-called *identifiability analysis*. We will use this theory in Section 3.3.2 to show that T_1 estimation using the signal equation and GFR estimation using the Patlak-Rutland model are usually well-posed problems, whereas CBF estimation using the Meier-Zierler model is ill-posed and multiple parameters might explain the same measured data.

The identifiability analysis will be carried out using a stochastic approach described in [125]. The main result of this section is Theorem 3, which shows that for a smooth parametric model $M(p)$ and a Gaussian noise model, a parameter vector p is identifiable if and only if $I(p) := \nabla M(p)^\top \nabla M(p)$ has full rank (cf. Lemma 7 and Theorem 3).

We start this section by recalling that we denote the probability density function of expected measurements d given some parameter p by $f(d|p)$, see also Section 3.2.1. The next definition clarifies what identifiability means. In a nutshell, parameters $p \in \mathbb{R}^m$ are called identifiable, if there are no other parameters which create the same expected distribution of measurements.

Definition 2.

1. Two parameter vectors $p_1, p_2 \in \mathbb{R}^m$ are called *observationally equivalent* if and only if $f(d|p_1) = f(d|p_2)$ for all d.

2. A parameter vector $p \in P \subseteq \mathbb{R}^m$ is called *identifiable in P*, if and only if there is no other $\tilde{p} \in P$ such that p and $\tilde{p}$ are observationally equivalent.

3. A parameter vector $p \in \mathbb{R}^m$ is called *locally identifiable*, if and only if there is some open neighborhood U of p such that p is identifiable in U.

In Section 3.3.1 we will present criteria to determine if parameters are locally identifiable. Note that for arbitrary models $M(p)$, results for global identifiability are much harder to obtain, see [125, 92] for some criteria. However, if noise is not too severe and measurements stay in the area around the expected measurement, we expect the local version to be a sufficient criterium to ensure that reconstruction is possible. In order to derive bounds on the size of this area, approaches originating from differential geometry might be used [7, 57]. These approaches analyze the curvature of the sample space $\{M(p)\}$ and relate it to the curvature of an approximating surface. However, these approaches are out of the scope of this work and will not be covered.

3.3.1 A Criterium for Local Identifiability

In order to derive a criterium for local identifiability we start this section by making some assumptions on the probability distribution $f(d|p)$:

Assumption 5. Let us assume that

A1) $P \subseteq \mathbb{R}^m$ is open,

A2) $f(\cdot|p)$ are a proper density functions, in particular $f \geq 0$ and $\int f(d|p) \, \mathrm{d}d = 1$ for all $p \in P$,

A3) the sample space $D := \{d \in \mathbb{R}^n : \ f(d|p) > 0\}$ is independent of p,

A4) f is smooth in p. Specifically for all $p \in P$ the functions $f(d|p)$ and $\log f(d|p)$ continuously differentiable with respect to p.

Theorem 3 will show that local identifiability is closely connected to another stochastic concept: The so-called Fisher information [27]. In motivating Fisher information, we closely follow [27]: For sake of simplicity let us assume that $p \in \mathbb{R}$. Following the theory derived in Section 3.2.1, we will assume that given a measurement d, parameters p are estimated by maximize the log-likelihood function $g(p) := \log f(d|p)$. However, if $g(p)$ is flat it would mean that small changes in p will only yield small changes in the likelihood and the reconstruction might not be trusted. Fisher information uses in this case the expected value of $(g'(p))^2$ with respect to all measurements d as an indicator of how well the parameters can be estimated.

Definition 3. Let $f(d|p)$ fulfill Assumption 5, let $p \in \mathbb{R}^m$ and let $g(p) := \log f(d|p)$. We define the Fisher information matrix of f at p as

$$I(p) := (I_{i,j}) \in \mathbb{R}^{m,m} \text{ where } I_{i,j}(p) := E_d \left(g_i(p) \cdot g_j(p) \right) = \int_D g_i(p) \cdot g_j(p) \, \mathrm{d}d,$$

for $g_i := \frac{\partial g(p)}{\partial p_i}$ and where $E_d(h)$ denotes the expected value of d with respect to the distribution h.

The following Theorem, which proof can be found in [125], shows that local identifiability can be expected if and only if the Fisher information matrix has full rank. As a technical prerequisite, we will additionally need to assume that $\bar{p}$ is a regular point of I, meaning that there is an open neighborhood of $\bar{p}$ such that $I(p)$ has constant rank.

Theorem 3 (See [125, Theorem 1]):
Let $f(d|p)$ fulfill Assumption 5 and let $\bar{p} \in \mathbb{R}^m$ be a regular point of $I(p)$. Then $\bar{p}$ is locally identifiable if and only if $I(\bar{p})$ has full rank.

PROOF (SEE [125]).

We will see in Section 3.4 that the same condition will be necessary to determine the sensitivity of reconstruction methods. Since determining the Fisher-information might be tedious, the following lemma shows that for twice differentiable functions Fisher information can also be expressed in terms of the curvature of g with respect to p. For $p \in \mathbb{R}$ it shows that indeed $I(p) = -E_d\left(g''\right)$:

Lemma 6. Let $f(d|p)$ fulfill Assumption 5, let $p \in \mathbb{R}^m$ and assume that g is twice differentiable with respect to p. Then

$$I(p) \in \mathbb{R}^{m,m} \text{ is given by } I_{i,j}(p) = -E_d\left(\frac{\partial^2 g(p)}{\partial p_i\,\partial x_j}\right).$$

PROOF (SEE [27]).

Since we are typically modeling noise with Gaussian character, the lemma derives the Fisher information for differentiable model M and Gaussian noise:

Lemma 7. Let $p \in P \subseteq \mathbb{R}^m$, $d \in D \subseteq \mathbb{R}^k$, let $M : P \to D$ be twice continuously differentiable and
$$f(d|p) := \frac{1}{\sqrt{2\pi\sigma^{2k}}}e^{-\frac{1}{2\sigma^{2k}}|M(p)-d|^2}.$$
Then the Fisher information matrix is given by
$$I(p) = \frac{1}{\sigma^{2k}}\left(\nabla M(p)^\top \nabla M(p)\right).$$

PROOF. To see this, let us first consider the likelihood function and its derivatives

$$g(p) = -\tfrac{1}{2} \cdot \log(2\pi\sigma^{2k}) - \frac{1}{2\sigma^{2k}}|M(p) - d|^2,$$

$$\nabla g(p) = -\frac{1}{\sigma^{2k}}(M(p) - d)^\top \nabla M(p),$$

$$\nabla^2 g(p) = -\frac{1}{\sigma^{2k}}\left(\nabla M(p)^\top \nabla M(p) + (M(p) - d)\nabla^2 M(p)\right).$$

where $\nabla^2 M(p) \in \mathbb{R}^{m,m,k}$ is a tensor containing second derivatives. Let us now calculate $I_{i,j}$. Dropping p for better readability and denoting derivatives by $g_i := \partial g/\partial p_i$ and $g_{i,j} := \partial^2 g/(\partial p_i \partial p_j)$ yields

$$I_{i,j} = -E_d\left(g_{ij}\right) = E_d\left(\frac{1}{\sigma^{2k}}\left(M_i^\top M_j + \sum_{i=1}^{k}(M_{ij})_k(M - d)_k\right)\right)$$

$$= \frac{1}{\sigma^{2k}}M_i^\top M_j + \sum_{i=1}^{k}(M_{ij})_k E_d((M - d)_k)$$

$$= \frac{1}{\sigma^{2k}}M_i^\top M_j$$

Here the last equality follows since it holds by construction that $E_d(d) = M$. □

3.3.2 Identifiability Results for the Proposed Models

In this section we will use the theory on identifiability outlined in the previous section to analyze the models which were introduced in Chapter 2. Specifically we will show that for all but the parameters of the deconvolution model are usually identifiable. The deconvolution model is a special case: Here reconstruction is maximally sensitive with respect to noise and parameters are usually not identifiable. This means that parameter estimation of CBF is delicate and will require special methods.

We start by summarizing our assumptions on the reconstruction:

Assumption 6. Let us assume that

1. The distributions are given by a Gaussian Noise model with constant variance σ^2, i.e.

$$f(d|p) := \frac{1}{\sqrt{2\pi\sigma^{2k}}}e^{-\frac{1}{2\sigma^{2k}}|M(p) - d|^2}.$$

2. For a measurement $y \in D$ parameter estimation is performed by minimizing the negative log-likelihood

$$\text{Minimize: } J(p) := |M(p) - d|^2.$$

Identifiability for the Meier-Zierler Model

Equation (3.3) shows that the deconvolution problem can also be written as a matrix equation:

$$\text{Find } I \in \mathbb{R}^k \text{ such that } LI = C.$$

For a regularly spaced timeline the discrete convolution matrix $L \in \mathbb{R}^{k,k}$ is given by

$$L := \Delta t \begin{pmatrix} c_{\mathrm{a},1} & & & \\ c_{\mathrm{a},2} & c_{\mathrm{a},1} & & \\ \vdots & \ddots & \ddots & \\ c_{\mathrm{a},k} & \cdots & c_{\mathrm{a},2} & c_{\mathrm{a},1} \end{pmatrix}.$$

Note that by construction the first component of the arterial input function $c_{\mathrm{a},1}$ is zero. This means that deconvolution is in this case indeed ill-posed. However, we will show in Section 3.4 that even if $c_{\mathrm{a}} \neq 0$ the sensitivity of the reconstruction will generally be very large: Small errors in the data may lead to large errors in the reconstruction.

Lemma 8. Let Assumption 6 hold an let $c_{\mathrm{a}} \in \mathbb{R}^k$. Then parameters $I \in \mathbb{R}^k$ are identifiable if and only if $c_{\mathrm{a},1} \neq 0$.

PROOF. Again note that the Fisher-Information matrix is given by $1/\sigma^{2k} L^\top L$ according to Lemma 7. Since $\det(A) = c_{\mathrm{a},1}^k$, it follows that L has full rank if and only if $c_{\mathrm{a},1} \neq 0$. The claim follows again from Theorem 3. □

Identifiability for the Patlak-Rutland Model

In this section we will derive identifiability results for GFR Estimation using the Patlak-Rutland Model. As the following Lemma shows that parameters are always identifiable, assuming that the arterial input function does not correspond to the analytic Dirac delta.

Lemma 9. Let Assumption 6 hold, let $0 \neq c_{\mathrm{a}} \neq (0, \ldots, 0, 1)$ and let $L \in \mathbb{R}^{k,k}$ be given by

$$L = \Delta t \begin{pmatrix} 1 & & \\ \vdots & \ddots & \\ 1 & \cdots & 1 \end{pmatrix}.$$

> Then the parameters $p = (v_a, K^{\text{trans}})$ are locally identifiable by the Patlak-Rutland Model.

PROOF. First note that according to Lemma 7, the Fisher-Information matrix is given by

$$I(p) = 1/\sigma^{2k} B^\top B \text{ for } B := (c_a, Ac_a) \in \mathbb{R}^{k,2}.$$

By construction L has only one eigenvalue Δt which has algebraic multiplicity k, geometric multiplicity 1 and corresponding eigenvector $e_k = (0, \ldots, 1)^\top$. This means that the matrix $B^\top B$ has full rank if and only if $c_a \neq 0$ and $c_a \neq e_k$ and the claim follows with Theorem 3. $\qquad\square$

Identifiability for the Signal Equation

In this section we will show that both the nonlinear and the linear parameter fit for T_1 reconstruction are usually well-posed:

> **Lemma 10.** Let Assumption 6 hold. Then (T_1, M_0) are identifiable by the linear and the nonlinear fit if and only if $k \geq 2$, $T_1 \neq \infty$ and $M_0 \neq 0$.

PROOF. We start by showing the claim for the linear model: A direct calculation shows that the Fisher-Information matrix $I(p)$ is given by

$$I(p) = 1/\sigma^{2k} A^\top A \text{ where } A := \begin{pmatrix} d_1/\tan\alpha_1 & 1 \\ \vdots & \vdots \\ d_k/\tan\alpha_k & 1 \end{pmatrix}.$$

According to Theorem 3 parameters are identifiable if and only if $I(p)$ and hence A has full rank. Let us check the columns of A for linear dependency: Assume that there is $\lambda, \mu \in \mathbb{R}$ such that $\lambda \tan\alpha_i - \mu d_i = 0$ for all i. In this case

$$\lambda \frac{\sin\alpha_i}{\cos\alpha_i} = \mu \left(M_0 \sin\alpha_i \frac{1 - E_1}{1 - \cos\alpha_i E_1} \right)$$

$$\frac{\lambda}{\mu M_0(1 - E_1)} = \frac{\cos\alpha_i}{1 - \cos\alpha_i E_1}$$

However, if $E_1 \neq 1$ and $M_0 \neq 0$ the last equality is never fulfilled for all i since the right hand side is independent of the flip-angle. It follows by assumption that A has full rank.

For the nonlinear case we have

$$\nabla S = \begin{pmatrix} d_1 \ s(\alpha_1) & d_1/M_0 \\ \vdots & \vdots \\ d_k \ s(\alpha_k) & d_k/M_0 \end{pmatrix} \text{ where } s(\alpha_i) := \frac{(\cos \alpha_i - 1)(\partial E_1/\partial T_1)}{(1 - \cos \alpha_i E_1)(1 - E_1)}.$$

Since $d \neq 0$ and since there are at least two measurements with different flip-angles it follows that ∇S has full rank and hence (T_1, M_0) is identifiable. $\qquad\square$

3.3.3 Discussion

Given a parametric and an expected noise model, we have derived both a sufficient and a necessary criterium to determine if a set of parameters can be recovered from multiple measurements. Theorem 3 shows in connection with Lemma 7 that for Gaussian noise, parameters $\bar{p} \in \mathbb{R}^m$ are identifiable if and only if the Fisher information matrix $I(\bar{p}) = [\nabla M(\bar{p})]^\top [\nabla M(\bar{p})]$ has full rank.

It is worth noting that this analysis is carried out with respect to the parameters and not with respect to the data. It thus cannot be used to decide if parameter results from a given curve are meaningful. However, it nevertheless brings deeper understanding of the models: In the case of T_1 estimation we have seen, that parameters which are not identifiable are associated with zero-data, i.e. $d = 0$. This means that recovery from data with low signal intensities cannot be trusted. For CBF estimation using the Meier-Zierler Model we have seen that estimation is always ill-posed. This follows because this specific problem can be formulated as a linear problem with a singular matrix. Only for the Patlak-Rutland model parameter estimation is usually well-posed, if extreme limit cases are ignored.

We stress, that identifiability is a very basic criterium: It refers to the theoretical possibility to recover a given pair of parameters from multiple observations of associated dynamic curves. Identifiability states only that recovery is possible but not how well this is the case. Indeed, in practice $I(\bar{p})$ might be close to singular, requiring a possibly large number of measurements to recover $\bar{p}$. This is especially crucial since in medical imaging often only one dynamic curve is given. We will address this issue, the so-called sensitivity, in the next section.

3.4 Sensitivity Analysis for Parametric Models

Whereas the last section was concerned with establishing theory which describes *if* parameters can be reconstructed from measurements, this section will be concerned with the question of *how well* this recovery is possible. It will be following the results the author presented in [S3] and the more general approach presented in [51]. The main result of this section is Theorem 6, which gives an expression for both the expected value and the variance of the estimated parameters if they are recovered from noisy measurements of a dynamic curve. These results will be basic for novel method for T_1 estimation, which will be introduced in Section 3.5, where we will construct an estimation technique which has minimal sensitivity.

We start with a short, motivating part for the sensitivity analysis, where we will not go into technical detail. Specifically, we will estimate the sensitivities of parameter estimation techniques $\phi : D \to P$, which are mapping data d to estimated parameters p. As a practical example, $\phi(d)$ could be given by a local minimizer of the negative log-likelihood function, i.e.

$$\phi : D \to P,$$
$$d \mapsto \operatorname*{argmin}_{p \in P} \left(- \log f(d|p) \right).$$

where $D \subseteq \mathbb{R}^k$, $P \subseteq \mathbb{R}^m$ and for motivational purposes we will assume that the minimum is unique. In this section we will derive results concerning the distribution of the recovered parameters $p = \phi(d)$. Specifically we are interested in variance of the distribution of p if it is reconstructed from noisy measurements of some ground-truth $\bar{d}$. Note that the variance gives us a measure of sensitivity: If we find that the variance of p is large, we know that the estimated $\bar{p} := \phi(\bar{d})$ cannot be trusted since small noise on $\bar{d}$ might yield large errors in the recovered p. Theorem 6 gives an explicit expression for the variance of p if it is estimated by some function $\phi : D \to P$. Since ϕ will often be highly nonlinear, we use linearization techniques to ease the analysis. For a given point $\bar{d}$ we will derive variance and mean of the distribution of $p = \phi_L(d) := \phi(\bar{d}) + \nabla\phi(\bar{d})(d - \bar{d})$ where it will be assumed that ϕ is differentiable. The following Theorem shows how the variance and mean of the linearized functional can be calculated:

Theorem 4 (Mean and Variance of ϕ_L, see [51]):
Let $P \subseteq \mathbb{R}^m$ and $D \subseteq \mathbb{R}^k$. Assume that $d \in D$ is distributed with mean $\bar{d}$ and covariance $\operatorname{Cov}(d) \in \mathbb{R}^{k,k}$. Finally let $\phi \in C^1(D, P)$, $\phi(\bar{d}) = \bar{p}$ and

$$p = \phi_L(d) := \phi(\bar{d}) + \nabla\phi(\bar{d})(d - \bar{d}).$$

Then
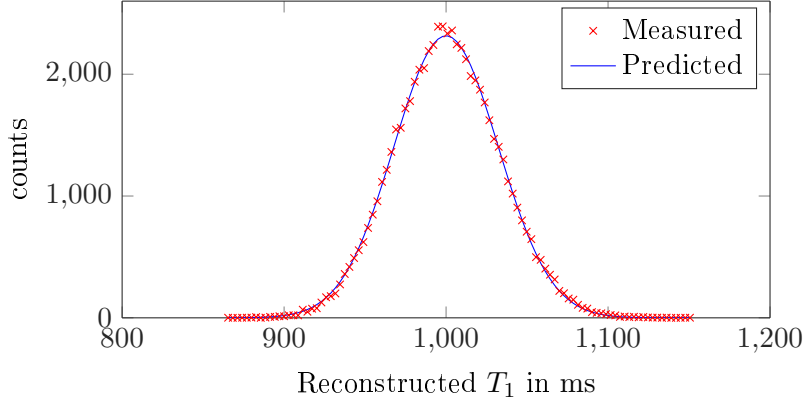
Figure 3.3: Example for Sensitivity Analysis: Measured and theoretical distribution of T_1 recovered from 65.000 noisy measurements using the nonlinear data-fit. The data was simulated with $T_1 = 1000$ms, $M_0 = 1000$, TR $= 800$ms and flip-angles $\alpha_1 = 10°$ to $\alpha_{10} = 100°$ in increments of $10°$. White gaussian noise had a standard deviation $\sigma_d = 10$ms. In this experiment the measured and predicted standard deviation both were $\sigma_{T_1} = 32.3$ms on an error level of 10^{-1}ms.

1. $E_d(p) = \bar{p}$, where $E_d(p)$ is the expected value of d with respect to the distribution p and

2. $\text{Cov}(p) = [\nabla\phi(\bar{d})]\,\text{Cov}(d)[\nabla\phi(\bar{d})]^\top$.

PROOF. The proofs are following directly:

1. $E_d(p) = E_d\left(\phi(\bar{d}) + \nabla\phi(\bar{d})(d - \bar{d})\right) = \phi(\bar{d})$, since $E_d(d) = \bar{d}$.

2. Follows also directly with the identities $\text{Cov}(Ap) = A\,\text{Cov}(p)A^\top$ and $\text{Cov}(p+c) = \text{Cov}(p)$ for constant $A \in \mathbb{R}^{m,k}$ and $c \in \mathbb{R}^k$. That these identities hold can be seen from

$$\text{Cov}(Ap) = E_p\left((Ap - E_p(Ap))(Ap - E_p(Ap))^\top\right)$$
$$= E_p\left(A(p - E_p(p))(p - E_p(p))^\top A^\top\right)$$
$$= A\,\text{Cov}(p)A^\top$$

due to the linearity of the expected value. The other identity follows from

$$\text{Cov}(p + c) = E_p\left(((p + c) - E_p(p + c))\left((p + c) - E_p(p + c)\right)^\top\right) = \text{Cov}(p)$$

since $E_p(c) = c$. □

3 Uncoupled Dynamic Models: State of the Art Solution Strategies, Analysis and Approximation

Theorem 4 gives a simple, explicit expression for both the variance and expected value of the linearized reconstruction function ϕ at point $\bar{d}$. However, there are many applications where ϕ is not given parametrically and the derivative $\nabla\phi$ is not straightforward to compute, as the following example shows: In the case of a nonlinear model-function $M(p)$ and Gaussian noise, ϕ could be given by

$$\phi(d) = \operatorname*{argmin}_{p \in P} g(p, d) \quad \text{where } g(p, d) := |M(p) - d|^2.$$

where for motivational purposes it is assumed that a unique minimum exists. Theorem 6 shows that in this case so-called implicit-function theorem [83, p.112] can be used to nevertheless calculate the derivative $\nabla\phi$.

Again let us briefly not go into technical detail and let us motivate the proof of Theorem 6 as follows: Key to an explicit expression for $\nabla\phi$ is the following reformulation of ϕ:

$$\phi(d) = \{p \in P : \ \nabla_p g(p, d) = 0\}$$

The implicit function theorem states that if g fulfills regularity assumptions to be specified later, then there is a function h which parametrizes the zero-levelset of $\nabla_p g(p, d)$, i.e. $h(d) = p \iff \nabla_p g(p, d) = 0$. Note that this implies that $\phi(d) = h(d)$, since h maps each measurement d on corresponding parameters p which minimize $g(\cdot, d)$. In this case a straight-forward application of the chain rule furthermore yields the following expression for the derivative $\nabla\phi(d)$:

$$\nabla_p g(\phi(d), d) = 0$$
$$\implies \ [\nabla_p^2 g(\phi(d), d)]\nabla\phi(d) + [\nabla_{dp} g(\phi(d), d)] = 0$$
$$\iff \ \nabla\phi(d) = -[\nabla_p^2 g(p, d)]^{-1}[\nabla_{dp} g(p, d)].$$

where $p := \phi(d)$, ∇_{dp} denotes differentiation with respect to both d and p and it is assumed that $[\nabla_p^2 g(\bar{p}, \bar{d})]$ is invertible.

If we define $f(p, d) := \nabla_p g(p, d)$, the Implicit Function Theorem states the following:

Theorem 5 (Implicit Function Theorem, [83, p.112f]):
Let $f \in C^1(U, \mathbb{R}^m)$ in a neighborhood $U \subseteq \mathbb{R}^m \times \mathbb{R}^k$ of a root $(\bar{p}, \bar{d})$ of f. Furthermore assume that $\nabla_p f(\bar{p}, \bar{d})$ is invertible. Then there are neighborhoods $P \subseteq \mathbb{R}^m$ of $\bar{p}$, $D \subseteq \mathbb{R}^k$ of $\bar{d}$ and a function $h \in C^1(D, P)$, with $\bar{p} = h(\bar{d})$ such that the zero-levelset of f is the graph of h, i.e.

$$f(p, d) = 0, \ (p, d) \in P \times D \iff p = h(d), \ d \in D.$$

Furthermore the derivative $\nabla h(\bar{d})$ is given by

$$\nabla h(\bar{d}) = -[\nabla_p f(\bar{p}, \bar{d})]^{-1}[\nabla_d f(\bar{p}, \bar{d})].$$

PROOF. Given in [83, p.112f]. □

The following theorem summarizes, how sensitivities can be computed for a standard maximum-likelihood reconstruction in the presence of Gaussian Noise:

Theorem 6:
Let $M \in C^2(\mathbb{R}^m, \mathbb{R}^k)$ be a parametric model and assume that for constant variance σ^2 the expected measurements are distributed as

$$f(d|p) := \frac{1}{\sqrt{2\pi\sigma^{2k}}} e^{-\frac{1}{2\sigma^{2k}}|M(p)-d|^2}.$$

Furthermore assume that $M(\bar{p}) = \bar{d}$, that $H(\bar{p}) := [\nabla M(\bar{p})]^\top [\nabla M(\bar{p})]$ is invertible and that

$$\phi(d) := \underset{p \in \mathbb{R}^m}{\operatorname{argmin}} |M(p) - d|^2$$

Then reconstructions $p = \phi(d)$ from a neighborhood of $\bar{d}$ are approximately distributed with

$$E_d(p) = \bar{p} \quad \text{and} \quad \operatorname{Cov}(p) = \sigma^2(AA^\top)$$

for $A := [H(\bar{p})]^{-1}[\nabla M(p)]^\top$.

PROOF. We follow the general outline as described in the beginning of the section and calculate mean and expected value of the linearized functional $\phi_L(d) := \phi(\bar{d}) + \nabla\phi(\bar{d})(d - \bar{d})$. We start by giving an explicit expression for $\nabla\phi(\bar{d})$ by using the implicit function Theorem with $g(p, d) := \nabla_p |M(p) - d|^2$:

$$g(p, d) = [\nabla M(p)]^\top (M(p) - d) \qquad\qquad \in \mathbb{R}^m$$
$$\nabla_p g(p, d) = [\nabla^2 M(p)](M(p) - d) + [\nabla M(p)]^\top [\nabla M(p)] \in \mathbb{R}^{m,m}$$
$$\nabla_d g(p, d) = -[\nabla M(p)]^\top \qquad\qquad\qquad \in \mathbb{R}^{m,k}$$

Here $\nabla M(p) \in \mathbb{R}^{k,m}$ is the Jacobian of M and $[\nabla^2 M(p)]$ is a $m \times m \times k$ tensor containing second derivatives of M. Note that by assumption the implicit function theorem is applicable. Using it at location $(\bar{p}, \bar{d})$ yields the following expression for $\nabla\phi(\bar{p})$:

$$\nabla\phi(\bar{p}) = [H(\bar{p})]^{-1}[\nabla M(\bar{p})]^\top \in \mathbb{R}^{m,k}$$

where $H(\bar{p}) := [\nabla M(\bar{p})]^\top [\nabla M(\bar{p})]$.

We can now use this expression directly with the results which were obtained in Theorem 4: The claim $E_d(p) = \bar{p}$ follows by construction from the properties of f. For the expression for the covariance $\operatorname{Cov}(p)$ it follows:

$$\operatorname{Cov}(p) = \sigma^2 [\nabla\phi(\bar{p})][\nabla\phi(\bar{p})]^\top \in \mathbb{R}^{m,m} \qquad\qquad\qquad □$$

The following lemma shows, that the covariance matrix can be expressed in terms of an SVD of ∇M and depends quadratically on the inverse singular values of ∇M.

Lemma 11. Let $M \in C^2(\mathbb{R}^m, \mathbb{R}^k)$, $\sigma > 0$, assume that $\nabla M(\bar{p})$ has full rank and let $\mathrm{Cov}(p)$ be the covariance matrix of ϕ_L, see Theorem 6. Let furthermore $USV^\top = \nabla M(\bar{p})$ be an SVD of $\nabla M(\bar{p})$. Then

$$\mathrm{Cov}(p) = \sigma^2(VS_m^{-2}V^\top)$$

for $S_m^{-2} := \mathrm{diag}(1/s_1^2, \dots 1/s_m^2) \in \mathbb{R}^{m,m}$ and singular values $s_i > 0$ of ∇M.

PROOF. For ease of notation let us drop $\bar{p}$. Theorem 6 shows that $H = [\nabla M]^\top[\nabla M]$ and hence $H^{-1} = V(S^\top S)^{-1}V^\top$. Note that H is invertible by assumption. It follows that

$$A = H^{-1}[\nabla M]^\top = V(S^\top S)^{-1}S^\top U^\top = VS_{mk}^{-1}U^\top$$

where $S_{mk}^{-1} \in \mathbb{R}^{m,k}$ has diagonal elements $1/s_i$ for the singular values s_i. It follows from Theorem 6 that

$$\mathrm{Cov}(p) = \sigma^2(AA^\top) = \sigma^2(VS_m^{-2}V^\top)$$

where $S_m^{-2} := \mathrm{diag}(1/s_1^2, \dots 1/s_m^2) \in \mathbb{R}^{m,m}$. $\qquad\qquad\square$

3.4.1 Conclusion

Theorem 6 gives an explicit expression for the sensitivities of reconstruction methods, which are not given parametrically but as minimizers of an objective function. Since parameter estimation is often expressed as a fitting problems [6], this formulation covers a large class of applications. The theory of this section also provides a link to the identifiability analysis performed in the previous section: It can readily be seen that both Theorem 3 and Theorem 6 are fundamentally based on the criterium that $[\nabla M(p)]^\top[\nabla M(p)]$ is invertible in a neighborhood of $\bar{p}$. This property has a simple motivation: If the model can be linearized as $M(p) = M(\bar{p}) + \nabla M(\zeta)(p - \bar{p})$ for some $\zeta \in [\bar{p}, p)$ and $\nabla M(\zeta)$ has not full rank in a neighborhood of $\bar{p}$, reconstruction will certainly be ill posed. From this point of view the additional assumption that M is twice differentiable seems to be of a technical nature and it might be possible to relax it even more. However, since the models we have introduced are smooth enough, the current criterium is sufficient. Additionally Lemma 11 shows that small singular values of ∇M have adverse influence on the stability of the recovery: Again, this is in-line with the previous motivation, since for small singular values of M neighboring parameters might be mapped to very similar data, leading to unstable parameter estimation.

The next section will show how the obtained expression for the sensitivity can be used to improve T_1 estimation.

3.5 Linear Approximations in T_1 Estimation

In this section we will introduce a novel method for T_1 estimation from variable flip-angle data which combines the speed and robustness of a linear approach with the robustness of a nonlinear approach and has been published in [S3]. To do this, we will introduce a family of estimation approaches and show that both the linear and the nonlinear fit belong to this family. After that, we will construct a third member of this family which is similar to the linear fit, but more robust. Experimental evaluation of the novel method on phantom data can be found in Section 3.6, real data results in Section 7.1. In these sections we will show that the novel, linear recovery method yields both comparable results to a nonlinear parameter fit and improved results as compared to an established linear method.

We start this section with a brief summary of established T_1 estimation from variable flip-angle data [154, 33, 43]. For a more detailed outline we refer to Section 2.3. We will assume that multiple MR measurement $(d_1, \ldots, d_k) \in \mathbb{R}^k$ of an object were obtained with different flip angles $(\alpha_1, \ldots, \alpha_k)$ but otherwise constant parameters and are interested in an estimation of $p := (T_1, M_0)$. Both the linear and the nonlinear estimation approaches are based on the Signal Equation for Spoiled Gradient Echo Sequences (2.19):

$$M(M_0, T_1) = M_0 \sin \alpha \frac{1 - e^{-\mathrm{TR}/T_1}}{1 - \cos \alpha e^{-\mathrm{TR}/T_1}} = \frac{N \sin \alpha}{1 - \cos \alpha E_1}.$$

Here (M_0, T_1) and (E_1, N) are related by a change of coordinates

$$E_1 = e^{-\mathrm{TR}/T_1}, \qquad\qquad T_1 = -\mathrm{TR}/\log(E_1),$$
$$N = M_0(1 - e^{-\mathrm{TR}/T_1}), \qquad\qquad M_0 = N/(1 - E_1).$$

In the *Nonlinear Fit*, $p := (T_1, M_0)$ is reconstructed by

$$p = \operatorname*{argmin}_{p} |M(p) - d|^2$$

The *Linear Fit* estimates $p := (E_1, N)$ as the solution of the linear system:

$$A\bar{p} = b, \text{ for } A_i := [d_i/\tan \alpha_i \ , \ 1] \text{ and } b_i := d_i/\sin \alpha_i.$$

with $A \in \mathbb{R}^{k,2}$ and $b \in \mathbb{R}^k$.

3.5.1 A Family of Weighted T_1 Estimation Techniques

We will now introduce a family of weighted reconstruction approaches for T_1 estimation. To do this, we first recognize that the nonlinear residue can be written in terms of the linear residue as follows:

$$\frac{N \sin \alpha_i}{1 - \cos \alpha_i E_1} - d_i = \omega_i^{\mathrm{NL}}\Big(d_i/\tan \alpha_k E_1 + N - d_i/\sin \alpha_i\Big) = \omega_i^{\mathrm{NL}}(A_i p - b_i),$$

where $\omega_i^{\mathrm{NL}} := \sin \alpha_i/(1 - \cos \alpha_i E_1)$.

A straightforward generalization is hence to compute the unknowns $\bar{p}$ by solving a weighted k-by-2 least squares problem [87]:

$$\phi(d) = \operatorname*{argmin}_p |Ap - b|_W^2 = \sum_{i=1}^{k} \omega_i^2 (A_i p - b_i)^2. \tag{3.7}$$

We can interpret the linear approach by Gupta [66] as a choice $\omega^{\mathrm{Lin}} := (1, \ldots, 1)^\top \in \mathbb{R}^k$, which results in a standard linear least squares problem with an explicit solution formula, cf. Section 3.2.4. Note, however, that lower errors in reconstruction can be achieved using nonlinear weighting, as noise influences A_k as well as b_k [154, 153, 33]. Based in the Sensitivity Analysis in Section 3.4, we will now introduce linear weights which approximate the nonlinear weights. As the linear weighting method comes with a closed form solution formula, it is preferable from a practical point of view.

3.5.2 Linear Approximations of Nonlinear Weights

In this section we will introduce a method to construct linear weights, which approximate the nonlinear weights. To find the weights, we first calculate the sensitivity of recovered T_1 in dependency of the weights ω. Indeed, for the nonlinear and the linear case our findings coincide with the results obtained in [154, 43].

Lemma 12 (Sensitivity for linear weights). Let $\phi(d)$ be given by (3.7) and assume that the weights w are independent of the data d. Furthermore assume that $\bar{p} := (\overline{T_1}, \overline{M_0})$ is identifiable and that $\bar{p}$ and $\bar{d}$ are connected by $\bar{d} := M(\bar{p}) \in \mathbb{R}^k$. Finally let $d \propto \mathcal{N}(\bar{d}, \sigma)$ for the normal distribution $\mathcal{N}$ with mean $\bar{d}$ and variance σ^2.

Then the variance $\sigma_{E_1}^2$ of the reconstructed E_1 and the variance $\sigma_{T_1}^2$ of the reconstructed T_1 are given by

$$\sigma_{E_1}^2 = \sigma^2 |\nabla \phi_1(\bar{d})|^2 \quad \text{and} \quad \sigma_{T_1}^2 = \sigma_{E_1}^2 \cdot \left(\overline{T_1}^4/(\overline{E_1}\mathrm{TR})^2\right). \tag{3.8}$$

Here $\phi(d) = (\phi_1(\bar{d}), \phi_2(\bar{d}))^\top$ and

$$\nabla\phi(\bar{d}) = (A^\top W^2 A)^{-1} A^\top W$$

and the matrices $W := \mathrm{diag}(\omega_1, \ldots, \omega_k)$ and A are defined as in (3.7).

PROOF. Following Theorem 6, the covariance $\mathrm{Cov}(p)$ with respect to (E_1, N) is given by

$$\mathrm{Cov}(p) = \sigma^2 [\nabla\phi(\bar{d})][\nabla\phi(\bar{d})]^\top.$$

where $\nabla\phi(\bar{d})$ is as given above. Since $\mathrm{Cov}(p)_{1,1}$ is the variance of E_1, we have $\sigma^2_{E_1} = \sigma^2 |\nabla\phi_1(\bar{d})|^2$. To obtain the variance $\sigma^2_{T_1}$, we need to account for a change of coordinates and multiply with

$$\frac{dT_1}{dM_0} = \frac{\mathrm{TR}}{(\log E_1)^2 E_1} = \frac{T_1^2}{E_1 \mathrm{TR}}.$$

This yields the claim. $\qquad\square$

The obtained expression for $\sigma^2_{T_1}$ is a nonlinear function both in the weights ω and the parameters $\bar{p}$: Given some ground-truth parameters $\bar{p} := (\overline{T_1}, \overline{M_0})$ and a set of weights ω, the value $\sigma^2_{T_1}$ describes the variance of the recovered T_1. Naturally, this value should be as small as possible. Given a set of expected parameters $p_1, \ldots, p_L$, we propose to solve the following minimization problem to obtain a set of optimized weights $\omega \in \mathbb{R}^k$ which minimizes the joint variance for $p_1, \ldots, p_L$:

$$\omega^{\mathrm{Opt}} = \operatorname*{argmin}_{\omega \in \Sigma} \left\{ \sum_{j=1}^{L} f_j \sigma^2_{T_1}(p_j, \omega) \right\}. \tag{3.9}$$

Here $f_j \in \mathbb{R}_0^+$ are user-defined factors which can be used to highlight particular ranges of parameters p_j in the reconstruction. The optimized weights can be used to reconstruct T_1 by solving a linear system but with improved sensitivity of the reconstruction. Note that $\sigma_{T_1}(\bar{p}, \omega) = \sigma_{T_1}(\bar{p}, s\omega)$ for any $s > 0$, since a uniform scaling of the weights will not change the sensitivity. The objective function is hence invariant with respect to scaling and minimization ill-posed. To cope with this issue, we define the set of admissible weights $\Sigma := \{\omega \in \mathbb{R}^L, \ \omega \geq 0, \ \sum \omega_i = 1\}$.

To illustrate the setting, we visualize the landscape of standard deviations σ_{T_1} in dependency of the weights for $\overline{T_1} = 1000\,\mathrm{ms}$ and $\overline{M_0} = 1000$ in Figure 3.4. We then solved problem (3.9) to obtain a set of weights with minimal variance. Results are showing ω^{Opt} different from ω^{NL}, but with the same standard deviation.
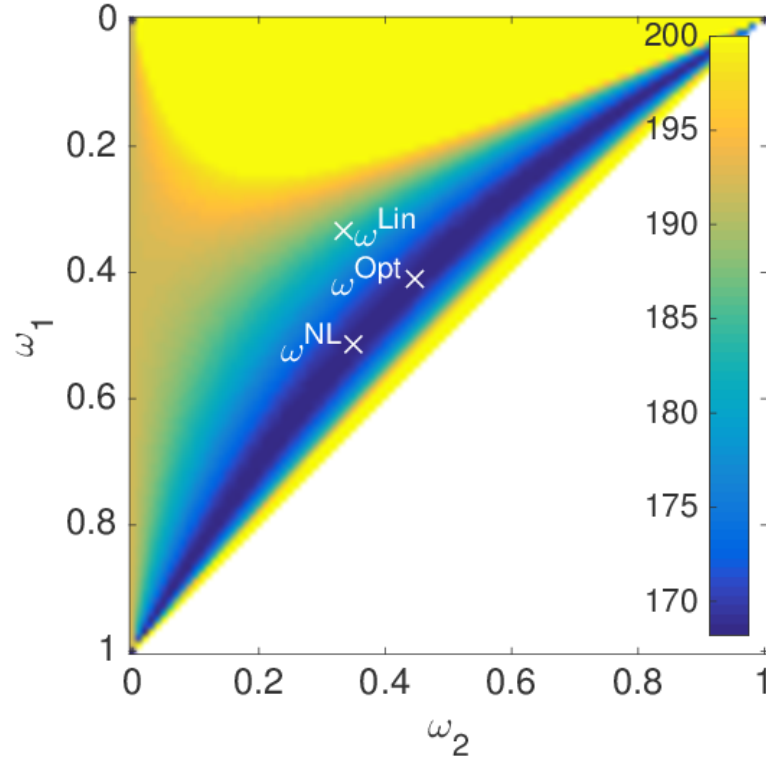
Figure 3.4: Displayed is σ_{T_1} for different weights $\omega = (\omega_1, \omega_2, 1 - \omega_2 - \omega_3) \in \mathbb{R}^3$ and fixed T_1 and M_0. Control parameters were $T_1 = 1000\,\text{ms}$, $M_0 = 1000$, $\alpha = (5°, 8°, 15°)$ and $\text{TR} = 5\,\text{ms}$. Highlighted by a white $\times$ are the nonlinear weights, linear weights and the optimized weights. Standard deviations for the optimized weights and for the nonlinear weights were $\sigma_{T_1}^{\text{Opt}} \approx \sigma_{T_1}^{\text{NL}} \approx 168.21\,\text{ms}$ with $\sigma_{T_1}^{\text{Opt}} - \sigma_{T_1}^{\text{NL}} = 2.05 \times 10^{-11}\,\text{ms}$. Standard deviation for the linear weight was $\sigma_{T_1}^{\text{Lin}} = 184.34\,\text{ms}$. For this configuration σ_{T_1} can be reduced by 9% using the optimized or the nonlinear weighting scheme instead of the standard linear one.

3.6 Results: Linear Approximations for Software Phantom Data

We will now demonstrate how T_1 estimation can be improved by using the optimized weights, which were introduced in the previous section. Specifically, we will compare the following three different choices of weights:

$$\omega^{\text{Lin}} := (1, \dots, 1) \in \mathbb{R}^k, \quad \omega^{\text{NL}} := \sin \alpha_i / (1 - \cos \alpha_i e^{-\text{TR}/T_1}), \quad \omega^{\text{Opt}} \text{ given by } (3.9).$$

In this section we will present phantom data which indicates that as compared to the standard linear weights, the optimized weights are yielding improved T_1 results which are similar to the ones obtained with the nonlinear weights. However, note that the optimized weights allow to recover T_1 using a standard linear parameter fit and are computationally much more efficient. Real data experiments which support this assertion will be presented in Section 7.1.

We start by outlining a software phantom for the human kidney which served as ground-truth for our experiments on synthetic data.

3.6.1 An XCAT Software Phantom for the Human Kidney

In this section we will describe the construction tailored software phantom for the human kidney. A simulated 2D MRI scan of the human kidney was set up as follows: The anatomy (cortex, medulla, background) was obtained from the XCAT-Phantom [133]. Estimates of T_1 and T_2^* of the mentioned structures were taken from literature [8]. The second constant M_0 was determined by $M_0 = M_{00}\exp(-\text{TE}/T_2^*)$ for a constant M_{00}, which was roughly approximated from our own measurements. Note that since no values for T_2^* were available, we approximated T_2^* with T_2. More precisely, we chose for the kidney cortex $T_1 = 966\text{ms}$, $T_2 = 87\text{ms}$ and $M_{00} = 3100$. The medulla was set up with parameters $T_1 = 1412\text{ms}$, $T_2 = 85\text{ms}$ and $M_{00} = 3500$ and for the background we chose $T_1 = 110\text{ms}$, $T_2 = 85\text{ms}$ and $M_{00} = 2000$, [8]. After that, partial volume effects were simulated by smoothing the parameter maps M_0 and T_1 with a Gauss-filter of width 3 voxel and standard-deviation of 0.7 voxel. Finally, we used the signal-equation (2.19) to simulate (perfect) MR signals with scan parameters taken from different publications on T_1 estimation. Sequence S1 was designed to yield robust reconstruction over a large range of T_1 [33], S2 and S3 were chosen from a publication on the nonlinear reconstruction method [153] and S4 was used to acquire the real data presented in this work. Examples of the phantom images are displayed in Figure 3.5

In order to add more realistic conditions, Rician noise was simulated by adding uniform white Gaussian noise to the k-space data, see Section 3.1. After that, the noise intensity was measured point wise as the fraction of noise of the undisturbed signal and then averaged over the complete dataset. Similar to [153], in our experiments effective noise intensities of $3\%, 5\%, 7\%$ and 9% were simulated.

Identifier	Flip-Angles	TR	TE	Source
S1	$(2°, 9°, 19°)$	4.4ms	1.1ms	[33]
S2	$(2°, 10°, 20°, 30°)$	2.7ms	1ms	[153]
S3	$(5°, 10°, 15°)$	4ms	1.67ms	[153]
S4	$(5°, 8°, 15°, 25°)$	2.51ms	0.89ms	[74]

$$\alpha_1 = 5° \qquad \alpha_2 = 8°$$
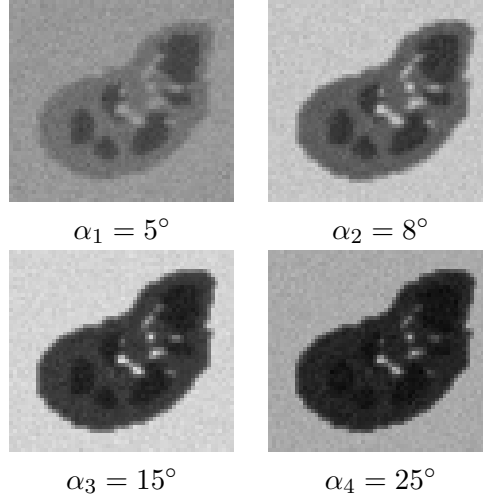

$$\alpha_3 = 15° \qquad \alpha_4 = 25°$$

Figure 3.5: MR sequence parameters sequences and corresponding sources from literature. Also given is the software phantom for S_4 with a noise intensity of 7%.

Table 3.1: Table displaying the predicted sensitivities of the different weighting schemes for the sequences described in Figure 3.5. Given is the standard deviation in reconstructed T_1 with respect to $\sigma_d = 1$ms, see (3.8). Factors f_j were selected proportional to the expected amount of cortex,medulla and background. Results from the nonlinear and the optimized weighting scheme are in close agreement.

	Cortex			Medulla			Background		
	ω^{Lin}	ω^{NL}	ω^{Opt}	ω^{Lin}	ω^{NL}	ω^{Opt}	ω^{Lin}	ω^{NL}	ω^{Opt}
S1	18.29	17.70	17.70	40.39	37.12	37.12	2.71	1.61	1.92
S2	24.84	20.65	20.66	62.38	48.53	48.53	1.99	1.46	1.68
S3	28.61	27.57	27.57	80.28	76.09	76.09	2.10	1.99	2.00
S4	57.30	45.82	45.82	182.38	139.28	139.28	1.31	1.28	1.37

3.6.2 Experimental Comparison of Different Weighting Strategies

We will now show, how the proposed weighting strategies can improve T_1 estimation. It is well-known that the sensitivity of the reconstruction does not only depend on the reconstruction method and the expected (T_1, M_0), but also on the employed sequences parameters [33, 154, 43]. The proposed weighting strategies (Nonlinear, Linear, Optimized) were evaluated with respect to their theoretical sensitivities and with respect to the mean error on simulated as well as real data.

We calculated the standard deviation σ_{T_1} for the three proposed strategies (Nonlinear, Linear, Optimized) and for the sequences described in Figure 3.5. Results are given in Table 3.1, showing predicted sensitivities which are in close agreement to the nonlinear weights. However, note that these sensitivities are given with respect to a purely Gaussians noise model, whereas in MRI we expect Rician noise [65]. To compare the three strategies with Rician Noise, parameter estimation was performed on phantom data which was degraded by Rician Noise. Results are given in Table 3.2. Here it can be seen, that also for Rician Noise the optimized weights improve the errors in reconstructed T_1 drastically.

Also the improvement depends largely on the difference in sensitivities between the nonlinear weights and the linear weights: If these are very similar, as it is the case for S1, only minor influence on the reconstruction can be observed.

As expected the results depend on the employed sequence: For Sequences S2 and S4 we can observe large positive impact of weighting strategies on reconstruction with reductions in the standard deviation of up to 24%, for sequences S1 and S3 improvements are only minor. In our experiments, we found that the optimized weights can perform similar to the purely nonlinear weights with a fraction of the computational overhead. The phantom data results are confirmed by the real data results, where differences between the nonlinear reconstruction and the linear reconstruction with optimized weights were in the range of 1%, see Section 7.1.

Table 3.2: Table displaying improvement of the mean-relative error in reconstructed T_1 by use of the weighting schemes. Experiments were conducted with different sequences and different noise-levels on phantom data. Given is the mean relative error in percent, calculated in 5000 experiments and evaluated over Cortex and Medulla. Factors f_j were selected as in Tab. 3.1. The nonlinear parameter fit yields the smallest and the standard linear fit the largest errors. Results of the optimized weights are in close agreement with the nonlinear weights for small noise. For high noise level, results of the optimized weights and the nonlinear weights begin to differ, possibly due to the less gaussian character of the noise.

		3%	5%	7%	9%
	ω^{Lin}	5.48	9.06	12.68	16.74
S1	ω^{NL}	5.23	8.62	12.03	15.75
	ω^{Opt}	5.24	8.66	12.13	16.05
	ω^{Lin}	4.09	6.81	9.47	12.37
S2	ω^{NL}	3.34	5.56	7.72	10.05
	ω^{Opt}	3.34	5.58	7.76	10.15
	ω^{Lin}	10.76	18.59	43.45	96.91
S3	ω^{NL}	10.20	17.29	25.43	35.71
	ω^{Opt}	10.38	18.01	38.22	89.31
	ω^{Lin}	11.43	51.41	89.74	182.46
S4	ω^{NL}	8.72	22.72	24.79	34.86
	ω^{Opt}	8.81	21.30	38.53	113.57

4 Spatial Coupling for Dynamic Models with Multiple Parameters

Since in dynamic imaging acquisition times need to be low, we have argued in Section 1.2.4 that low SNR is one of the main problems estimate high-quality parameter maps. In order to improve parameter estimation in the presence of noise, in this chapter we will introduce a way to include spatial information into the estimation process. Up to now we have only addressed the recovery of parameters from single dynamic curves. However, as the example of T_1 estimation or of CBF estimation shows, there are many applications where the recovery needs to be performed voxel-wise, leading to a large amount of estimation problems. Since in clinical practice experiments are typically only performed once, it is necessary that all parameter estimation problems yield accurate results. If, for some reason, data is degraded by noise or imaging artifacts, recovered parameters might be ultimately inexact with no further way to improve them. To cope with this issue, either the imaging process or the estimation process can be improved. In this work we will focus on improving the estimation model.

We will do this by including prior information on the expected spatial structure of the recovered parameter maps by so-called *spatial coupling*. In contrast to uncoupled recovery, where parameters for each voxel are estimated individually, in coupled recovery, parameter values of different curves will depend on each other. The specific form of the connection follows an expectation on the spatial layout of the parameter map. Our focus will lie on local coupling, where parameter values of adjacent voxels are connected. In the simplest case, one might expect that these should not vary largely. Specifically, we will introduce a novel class of coupling terms to parameter estimation, which is originally stemming from RGB denoising. These coupling terms are tailored to stabilize the recovery of multiple parameters from data which with clear edges. In Section 4.4 we will demonstrate advantages for the case of T_1 estimation. We will demonstrate on phantom data, that using these coupling terms can improve recovery up to 8% as compared to established approaches.

We start in Section 4.1 by describing established techniques for spatial coupling. We will distinguish between coupling by denoising and fully coupled approaches. By coupling by denoising we will refer to pipeline approaches, where first an unstabilized parameter estimation is performed and spatial information is added subsequently. In contrast, fully coupled approaches are combining these two steps: As we will see, these can be regarded as standard estimation approaches with the constraint that the recovered parameter maps are fulfilling a user-defined spatial regularity property. We

will introduce three of the most common coupling terms: So-called Tikhonov coupling, Tikhonov-Phillips coupling and Total Variation coupling. Analytical results are presented which show that Tikhonov coupling will result global scaling, Tikhonov-Philips coupling will yield smooth solutions and that Total Variation coupling allows the image to admit edges. Note that especially the last property is crucial for crisp parameter maps.

In Section 4.2 we will describe extended approaches for the case that multiple parameters are recovered per voxel, as e.g. in the case of (T_1, M_0) estimation. As we will see, established approaches for coupled T_1 estimation from variable flip angle data can be divided into two classes: either they are addressing the spatial regularity of only one parameter map and are recovering other parameters without any assumptions on their structure [155], or they are introducing regularity for each parameter map separately [114]. In contrast we will introduce a family of coupling terms, which are not only enforcing spatial regularity of one parameter map, but are coupling all estimated maps. These terms have been studied for for RGB denoising extensively [61, 88, 105] and are currently increasingly used for image reconstruction [148, 81]. However, we have not found indications that they have yet been proposed for T_1 estimation or even parameter estimation in general.

This chapter culminates in Section 4.3 where we will use the above modeling to improve T_1 estimation. The section will extend results which have been published by the author in [S3]. In phantom experiments described in Section 4.4 we will show that the novel coupling terms can improve errors in T_1 up to 8%. Results on real data support this claim and can be found in Section 7.2.

We conclude this chapter with Section 4.5, where we will focus on numerical aspects of the above modeling. The problem formulations of the previous sections will require non-smooth numerical methods, since the objective functions will be non differentiable. However, as we will see the proposed objective functions will still be convex and hence exhibit a large amount of structure which can be exploited by an efficient numerical scheme. In Section 4.5 we will briefly describe a Primal Dual Hybrid Gradient Algorithm (often called the Chambolle-Cremers-Pock Algorithm) [28] which is state-of-the-art algorithm to solve such problems.

Notation for Multiple Voxels

We start by introducing notation for parameter estimation for multiple voxels. To this end we assume that the data d consists of k three-dimensional images d^i, where each image is presented on a cell-centered l_1-by-l_2-by-l_3 grid of $n := l_1 l_2 l_3$ points x_j, $j = 1, \ldots, n$. Thus $d = (d^1, \ldots, d^n)^\top d \in \mathbb{R}^{nk}$ for $d^i \in \mathbb{R}^n$. Similarly, we assume that the unknowns p consist of m three-dimensional parameter maps p^i and hence $p = (p^1, \ldots, p^m)^\top \in \mathbb{R}^{nm}$ for $p^i \in \mathbb{R}^n$. For the model we will continue to use the

established notation and obtain $M : \mathbb{R}^{nm} \to \mathbb{R}^{nk}$, where we assume that the ordering of $M(p)$ corresponds to the ordering of the data d.

4.1 Traditional Techniques for Spatial Coupling

In this section we will introduce different techniques to include spatial information into a parameter estimation process. We will start by introducing coupling terms $\mathcal{C}$ to the reconstruction process, which penalizes parameter maps which do not exhibit an expected spatial regularity property. Novel techniques will be introduced based on this general modeling.

After that we will introduce conventional choices for such coupling terms. Specifically we will consider the following three terms: Tikhonov coupling which penalizes large variations, Tikhonov-Philips coupling which penalizes large gradients of the recovered parameter maps and Total Variation coupling which also penalizes large gradients but can be shown to act edge-preserving. We will further present mathematical results which show characteristic properties of solutions corresponding to the different choices. Specifically we will show that the use of Tikhonov coupling will only yield scaling of recovered parameter maps and that Tikhonov-Philips coupling will yield solutions which are significantly smoother than the noisy image. For Total Variation coupling we will provide an intuitive motivation why locally constant parameters maps are recovered.

4.1.1 A General Framework for Spatial Coupling

In this section we will introduce a mathematical framework to include spatial information into a parameter-estimation process. Specifically, we will motivate an approach to recover parameters by solving a problem of the type

$$\text{Minimize: } J(p) := \|M(p) - d\|^2 + \lambda \, \mathcal{C}(p)$$

where $J : \mathbb{R}^{nm} \to \mathbb{R}$ and $\mathcal{C} : \mathbb{R}^{nm} \to \mathbb{R}$ is a cost function which penalizes parameter maps which are not fulfilling some expected property, as e.g. spatial smoothness. In this work we will concentrate on so-called local coupling terms, where information of adjacent voxels is taken into account. Since discrete gradients are capturing such information, continuous mathematical theory might be used for analysis. However, note that there are various other approaches to include prior knowledge on the spatial layout. As an example, non-local formulations might be used to include information of voxels which are in some way similar but are not in a direct neighborhood [21, 48]. However, we will stick with local approaches, since the can often be motivated directly from physical models [29]. We now start by introducing our framework for parameter estimation from an optimization point of view.

Fully Coupled Approaches

Fully coupled approaches [114, 155] can be motivated as follows: The goal is to find parameters $p \in \mathbb{R}^{nm}$ which are explaining the data with reasonable accuracy but are at the same time fulfilling a spatial regularity assumption, which can be measured by a cost-function $\mathcal{C} : \mathbb{R}^{nm} \to \mathbb{R}_0^+$. A classic example for such a cost function is $\mathcal{C}(p) := \sum \|\nabla^h p^i\|_2^2$, where the matrix ∇^h is a discrete gradient operator. Small values of $\mathcal{C}(x)$ would hence correspond to a smooth parameter map, large values to large variations. Different choices for $\mathcal{C}$ will be covered in the next sections.

For a Gaussian distribution of parameter maps, the resulting optimization problem can be phrased as [107]

$$\text{Minimize } \mathcal{C}(p) \\ \text{s.t. } \frac{1}{2\sigma^2}\|M(p) - d\|_2^2 \leq \varepsilon \tag{4.1}$$

Following [6], rewriting this problem using the KKT conditions and a dual variable $\mu \geq 0$ [107] yields the equivalent reformulation as the saddle-point problem:

$$\underset{p\in\mathbb{R}^{nm}}{\text{Minimize}} \, \underset{\mu \geq 0}{\text{Maximize}} \quad L(p, \mu)$$

where the *Lagrangian* $L(p, \mu)$ is given by

$$L(p, \mu) := \mathcal{C}(p) + \mu(\|M(p) - d\|_2^2 - 2\varepsilon\sigma^2)$$

Note that in the above optimization problem the choice of ε depends on the user. To simplify matters, let us assume that instead we fix some value $\mu \geq 0$. Indeed we will see that in this case the problem will be independent of ε and we end up with the optimization problem:

$$\text{Minimize: } L(p) := \mathcal{C}(p) + \mu(\|M(p) - d\|_2^2 - 2\varepsilon\sigma^2)$$
$$\iff \text{Minimize: } J(p) := \|M(p) - d\|_2^2 + \lambda\mathcal{C}(p) \text{ for } \lambda := 1/\mu. \tag{4.2}$$

Here $\lambda \geq 0$ controls the strength of influence of the stabilizer: If λ is large, $\mathcal{C}$ will have large influence and we will obtain parameter maps which are fulfilling the regularity assumptions heavily. If on the other hand λ is small, the distance-term becomes the dominant term in (4.2) and the data fit becomes more prominent. However, note that a good choice for λ is not trivial: Although there are strategies to choose λ automatically, such strategies often come with considerable numerical overhead since optimization has to be performed with respect to λ [70]. In this work we will hence confine to the straight-forward approach to determine λ experimentally.

Coupling by Denoising

Compared to the fully coupled approach, coupling by denoising can be regarded as a sequential approach to parameter recovery. In coupling by denoising, first an uncoupled voxel-wise parameter estimation is performed, yielding estimated parameter-maps $p = (p^1, \ldots, p^m)^\top \in \mathbb{R}^{nm}$. After that, one or more of the maps are denoised, which introduces the desired coupling between parameters at different locations.

Given the recovered, unstabilized parameters $p \in \mathbb{R}^{nm}$ the denoising problem can be expressed similarly to (4.1) as

$$\text{Minimize } \mathcal{C}(q)$$
$$\text{s.t. } \tfrac{1}{2\sigma^2}\|q - p\|_2^2 \leq \varepsilon$$

The same argumentation as earlier leads to the problem:

$$\text{Minimize: } J(q) := \|q - p\|_2^2 + \lambda\,\mathcal{C}(q). \tag{4.3}$$

Discussion

The approach *coupling by denoising* can be regarded as a crude alternating direction approach to solve (4.2): In order to decouple the system, first parameters are estimated from data and then the coupled parameter map is obtained subsequently. Advantages of denoising approaches as for any pipeline approach are that available implementations for the respective sub-problems can be used at will. Note, however, that many authors have found that results using the simultaneous method are superior to denoising techniques [114, 155]. The fully coupled approach on the other hand often comes with some numerical overhead, since implementations need to be performed for each model and coupling term individually. However, since it is not clear which of the two proposed modelings will yield superior results, our experiments in Section 4.4.1 will cover both choices.

4.1.2 Established Choices for Coupling Terms

Having introduced a general framework to introduce spatial coupling into parameter recovery, we will now describe three established choices for the coupling term $\mathcal{C}$ and their properties. Specifically we will introduce Tikhonov coupling [147], Tikhonov-Phillips coupling [147] and Total Variation coupling [126]. We will present mathematical results which show that $\mathcal{C}_{\mathrm{TP}}$ will smear edges and that $\mathcal{C}_{\mathrm{T}}$ will have only a scaling effect on the solution. On the other hand, we will motivate that Total Variation coupling allows the solutions to admit edges. A practical comparison of the three coupling terms on real T_1 maps can be found in Figure 4.2. This section will follow the outline given in [29].

Let us start by switching to a continuous setting, where we assume for ease of presentation that $\Omega := (0,1)^2 \subseteq \mathbb{R}^2$ and $p : \Omega \to \mathbb{R}$. We start by giving explicit expressions for the stabilizers under investigation:

$$\mathcal{C}_{\mathrm{T}}(p) := \int_\Omega p^2 \, \mathrm{d}x, \qquad\qquad \text{for } p \in L^2(\Omega),$$

$$\mathcal{C}_{\mathrm{TP}}(p) := \int_\Omega |\nabla p|^2 \, \mathrm{d}x, \qquad\qquad \text{for } p \in H^1(\Omega),$$

$$\mathcal{C}_{\mathrm{TV}}(p) := \sup_{\substack{\xi \in C_0^\infty(\Omega;\mathbb{R}^2), \\ \||\xi|_2\|_\infty \leq 1}} \int_\Omega \langle p, \nabla\xi \rangle \, \mathrm{d}x, \text{ for } p \in L^2(\Omega).$$

Given a function $p \in L^2(\Omega)$, in this section we will focus on the following three model-problems:

Minimize:
$$J_i(p) := \|p - p_0\|_2^2 + \lambda\mathcal{C}_i(p) \tag{4.4}$$
$$\text{s.t. } p \in X_i$$

where

$$\mathcal{C}_1 := \mathcal{C}_{\mathrm{T}}, \qquad \mathcal{C}_2 := \mathcal{C}_{\mathrm{TP}}, \qquad \mathcal{C}_3 := \mathcal{C}_{\mathrm{TV}},$$
$$X_1 := L_0^2(\Omega), \quad X_2 := H_0^1(\Omega), \quad X_1 := L_0^2(\Omega).$$

We will now derive regularity results for the solutions of these problems, demonstrating the smoothness properties which can be expected.

We start with the simplest case, with Tikhonov coupling. Here the following Lemma shows that a solution will be just a uniformly scaled version of p_0 and is hence ill-suited to remove outliers.

Lemma 13. For Tikhonov coupling (4.4) admits a unique solution $p : \Omega \to \mathbb{R}$ which is given by
$$p_{\mathrm{T}}(x) = \frac{p_0(x)}{1 - \lambda}.$$

PROOF. Existence follows from [39, Theorem 3.3]. Further, solving the Euler-Lagrange Equations shows that the minimizer is given by the above expression. Uniqueness follows from the strict convexity of the squared L^2 norm. $\qquad\square$

However, things are different for Tikhonov-Phillips coupling. As the next Lemma shows, solutions of (4.4) will be significantly smoother inside Ω than the function p_0:

Lemma 14. For Tikhonov-Phillips coupling (4.4) admits a unique solution $p \in H_0^1(\Omega) \cap H_{\text{loc}}^{m+2}(\Omega)$.

PROOF. Let us consider the equation

$$\int_\Omega (p - p_0)\phi + \langle \nabla p, \nabla \phi \rangle \, \mathrm{d}x = 0 \tag{4.5}$$

for all $\phi \in C_0^\infty(\Omega)$. A straight-forward application of the Lax-Milgram Theorem [49, Section 6.2.1, Theorem 1] shows that (4.5) has a solution $\overline{p} \in H_0^1(\overline{\Omega})$ and [49, Section 6.3.1, Theorem 2] shows that the the solution is indeed in $H_{\text{loc}}^{m+2}(\Omega)$. However, since (4.5) is the Euler-Langrange Equation of the strictly-convex equation (4.4), it is automatically is a solution of (4.4) due to [39, Theorem 3.11]. Since (4.4) is strictly convex, the solution is unique. □

This lemma shows that for the coupling term $\mathcal{C}_{\text{TP}}$ we can expect the solution to smear edges: Even if p_0 has discontinuities, we can expect the solution p_{TP} to be smooth in the inside of Ω. Note that this might be desirable behavior for coarse resolutions, where smoothness due to partial volume effects is expected. However, for fine resolutions and tissue with clear structure we expect the image to admit discontinuities which cannot be captured by this coupling term. As the following Lemma motivates, for Total Variation coupling solutions of (4.4) on the other hand can exhibit discontinuities:

Lemma 15 (Regularity of Minimizers, see [29]). Let $\Omega := (0,1)$, $p_0 := \chi_{(1/2,1)}$, i.e. $p_0(x) = 0$ for $x \in (0, \frac{1}{2}]$ and $p_0(x) = 1$ for $x \in (\frac{1}{2}, 1)$. Then there is a minimizing sequence of (4.4) which converges to

$$p_{\text{TV}} := 1/2 \text{ for } \lambda \geq 1/2 \text{ and } p_{\text{TV}} := \begin{cases} \lambda & \text{for } x \in [0, \frac{1}{2}] \\ 1 - \lambda & \text{for } x \in (\frac{1}{2}, 1] \end{cases} \text{ for } \lambda < \frac{1}{2}$$

in the $L^2(\Omega)$ sense.

PROOF. Can be found in [29]. □

To further illustrate the properties of solutions to Total Variation coupling, let us highlight the Total Variation term from the perspective of smooth functions. The

following Lemma shows that for smooth u the Total Variation of u can be expressed in a more simple way:

Lemma 16. Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain with smooth boundary and $u \in C^1(\Omega)$. Then
$$\mathcal{C}_{\mathrm{TV}}(u) = \int_\Omega |\nabla u(x)|_2 \, \mathrm{d}x.$$

PROOF. We fix $\xi \in C_0^\infty(\Omega; \mathbb{R}^2) \cap \{\|\|\xi\|_2\|_\infty \leq 1\}$ and use the divergence theorem to obtain:
$$-\int_\Omega u \operatorname{div}(\xi) \, \mathrm{d}x = -\int_{\partial\Omega} \langle u\xi, \nu \rangle \, \mathrm{d}_\sigma x + \int_\Omega \langle \nabla u, \xi \rangle \, \mathrm{d}x = \int_\Omega \langle \nabla u, \xi \rangle \, \mathrm{d}x$$

since $\xi = 0$ on $\partial\Omega$. The claim follows now from Theorem 8, which shows that for each point $x \in \Omega$ it holds that
$$\sup_{\substack{\zeta \in \mathbb{R}^d \\ |\zeta|_2 \leq 1}} \langle \nabla u(x), \zeta \rangle = |\nabla u|_2.$$
$\square$

This allows the following interpretation of Total Variation for smooth functions $u : \mathbb{R} \to \mathbb{R}$: Since u can be split into monotonous segments, on each segment one can see by the fundamental theorem of calculus that $\mathcal{C}_{\mathrm{TV}}(u)$ indeed measures the absolute height of the jump, since $\int_a^b |u'| \, \mathrm{d}x = |u(b) - u(a)|$. This is illustrated by Figure 4.1. Using Total Variation as a coupling term will hence lead to solutions which try to jump as seldom and as low as possible. This has the advantage that solutions can exhibit discontinuities but the disadvantage that even in non-monotonous areas staircasing artifacts will be necessarily introduced as demonstrated in the example of Figure 4.2, see also [29, Page 29].

We conclude this section with Figure 4.2, which shows an illustrative comparison of the different coupling approaches for real T_1 data.

4.2 Spatial Coupling for Vector Fields

In this section we will introduce tailored approaches for the coupling terms in vector-field estimation. Note that the coupling approaches which were introduced in the previous Section 4.1 were tailored for scalar-valued functions $u : \Omega \to \mathbb{R}$. However, in parameter estimation one typically aims to recover more than one parameter per voxel. There have been multiple approaches to introduce spatial coupling into parameter
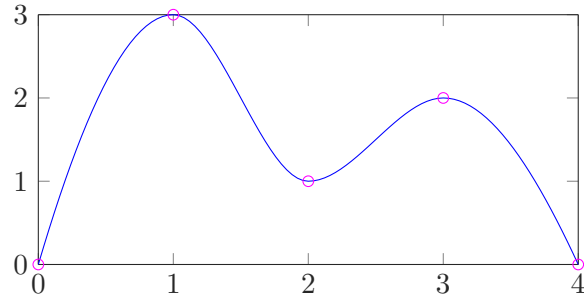
Figure 4.1: Figure illustrating Total Variation in the 1D Case. By the fundamental-theorem of calculus $\int |p'| dx = 3 + 2 + 1 + 2 = 8$. The Term $\mathcal{C}_{\mathrm{TVS}}(u)$ hence expresses the total height of all jumps of p.

| No Stab. | $\mathcal{C}_{\mathrm{T}}$ | $\mathcal{C}_{\mathrm{TP}}$ | $\mathcal{C}_{\mathrm{TV}}$ |
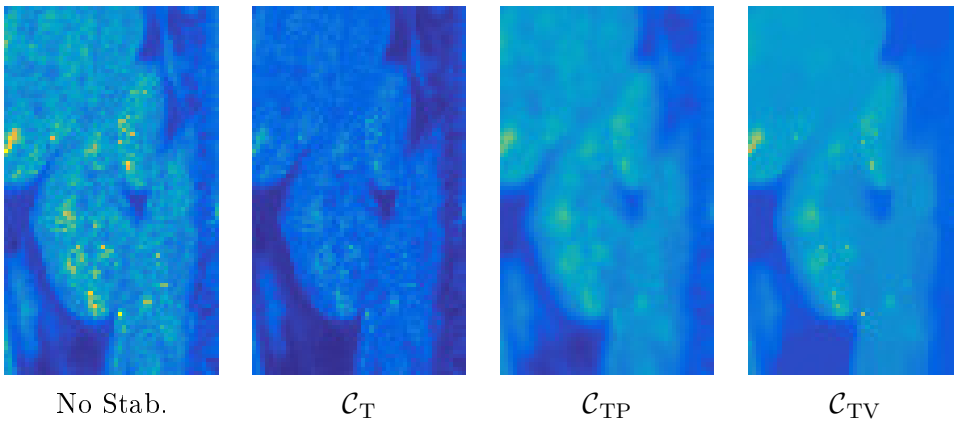
Figure 4.2: Figure showing solutions to (4.3) for a real data T_1 map of a right human kidney with different choices for the coupling term $\mathcal{C}$. p_{NL} is an uncoupled reconstruction result for a real-data T_1-map. It can be seen that the usage of $\mathcal{C}_{\mathrm{T}}$ results in just a scaled version of p_{NL} and that the usage of $\mathcal{C}_{\mathrm{TP}}$ smears the edges. Although edges are preserved for $\mathcal{C}_{\mathrm{TV}}$, so-called staircasing artifacts can be observed in constant areas.

estimation [114, 155, 135]. In the case of T_1 estimation, some of these approaches are only introducing coupling with respect to the parameter of interest, i.e. T_1 [155]. However, since the unknowns might be more than one parameter map, a natural extension is to introduce the coupling for the complete vector field $p : \Omega \to \mathbb{R}^m$ instead [114, 135]. This opens the question which coupling approaches for vector-fields are suitable. Established approaches for parameter estimation [114, 135] are using straight-forward extensions $\mathcal{C}(p) := \sum_{i=1}^m \mathcal{C}_i(p_i)$ where $\mathcal{C}_i$ are traditional coupling terms for scalar-valued functions. Using such approaches yields parameter maps which are admitting only individual regularity properties. However, in many applications of parameter estimation one would expect to see additional coupling between the individual parameter maps: If one map changes due to changes in the tissue, one would expect the other map to change, too.

In Section 4.2.1 we will motivate a family of coupling terms which admit such properties and are based on singular values of the Jacobian. The introduced family was originally designed for color image denoising [130, 61, 88, 105], which is a classic application for coupled recovery of vector fields. The proposed approaches will be based on the Schatten-p-Norm of the Jacobian ∇u, which is the regular l^p norm of its singular values: $\sigma_1, \ldots, \sigma_m$

$$|\nabla u(x)|_{\mathcal{S}_p} := \left(\sum_{i=1}^m |\sigma_i(x)|^p \right)^{1/p},$$

see also Theorem 2. The main contribution of this section can be found in Section 4.2.2, where we will introduce common choices for Schatten-p-norms for vector-field estima-tion and give simple geometric interpretations. As we will see, this class of coupling terms will yield parameter maps which are strongly coupled and similar to TV display clear edges.

4.2.1 Total Variation and Singular Values of the Jacobian

Let us start by motivating to use l^p norms of the singular values of $\nabla p(x)$ from a geometrical point of view: We have seen in Section 4.1.2 that a large class of coupling terms for scalar-valued functions relies on different norms of the gradient. Since the magnitude of the gradient is a natural and straight-forward measure to capture the amount of local variations of a function, this is an obvious choice. However, for vector-fields the gradient is not vector-valued but instead matrix-valued. If gradient-based coupling terms are to be extended to vector fields, this additional structure needs to be taken into account. Specifically, one desired property of a norm $|.|$ of the Jacobian would be invariance with respect to rotation, meaning that both $|\nabla p(x)| = |\nabla p(x)V^\top|$ and $|\nabla p(x)| = |U\nabla p(x)|$ for any orthonormal matrices $V \in \mathbb{R}^{d,d}$ and $U \in \mathbb{R}^{m,m}$. Geometrically the first property means that if the channel-gradients are uniformly rotated, the norm is not supposed to change. The second property means that also rotation in p does not alter the norm of the Jacobian, since $\nabla(Up(x)) = U(\nabla p(x))$.

A natural approach to characterize matrices from such a perspective is the singular value decomposition of a matrix, see Theorem 2. Indeed this theorem shows that if the Jacobian of a vector-field $p : \Omega \to \mathbb{R}^m$ is decomposed as $\nabla p(x) = U(x)\Sigma(x)V(x)^\top$, the diagonal matrix $\Sigma(x)$ is invariant under the any of the orthonormal transformation considered above. This property can be regarded as simple geometric motivation to consider standard p-norms of the singular values $\mathrm{diag}(\Sigma)$ of ∇p as stabilizers, the so-called *Schatten p norms* [88, 105]:

Definition 4. Let $1 \leq p \leq \infty$ and let $J = U\Sigma V^\top$ be an SVD of $J \in \mathbb{R}^{m,d}$, see Theorem 2. The *Schatten-p norm* of J defined as

$$|J|_{\mathcal{S}_p} := |\mathrm{diag}(\Sigma)|_p.$$

A different motivation to use Schatten-norms comes from geometric measure theory and shows an intimate connection to Total Variation. This connection is established by the so-called coarea formula (Theorem 7), which will be introduced at the beginning of this section. The coarea formula gives a closed-form expression for the total volume of the level-sets of a (continuous) vector-valued function $p : \Omega \to \mathbb{R}^m$, where $\Omega \subseteq \mathbb{R}^d$. We will then show in Lemma 17 that for the special case $m = 1$ the length of the levelsets corresponds to the Total Variation of the function. For $m > 1$, the total volume of all level sets could hence be regarded as an extension of Total Variation to vector valued functions [61]. Indeed we will show that this quantity is closely connected to the product of the singular values of the Jacobian ∇u. However, since the product of singular values is highly non-convex function, we will confine to consider Schatten-Norms to ease the analysis.

Singular Values of the Jacobian and the Coarea Formula

Let us begin by introducing the coarea formula, which established a relationship between derivatives and levelsets of vector fields. Note that statement and implications of this theorem will be clarified in the following paragraphs.

Theorem 7 (Coarea Formula, [94]):
Let $\Omega \subset \mathbb{R}^d$, $u : \Omega \to \mathbb{R}^m$ be $C^1(\Omega)$ and assume that $m \leq d$. Then

$$\int_\Omega |J_m u(x)| dx = \int_{\mathbb{R}^m} \left(\int_{u^{-1}(t)} \mathrm{d}\, \mathcal{H}^{d-m} \right) dt. \qquad (4.6)$$

where $\mathcal{H}^{d-m}$ denotes the $d - m$-dimensional Hausdorff measure and $|J_m u(x)|$ denotes the m-dimensional Jacobian of u at location x, which is defined as

$$|J_m u(x_0)| = \sup\{\mathcal{H}^m \left(\nabla u(x_0)(C) \right) : \ C \subset \mathbb{R}^d \text{ is an } m\text{-dimensional unit cube}\}.$$

Specifically we mean

$$\nabla u(x_0)(C) := \{[\nabla u(x_0)]x : \ x \in C\} \subseteq \mathbb{R}^m \text{ and}$$
$$C := OC_m \text{ for } C_m := [0,1]^m \subseteq \mathbb{R}^m \text{ and } O : \mathbb{R}^m \to \mathbb{R}^d \text{ with } O^\top O = \text{id}_m.$$

To illustrate the connection between Total Variation and (4.6), let us first look closer on the m-dimensional Jacobian $|J_m u|$ of u. This is done in the following Lemma.

Lemma 17. Let $\Omega \subset \mathbb{R}^d$ be a bounded domain, let $u \in C^1(\Omega, \mathbb{R}^m)$ and let $m \leq d$. Then

$$\int_\Omega |J_1 u(x)| dx = \int_\Omega \sigma_1(x) dx,$$
$$\int_\Omega |J_m u(x)| dx = \int_\Omega \sigma_1(x) \cdot \ldots \cdot \sigma_m(x) dx.$$

where $\sigma_i(x)$ is the i-th singular value of $\nabla u(x)$. More specifically it holds for the special cases $m = 1$ and $m = d$

$$\int_\Omega |J_1 u(x)| dx = \int_\Omega |\nabla u|_2 dx \qquad \text{for } u \in C^1(\Omega, \mathbb{R})$$
$$\int_\Omega |J_d u(x)| dx = \int_\Omega |\det(\nabla u)| dx \qquad \text{for } u \in C^1(\Omega, \mathbb{R}^d).$$

PROOF (SEE [61]). Let $u \in C^1(\Omega, \mathbb{R}^m)$ be arbitrary for $m \leq d$. As we want to characterize $J_1 u(x)$, we need to characterize the 1-dimensional unit-cubes of $\mathbb{R}^d$. These are exactly the line-segments $[0, v] \subset \mathbb{R}^d$ for $|v|_2 = 1$. We furthermore denote the Jacobian of u at location x by $G := \nabla u(x) \in \mathbb{R}^{m,d}$. As $G : \mathbb{R}^d \to \mathbb{R}^m$ is linear, $G([0, v]) = [0, Gv]$ and hence $\mathcal{H}^1([0, v]) = |Gv|_2$. It follows

$$|J_1 u(x)| = \sup\{\mathcal{H}^1(G(C))) : \ C = [0, v] \text{ for } |v|_2 = 1\} = \sup_{|v|_2=1} |Gv| = \sigma_1(G).$$

Let us now characterize $J_m u(x)$. The SVD of $G = \nabla u(x)$ is given by $G = U \Sigma V^\top$ for orthogonal $U \in \mathbb{R}^{d,d}, V \in \mathbb{R}^{m,m}$ and diagonal $\Sigma \in \mathbb{R}^{d,m}$. It now follows directly from the properties of the SVD (Theorem 2) that the vertices of the maximizing m-dimensional unit-cube are given by $V = (v_1, \ldots, v_m)$ and its volume by $\det(G) = \sigma_1 \cdots \sigma_m$. $\qquad\square$

Thus for a function $u : \Omega \to \mathbb{R}$ the 1-dimensional Jacobian corresponds to the Total Variation of u. For transformations $u : \Omega \to \mathbb{R}^d$ the d-dimensional Jacobian is just the Jacobi determinant. Note that both cases can be expressed in terms of the singular values of the Jacobian.

The Coarea Formula and Level-Sets

To understand Total Variation from a level-set point of view, let us now elaborate on the right integral of (4.6). If we assume that $u \in C^1(\Omega, \mathbb{R})$ and use Lemma 17, equation (4.6) simplifies to

$$\int_\Omega |\nabla u(x)| dx = \int_\mathbb{R} \left(\int_{u^{-1}(t)} d\,\mathcal{H}^{d-1} \right) \, \mathrm{d}t.$$

We can interpret the right integral as an integration over the lengths of the level-sets of u: For each $t \in \mathbb{R}$ the levelset function $\int_{u^{-1}(t)} d\,\mathcal{H}^{n-1}$ measures the size of the pre-image of t under u. An illustration of the situation is given in Figure 4.3.

Total variation coupling thus produces solutions which have both few and short levelsets. Avoiding oscillations, merging multiple function values into one and thus producing locally constant solutions is thus preferable from a Total Variation point of view. In the other extreme case that $u : \Omega \to \mathbb{R}^d$ is a diffeomorphism, equation (4.6) becomes the well-known transformation formula

$$\int_\Omega |\det(\nabla u)| dx = \int_{u^{-1}(\Omega)} dx.$$

Schatten-p-Norms

A straight-forward extension of Total Variation to vector-valued functions from $u : \Omega \to \mathbb{R}^m$ would hence be to consider the term $|J_m u(x)| = \sigma_1(x) \cdots \sigma_{\min\{m,d\}}(x)$. However, the product of the singular values is highly nonlinear and non-convex, which significantly complicates the analysis and numerical evaluation. Additionally, the product is zero (and the energy thus minimal), if foldings along any dimension occur. For our means, we want to enforce foldings in as many dimensions as possible, leading to low-rank solutions, see Section 4.2.2. To cope with these problems, multiple approaches to stabilize recovery are considering Schatten-p-Norms of the Jacobian as a remedy.
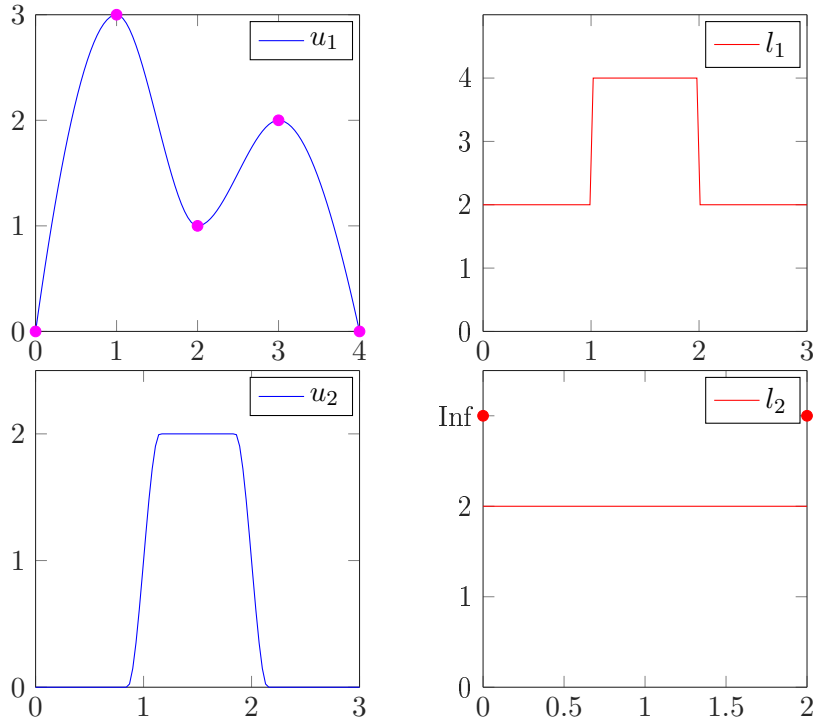
Figure 4.3: Figure illustrating the connection between Total Variation and levelsets. Given are two functions $u_i(x)$ with corresponding levelset functions $l_i(t) := \int_{u_i^{-1}(t)} d\,\mathcal{H}^0$. By the fundamental-theorem of calculus $\int |u_1'| dx = 3 + 2 + 1 + 2 = 8$ and $\int |u_2'| dx = 2 + 2 = 4$, corresponding to the integrals of l_1 and l_2.

Conclusion

In this section we have introduced two different motivations to consider Schatten-p-Norms of the Jacobian as coupling terms: First, we have given a geometric intuition: We have argued that since singular values are invariant under rotation, they are fulfilling basic properties for well-designed coupling term. Specifically we have shown that the singular values of the Jacobian will not change if the gradients or the vector field is rotated. As an additional motivation, we have established a connection between the singular values of the Jacobian and Total Variation: We have shown, that for functions $p : \Omega \to \mathbb{R}^m$ and $m = 1$ Total Variation measures the total length of the levelsets of the function. As an extension, we have found that for $m > 1$ the integral over the product of the singular values measures the total volume of all levelsets. Since the product is a highly non-convex function and will additionally be minimal if only few foldings occur, we have argued that l^p norms will be a more suitable choice, both from an analytical and a practical perspective.

4.2.2 Common Choices for Coupling Terms for Vector Field Estimation

We will now introduce common conventional choices for coupling terms for vector field estimation. Additionally, we will show how many approaches for vector-field coupling can be formulated in terms of the Schatten-p-norms. We will see in Section 4.3 that for T_1 estimation the choices $p = 1$ and $p = 2$ will yield the best results. The different approaches will be introduced for the example of color image denoising. A direct comparison on simple color data can be found in Figure 4.4.

Color Image Denoising

In this section we will consider the problem to recover an RGB image $u_{\mathrm{GT}} : \Omega \to \mathbb{R}^3$ from a measurement $u_0 = u_{\mathrm{GT}} + n$, for white gaussian noise n. Analogous to the previous modeling, this will be done by finding a minimizer $u : \Omega \to \mathbb{R}^3$ of

$$\mathcal{J}(u) := \int_\Omega |u - u_0|_2^2 dx + \lambda \int_\Omega g(\nabla u) dx \tag{4.7}$$

where $g : \mathbb{R}^{3,d} \to \mathbb{R}$ depends on the Jacobian of u. Since this problem is well-studied for color image denoising, we will consider established choices for g which can be found in [130, 61, 88, 105]. Note, however, that to the best of the authors knowledge these
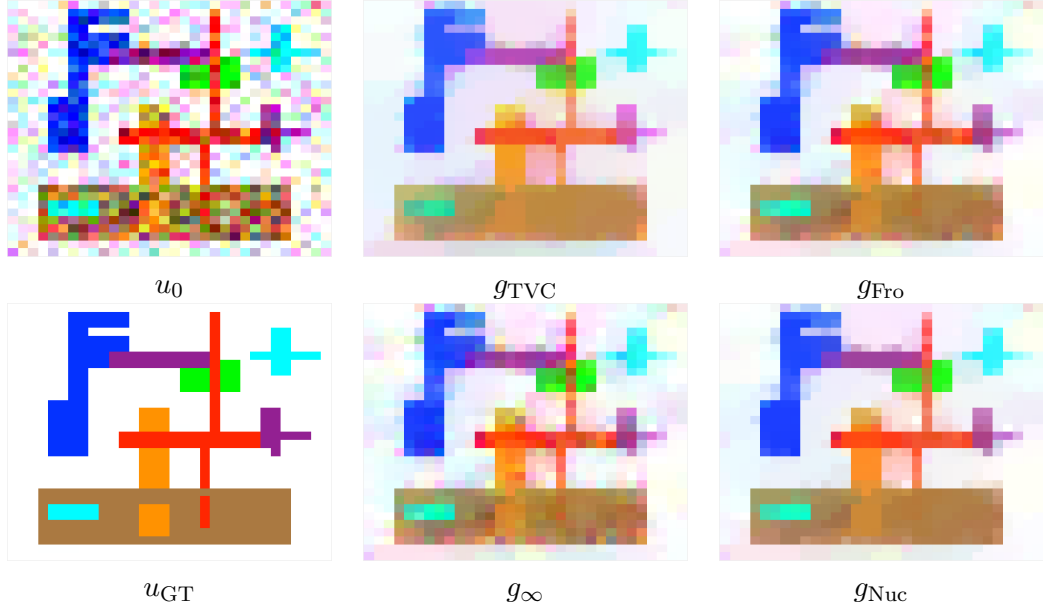
Figure 4.4: Motivating example displaying solutions of (4.7) with different coupling terms. Reconstruction was performed from u_0, which was obtained by adding white gaussian noise with standard deviation $\sigma = 50$ to the ground-truth image u_{GT}. It can be observed that although all coupling terms allow edges, results with fewest off-colors are obtained using g_{Nuc}. Implementation details can be found in Section 3.2.

terms have not been used for the proposed cases parameter estimation.

$$
\begin{aligned}
g_{\text{TP}}(\nabla u) &:= |\nabla u|_{\mathcal{S}_2}^2, && [114, 135, \text{H2}] \\
g_{\text{Fro}}(\nabla u) &:= |\nabla u|_{\mathcal{S}_2}, && [130] \\
g_{\text{Nuc}}(\nabla u) &:= |\nabla u|_{\mathcal{S}_1}, && [88] \\
g_{\infty}(\nabla u) &:= |\nabla u|_{\mathcal{S}_\infty}, && [61] \\
g_{\text{TVC}}(\nabla u) &:= \sum |\nabla u_i|_2, && [15].
\end{aligned}
$$

Geometric Interpretation of the Approaches

As it turns out, there are simple geometric interpretations which allow a deeper understanding of most of the coupling terms. Before we consider the terms in detail, we start with a key relationship which shows that the Schatten-2-norm corresponds to the standard Frobenius norm of a matrix and can hence be calculated with having access to an SVD of the matrix:

Lemma 18. Let $A \in \mathbb{R}^{n,m}$ be arbitrary with singular value decomposition $A = U\Sigma V^\top$. Then

$$|A|_{\mathcal{S}_2}^2 = \sum_{i,j} a_{i,j}^2$$

PROOF. Follows directly from the considerations

$$|A|_{\mathcal{S}_2}^2 = \text{trace}(\Sigma^2) = \text{trace}(V\Sigma^2 V^\top) = \text{trace}(A^\top A) = \sum_{i,j} a_{i,j}^2,$$

where the second identity holds since for any matrices $\text{trace}(AB) = \text{trace}(BA)$. $\square$

Keeping this relationship in mind, we now give simple geometric interpretations of the different coupling terms:

g_{TP} Since it follows from Lemma 18 that that $|\nabla u(x)|_{\mathcal{S}_2}^2 = \sum |\nabla u_i(x)|_2^2$, we can interpret the squared Schatten-2-norm as the sum of standard Tikhonov-Philips coupling terms. This term will hence generally yield smooth solutions, cf. Section 4.1.2.

g_{TVC} This is sometimes referred to as *channel-by-channel TV* [15]. Each channel will be smoothed with TV, but no coupling between the channels is present since (4.7) still decouples.

g_{Fro} This stabilizer is a special case of a class of stabilizers introduced in [130]. Since $|\nabla u|_{\mathcal{S}_2} = \sqrt{\sum |\nabla u_i|_2^2}$ (Lemma 18), it can be regarded as another straight-forward extension from TV to multiple channels. However, in contrast to g_{TVC} the channels are now coupled.

g_{Nuc} This Schatten-1 norm is also referred to as the *nuclear norm*. In compressed sensing, it is employed in rank minimization problems [120]. For color-image denoising this means that it enforces solutions where the Jacobian has low rank, implying that channel gradients will be linearly dependent. This is depicted in Figure 4.5. Usage of this term usually yields a stronger coupling of the channels than g_{Fro}, cf. [88, 105]. In medical imaging nuclear norm minimization has also recently been introduced for image fusion [91].

g_∞ Using the Schatten-∞-norm for stabilization has been introduced in [61]. By minimizing the largest singular value, the other singular values will go to zero to likewise.

In order to illustrate the properties of the different norms, we used them for denoising of color images. Results are given in Figure 4.4.
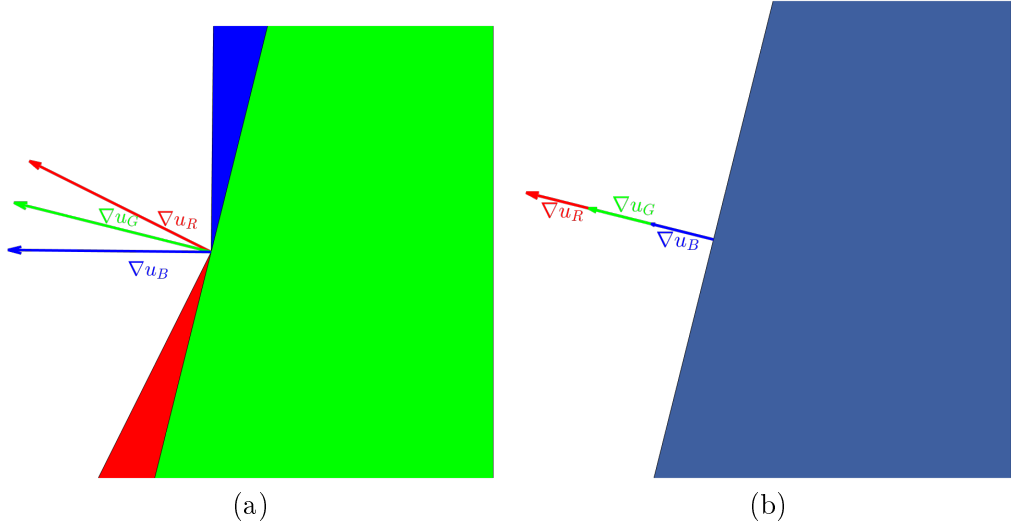
Figure 4.5: Figure demonstrating different coupling approaches for vectorial images. For color images, (b) is preferable to (a) as channels are aligned. Note that for g_{TVC} (a) and (b) have the same energy.

4.3 Spatial Coupling for T_1 Estimation

To demonstrate advantages of the coupling terms for vector field recovery, we introduce these terms for T_1 estimation from variable flip angle data. Note that parts of these results have been accepted for publication in [S3]. In this section we will only present the models, results are given in Section 4.4

We start by outlining established coupling techniques for T_1 estimation [114, 155]. For ease of presentation, let us again stay in the functional setting, i.e. $p : \Omega \to \mathbb{R}^2$ with $p(x) = (T_1(x), M_0(x))$. In [114] diffusive coupling was introduced for inversion time estimation, i.e.

$$\mathcal{C}_{\mathrm{TP}}(p) := \int_\Omega |\nabla T_1|^2 + |\nabla M_0|^2 \, \mathrm{d}x.$$

An advantage of this method is that it is easy to implement, as standard nonlinear-least squares methods can be used for optimization [87, 107]. A downside of this approach is that it is usually not edge preserving. The idea to include Total Variation (TV) to stabilize T_1 estimation has been introduced in [155]: Here the proposed coupling term was

$$\mathcal{C}_{\mathrm{TV}}(p) := \int_\Omega |\nabla T_1| \, \mathrm{d}x$$

With TV stabilized methods, edges in the T_1 map are preserved. However, note that in the proposed approach no coupling with respect to M_0 is performed. Additionally, since Total Variation is non-differentiable, the numerical minimization will require non-standard methods and will be described in more detail in Section 3.2.

We hence propose to extend these techniques. Specifically we will show that more accurate estimation can be obtained by solving the problem

Minimize:
$$J(p) := \|M(p) - d\|_2^2 + \lambda\, \mathcal{C}_{\text{Nuc}}(p).$$

In Section 4.4 we will conduct different experiments to support this claim. First, we will clarify if the problem should be formulated in the (T_1, M_0) [114, 155] or in the (E_1, N) [153, H2] realm, cf. Section 3.2.4. Second, we will compare denoising approaches with simultaneous approaches. Third, we will compare traditional approaches to stabilize only T_1 [114, 153] with the proposed vector-field coupling techniques.

4.4 Results: Spatial Coupling for T_1 Estimation for Software Phantom Data

In this section we will show how spatial coupling can improve T_1 estimation from variable flip angle data. This section will cover phantom data, real data experiments can be found in Section 7.2.

Note that we have introduced several novel aspects to the existing coupling approaches for T_1 estimation, which were introduced in [8, 155]: First, we have proposed to use coupling in (E_1, N) instead of (T_1, M_0) and second we have proposed to use simple denoising approaches instead of the simultaneous approaches. We hence start with Section 4.4.1, where we compare the existing approaches with respect to the above criteria. We will find that best results for T_1 estimation can be achieved with Total Variation coupling in the (T_1, M_0) coordinate system using the fully coupled approach for both parameters. This extends the existing results published in [155], where TV coupling was performed only with respect to T_1. These results have been published by the author in [S3].

Since the author made the above observations chronologically earlier than the considerations about vector fields, we proceed to compare different coupling terms for vector fields in Section 4.4.2. Specifically, we will consider the following choices for the coupling terms $\mathcal{C}$ for $p = (T_1, M_0)$:

$$
\begin{aligned}
\mathcal{C}_{\text{TVC}}(p) &:= \lambda_1 \int_\Omega |\nabla T_1|\, \mathrm{d}x & &+ \lambda_2 \int_\Omega |\nabla M_0|\, \mathrm{d}x, \\
\mathcal{C}_{\text{TP}}(p) &:= \lambda_1 \int_\Omega |\nabla T_1|^2\, \mathrm{d}x & &+ \lambda_2 \int_\Omega |\nabla M_0|^2\, \mathrm{d}x, \\
\mathcal{C}_{\text{Fro}}(p) &:= \lambda \ \ \int_\Omega |\nabla p|_{\text{Fro}}\, \mathrm{d}x, \\
\mathcal{C}_{\text{Nuc}}(p) &:= \lambda \ \ \int_\Omega |\nabla p|_{\mathcal{S}_1}\, \mathrm{d}x.
\end{aligned}
$$

Note that the functionals $\mathcal{C}_{\text{TV}}$ and $\mathcal{C}_{\text{TP}}$ differ from the original definition given in Section 4.2.2 since they depend on two control parameters $\lambda_1, \lambda_2 \in \mathbb{R}$. This additional degree of freedom gives us an range for fine-tuning of the results and to additionally

Table 4.1: Table displaying the impact of spatial coupling on phantom data for sequence S1 and noise-level 7%. Results are given as mean relative error calculated from 5000 experiments on cortex and medulla and are given in percent. It can be observed that for phantom data fully coupled approaches are only superior to denoising approaches if both parameters are regularized. Also, coupling in (T_1, M_0) is superior to coupling in (E_1, M_0).

| | Simul. | | | | Denoising | | |
| | $\mathcal{C}_{\mathrm{TV}}$ | | $\mathcal{C}_{\mathrm{TP}}$ | | $\mathcal{C}_{\mathrm{TV}}$ | $\mathcal{C}_{\mathrm{TP}}$ | No Stab. |
# Parameters	1	2	1	2	1	1	
(T_1, M_0)	6.76	5.96	7.36	6.59	6.31	6.74	12.91
(E, N)	6.48	6.49	11.32	11.32	6.58	9.71	12.91

exploit benefits of coupling terms. Here we find that best results can be achieved using $\mathcal{C}_{\mathrm{Nuc}}$ or $\mathcal{C}_{\mathrm{Fro}}$, which outperform $\mathcal{C}_{\mathrm{TVC}}$ by approximately 8%.

4.4.1 Analysis of existing coupling strategies

We simulated multiple 2D scans of the human kidney using the software phantom which is described Section 3.6.1, with sequence parameters S1 and a modest noise-level of 7%. Since S1 yielded the lowest reconstruction error in Section 3.6, the experiments demonstrate the advantages of coupling even for well-designed sequences. After each T_1 estimation, the quality of the estimated T_1 map was assessed using the mean relative error, which was evaluated over cortex and medulla of the kidney. The number of experiments was determined experimentally to be 5000, since for this number the mean relative error was approximately stable up to 10^{-5}. All coupling parameters γ were determined experimentally by sampling the parameter space.

Results are presented in Table 4.1. It can readily be seen that all discussed coupling methods are capable to drastically reduce the relative error as compared to no coupling. Also, coupling in both T_1 and M_0 has advantages as compared to coupling only in T_1. However, these effects can not be observed for coupling in the (E_1, N) coordinate system. In this case, no large difference in coupling of (E_1, N) as compared to coupling only in E_1 can be observed. This might be due to the exponential scaling, as variations in high T_1 are less accurately detectable in E_1. Comparing the fully coupled approaches with the denoising approaches, one can see the fully coupled approaches are only superior if both parameters are coupled. This establishes the need for fully coupled approaches considering both variables M_0 and T_1.

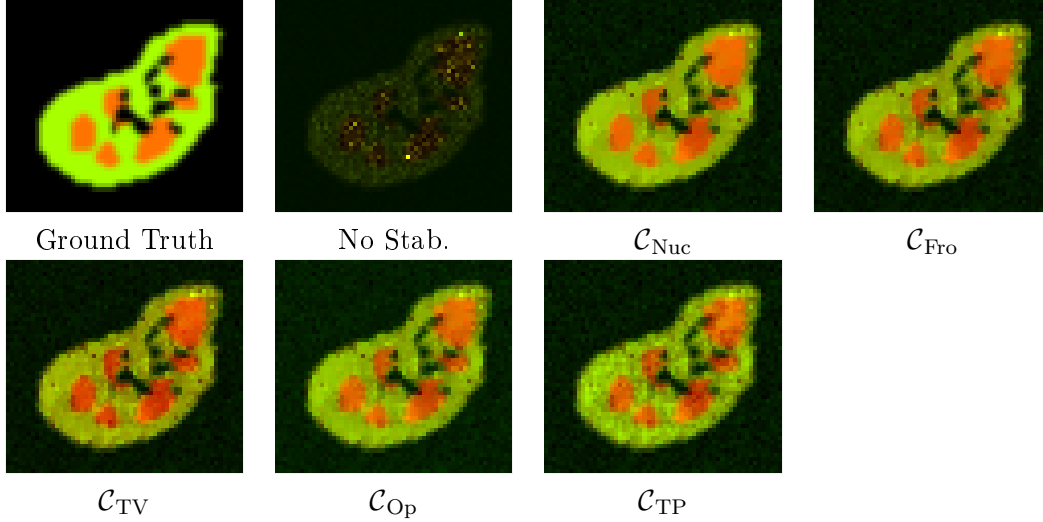| Noise	$\mathcal{C}_{\text{NL}}$	$\mathcal{C}_{\text{TP}}$	$\mathcal{C}_{\text{Op}}$	$\mathcal{C}_{\text{TV}}$	$\mathcal{C}_{\text{Fro}}$	$\mathcal{C}_{\text{Nuc}}$
3%	9.17	4.26	3.95	3.86	3.79	3.45
6%	18.72	6.70	6.18	6.94	5.99	5.57
9%	34.56	9.53	8.87	9.00	8.67	8.25

Figure 4.6: Parameter estimation results for vector field coupling for Sequence S4 and various noise levels. Given is the mean relative error over cortex and medulla in 5000 experiments in percent. The table shows that the lowest relative error can be achieved using $\mathcal{C}_{\text{Fro}}$. Also given are samples of estimated (T_1, M_0) maps from data with 6% noise, visualized as RGB images. The red channel was T_1, the blue channel M_0. This visualization allows a better comparison of common properties of the recovered parameter maps.

4.4.2 Analysis of Coupling Strategies for Vector Fields

Since the results in Section 4.4.1 indicated superiority of coupling approaches for (T_1, M_0) in the simultaneous setting, we proceeded to compare the different coupling techniques for vector fields. This time 2D scans were simulated using the parameters of sequence S4, which corresponded to the real data. As in the previous section we measured the mean relative error on cortex and medulla over 5000 experiments. Results are show in Figure 4.6. Results indicate that vector field coupling is capable to improve T_1 recovery on the given phantom data. Also the visual comparison shows that most consistent results can be achieved with vector field coupling terms.

4.4.3 Discussion

The results demonstrates that spatial coupling can be used to improve T_1 estimation results from low-quality data. Even simple coupling terms as e.g. standard Tikhonov-Philips can be used to obtain clearer and more distinct parameter maps. It can additionally be seen that stabilization of all parameters is crucial: In our experiments we have found that stabilization of only one parameter as e.g. proposed in [155] leads to only minor improvements in T_1 estimation. Our phantom data results furthermore show that coupled estimation is able to additionally improve the parameter recovery, since it leads to spatially aligned T_1 and M_0 maps. However, note that the impact of coupling naturally depends on the expected quality of the data: For data with high SNR we expect only minor advantages introduced by coupling. Additionally it needs to be pointed out, that by design the developed software phantom has some bias for locally constant parameter maps, favoring the non-differentiable coupling terms. This comes from the assumption that the parameters T_1 and M_0 are constant on distinct anatomical structures. The experimental findings on software phantom data will be extended to real data in Section 7.2.

4.5 Numerical Aspects of Spatial Coupling

In this section we will describe the numerical methods which were used to solve the minimization problems outlined in the previous sections. All implementations were performed in Matlab [97] and run on a standard PC (3 GHz Intel Core i7, 16 GB RAM). The optimization methods were tailored to fit the respective objective functions. The employed algorithms can be divided in two groups: Smooth and non-smooth optimization methods. For differentiable objective functions, optimal solutions were obtained by either discretizing the first-order necessary condition and solving the resulting system directly or by employing a constrained Gauss-Newton solver for minimization [9]. For the cases that the objective function J was not differentiable, we exploited the specific structure of J, which often was the sum of two convex functions. Specifically, we used the so-called primal-dual-hybrid-gradient algorithm [28], which is described in Section 4.5.1 in more detail.

The problems were usually solved in a Gauss-Newton fashion: For nonlinear models M the data-term $\|M(p)-d\|$ is often non-convex and might admit multiple local minima. In this case optimization is thus not straight-forward [107]. To cope with this issue, a standard strategy which exploits higher order derivatives is the so called Gauss-Newton method [107]. Instead of linearizing the objective function and obtaining a first-order algorithm, in Gauss-Newton methods only the residue $r(p) := M(p) - d$ is linearized. This leads to a convex and nonlinear data term which can then be optimized using second order methods as e.g. Newton's method. Specifically, the objective function

$J(p) = \|M(p) - d\|^2 + \mathcal{C}(p)$ is approximated by

$$J_{\text{Lin}}^k(p) := \|A^k p - b^k\|^2 + \mathcal{C}(p)$$
$$\text{for } A^k := \nabla M(p^k) \text{ and } b^k := d + \nabla M(p^k)p_k - M(p^k).$$

Updated guesses p^{k+1} is obtained iteratively by minimizing J_{Lin}^k until $\{p^k\}_k \in \mathbb{N}$ converges, see [107] for a detailed analysis of such methods. As compared to optimizing second-order Taylor approximations of the objective function (=Newton's method), this strategy has the advantage that the obtained approximation will always be convex. This implies that the resulting linear system is symmetric, positive semidefinite and can be solved efficiently using e.g. a conjugate gradient algorithm [107]. Note that in the case of non-differentiable coupling terms, established optimization algorithms often require convex objective functions. This property additionally makes the Gauss-Newton method a strong candidate for this class of objective functions.

4.5.1 A Primal-Dual Hybrid Gradient Algorithm for Convex Functions

In this section we will show how the problems described in Section 4.3 can be implemented numerically. The main tool to solve the problems is the so-called Primal Dual Hybrid Gradient (PDHG)-Algorithm (also called Chambolle-Pock or Chambolle-Cremers-Pock Algorithm) [28], which is a solution strategy for such problems and is famous for being quite simple to implement. As it turns out, easy to implement does not coincide with easy to understand. In this section we will describe why and how the PDHG Algorithm works.

The PDHG-algorithm is tailored to solve problems of the type

$$\text{Minimize } G(x) + F(Kx). \tag{4.8}$$

Here $F : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $G : \mathbb{R}^m \to \overline{\mathbb{R}}$ are proper, convex, lower semi-continuous functions and $K : \mathbb{R}^n \to \mathbb{R}^m$ is a linear operator. Here property *proper* is of a technical nature and means that the functions does not take $-\infty$ as a function value and is finite in at least point. Roughly this property means that the function indeed admits a minimum which is not trivial to obtain. We will show later that the minimization problems which were introduced in the previous section fall exactly into this scheme. In our setting $F(x) := \|x\|_p$ will be a (non-differentiable but convex) norm, $K := \nabla^d \in \mathbb{R}^{n,dn}$ will be a discrete gradient operator and $G(x) := \|Ax - b\|_2^2$ will be both differentiable and convex. However, since F will generally be non-differentiable, smooth optimization algorithms are no option to tackle the problem [29].

Key to understanding the PDHG-Algorithm are two insights: First, that convex functions can be equivalently described in terms of their so-called convex conjugates and second, that the convex-conjugates of many convex functions can be minimized much more efficiently than the function itself.

Let us start by elaborating on the first insight: Formally the convex-conjugate F^* : $\mathbb{R}^n \to \overline{\mathbb{R}}$ of a function $F : \mathbb{R}^n \to \overline{\mathbb{R}}$ is given by:

$$F^*(y) := \sup_x \langle x, y \rangle - F(y).$$

Indeed [123, Corollary 12.2.1] shows that if F is convex and proper it holds that

$$F(y) = F^{**}(y) = \sup_x \langle x, y \rangle - F^*(x) \tag{4.9}$$

This means, that any convex F can be described equivalently using the dual function F^*. The following theorem and the example show that convex-conjugates of p-norms are just characteristic functions of convex sets and thus admit a very simple structure.

Theorem 8:
Let $1 \le p \le \infty$ and $v \in \mathbb{R}^n$ with $|v|_p < \infty$. Then it holds for $q := p/(1-p)$ that

$$|v|_p = \sup_{\substack{w \in \mathbb{R}^n, \\ |w|_q \le 1}} \langle v, w \rangle = \langle v, w \rangle + \chi_q(w)$$

where $\chi_q(w) = 0$ for $|w|_q \le 1$ and $\chi_q(w) = \infty$ for $|w|_q > 1$.

PROOF. For $w \in \mathbb{R}^n$ with $|w|_q \le 1$ it follows directly from Hölders inequality that

$$\langle v, w \rangle \le |v|_p |w|_q \le |v|_p.$$

And hence $\sup \langle v, w \rangle \le |v|_p$. For the other direction let us first consider the case $1 < p < \infty$. Let $w \in \mathbb{R}^n$ with

$$w_i := \begin{cases} |v|_p^{1-p} \frac{|v_i|^p}{v_i} & \text{for } v_i \ne 0 \\ 0 & \text{for } v_i = 0. \end{cases}$$

Then it follows that

$$\langle v, w \rangle = |v|_p^{1-p} \sum |v_i|^p = |v|_p^{1-p} |v|_p^p = |v|_p,$$

$$|w|_q^q = |v|_p^{-p} \sum |v_i|^p = |v|_p^{-p} |v|_p^p \le 1$$

and hence $\sup \langle v, w \rangle \ge |v|_p$.

It remains to show the cases $p = 1$ and $p = \infty$. If $p = 1$ let $w_i := \mathrm{sgn}(v_i)$. Then $\langle v, w \rangle = |v|_1$ and $|w|_\infty \le 1$ and the claim follows. If $p = \infty$ let $w_i := 0$ for $i \ne \mathrm{argmax}_i |v_i|$ and $w_i := \mathrm{sgn}(v_i)$ for $i = \mathrm{argmax} |v_i|$. Once again $\langle v, w \rangle = |v|_\infty$, $|w|_1 = 1$ and the claim follows. $\qquad\square$

Example 1. In this example we will derive the convex-conjugate to the p-norm $F : \mathbb{R}^n \to \mathbb{R}$, $F(x) := |x|_p$. We first note that all norms are convex functions since $|\lambda x + (1-\lambda)y| \le \lambda|x| + (1-\lambda)|y|$ due to the triangular inequality and homogeneity. Lemma 8 shows that $|y|_p = \sup_x \langle x, y \rangle + \chi_B(x)$, where $B := \{x \in \mathbb{R}^n : |x|_q \le 1\}$ and $\chi_B(x) := 0$ for $x \in B$ and $\chi_B(x) := \infty$ for $x \notin B$. Comparing this with equation (4.9) yields $F^*(x) = -\chi_B(x)$.

Let us now elaborate on the second insight and describe why the optimization of convex conjugates can often be carried out easier than optimization of the convex function itself. To see this, we note that for smooth functions a backward gradient step can also be calculated by solving a minimization problem:

Lemma 19. Let $F \in C^1(\mathbb{R}^n)$ be convex and $\tau > 0$. Then

$$x_{k+1} = \operatorname*{argmin}_{x \in \mathbb{R}^n} \tfrac{1}{2\tau}|x - x_k|^2 + F(x) \iff \frac{x_{k+1} - x_k}{\tau} = -\nabla F(x_{k+1})$$

PROOF. First note that since $|\cdot|^2$ is strictly convex and F is convex, the left hand side problem indeed admits a unique minimizer. Now the claim follows from a direct application of the first-order necessary condition for a minimizer. $\square$

This lemma can be relaxed from differentiable to proper, convex functions F, see [123]. Since this would imply extensive introduction to convex analysis, we will not show this extension in this work. Nevertheless the importance of the above reformulation is summarized in the following definition (see [28]):

Definition 5 (Prox-Operator). Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper and convex and let $\tau \in \mathbb{R}^+$. The *Proximity-Operator for f with step-width τ* is defined as

$$\operatorname{Prox}_F^\tau(y) := \operatorname*{argmin}_{x \in \mathbb{R}^n} \tfrac{1}{2\tau}|x - y|_2^2 + F(x).$$

An evaluation of the Prox-Operator at y can hence be regarded as calculating a backward gradient step for f with step-length $\tau > 0$. As the following example shows, the Prox-Operator for convex-conjugates of p-norms corresponds to simple projections onto unit balls:

Example 2. In this example we describe $\text{Prox}_{F^*}^{\tau}(y)$ for $F : \mathbb{R}^n \to \mathbb{R}$, $F(x) := |x|_p$. Using the results from Example 1 yields

$$\text{Prox}_{F^*}^{\tau}(y) = \operatorname*{argmin}_{x \in \mathbb{R}^n} \tfrac{1}{2\tau} |x - x_k|^2 + \chi_q(x)$$

for $\chi_B(x)$ as in Example 1. This shows that

$$\text{Prox}_{F^*}^{\tau}(y) = \operatorname*{argmin}_{x \in B} \tfrac{1}{2\tau} |x - x_k|^2 = \mathcal{P}_B(x_k)$$

where $\mathcal{P}_B(x_k)$ denotes projection of x_k onto the set B with respect to the Euclidian norm.

These two connections are enough to introduce the PDHG-Algorithm [28]. Here Problem (4.8) is tackled as follows: First, the functions $F(y)$ and Kx are decoupled by using the identity $F^{**} = F$ [123, Corollary 12.2.1]. Now the problem can be reformulated as follows:

$$
\begin{aligned}
& \min_{x} G(x) + F(Kx) \\
= \quad & \min_{x} G(x) + F^{**}(Kx) \\
= \quad & \min_{x} G(x) + \max_{y} \langle Kx, y \rangle - F^*(y) \\
= \min_{x} \max_{y} \quad & G(x) + \langle Kx, y \rangle - F^*(y).
\end{aligned}
\tag{4.10}
$$

Note that in the given situation min and max can be interchanged due to [123, Theorem 36.3]. The minimization problem (4.8) thus becomes a saddle-point problem (4.10), where a minimization with respect to one variable and maximization with respect to the other variable has to be performed. In order to solve (4.10), alternating backward gradient steps are performed: For x in negative gradient direction and for y in positive gradient direction. In terms of the Prox-Operator these become:

$$
\begin{aligned}
y_{k+1} &= \text{Prox}_{F^*}^{\sigma}(y_k + \sigma K x_k) \\
x_{k+1} &= \text{Prox}_{G}^{\tau}(x_k - \tau K^{\top} y_{k+1})
\end{aligned}
$$

In [28] a slight modification was added to this algorithm: Do gain speed, y is updated not with the last iteration of x_k, but with the more optimistic guess $\bar{x}_k := x_k + (x_k - x_{k-1})$, yielding Algorithm 1. In [28, Theorem 1] it is shown, that the algorithm indeed converges if the step-sizes fulfill the condition $\sigma \tau \leq |K|_{\mathcal{S}_\infty}$.

Algorithm 1 $(\overline{x}, \overline{y}) = \mathrm{PDHG}(x_0, K, G, F^*, \lambda)$

1: Estimate operator norm $L := \sup_x |Kx|/|x|$
2: Choose step-width $\tau = \lambda/L$ and $\sigma = 1/(L\lambda)$
3: Initialize $\overline{x}_0 := x_0$ and $y_0 := Kx$
4: **repeat**
5:     $y_{k+1} = \mathrm{Prox}_{F^*}^{\sigma}(y_k + \sigma K \overline{x}_k)$
6:     $x_{k+1} = \mathrm{Prox}_{G^*}^{\tau}(x_k - \tau K^{\top} y_k)$
7:     $\overline{x}_{k+1} = 2x_{k+1} - x_k$
8: **until** convergence
9: **return:** $\overline{x} := x_{k+1}$ and $\overline{y} := y_{k+1}$

4.5.2 Extended Notation for Mixed Norms

As the formulations in Section 4.2 are for continuous functions $d : \Omega \to \mathbb{R}^k$ and $p : \Omega \to \mathbb{R}^2$, we introduce some notation for the discrete setting. The notation closely follows the one introduced in the introduction of this chapter. However, as additional derivatives are introduced, we need to extend the notation by parts. Note that this part we will make use of $3D$ arrays $A \in \mathbb{R}^{l,m,n}$.

As introduced in the beginning of this chapter, we identify an image $u \in \mathbb{R}^{l_1,l_2,l_3}$ with its vectorized form $u \in \mathbb{R}^n$ for $n := l_1 l_2 l_3$. Correspondingly, we identify k data images with $d \in \mathbb{R}^{n,k}$ and the parameter maps with $p \in \mathbb{R}^{n,2}$. A discrete gradient operator, which maps a parameter map p to the pointwise Jacobian matrices, is denoted by $\nabla_d : \mathbb{R}^{n,2} \to \mathbb{R}^{n,2,2}$. We furthermore introduce several matrix- and vector norms. Generally, will denote a norm for short matrices or vectors by $|.|_X$ and for long vectors by $\|.\|_Y$. We also introduce the following mixed norms and the following scalar products:

Definition 6. Let $A \in \mathbb{R}^{n,l_1,l_2} = [A_1, \ldots, A_n]$ for $A_i \in \mathbb{R}^{l_1,l_2}$ and let $1 \le p \le \infty$ and $1 \le q \le \infty$. We define the mixed-norms $\|.\|_{p,q}$ by:

$$\|A\|_{p,q} := \left(\sum_{i=1}^{n} |A_i|_{\mathcal{S}_p}^q \right)^{1/q} \quad \text{for } 1 \le q < \infty,$$

$$\|A\|_{p,\infty} := \max_i |A_i|_{\mathcal{S}_p} \quad \text{for } q = \infty.$$

Definition 7. Let $A, B \in \mathbb{R}^{l_1, l_2, l_3}$. We define the scalar product of A and B as

$$\langle A, B \rangle := \sum_{i,j,k} a_{i,j,k} b_{i,j,k}.$$

4.5.3 Primal-Dual formulation of $\mathcal{C}_{\mathrm{Nuc}}$

We will now describe an example of how the objective functions were discretized and optimized using the PDHG Algorithm. Since optimization for Total Variation is standard and can be found e.g. in [28], we will describe the more advanced minimization of the Nuclear norm, following the ideas presented in [88].

We start by giving the discrete version of $J(z)$ for 2D data $d \in \mathbb{R}^{n,k}$, where for each voxel v_i, $i \in \{1, \dots, n\}$ a discrete dynamic curve curve $d_i \in \mathbb{R}^k$ is given. For parameter maps $p \in \mathbb{R}^{n,2}$, a linear operator $A : \mathbb{R}^{n,2} \to \mathbb{R}^{n,k}$, $b \in \mathbb{R}^{n,k}$ and a discrete Jacobian $\nabla^d : \mathbb{R}^{n,2} \to \mathbb{R}^{n,2,2}$ the objective function is given by:

$$\text{Minimize } J(p) := \underbrace{\tfrac{1}{2}\|Ap - b\|_{2,2}^2}_{=:G(p)} + \underbrace{\lambda\|\nabla^d p\|_{1,1}}_{=:F(Kp)} \tag{4.11}$$

We now use the relationship established in Theorem 8 to reformulate $F(Kp)$ as

$$\lambda\|\nabla^d p\|_{1,1} = \max_{\substack{w_1, \dots, w_n \in \mathbb{R}^{2,2}, \\ |w_i|_{\mathcal{S}_\infty} \leq \lambda \text{ for all } i.}} \sum_{i=1}^n \left\langle (\nabla^d p)_i, w_i \right\rangle = \max_{\substack{w \in \mathbb{R}^{n,2,2}, \\ \|w\|_{\infty,\infty} \leq \lambda}} \left\langle \nabla^d p, w \right\rangle$$

We include the constraints on w by means of an indicator function ϕ_U for $U := \{w \in \mathbb{R}^{n,2,2} : \|w\|_{\infty,\infty} \leq \lambda\}$ with $\phi_U(w) = 0$ for $w \in U$ and $\phi_U(w) = \infty$ everywhere else. We can hence rewrite (4.11) as the saddle-point problem

$$\min_p \max_w \underbrace{\tfrac{1}{2}\|Az - b\|_{2,2}^2}_{G(p)} + \underbrace{\langle \nabla_d p, w \rangle}_{\langle Kz, w \rangle} - \underbrace{\phi_U(w)}_{F^*(w)}.$$

Let us formulate the respective proximity operators. Following Example 2, we see that

$$\text{Prox}_G^\tau(y) = \hat{z} \text{ with } (\text{id}_n + \tau A^\top A)\hat{z} = y + \tau A^\top b$$

$$\text{Prox}_{F^*}^\sigma(w) = (\tilde{w}_1, \dots, \tilde{w}_n) \in \mathbb{R}^{n,2,2} \text{ with } \tilde{w}_i := U_i \mathcal{P}(\Sigma_i)V_i \in \mathbb{R}^{2,2}$$

here $U_i \Sigma_i V_i = w_i$ is an SVD of w_i and $\mathcal{P}$ is a projection operator which cuts off pointwise at λ. Note that this formulation requires to solve an SVD for each voxel. To improve speed we implemented the SVD for 2×2 matrices as given in [13] explicitly.

4.5.4 Smooth Optimization

To optimize differentiable objective functions, we used a non-standard modification of the Gauss-Newton Method described in [9], which allows to incorporate hard-constraints. Although this method is interesting from an optimization perspective, we will not go into technical detail. The modification allowed to include the hard constraints $1\text{ms} \leq T_1 \leq 5000\text{ms}$ and $M_0 \geq 1$, which added additional stability to the method. Note that bounds $M_0 > 0$ and $0 < T_1 < \infty$ are necessary to guarantee identifiability of the parameters, see also Lemma 10. Standard stopping criteria based on relative update size of p^k, $J(p^k)$ and $|\nabla J(p^k)|$ were used [58, Section 8.2.3.3].

4.5.5 Conclusion

In this section we haver presented different numerical methods to tackle the optimization problems. We have concentrated on the Primal-Dual-Hybrid Gradient algorithm (PDHG) [28] for two reasons: First, it demonstrates that optimization of non-differentiable objective functions is a complex task which requires tailored approaches, second the PDHG Algorithm is at the current time one of the most popular algorithms in image processing (Google Scholar currently gives 1532 citations, retrieved 26.11.2016, 20:16.). Without diving into technical detail on convex analysis, we given a brief motivation of why the PDHG Algorithm works. Additionally, we have given a detailed discretization how it can be used to minimize the nuclear norm.

However, note that there are various points where the numerical method can be refined even further: First, the PDHG Algorithm can be extended to employ hard constraints for the variables. This is possible by adding additional characteristic functions to the objective function, but comes with some additional numerical overhead. Second, more sophisticated algorithms than Gauss-Newton could be used to cope with the nonlinear data term. An interesting candidate to do this could be the iPiano Algorithm [109], which is based on the heavy-ball method. iPiano extends the gradient descent for the primal variable x by an additional inertial term, which theoretically causes the algorithm to jump over small local minima. However, since the current numerical results are convincing, we only expect minor improvements by employing these adapted methods.

5 Joint Methods and Coupled Dynamic Models

In this chapter we show, how the structure of dynamic models can be exploited by adding additional parameters the estimation. Specifically, we introduce a joint approach for T_1 estimation which combines estimation of T_1 with a simultaneous motion correction of the data. Based on ideas which were originally employed to register a time-series of DCE-MRI images, we will use guesses of the parameters (T_1, M_0) to simulate a motion-less dataset, which can itself be used to improve the deformation parameters. To cope with noise effects, we additionally incorporate smoothness constraints for the parameter maps. Additionally to this iterative approach, we will propose an approach where estimation of dynamic parameters and deformation parameters is performed simultaneously. Results for simple affine deformations have been published in [H2]. In this work we will additionally describe an extension of this approach to nonlinear deformations. We will also present results, where incorporation of spatial coupling improves parameter estimation results by up to 30%. As another example of how additional parameters can be included in the estimation, we will show how the Patlak-Rutland Model can be used to estimate not only pharmacokinetic parameters but also the arterial input function. These results have been published in [H1]. Note that some results of this chapter have preliminary character, as the uneven impact of the different variables often introduced instabilities during the optimization.

5.1 Joint Motion Correction and T_1 Estimation

As outlined in Section 1.2.4, a main source of error in T_1 estimation for abdominal imaging is spatial misalignment of the data. Image registration might hence be a key factor to improve T_1-estimation [69]. However, registration of variable flip angle data is a challenging task since image contrast might differ significantly between different flip-angles, see Figure 5.1 for examples of such data. Conventional approaches for registration would hence aim to first register the data, using a contrast-invariant distance measure like Mutual Information or Normalized Gradient Fields [103, 102] and then obtain T_1 using a parameter fit. Since usage of dynamic models would give us additional means to determine how image intensities change, we will include this knowledge, which would be discarded in standard approaches, in the registration process.

Note that the problem to register variable flip angle data is structurally similar to the registration of of DCE-MRI images. Here approaches which couple parameter estimation and registration are capable to improve results of conventional approaches with respect to deformation as well as parameter errors. However, to the best of the author's knowledge the employment of such algorithms is not wide-spread in the community. In the case of DCE-MRI, two main challenges might be responsible for that: First, to create a model-simulation, at least two different models need to be concatenated: First, dynamic-parameters need to be converted to CA concentrations and second, CA concentrations need to be converted to MRI signal intensities. Since both of these models are frequently exponential [136], this can lead to numerical instabilities which are difficult to handle. Additionally, error analysis is complicated since a failure of the method needs to be traced back to failure of one or both of the two models. A second problem is that pharmacokinetic models are usually only valid for very specific kinds of tissue [138, 137]. If they are to be used to simulate a complete 3D dataset, this requires either prior segmentation and accurate knowledge of corresponding pharmacokinetic models or a strategy to discard the effects of tissue with unknown pharmacokinetic response. To cope with the latter issue, established approaches often assume a broad validity of a single mode [23, 11]. Although this might work for specific body regions, it is not clear if this approach can be generalized arbitrarily.

In order to better understand the behavior of joint approaches, we have thus decided to confine to a simpler setting: In this section, we propose a joint approach for T_1 estimation from variable flip angle data. Note that as compared to registration of DCE-MRI images, this is a simpler task. First, T_1 estimation requires the usage of only one model. Second, the model is valid on the complete image. In this work we will take advantage of these simplifications to point out the positive effects of additional spatial coupling, which might have been lost in the more complicated setting. A joint approach for T_1 estimation has been proposed previously for affine deformations [69]. Since those results had been published only few months earlier than ours, our first approaches for joint parameter estimation and registration were developed independently. Among minor technical differences, the most important additional feature of our method is the introduction of a spatial coupling term, which additionally stabilizes the parameter estimation. As our experimental results indicate, this term can have significant positive effect on both parameter estimation and registration.

This section is structured as follows: We start by introducing a framework for the joint approach, which will be called Coupled Model-based Parameter Estimation and Registration (COPTER). Since this modeling will lead to a large-scale optimization problem, we will subsequently discuss different strategies to solve the problem. Experimental evaluation on software phantom data, where ground-truth is available, will be performed in Section 5.2. Real data Experiments can be found in Section 7.3.
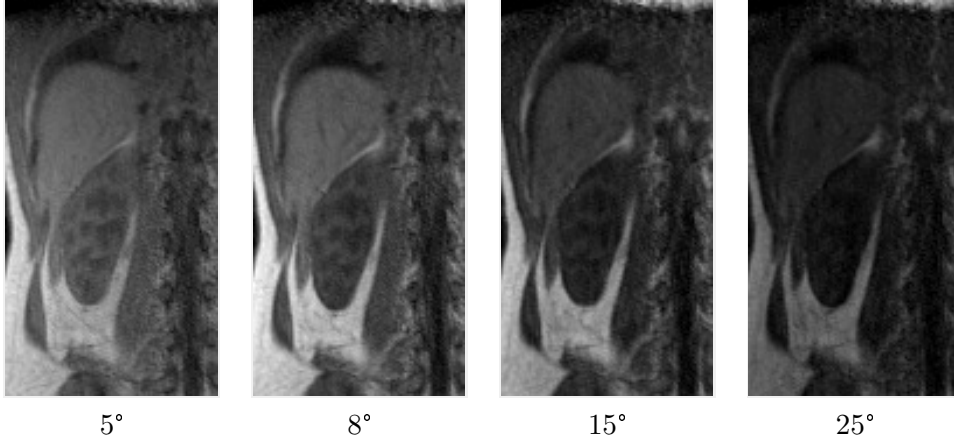
Figure 5.1: Figure showing MRI scans of a right human kidney obtained with variable flip angles. Scan sequence was TWIST with parameters TR = 2.51 ms, TE = 0.83 ms, isotropic voxel size 1.8 mm and matrix size [256,192]. Large differences in contrast can be observed. Since the scan was obtained during multiple breathholds, only small motion artifacts along the diaphragm can be observed.

Algorithm 2 Conventional Approach $(y, p) = \text{ConventionalApproach}\,(d, \alpha, \text{TR}, \lambda, \mu)$

1: Initialize first deformation as identity, i.e. $y_1(x) = x$.
2: **for** $i = 2$ to k **do**
3:    Estimate transformations $y_i = \text{argmin}_y \, \mathcal{D}(d_{i-1}(y_{i-11}), d_i(y)) + \mu \mathcal{R}(y)$
4: **end for**
5: Apply deformations to dataset $dy = (d_1(y_1), d_2(y_2), \ldots, d_k(y_k))$
6: Estimate parameters $p = \text{argmin}_p \, \mathcal{D}_{\text{SSD}}(M(p), dy) + \lambda \mathcal{C}(p)$
7: Concatenate $y = (y_1, \ldots, y_k)$
8: **return** (y, p).

5.1.1 A Brief Introduction to Image Registration

We start with a brief introduction to image registration, closely following [103, 102]. Let us assume that $R, T : \Omega \to \mathbb{R}$ are two (continuous) images, which are defined on a domain $\Omega \subset \mathbb{R}^3$. Let us additionally assume that both R and T display the same object in the presence of motion. Image registration can be formulated as the task of finding a suitable transformation $y : \Omega \to \mathbb{R}^3$ such that the energy defined by

$$\mathcal{D}(T \circ y, R) \tag{5.1}$$

is minimal. Here $\mathcal{D}$ is a distance measure, which evaluates the similarity between the two images and $T \circ y$ denotes the transformed image T. Concerning the distance measure $\mathcal{D}$, a straight-forward extension of the modeling described in Section 3.2.1 leads to sum-of-squared differences

$$\mathcal{D}_{\mathrm{SSD}}(T \circ y, R) := \int_\Omega (T \circ y(x) - R(x))^2 \, \mathrm{d}x.$$

An inherent assumption of this metric is that $T \circ y$ and R admit comparable image intensities. In the case, that R and T are obtained using different modalities or scan-parameters, several other metrics exist, see [103] for an overview.

Since the above problem is ill-posed [45], suitable regularization of the deformation field has to be established. One way to do this is to restrain it to simple linear models where $y(x) = Ax + b$ for $A \in \mathbb{R}^{3 \times 3}$ and $b \in \mathbb{R}^3$, which yields a better conditioned problem. This is the case for rigid or the more general affine transformations. As the example of motion in the abdominal area shows, there are many scenarios where rigid or affine transformations do not need to apply. In this case, non-linear registration can be a remedy for this problem. In non-linear registration, the transformation y is estimated voxel-wise. To make the problem well-posed, (5.1) can be extended by a regularizer $\mathcal{R}$, which measures the energy of the transformation according to some deformation model. This can be formulated as the following problem:

Minimize

$$J(y) := \mathcal{D}(T \circ y, R) + \mathcal{R}(y - y^{\mathrm{ref}}),$$

where y^{ref} enables a bias towards a particular solution, e.g. $y^{\mathrm{ref}}(x) = x$. A popular choice for $\mathcal{R}$ is for example elastic regularization as proposed by Broit [19], which is given by

$$\mathcal{R}_{\mathrm{elas}}(z) = \frac{1}{2} \int_\Omega \mu \langle \nabla z, \nabla z \rangle + (\lambda + \mu) \left(\mathrm{div}(z)\right)^2 \, \mathrm{d}x.$$

Here $\lambda, \mu \geq 0$ are the so-called Lamé constants and $\mathcal{R}_{\mathrm{elas}}$ approximates the elastic energy of the transformation. A mechanical interpretation of this energy and extensive additional information on image registration can be found in [103]. Various extensions of this energy exist, which can e.g. enforce invertible transformations and additionally allow control on volume and surface changes introduced by the transformation [24, 128]. However, since we expect transformations to be limited in scale, we will again keep the setup as simple as possible and confine to the standard elastic model.

5.1.2 COPTER: Coupled Model-based Parameter Estimation and Registration

In this section we will model joint approaches for T_1 estimation from variable flip-angle (VFA) data. We will assume that intra-image motion artifacts are negligible or have been included in the reconstruction process. Instead, our modeling will account for motion occurring between different breath-holds [69, 23, 11].

We start by introducing the modeling of the motion for the case of T_1 estimation. For ease of presentation, let us assume a continuous setting in this section. Let $\bar{p} = (T_1, M_0) : \Omega \to \mathbb{R}^2$ denote the ground-truth parameter maps for the image domain $\Omega \subseteq \mathbb{R}^3$. Corresponding MRI measurements with k flip-angles will be denoted by $\bar{d} = M(\bar{p})$ for $\bar{d} : \Omega \to \mathbb{R}^k$, for details see Section 2.3. As an extension of the previous modeling, in this section we will assume that $\bar{d}$ is not only degraded by noise, but instead is influenced by both noise and motion effects. Following [23, 11, 69] we may assume that for each measurement there is a transformation $y_i : \mathbb{R}^3 \to \mathbb{R}^3$ such that $d_i = \bar{d}_i(y_i) + n$, where $i \in \{1, \dots, k\}$. To simplify notation, we will use the short-hand notation $d(y) \coloneqq (d_1(y_1), \dots, d_k(y_k))$ and $y \coloneqq (y_1, \dots, y_k)$. We will estimate the quality of a given set of parameters p and transformations y by

$$\mathcal{D}_{\mathrm{SSD}}(M(p), d(y)) \coloneqq \|M(p) - d(y)\|_2^2. \tag{5.2}$$

Based on the proposed distance measures and the modeling introduced in Chapter 4 and [23, 11, 69], we propose to recover parameters as well as deformations by solving the following joint problem, Coupled Model-based Parameter Estimation and Registration (COPTER):

Minimize:
$$J(p, y) \coloneqq \mathcal{D}_{\mathrm{SSD}}(M(p), d(y)) + \lambda\, \mathcal{C}(p) + \mu\, \mathcal{R}(y) \tag{5.3}$$

where $\mathcal{D}$ is the datafit term proposed previously, $\mathcal{C}$ is a coupling functional for the parameter maps proposed in Chapter 4 and $\mathcal{R}$ is a regularization term for the transformations, see Section 5.1.1. Again the parameters $\lambda, \mu \geq 0$ are control parameters which determine the amount of regularity of the deformations and parameter maps. A variant of this approach has been published previously for the choice $\lambda = 0$, an affine deformation model and the linear datafit term in [69]. The choice of the regularization parameters λ, μ is delicate: Large values often yield fast convergence and solutions which fulfill the imposed regularity assumptions heavily. However, these parameters also often coincide with a less accurate datafit and thus fail to explain the data. In order to determine optimal parameters, there are various automatic strategies [70, 67]. However, since for our the phantom experiments ground-truth data will be available, we will determine these parameters experimentally.

Let us consider problem (5.3) more thoroughly: We can interpret the joint reconstruction as a parameter fitting problem with the possibility to adjust not only the parameters p, but also to rearrange the datapoint locations to improve the fitting.

Instead of fitting $M(p(x))_i$ to $d_i(x)$, the fitting might also be performed with respect to nearby data point $d_i(y)$. Since this problem naturally becomes ill-posed, regularity of both the parameter maps and the transformations are controlled by the coupling and regularization terms $\mathcal{C}$ and $\mathcal{R}$.

5.1.3 Solution Strategies for COPTER

We will now discuss various solution strategies to obtain a solution of (5.3). Most notably, we will introduce a *fully coupled* approach, where the objective function in (5.3) is optimized with respect the variable $x := (y, p)$, and an *alternating approach*, where subsequent optimizations with respect to $y_2, \ldots, y_k$ and p are performed. An experimental comparison of the approaches on phantom data will be performed in Section 5.2, showing convergence of both approaches. However, since in the fully coupled approach it takes a significant amount of time to solve a linear system, we will use the alternating approach in our later applications.

Fully Coupled Approach

In the fully coupled approach, the variables y and p are concatenated and optimization of (5.3) is performed with respect to $x := (y, p)$. For optimization standard methods like steepest descent or Gauss-Newton methods might be used. Pseudo-code for the fully coupled optimization approach can be found in Algorithm 3. We will demonstrate in Section 5.2 that this approach indeed converges quickly to a solution of the problem. However, note that for second order methods this approach can come with significant computational overhead. First, the dataterm $\mathcal{D}_{\text{SSD}}(M(p), d(y))$ introduces a coupling between the deformations and the parameter maps, causing the need to solve the Hessian system as a whole and rendering decoupling methods impractical. Additionally, x has size $(k-1)n_{\text{Def}} + m \cdot n$, where $n_{\text{Def}} \in \mathbb{N}$ is the number of parameters of a deformation field, $m \in \mathbb{N}$ is the number of parameter maps and $n \in \mathbb{N}$ is the number of voxels. Especially in the case of nonlinear deformations, this leads to large linear systems and a slow solution of the systems, also see Section 5.2. However, as it turns out, for linear models and quadratic coupling terms there is an efficient method to reduce the number of free variables and to perform the optimization only with respect to y, the so-called Variable Projection Method.

The Variable Projection Method

The Variable Projection Technique [62, 36] is a method to optimize objective functions which are quadratic in one set of variables and both non-quadratic and nonlinear in another. As it turns out, in such cases there is a way to reduce the dimensionality of the problem by optimizing the objective function on a sub-manifold of the complete variable space which is simple to parametrize. Optimization can then be performed

with respect to a smaller set of variables, thus yielding a lower dimensional, but also a possibly highly-nonlinear problem. In [36], the variable projection technique was experimentally compared to a standard, fully coupled approach for super-resolution, which is a methodologically closely related problem. Since results have shown an improved convergence speed, we will use VarPro to to optimize (5.3) for the linear model for T_1 estimation.

To explain the variable projection technique, let us assume that the objective function is quadratic in the parameter maps p. The main leverage of the variable projection technique is that in this case, given transformations y, optimal dynamic parameters $\bar{p} =: p(y)$ can be determined efficiently by solving a simple least-squares problem. This means that instead of optimizing $J(y, p)$ we can instead optimize $\tilde{J}(y) := J(y, p(y))$, which depends only on the transformations. On the first glance, it looks like this strategy comes with a additional technical difficulties: Gradient-based optimization strategies will require access to the derivative $\nabla_y p(y)$, which might be difficult to obtain. However, the following calculation shows that indeed $\nabla \tilde{J}(y) = \nabla_y J(y, p(y))$, meaning that the derivative of $\tilde{J}$ is equal to the derivative of J with respect to y:

$$\nabla \tilde{J}(y) = \nabla_y J(y, p(y)) + [\nabla_y p(y)]^\top [\nabla_p J(y, p(y))] = \nabla_y J(y, p(y)),$$

since by construction $\nabla_p J(y, p(y)) = 0$. As proposed in [36], we used a Gauss-Newton approximation of the Hessian to obtain a second order method. Pseudo-code for the variable projection technique can be found in Algorithm 4.

Alternating Optimization

Another way to solve (5.3) is to use an alternating optimization strategy: If some estimate y^l of the transformations is known, an estimate p^l of the parameters can be obtained by solving $p^l = \operatorname{argmin}_p \mathcal{D}_{\mathrm{SSD}}(p, y^l) + \lambda\,\mathcal{C}(p)$. Solving for p hence becomes a standard parameter estimation problem for the transformed data $d(y^l)$. Likewise, if some estimate p^l of the parameters is available, y can be updated by solving $y^l = \operatorname{argmin}_v \mathcal{D}_{\mathrm{SSD}}(M(p^l), d(y)) + \mu\,\mathcal{R}(y)$. This time the resulting problem becomes a standard registration problem where the data d is registered to the simulations $M(p^l)$. Also alternating optimization strategies usually are considerably slower convergence than coupled optimization strategies [10], we will see in Section 5.2 that the subproblems can indeed be solved faster yielding an overall improved convergence speed. Pseudo-code for the alternating optimization approach can be found in Algorithm 5.

Further Implementation Details

To increase robustness of the registration and to avoid local minima, which might prevent large deformations in the transformation field, we added a multi-level approach from coarse to fine: We start with down sampled data at a coarse resolution, which

allows us to capture large deformations and to initialize the parameter fit in the approximate range. We then subsequently refine the resolution until we reach the image resolution with our final results. For differentiable objective functions, we again used a variant of the Gauss-Newton algorithm which incorporate the constraints $1 \leq T_1 < \infty$ and $1 \leq M_0$, which ensured a well-posed parameter fit (cf. Section 3.3.2). Also note, that only deformation parameters $y = (y_2, \ldots, y_k)$ will be estimated, since the problem becomes ill-posed for $y = (y_1, \ldots, y_k)$.

Conclusion

We have proposed different optimization schemes, which might be applied to obtain a solution of (5.3). Most prominently, we have distinguished between a fully coupled approach, where optimization is performed with respect to $x := (y, p)$ and an alternating approach, where optimization is performed iteratively with respect to $y_2, \ldots, y_k$ and p. For the fully coupled approach and a linear model formulation, we have additionally presented a technique to reduce the dimensionality of the problem (variable projection technique). In the following sections we will present evidence which clarifies, under which circumstances COPTER is superior to conventional techniques and which optimization scheme is most suitable for practical applications.

Algorithm 3 Joint Optimization $(y, p) = \text{COPTER\_Joint}(d, \alpha, \text{TR}, \lambda, \mu)$

1: Setup multilevel data $d^{\text{minLevel}}, \ldots, d^{\text{maxLevel}}$.
2: Initialize $p^{0,\text{minLevel}}$ by a nonlinear parameter fit to the unregistered data.
3: Initialize $y^{0,\text{minLevel}} := (y_2^0, \ldots, y_k^0)$ as identity transformations.
4: Concatenate $x^0 := (p^{0,l}, y^{0,l})^\top$.
5: **for** $l = \text{minLevel}$ to maxLevel **do**
6:     Initialize objective function J^l for current level:
7:     $J^l(x) := \mathcal{D}_{\text{SSD}}{}^l(M(p), d(y)) + \lambda \, \mathcal{C}^l(p) + \mu \, \mathcal{R}^l(y)$.
8:     **repeat**
9:         Get Gradient dJ^i at x^i
10:         Get Gauss-Newton approximation H^i of the Hessian at x^i
11:         Solve $H^i u^i = -dJ^i$
12:         $\beta^i = \text{ArmijoLinesearch}(u^i, x^i, J)$
13:         $x^{i+1} = x^i + \beta^i u^i$
14:     **until** convergence
15:     Prolong to next level $x^{0,l+1} \leftarrow x^{i+1,l}$
16: **end for**
17: **return** $(y^{i+1,\text{maxLevel}}, p^{i+1,\text{maxLevel}})$.

Algorithm 4 Variable Projection Technique $(y, p) = $ COPTER_VarPro$(d, \alpha, \mathrm{TR}, \lambda, \mu)$

1: Setup multilevel data $d^{\mathrm{minLevel}}, \ldots, d^{\mathrm{maxLevel}}$.
2: Initialize $y^{0,\mathrm{minLevel}} := (y_2^0, \ldots, y_k^0)$ as identity transformations.
3: **for** $l = \mathrm{minLevel}$ to $\mathrm{maxLevel}$ **do**
4:     Initialize function $p^l(y)$ which maps y onto optimal p.
5:     Initialize objective function $\tilde{J}^l$ for current level:
6:     $\tilde{J}^l(y) := \mathcal{D}_{\mathrm{SSD}}{}^l(M(p(y)), d(y)) + \lambda \mathcal{C}^l(p(y)) + \mu \mathcal{R}^l(y)$.
7:     **repeat**
8:         Get Gradient $d\tilde{J}^i$ at y^i
9:         Get Gauss-Newton approximation H^i of the Hessian at y^i
10:        Solve $H^i u^i = -d\tilde{J}^i$
11:        $\beta^i = \mathrm{ArmijoLinesearch}(u^i, y^i, \tilde{J})$
12:        $y^{i+1} = y^i + \beta^i u^i$
13:     **until** convergence
14:     Prolong to next level $y^{0,l+1} \leftarrow y^{i+1,l}$
15: **end for**
16: **return** $(y^{i+1,\mathrm{maxLevel}}, p^l(y^{i+1,\mathrm{maxLevel}}))$.

Algorithm 5 Alternating Optimization $(y, p) = $ COPTER_Alt$(d, \alpha, \mathrm{TR}, \lambda, \mu)$

1: Setup multilevel data $d^{\mathrm{minLevel}}, \ldots, d^{\mathrm{maxLevel}}$.
2: Initialize identity transformations $y^{0,\mathrm{minLevel}} := (y_2^0, \ldots, y_k^0)$.
3: **for** $l = \mathrm{minLevel}$ to $\mathrm{maxLevel}$ **do**
4:     Initialize functions $\mathcal{D}_{\mathrm{SSD}}{}^l(M(p), d(y))$, $\mathcal{R}^l(y)$, $\mathcal{C}^l(p)$ for current level.
5:     **repeat**
6:         Estimate parameters $p^{i+1,l} = \mathrm{argmin}_p \mathcal{D}_{\mathrm{SSD}}{}^l(M(p), d(y^{i,l})) + \lambda \mathcal{C}^l(p)$
7:         Estimate transformations $y^{i+1,l} = \mathrm{argmin}_y \mathcal{D}_{\mathrm{SSD}}(M(p^{i+1,l}), d(y)) + \mu \mathcal{R}(y)$
8:     **until** convergence
9:     Prolong deformations to next level $y^{0,l+1} \leftarrow y^{i+1,l}$.
10:    Prolong parameter maps to next level $p^{0,l+1} \leftarrow p^{i+1,l}$.
11: **end for**
12: **return** $(y^{i+1,\mathrm{maxLevel}}, p^{i+1,\mathrm{maxLevel}})$.

5.2 Results: COPTER for Software Phantom Data

We now proceed to evaluate COPTER on phantom data. Real data experiments can be found in Section 7.3 We start by clarifying the boundaries, where joint parameter estimation and registration can be expected to be superior to standard approaches, where these steps are combined. Our experiments indicate that for both simple data and simple deformations, only minor advantages of the joint approach can be expected. However, note that for slightly more complex data results of [69] already show significant improvements by the use of joint approaches. We will find that although joint optimization yields minor improvements in parameter errors, results from alternating optimization are comparable. For the following experiments we hence proceed to use the alternating approach. Finally, we present results on phantom data which indicate that employment of spatial coupling is capable to improve both T_1 estimation and motion correction as compared to uncoupled methods by about 30%.

5.2.1 Comparison of COPTER with Conventional Approaches on Simple Data

In order to analyze joint approaches for parameter estimation an registration, we start by comparing conventional approaches, where data is registered first and parameter estimation is performed subsequently, with joint approaches, which combine these steps. In this section we will demonstrate, that for simple data and simple deformations conventional and joint approaches are yielding comparable results. Note that this is promising especially from a practical perspective, since standard software might be used to estimate parameters and deformations separately. However, it has been shown in [69] that for more complicated data joint approaches can significantly improve results as compared to conventional approaches. Indeed, we will demonstrate in the next section that the results of [69] can be improved even further if coupling terms for parameter maps are introduced into the parameter estimation. However, note that parts of the employed modeling in this section were early drafts which later turned out to be sub-optimal for more general problems. Specifically, in this section coupled parameter estimation will be performed in the (E_1, N) coordinate system, which later was determined unsuitable due to exponential effects, see Section 4.4. Also, in this section we will consider an additional stabilization term, which is based on L_2 distances of different images and thus ill suited for different flip angles. We nevertheless found positive effects by including the latter in our setting. These parts will be improved in the next section. Results of this section have been published in [H2].

Objective Function

We start by clarifying the estimation model. Based on the linear reformulation of T_1 estimation (see Section 3.2.4) and an affine transformation model, in this section

we will study the following problem to recover $p = (E_1, N)$ from motion and noise degraded data:

Minimize:
$$J(y, p) := \mathcal{D}_{\text{SSD}}(A(y)p, b(y)) + \mu \, \mathcal{R}(y) + \lambda \, \mathcal{C}(p),$$
such that

$$\mathcal{R}(y) = \sum_{i=2}^{k} \mathcal{D}(d_i(y_i), d_{i-1}(y_{i-1})), \tag{5.4}$$

$$\lambda \, \mathcal{C}(p) = \lambda_1 \|\Delta_h p_1\|^2 + \lambda_2 \|\nabla_h p_2\|^2,$$

$$y_i(x) = T_i x + s_i \text{ for } T_i \in \mathbb{R}^{3,3} \text{ and } s_i \in \mathbb{R}^3.$$

As described in Section 3.2.4, we can then recover $T_1 = -\log(E_1)/\text{TR}$ from p. A straight forward comparison with (5.3) shows that the above problem falls in the class of joint approaches with a specific choice for the regularizer $\mathcal{R}(y)$: The data-terms $\mathcal{D}(d_i(y_i), d_{i-1}(y_{i-1}))$ can be regarded as stabilization terms for the transformations, since they are forcing additional constraints for the deformations. However, note that due to the intensity differences usage of an SSD distance is sub-optimal and will be discarded in later experiments. Nevertheless we have found that for our data this term has had positive effects on the parameter estimation. For the coupling term $\mathcal{C}(p)$, we denote by $\nabla_h \in \mathbb{R}^{n,n}$ a discrete Gradient- and by $\Delta_h \in \mathbb{R}^{n,n}$ discrete Laplacian Operator. the Laplacian stabilization in E_1 is chosen to account for the bias introduced by the exponential scaling. Optimization is performed using the Variable-Projection technique described in Section 5.1.3.

A Software Phantom with Affine Deformations

We used the XCAT phantom [133] to obtain ground truth data of an axial 2D scan of a human body with voxel size $1.8 \, \text{mm}^3$, matrix size 256×256 and TR $= 2.51 \, \text{ms}$. Flip angles $\alpha := (5°, 8°, 15°, 25°)$ were chosen to simulate signals the following areas using (2.19): Kidney (medulla/cortex/pelvis), spleen, liver, bone marrow, fat and air. T_1-values for these structures were taken from [8]. Since no data for the steady-state magnetization M_0 was at hand, we used the constant value $M_0 = 1500$ for all tissue types as previously proposed in [33], although this value is expected to vary with the tissue. The static phantom was then perturbed by small affine deformations, yielding a dataset d_{def} (cf. Figure 5.2). Finally Rician noise, as expected in MR data (cf. Section 6.2.1), was created by adding white Gaussian noise to the real and imaginary part of the Fourier-transformed data. We created multiple noise and motion corrupted datasets at various Signal to noise ratios, defined as SNR $:= \mu(d_{\text{def}})/\mu(d - d_{\text{def}})$ where μ denotes the mean value of the respective signals.

Reference Methods

We have compared the proposed algorithm to two conventional approaches to recover T_1-Maps. These consist of registration and subsequent parameter estimation. We begin by outlining the registration methods:

1. A *sequential registration.* Here the registration was performed sequentially $d^i \to d^{i-1}$ with an implementation from FAIR [102].

2. A *sequential stabilized registration.* Here the registration was performed by minimizing $J(v) := \sum_{i=2}^{k} \|d^i(y_i) - d^{i-1}(y_{i-1})\|^2$. The method was chosen to investigate possible advantages introduced by the coupling of registration parameters in (5.4).

For the subsequent parameter estimation, parameter maps (E_1, N) were calculated by solving

$$\text{Minimize: } J(z) := \|Az - b[y_{\text{opt}}]\|^2 + \alpha_1 \|\Delta_d E_1\|^2 + \alpha_2 \|\nabla_d n\|^2),$$

where y_{opt} denotes the optimized deformation parameters. Following the parameter estimation, T_1 was reconstructed from E_1 by calculating $T_1^i = -\text{TR}/\log(E_1^i)$. Since changing to logarithmic scale can introduce large outliers, the relative error in T_1 ($\text{real}T_1 := \|T_1^{\text{rec}} - T_1^{\text{true}}\|_2/\|T_1^{\text{true}}\|_2$) was calculated only on areas with constant T_1 (cf. Figure 5.2). We experimentally determined parameters $\alpha_1 = 10^5$, $\alpha_2 = 10^{-1}$ and $\beta = 1$ for all experiments including noise. For the experiment without noise we chose $\alpha_1 = \alpha_2 = 10^{-3}$.

Results

Evaluation was performed with respect to the maximal deformation error (maxDE), meaning the maximal error of all four deformations with respect to the mean deformation in pixels. Results for various SNRs are displayed in Table 5.1. It can be seen, that the parameter error in T_1 mildly profits from the joint approach. Relative errors in T_1 are reduced slightly from 0.4253 to 0.4049. However, improved reconstruction was possible regardless of larger deformation errors. Reasons for this behavior might be linked to smoothing effects resulting from the spatial coupling in (E_1, N).

5.2.2 An Extension of COPTER to Nonlinear Deformations

In this section we show how the stabilized reconstruction introduced in (5.3) can be used to achieve higher precession in T_1 reconstruction in the presence of nonlinear deformations. The main contribution of this work lies in the introduction of spatial coupling for the parameter fit. For the regularization of the deformation we chose an elastic regularizer for all experiments, see [103, 102] for details.
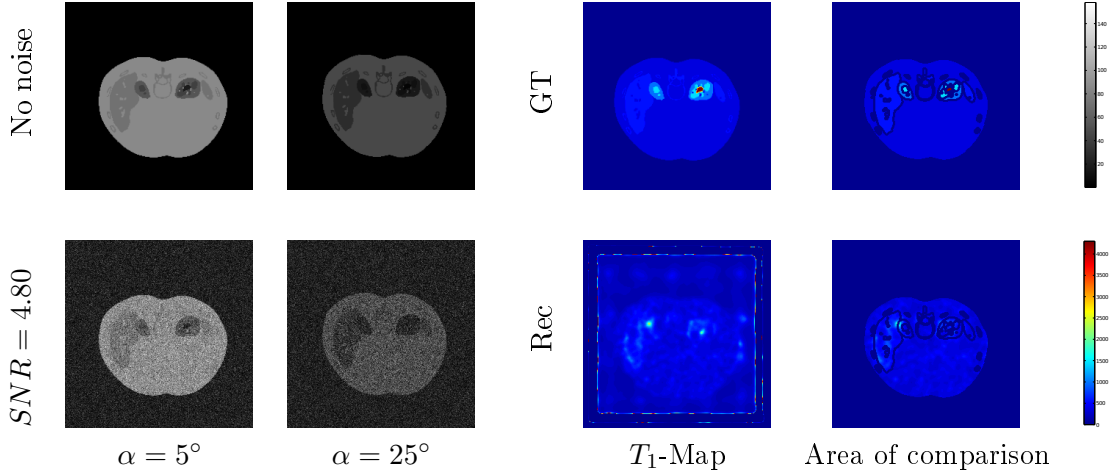
Figure 5.2: Phantom data and reconstruction results. Left images: Ground-truth signals and noisy signals for minimal and maximal flip-angles $\alpha_1 = 5°$ and $\alpha_4 = 25°$. Right images: Ground-truth (GT) and Reconstructed (Rec) T_1-Maps for SNR=4.80 as well as the area of comparison to calculate the relative error.

Table 5.1: Results for various SNRs. For the initial deformation we measured maxDE=9.6766 with respect to the identity-transformation. Deformation errors are expressed in pixels.

	Joint		Seq		Seq-Stab	
SNR	maxDE	reT1	maxDE	reT1	maxDE	reT1
∞	0.3149	0.0076	0.2256	0.0029	0.2552	0.0039
10.66	2.0685	0.3045	0.4710	0.3108	0.5262	0.2848
5.38	2.8197	0.3692	1.6124	0.4062	3.0644	0.3836
4.80	3.2040	0.4049	2.3348	0.4381	3.4063	0.4253

Let us begin by stating the model equations for the discrete setting: Assume that a discrete motion and noise corrupted dataset $d = (d_1, \ldots, d_k)^\top \in \mathbb{R}^{kn}$ is given, where $d_i \in \mathbb{R}^n$. We aim to reconstruct $p \in \mathbb{R}^{2n}$ and $y \in \mathbb{R}^{n_{\text{def}}, k-1}$, where n_{def} is the number of deformation parameters and for our model in the order of $3n$. In the experiments we will do this by minimizing one of the following functionals:

$$J_{\text{COPTER}}(p, y) := \mathcal{D}_{\text{SSD}}(M(p), d(y)) + \alpha\, \mathcal{C}_{\text{TP}}(z) + \beta\, \mathcal{R}_{\text{Elas}}(y), \qquad (5.5)$$

$$J_{\text{Ref}}(p, y) := \mathcal{D}_{\text{SSD}}(A(y)p, b(y)) \qquad\qquad + \beta\, \mathcal{R}_{\text{Elas}}(y). \qquad (5.6)$$

Note that the reference method J_{Ref} is the direct extension of the method originally presented in [69] to nonlinear transformations and is hence formulated for the linear model described in Section 3.2.4. Since we expect benefits from using the nonlinear
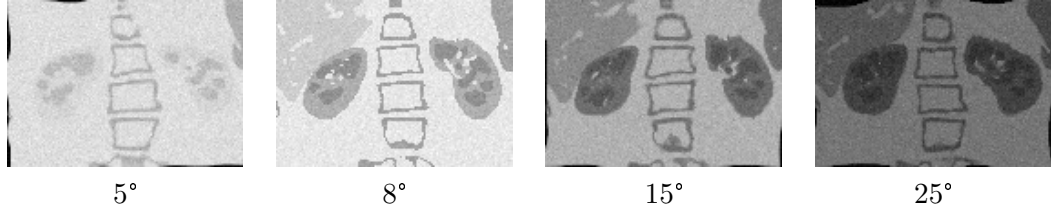
| 5° | 8° | 15° | 25° |

Figure 5.3: Figure showing a 2D slice of the 3D software phantom with nonlinear motion at 9% noise. Signals were simulated with flip angles $\alpha = (5°, 8°, 15°, 25°)$ and nonlinear motion. For details see text.

datafit as well as spatial coupling of the parameters, we propose to minimize the stabilized variant J_{COPTER} instead. Note that an extension to different stabilizers is straight-forward. However, due to the heavy computational load for 3D data, we confine to Tikhonov-Philips coupling.

A Software Phantom with Nonlinear Deformations

A ground-truth T_1 parameter map displaying the human kidney was set up as as described in Section 3.6.1. After that, $k - 1$ motion fields were generated by setting up spline transformation with three random coefficients per voxel (standard deviation was 0.015) [102]. The motion fields were then applied to $k - 1$ copies of T_1 and M_0 using linear interpolation to simulate partial volume effects. As we were mainly interested to study the impact of the spatial coupling, we obtained the final ground-truth deformation field by registering the undeformed T_1 map to the T_1 maps deformed with spline coefficients using an elastic regularizer. This ensures that the ground-truth deformation maps are physically meaningful deformations. Finally, we used the signal-equation to simulate MR signals with sequence parameters S4, which are the same as the real data (Section 3.6.1). After that Rician noise was added at with an average of 0%, 9% and 15% of the signal strength.

Results and Conclusion

The T_1 estimation was evaluated using the voxel-wise relative error in T_1, which was then averaged over the cortex and medulla of the kidney. The regularization parameters were determined manually for each noise level individually. Results are shown in Figure 5.4. One can see that the alternating optimization and the joint optimization strategy yield comparable results. For the highest noise level, results from the alternating approach are even superior to the fully coupled approach. This is possibly due to the simpler optimization: For the fully coupled approach we have found that it is difficult to cope with effects of different variables and the coupling. Here additional experiments are necessary to further improve the optimization and to find strategies

to deal with the different scaling of the variables: Changes in the parameter maps are expected to have a significantly larger effect on the objective function that changes in the deformation parameters. Since experiments for the affine deformation model have not shown these effects, we suspect the significantly reduced number of parameters to be responsible for this behavior. This effect might also be responsible for the non-zero error for the zero noise data.

As compared to the uncoupled approach where J_{ref} is minimized, one can see that parameter errors are significantly lower. This is also reflected in a visual inspection of the recovered T_1 maps, where large outliers for the recovery using J_{ref} can be observed. Note that concerning the estimation of the deformation field, growing structures can be observed in the unstabilized approach. In these regions the model parameters are emulating motion, leading to a inferior motion correction. Note that this behavior is reduced by employing spatial coupling. In conclusion, one can see that for the given data spatial coupling is indeed capable to improve both reconstruction and registration results. Nevertheless we suggest to perform additional experiments, where data is registered first and parameters are estimated subsequently. These experiments could help to determine causes for the non-zero errors in the presence of no noise, which might be linked to the large deformations appearing in the model. Real data results for these approaches can be found in Section 7.3.

5.3 Blind Parameter Estimation for the Patlak Model

In this section we will propose a novel method to compute a voxel-wise map displaying the local filtration of the kidney using the Patlak Rutland Model. Most notably, the described method will not require knowledge of the plasma concentrations c_{a}, the so-called arterial input function (AIF). A variant of this method has been published by the author in [H1].

A well-known challenge of this and other pharmacokinetic models for GFR estimation lies in the dependency of the model on the AIF [38, 99, 32]. As outlined in Section 2.2, the AIF is usually estimated directly from the DCE-MRI data by measuring the plasma concentration in a feeding vessel of the kidneys. However, since there is no standard protocol to select the AIF and GFR results are highly dependent on the choice of the region, this approach is difficult in practice and requires experienced users [99, 38]. Since these factors make AIF selection a nuisance, different approaches have been proposed to cope with this issue, most common the approach to use standardized AIFs [156, 113]. However, as the results from [113, 99] show, the AIFs between patients can differ significantly. Another class of approaches exploits explicit knowledge of parameters for specific tissue [85, 162, 146]. Here different methods have been employed recover the AIF from such tissue. In this work we will follow a similar approach which was first proposed in [121]. Here the AIF was introduced as a further unknown in the parameter estimation. It was assumed that multiple tissue curves were available

Noise	J_{Ref}	COPTER_Alt	COPTER_Joint
0%	0.1455	0.1020	0.0980
9%	0.1832	0.1208	0.1201
15%	0.2061	0.1379	0.1395
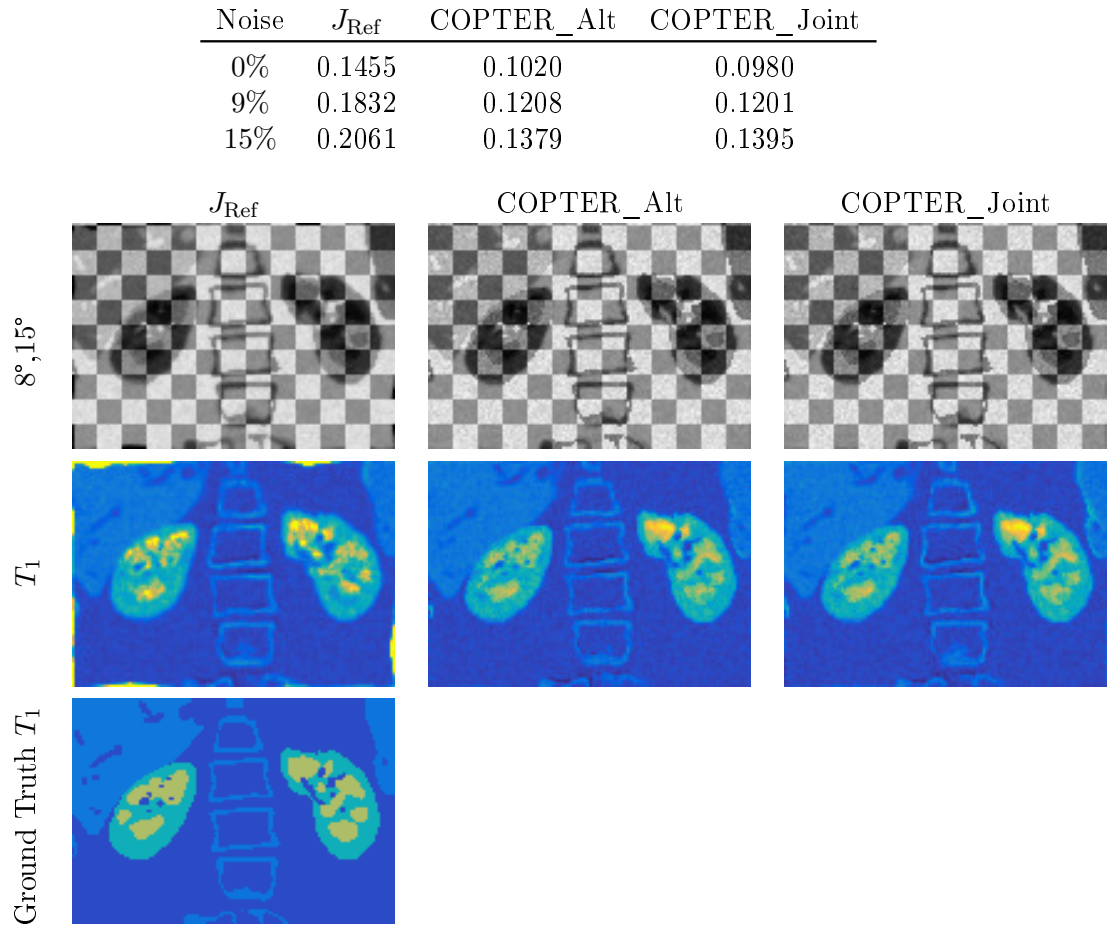
Figure 5.4: Table showing the mean relative error in T_1 for registration of $3D$ phantom data. Also shown are registration results for slice 35 of the phantom data at 9% noise and flip-angles $\alpha_2 = 8°, \alpha_3 = 15°$ as well as slice 35 of the recovered T_1-maps estimation. Results show that both the unstabilized and the stabilized approach accurately capture motion whereas improved results for parameter estimation are obtained in the stabilized method.

which all could be explained with the same AIF. This allows to formulate a joint optimization problem, where not only the pharmacokinetic parameters are estimated but also one AIF which couples all tissue curves. However, our experiments show that in the presence of noise and low-quality data the proposed method shows a lack of stability. In this paper, we suggest to extended the approach [121] by adding spatial coupling for the parameter maps and a normalization constraint on the AIF to cope with non-uniqueness of solution.

Our results indicate that this additional stabilization indeed yields improved results of up to 45 percent for K^{trans} in the presence of strong noise. However, we also find that stabilization in the coupled setting is difficult: Sine one parameter determines the other, regularity constraints on one parameter will impact on the other as well. Our experiments indicate that hence not only stabilization of both parameters (AIF and parameter maps) is necessary, but that additionally some estimate of one parameter needs to be available. We will thus assume that an estimate of the AIF is available, which can be given by noisy measurements or a population based average as proposed in [113]. Although our method can deal with a variety of pharmacokinetic models, for ease of presentation and computation we focus on the so-called Patlak-Rutland model [116, 115, 127]. We start this section by repeating the Patlak-Rutland Model which has been introduced in Section 2.2 and in Section 3.2.3. For a continuous arterial input function $\mathbf{c}_{\mathrm{a}} : [0, T] \to \mathbb{R}$ with discrete counterpart $c_{\mathrm{a}} := (\mathbf{c}_{\mathrm{a}}(t_1), \ldots, \mathbf{c}_{\mathrm{a}}(t_1))^\top$ this is given by

$$M(p, c_{\mathrm{a}}) = v_a \mathbf{c}_{\mathrm{a}}(t) + K^{\text{trans}} \int_0^t \mathbf{c}_{\mathrm{a}}(s) \, \mathrm{d}s$$
$$\approx p_1 c_{\mathrm{a}} + p_2 (L c_{\mathrm{a}})$$

where $L \in \mathbb{R}^{k,k}$ is a lower triangular matrix which describes the integration (see Section 3.2.3).

5.3.1 A Heuristic Motivation for Joint Approaches

We start by motivating the notion behind joint approaches from the perspective of linear systems: Given data $d \in \mathbb{R}^k$ our usual goal is to parameters $p \in \mathbb{R}^m$ such that $J(p) := \|M(p) - d\|^2$ is minimal. However, let confine to linear systems since our standard technique for parameter estimation is to use a Gauss-Newton technique where the residue is linearized:

$$\text{Minimize } J(p) := \|Ap - b\|^2, \text{ where}$$

$$A := [\nabla_p M(p_0)] \text{ and } b := \nabla_p M(p_0) p_0 + d - M(p_0).$$

In the following let us assume that $A \in \mathbb{R}^{k,m}$ has full rank, meaning that the p_0 is identifiable. Indeed, in this case standard theory implies that this problem has a unique solution if $k \geq m$, which in our situation means that there are more measurements

than parameters. However, in many cases it even holds that $k >> m$ and the resulting system is heavily overdetermined. This motivates the idea that it could be possible to estimate additional variables by extending the model. In the case of the Patlak Model we will include the arterial input function c_a as an additional variable with k unknowns in the parameter estimation. Combined with the standard parameters $p \in \mathbb{R}^2$ this yields a new vector of parameters $(p, c_\mathrm{a})^\top \in \mathbb{R}^{2+k}$. However, if only one data curve $d \in \mathbb{R}^k$ is available, the resulting system is underdetermined and hard to solve. If on the other hand more data curves $(d_1, \dots, d_n) \in \mathbb{R}^{nk}$ are available, we will have roughly kn equations for $2n + k$ unknowns. If $k > 2$, this system will still be overdetermined for $n > k/(k-2)$ data-curves.

5.3.2 Joint AIF and Parameter Estimation

We assume that discrete concentration values $d \in \mathbb{R}^{nk}$ are available. Following Section 3.2.3 and the introduction of this section, we need to find for each location i parameters $p_i := (v_{a,i}, K_i^{\mathrm{trans}})$ such that $M(p_i, c_\mathrm{a}) \approx d$, where $c_\mathrm{a} \in \mathbb{R}^k$ is the arterial input function. As for all compartment models, it is assumed that there is one global AIF for all compartments. Following the approach in [139], we neglect local delays in the arrival time and assume the AIF to be constant for all voxels. We will recover the parameters and the AIF by minimizing the following objective function $J : \mathbb{R}^{2n} \times \mathbb{R}^k \to \mathbb{R}$:

$$
\begin{aligned}
J(p, c_\mathrm{a}) &:= \mathcal{D}_{\mathrm{SSD}}(M(p; c_\mathrm{a}), d) + \alpha_1 \ \mathcal{C}(p) + \alpha_2 \ \mathcal{S}(c_\mathrm{a}), \\
&\text{s.t.} \quad p \geq 0, c_\mathrm{a} \geq 0 \text{ and } \langle e, c_\mathrm{a} \rangle = 1.
\end{aligned}
\tag{5.7}
$$

In our experiments we chose $\mathcal{C}(p) := \mathcal{C}_{\mathrm{TP}}(p)$ and $\mathcal{S}(c_\mathrm{a}) = \|W \cdot \mathcal{F}(c_\mathrm{a})\|^2$, where $W := 1/|c_{\mathrm{aref}}|$ are the inverse amplitudes of a Fourier transform of a reference AIF. The latter constraint gives an additional bias onto a pre-defined arterial input function, which might be corrupted measurement or a population-based AIF. The vector e is an integration vector, which forces a fixes scale of the model: To see the necessity for this constraint, we note that the Patlak-model is invariant with respect to scale $\sigma \in \mathbb{R}$, if both p and c_a are estimated. This can be seen from

$$
M(p, \sigma c_\mathrm{a}) = v_a \sigma c_\mathrm{a}(t) + K^{\mathrm{trans}} \int_0^t \sigma c_\mathrm{a}(s) \, \mathrm{d}s = M(\sigma p, c_\mathrm{a})
$$

This means that results of the approach will be independent of the scale-factor σ. Of course, various modifications of this model are possible.

The minimization of (5.7) with the modified Gauss-Newton optimization on a unit simplex, see [9]. Results for this reconstruction are displayed in Section 5.4.

5.4 Results: Joint AIF and Parameter Estimation for Software Phantom Data

In this section we will present results which show how joint methods can improve parameter estimation, if no or only a low-quality AIF is available. Specifically we will evaluate the method which was proposed in Section 5.3.

In a first step, we show results for a controlled environment, where the true solution is known. Here, we use a software phantom to create numerical data from given parameter maps and added various amounts of white Gaussian noise. Note that we have assumed a Gaussian Noise model for contrast agent concentrations, since to the best of the authors knowledge no analytic relationship between noisy data and contrast agent concentrations has been established yet. We compared our method to the constrained version of an unregularized method proposed in [121], which is most closely related. We find that in the presence of noise the spatial stabilization improves robustness of the parameter estimation.

5.4.1 Phantom Data

We simulated 11679 uptake curves with a software phantom using the AIF measured by Parker et al. [113] and the Patlak model. We set the parameters for the two regions (Cortex and Medulla) constantly to $p_1^{\mathrm{Cortex}} = 0.08$, $p_2^{\mathrm{Cortex}} = 0.25$ and $p_1^{\mathrm{Medulla}} = 0.01$, $p_2^{\mathrm{Medulla}} = 0.16$, respectively, thereby approximating real kidney parameters. To simulate partial volume effects, the parameter maps were smoothed with a Gaussian filter. The result is displayed in Figure 5.5. The uptake curves were generated in a time interval $[0\mathrm{s}, 40\mathrm{s}]$ at a time resolution of 0.1s. Afterwards the MR measurements were simulated by taking only every 25th value, which corresponds to a measurement every 2.5s. White Gaussian noise was added to the simulated uptake curves at various signal-to-noise ratios (SNRs). With the arithmetic mean $\mu(C)$ and the noise standard deviation σ_{Noise}, we use $\mathrm{SNR} := \mu(C)/\sigma_{\mathrm{Noise}}$ as an indicator of quality.

Results

We compared our results by the following relative error and relative standard deviation $\mathrm{RE}(q) := \|q_{\mathrm{true}} - q\|_2/\|q_{\mathrm{true}}\|_2$, $\mathrm{SD}(q) := \sigma_{(q_{\mathrm{true}}-q)}/\mu(q_{\mathrm{true}})$. The regularization parameters were determined experimentally and were fixed at $\alpha_1 = (1, 2), \alpha_2 = 1$. Using the novel method, stable reconstruction with a relative error less than 6% was possible up to a noise level of $\mathrm{SNR} = 6$. In order to remove the dependency on scale, we renormalized the integral of the estimated AIF after the calculation to be the same as the integral of the AIF of the ground truth. Results are displayed in Figure 5.5 and Table 5.2.

Table 5.2: Reconstruction results for software phantom data, $11,679$ spatial locations, 16 time points. The relative errors and standard deviations of the parameter maps p and q obtained by the method of [121] and the new regularized method are summarized.

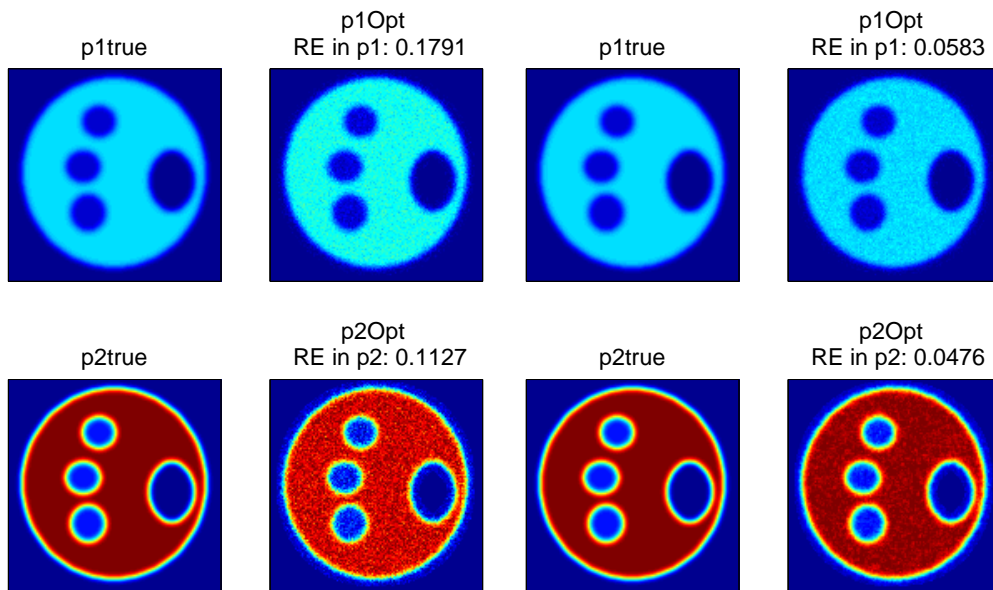|	uncoupled model [121]		coupled model	
	$\mathrm{RE}(p_1)$	$\mathrm{RE}(p_2)$	$\mathrm{RE}(p_1)$	$\mathrm{RE}(p_2)$
no noise	0.0317	0.0059	0.0307	0.0454
SNR = 6	0.1791	0.1127	0.0583	0.0476
SNR = 4	0.9510	0.6400	0.0971	0.0734

Figure 5.5: Reconstruction of p and q maps by the unregularized method [121] compared to the new regularized method. Phantom data with SNR = 6.

Conclusion

A simulation shows that it is indeed possible to estimate parameters for the Patlak Model, even if only a low-quality measurement of the AIF is possible and parameters are degraded by severe noise. Compared to the unregularized method [121], the relative error in noisy data is reduced by up to 45 percentage points for K^{trans}. However, note that that it was not possible to remove the dependency on the AIF completely and we still rely on a low-quality estimate of the AIF. If only the parameter maps were stabilized, we found that results were highly dependent on the stabilization parameter. We assume that reasons for this are linked to the coupling of c_{a} and p: Since in the double-blind setting one parameter determines the other, systematic errors in one lead to systematic errors in the other. In this case best results were achieved if all stabilization parameters were set to zero. This sensitivity needs to be taken into account especially for clinical data. Another problem is that due to the underlying model our method cannot yet estimate parameters uniquely. A proper scaling approach has still to be identified, but could exploit knowledge of some tissue parameters [146] or knowledge of the approximate integral of the arterial input function as in [53]. Note that this problem is delicate, since this unknown is a direct scaling constant. This problem affects especially the evaluation on clinical data, see Section 7.4.

However, we have found that if an approximations of the arterial input function is available, the introduction of spatial coupling yielded significantly reduced errors in parameter maps were observed on our test dataset. Also, our method has the advantage of being easily extendable to other compartment models, different regularization such as TV and time delay or smoothness assumptions on the AIF. Evaluation on real data will be performed in Section 7.4.

6 Perfusion in Coupled Systems

In this chapter we will demonstrate limitations of traditional dynamic models to re-cover perfusion in coupled systems. We will show that, although dynamic models are powerful concepts which have been used with great success, the inherently local design has limitations if blood flow is expected to pass through a coupled system. Specifically, we will show that traditional dynamic models will overestimate perfusion if they are applied to only parts of a such a system. In the experimental evaluation in Section 6.3 we will additionally present real data which indicates that such effects can also be found in coarse scale on real data. All results of this chapter have been submitted to the *IEEE Transactions on Biomedical Engineering* [S4].

We will start in Section 6.1 by introducing a novel continuous model for contrast agent propagation through a tissue patch with a highly developed capillary system. We will later show, that a discretization of this continuous model can equivalently be described by coupled traditional models. The simulation of the contrast-agent propagation will be performed in two steps, which are depicted in Figure 6.1: We start by modeling steady-state blood flow through a tissue patch under the assumption that the flow is mainly driven by pressure differences (Darcy's law). Having obtained a flow-field, we then model the contrast-agent propagation through the flow field in a second step. As compared to established approaches, which model contrast agent propagation through blood vessels [26], our modeling is based on porous media theory. It thus additionally incorporates the assumption that only fractional parts of the tissue are accessible for fluid, which is a basic assumption for live tissue.

We will then relate the novel, continuous model to the traditional Meier-Zierler Model, which describes the average contrast agent concentration in a tissue patch by

$$C(t) = P_a \int_0^t c_a(s) R(t-s) \, ds.$$

In Section 6.2.1 we will show that a discretization of our model can be equivalently de-scribed by coupled traditional dynamic models for perfusion. Specifically, we will find that each voxel of the continuous model can also modeled as a well-mixed compart-ment, see Section 2.1. We will also describe perfusion estimates of the deconvolution method, if it is applied to only parts of the system.

In Section 6.2.2 we will show that the medical concept of perfusion is ill-suited for strongly coupled systems as the one described above. We show that for the above case recovered perfusion scales with the inverse voxel volume and thus extend the results
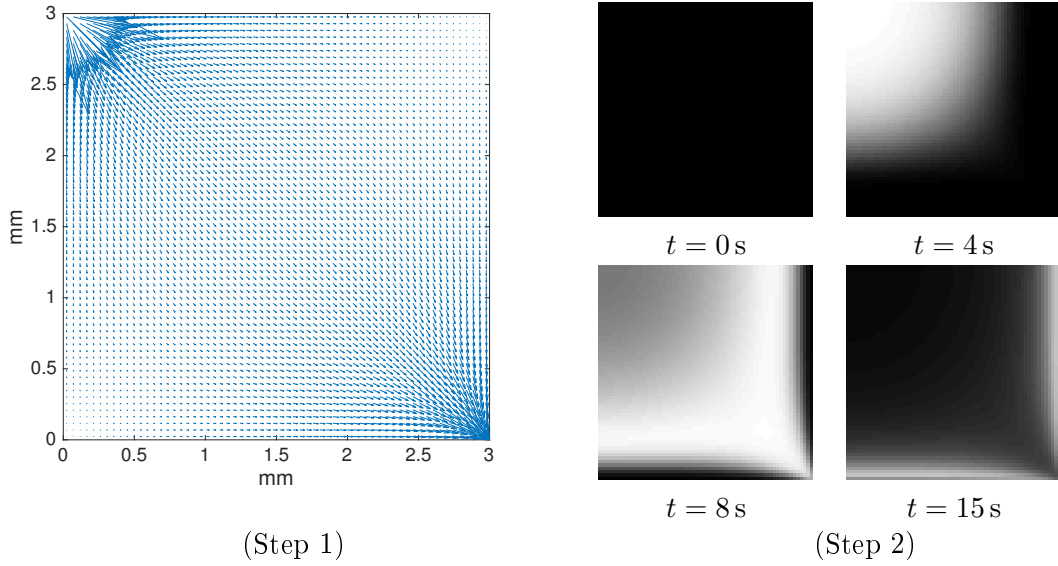
$t = 0\,\mathrm{s}$ $t = 4\,\mathrm{s}$

$t = 8\,\mathrm{s}$ $t = 15\,\mathrm{s}$

(Step 1)   (Step 2)

Figure 6.1: Displayed are the two steps for setting up the software phantom. First, a steady-state flow-field is simulated using Darcy's law and standard modeling for incompressible fluid. Second, contrast agent propagation through the flow-field from (Step 1) is simulated using porous media modeling.

of [72], where a generic argument for such overestimation is presented. Nevertheless we introduce a tailored notion of perfusion for the PDE model, which is in-line with the medical one. However, we predict overestimations of perfusion if the traditional models are applied to only parts of the system.

The experimental evaluation of this section can be found in Section 6.3.

6.1 A Synthetic Model for Capillary Flow

Structurally, the assumptions for the convolution model as well as the maximum slope model for CBF estimation presented in Section 2.1.4 are similar. The validity of both methods rely on a ROI having only one inlet and one outlet and that transition times are prescribed by some probability distribution. In fact, the assumption of one inlet and one outlet may easily be violated when we locally describe contrast agent propagation through a larger area with a highly developed capillary system. For this type of model system we expect instead a set of coupled equations where each voxel can be regarded as an inlet for surrounding voxels. Hence, in order to make a realistic synthetic model for capillary flow, we decided to describe the contrast agent propagation as a spatially coupled transport process, i.e. using partial differential equations (PDE) for transport. We will use this PDE field model for the validation of the traditional methods for perfusion estimation in such systems.

We start by introducing the modeling of a directional flow field in the simulated patch, see Figure 6.1 for such a flow-field. Within the porous capillary system we expect the blood flow to be driven mainly by pressure differences. We therefore set up a PDE flow model using the continuity equation as well Darcy's law [40]. Additional parameters represented in the model were blood viscosity and a possibly anisotropic permeability. For sake of argument, permeability was assumed to be isotropic and constant in our experiments. A major difference of the described flow-model model and traditional tracer kinetic modeling is the normalization of the flow field. In traditional tracer kinetic modeling normalized fluid flow is normalized with respect to volume and has thus units $[\mathrm{mm}^3/(\mathrm{s}\,\mathrm{mm}^3)]$ (*perfusion*). However, we will demonstrate in Section 6.2.2 that the concept of perfusion is discretization dependent. For the current flow simulations we instead used vector valued surface fluid flux $q = q(x)$ with units $[\mathrm{mm}^3/(\mathrm{s}\,\mathrm{mm}^2)]$, in agreement with geoscience and porous media simulation theory. The fluid flux is a vector field describing the volume of fluid per unit time flowing across a sliced unit area of the sample. A novel model to convert vector valued flux to scalar valued perfusion with units $[\mathrm{mm}^3/(\mathrm{s}\,\mathrm{mm}^3)]$ will be introduced in Section 6.2.2. Apart from the normalization with respect to surface, the assumptions of linearity and stationarity in the fluid flux are in complete agreement with standard pharmacokinetic modeling [137].

Having obtained a 3D directional flow field, we proceed to simulate contrast agent propagation proportional to the flow using linear transport equations. In the modeling of the transport we account for fractional blood volume as well as conservation of contrast agent mass. Since we assume that contrast agent propagates along the vessels, we decided to neglect diffusive effects.

Whereas propagation of contrast agent through major blood vessels based on Navier-Stokes equations has been performed earlier [26], our modeling additionally considers fractional cerebral blood volumes (porosities) different to one. Specifically, our modeling of the transport and fluid flow is in line with standard porous media theory [1]: We assume that within a tissue-patch $\Omega \subseteq \mathbb{R}^3$ not all areas are accessible for fluid. Specifically, we assume that Ω is partitioned by

$$\Phi : \Omega \to \{0, 1\}$$
$$x \mapsto \begin{cases} 1 & \text{if } x \text{ is accessible for fluid,} \\ 0 & \text{if } x \text{ is not accessible for fluid.} \end{cases}$$

Since our later modeling will be tailored to describe average concentrations on small subdomains (voxels) $V \subseteq \Omega$, we furthermore define the porosity for subsets of Ω,

$$\phi : \mathcal{P}(\Omega) \to [0, 1]$$
$$V \mapsto \int_V \Phi(x) \, \mathrm{d}x \ / \ \mathrm{Vol}(V). \tag{6.1}$$

We start by introducing the modeling of blood flow through a patch of tissue with a highly developed capillary system.

6.1.1 Modeling Capillary Blood Flow

For the time being, we will not consider the contrast agent concentrations, but rather the flow of fluid in general. The modeling will be tailored for porous media systems, where locally only fractions of the system are accessible for fluid.

Let $\Omega \subseteq \mathbb{R}^3$ be an open, bounded domain, through which a fluid with density $\rho : \Omega \times [0, T] \to \mathbb{R}$ [mg/mm^3] is propagated. We furthermore assume that fluid-transport follows a flow-field $\mathbf{q} : \Omega \to \mathbb{R}^3$ [mm^3/(s mm^2)]. We will describe the amount of fluid passing through an arbitrary control volume $\Omega_i \subseteq \Omega$ with porosity $\phi_{\Omega_i} := \phi(\Omega_i)$. Let $\rho_{\Omega_i}(t)$ [mg/mm^3] be the average density of the fluid in Ω_i at time point t. Assuming conservation of mass and a source/sink-term $\tilde{Q} : \Omega \to \mathbb{R}$ [mg/(mm^3 s)], where fluid is introduced into Ω_i, leads to the following modeling:

$$\int_{\Omega_i} (\phi_i \rho_i)_t \, \mathrm{d}x = - \int_{\partial \Omega_i} \langle \mathbf{q}\rho, \nu \rangle \, \mathrm{d}_\sigma x + \int_{\Omega_i} \tilde{Q} \, \mathrm{d}x,$$

where $\nu(x)$ is the outward unit normal of $\partial\Omega_i$. Assuming incompressibility of the fluid and stationarity of the flow-field $\mathbf{q}$ leads to

$$0 = - \int_{\partial \Omega_i} \langle \mathbf{q}, \nu \rangle \, \mathrm{d}_\sigma x + \int_{\Omega_i} Q \, \mathrm{d}x,$$

for the new source/sink term $Q := \tilde{Q}/\rho$ [mm^3/(s mm^3)], which is now independent of fluid an given as normalized volume fluid flow. Note that this equation is in local form consistent with

$$\left| \begin{array}{ll} \mathrm{div}(\mathbf{q}) = Q, & x \in \Omega \\ \langle \mathbf{q}, \nu \rangle = 0 & x \in \partial\Omega \end{array} \right| \tag{6.2}$$

where Q is a volume flux, only non-zero within the source or the sink locations. Note that Q is closely related to the perfusion P_Ω of the tissue, which has also units mm^3/(s mm^3). Since perfusion P_Ω by definition describes the amount of blood being delivered to Ω per unit time, we find that

$$P_\Omega = \frac{\int_{Q>0} Q \, \mathrm{d}x}{\mathrm{Vol}(\Omega)}.$$

For details on perfusion see Section 2.1.

Since (6.2) is ill-posed even for known Q, we now proceed to find further constraints to find an explicit expression for $\mathbf{q}$. Low velocity fluid flux in porous media is usually assumed to be driven by pressure differences and described by Darcy's law in the absence of gravitational forces [40]:

$$q = -\frac{\mathbf{k}}{\mu} \nabla p. \tag{6.3}$$

Here $\mathbf{k} : \Omega \to \mathbb{R}^{3,3}$ [mm^2] is the permeability tensor, $p : \Omega \to \mathbb{R}$ [Pa] is the pressure, and $\mu : \Omega \to \mathbb{R}$ [Pa s] is the viscosity of the fluid. For sake of argument we will assume that $\mathbf{k}(x)$ is a positive definite and diagonal tensor. Equations (6.2) and (6.3) can be combined, yielding the following elliptic partial differential equation in the pressure-field p,

$$\left| \begin{aligned} -\operatorname{div}\left(\frac{\mathbf{k}}{\mu}\nabla p\right) &= Q, & x \in \Omega \\ \langle p, \nu \rangle &= 0, & x \in \partial\Omega. \end{aligned} \right| \tag{6.4}$$

Here $p : \Omega \to \mathbb{R}^n$ is unknown, $\Omega \subseteq \mathbb{R}^n$ is a bounded domain, $\mathbf{k} : \Omega \to \in \mathbb{R}^{n,n}$ is a positive definite diagonal tensor field and $\mu : \Omega \to \mathbb{R}^+$ and $Q : \Omega \to \mathbb{R}$ are scalar functions. Note that the assumption $\mu > 0$ and $\mathbf{k}$ positive definite are ensuring that the flow is still following the negative gradient of p. Also note that the Neumann-Boundary condition was added to ensure an isolated system.

To illustrate (6.3), let us consider some properties of its solutions. From a physical perspective, one would expect that the total amount of fluid entering the system must equal the amount of fluid leaving the system, since Neumann-Boundary condition were assumed. The following Lemma verifies this guess and additionally shows, that we cannot expect a unique solution:

Lemma 20. Let $p \in C^1(\overline{\Omega})$ be a solution of (6.4). Then it holds that

1. $p(x) + c$ is a solution for any $c \in \mathbb{R}$

2. $\int_\Omega Q(x)\,\mathrm{d}x = 0$.

PROOF. Let $p \in C^1(\overline{\Omega})$ be a solution of (6.4).

1. Let $c \in \mathbb{R}$ be arbitrary and $\tilde{p}(x) := p(x) + c$. A straight-forward calculation shows that $\tilde{p}(x)$ is another solution of (6.4).

2. A direct application of the divergence theorem shows that

$$\int_\Omega Q\,\mathrm{d}x = \int_\Omega \operatorname{div}\left(\frac{\mathbf{k}}{\mu}\nabla p\right)\mathrm{d}x$$

$$= \int_{\partial\Omega} \left(\frac{\mathbf{k}}{\mu}\nabla p\right) \cdot \nu\,\mathrm{d}_\sigma x$$

$$= \int_{\partial\Omega} \frac{\mathbf{k}}{\mu} p_\nu\,\mathrm{d}_\sigma x = 0.$$

since $p_\nu = 0$ on $\partial\Omega$. $\qquad\square$
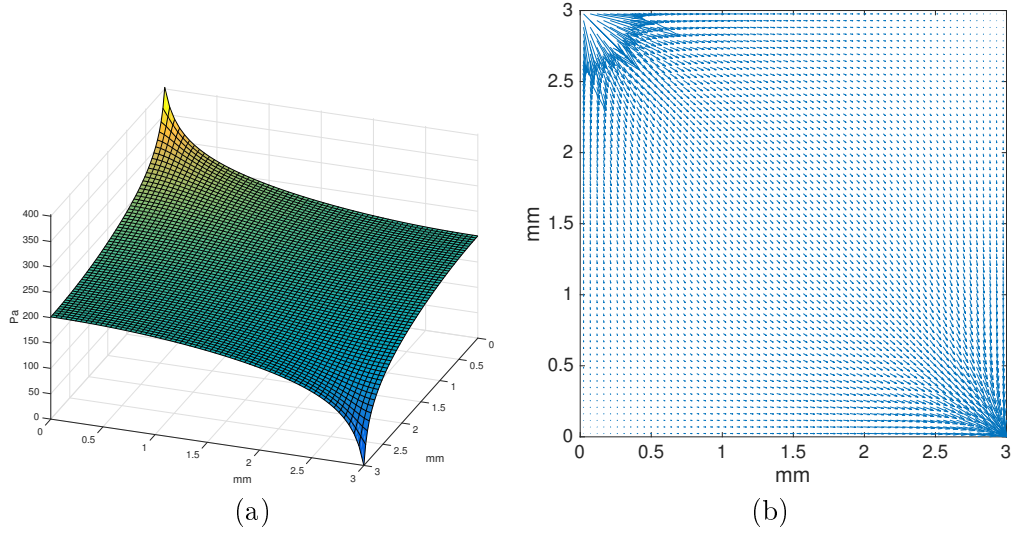
(a)   (b)

Figure 6.2: Porous media (PM) flow model with a source in the upper left corner and a sink in the lower right corner. (a) Pressure map p obtained by solving (6.4). (b) Flow $\mathbf{q} := -\mathbf{k}/\mu \nabla p$, given by (6.3).

Experiments

Equation (6.4) was solved numerically using two-point flux-approximation (TPFA), well known within porous media simulations [31, 50]. We simulated a tissue patch Ω with dimensions $3\,\mathrm{mm} \times 3\,\mathrm{mm} \times 1\,\mathrm{mm}$, within the same order as the characteristic length of the capillary bed or individual capillaries, ranging from $0.1\,\mathrm{mm}$ to $3\,\mathrm{mm}$ [34], or $0.25\,\mathrm{mm}$ to $850\,\mathrm{mm}$ [149]. The source term was assigned to the upper left voxel and the sink term was assigned to the lower right voxel. The source can be understood as the arterial compartment, the sink as the venous compartment, and the remaining field of view as the capillary system. We assumed that a total perfusion of $50\,\mathrm{ml/s/100ml}$ for Ω, which is within the average range reported for brain perfusion [108, 134]. Permeability was chosen to be isotropic and constant throughout the domain $\mathbf{k} = k\mathbf{I}$ for the identity $\mathbf{I}$ and $k = 5 \times 10^{-6}\,\mathrm{mm}^2$. Dynamic blood viscosity was chosen as $\mu = 5 \times 10^{-6}\,\mathrm{Pa\,s}$ according to [124]. Porosity (e.g. CBV) was assumed to be $\phi = 0.05$, in line with measured CBV of the brain [134]. Results for the pressure map p and the flow-field $\mathbf{q}$ are depicted in Figure 6.2

Conclusion

In this section we have derived a basic model for blood flow through a tissue patch with a highly developed capillary system. We have assumed that the patch can be modeled as a porous medium, where the flow is driven by pressure as expressed by Darcy's law. With the additional assumption of steady-state and incompressibility of the fluid, we

have shown that the flow-field follows the general modeling for incompressible fluid-flow with $\mathrm{div}(-\tilde{\mathbf{k}}\nabla p) = Q$, for a source/sink field Q and a pressure map p, although we were modeling flow in a porous medium. We now proceed to describe contrast agent propagation through the flow-field.

6.1.2 Modeling Indicator Dilution

In the previous section we have introduced a model which describes blood flow by fluxes and pressure fields. The introduced framework does not relate flow to propagation of a contrast agent. This section describes a model for contrast agent propagation in the tissue as it is dissolved in the evolving fluid. We assume that the contrast agent is entering the domain along with the fluid flowing in via the source, and similarly extracted at a sink. The resulting contrast agent concentration map is a simulation of the contrast agent concentration one would observe within real MRI measurements.

In order to define meaningful and continuous contrast agent concentrations, we first describe the average contrast agent concentration in an (arbitrarily) small tissue patch $\Omega_i \subseteq \Omega$. Assume that V_i is the volume of Ω_i and v_i the blood volume within Ω_i. By definition (6.1), porosity reflects the relative space within the vascular system, and is given by $\phi_i = v_i/V_i$. Let $C_i(t)$ denote the contrast agent concentration in Ω_i with respect to the whole volume V_i at time point t. The contrast agent concentration with respect to the blood volume v_i is denoted by $c_i(t)$. From the definition of c_i, C_i and ϕ_i we obtain the relation $C_i(t) = \phi_i c_i(t)$. Since we expect mainly transport and marginal diffusion, the change in tracer mass within Ω_i occurs only from advective flow and the source and sink field Q. Let us write the source- and the sink term as $Q = Q_{\mathrm{si}} + Q_{\mathrm{so}}$ where $Q_{\mathrm{si}} < 0$ is the sink and $Q_{\mathrm{so}} > 0$ is the source, and zero elsewhere. Both are assumed to be zero everywhere except at in the respective source and sink locations. Note that $\int_\Omega Q\,\mathrm{d}x = 0$, cf. Lemma 20. The change in contrast agent at time point t from fluid entering the control volume can be written as

$$-\int_{\partial\Omega_i} c_i(t)(q_i \cdot n)\,\mathrm{d}s + \int_{\Omega_i} c_{\mathrm{a}}(t)Q_{so,i}\,\mathrm{d}x + \int_{\Omega_i} c_i(t)Q_{si,i}\,\mathrm{d}x, \tag{6.5}$$

where n is the outward unit normal on $\partial\Omega_i$. In standard pharmacokinetic modeling, c_{a} is referred to as the arterial input function (AIF). From the principle of conservation of tracer molecules it must hold that

$$\int_{\Omega_i} \phi_i \frac{dc_i}{dt}\,\mathrm{d}x + \int_{\partial\Omega_i} c(t)(q_i \cdot n)\,\mathrm{d}s$$
$$= \int_{\Omega_i} c_{\mathrm{a}}(t)Q_{so,i}\,\mathrm{d}x + \int_{\Omega_i} c_i(t)Q_{si,i}\,\mathrm{d}x. \tag{6.6}$$

Now, let the contrast agent concentrations, porosity, volume fluxes, and surface flux be continuous functions of space and time. Letting the $|\Omega_i| \to 0$ shows that equation
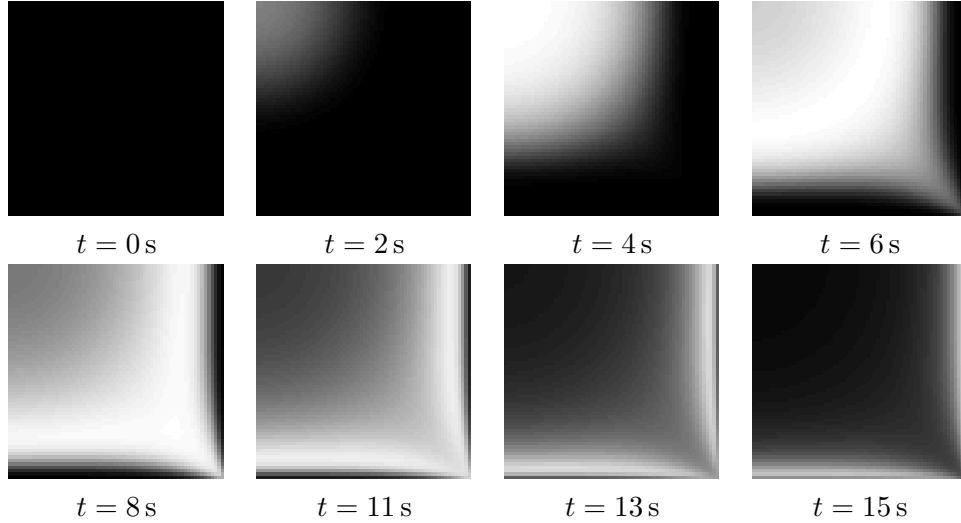
Figure 6.3: Figure showing simulated contrast agent propagation through the flow field **q** using (6.6). The actual simulation was carried for 30 s, out until all contrast agent had left the system.

(6.6) is consistent with the continuity equation on local form

$$
\left| \begin{aligned} \phi\frac{\partial c}{\partial t} + \nabla \cdot (cq) &= c_{\mathrm{a}}Q_{\mathrm{so}} + cQ_{\mathrm{si}} && x \in \Omega, \ t > 0, \\ c(x,t) &= 0 && x \in \Omega, \ t = 0. \end{aligned} \right| \tag{6.7}
$$

where we also added the initial value $c(x,0) = 0$ to ensure existence. Equation (6.7) is a linear transport equation in $c(x,t)$. Following [49], equation (6.7) admits a unique local solution.

Experiments

The transport of contrast agent described in (6.6) was implemented using first order upwinding [60, 89], yielding a discrete 2D+time contrast agent concentration map $C(x_i, t_j)$. For arterial input, we chose a standard function [111], the gamma-variate function $c_{\mathrm{a}}(t) := D_0(t-t_0)^\alpha e^{-(t-t_0)/\beta}$ for default parameters $\alpha = 3$, $D_0 = 1\,\mathrm{mmol/(1s)}$, $\beta = 1.5\,\mathrm{s}$ and $t_0 = 0\,\mathrm{s}$. Porosity (e.g. CBV) was assumed to be constant $\phi = 0.05$, in line with measured CBV of the brain [134]. Snapshots of the simulated contrast agent propagation can be found in Figure 6.3

Conclusion

Based on the assumption that pure capillary system can be modeled as a porous medium, we have introduced a model which describes contrast agent propagation

through such a system. Whereas a similar modeling has been applied to simulate contrast agent flow through the aorta, we additionally consider porosities other than one and thereby generalize the approach taken in [26]. Although our experiments made use of simple parameters, the modeling can be extended to a more complex structure easily. Since the analysis in the next section will be mostly independent of final structure of the model, we will confine to the simple setting described above. In the next section we will proceed to relate the novel model for blood flow with the traditional models, as e.g. the Meier-Zierler model.

6.2 Relating the Synthetic Model and Traditional Models

In this section we will describe connections between traditional models for perfusion and the novel continuous model which was introduced in the previous section. We start by showing that a discretized version of (6.7) can equivalently be described by coupled versions of traditional models for flow through a well-mixed compartment (see Section 2.1.3). Based on this understanding we will analyze which CBF results are to be expected if traditional deconvolution models are used on local patches of the PDE simulation. Here we will describe two cases: CBF results of the deconvolution model using the global AIF of the patch and CBF results of the deconvolution model using a local AIF.

We then proceed to discuss conceptual differences of physical flux and medical perfusion. We will demonstrate that the medical notion of perfusion is ill-suited for coupled systems, since normalization of the flow is performed with respect to volume and not to surface. Furthermore we will introduce a ground-truth for perfusion for coupled systems, which is in-line with the medical understanding of perfusion.

We conclude by showing that traditional models are accurately describing the blood volume in the complete system but will yield unrealistically high results in coupled systems.

6.2.1 Relating the PDE Model with the Meier-Zierler Model

In this section we will describe how the continuous model is related to the traditional deconvolution model. More specifically, we will show that in the continuous model each voxel can be described by a traditional model with arterial input determined by the adjacent upstream voxels. Additionally, we will consider the effect of deconvolving a voxel curve C_i of the continuous model with the global arterial input function c_a of the tissue. We will show that in this case the resulting residue function will depend on the perfusion of all upstream voxels. Note that this effect makes the application of standard deconvolution techniques difficult in coupled systems, since perfusion results will depend on the flow in feeding tissue.

Relating the Continuous Model with Traditional Models I

Let us start by modeling the contrast agent concentration in a given voxel using traditional models. This section will perform the modeling for continuous variables. To further illustrate the derivation, the next section will perform the same modeling for a discrete system. Let us consider the continuity equation for one voxel (6.6). For sake of simplicity let us assume that $Q_{\mathrm{so},i} = Q_{\mathrm{si},i} = 0$. Note that it is possible to extend the following approach also to voxels where $Q_{\mathrm{so},i} \neq 0$ or $Q_{\mathrm{si},i} \neq 0$. We will now show that (6.6) can equivalently be described as a traditional one-dimensional equation for a well-mixed compartment. First, let us define the areas of inflow and outflow over the boundary by $S_{\mathrm{in}} := \{x \in \partial\Omega_i : \ q_i(x) \cdot n(x) < 0\}$ and $S_{\mathrm{out}} := \{x \in \partial\Omega_i : \ q_i(x) \cdot n(x) > 0\}$ respectively. In order to define a single arterial input rather than an arterial input which depends on the location, we define c_{in} as a weighted average of the concentrations at the boundary:

$$c_{\mathrm{in}}(t) := \frac{\int_{S_{\mathrm{in}}} c(t)(q_i \cdot n)\,\mathrm{d}s}{\int_{S_{\mathrm{in}}} q_i \cdot n\,\mathrm{d}s} \tag{6.8}$$

We now define the perfusion of the voxel Ω_i as follows:

$$P_{\mathrm{in}} := \frac{-1}{\mathrm{Vol}(\Omega_i)} \int_{S_{\mathrm{in}}} q_i \cdot n\,\mathrm{d}s, \quad P_{\mathrm{out}} := \frac{1}{\mathrm{Vol}(\Omega_i)} \int_{S_{\mathrm{out}}} q_i \cdot n\,\mathrm{d}s \tag{6.9}$$

Let us now assume that the flow-field q is divergence-free and the amount of fluid entering the region is the same as the amount of fluid leaving it. Then it holds that $P_{\mathrm{in}} = P_{\mathrm{out}}$ and hence equation (6.6) can be reformulated as

$$(\phi_i c_i)'(t) = P_{\mathrm{in}}(c_{\mathrm{in}}(t) - c_{\mathrm{out}}(t)). \tag{6.10}$$

Note that in the proposed upwinding discretization described in Section 6.1.2, it is assumed that the contrast agent concentration at the outflow S_{out} equals the concentration within the voxel and that the concentration at the inflow is the concentration at the adjacent voxels. In this case In this case it follows that (6.10) reduces to standard equation for a well-mixed compartment $C_i'(t) = P_{\mathrm{in}}(c_{\mathrm{in}}(t) - c_i(t))$ with solution

$$C_i(t) = (I_i * c_{\mathrm{in}})(t) \text{ for } I_i(t) = P_{\mathrm{in}}e^{-P_{\mathrm{in}}/\phi_i t}. \tag{6.11}$$

The arterial input c_{in} is determined by (6.8), which recursively depends on all upstream voxels until the global arterial input is reached.

Relating the Continuous Model with Traditional Models II

To further illustrate this derivation, let us now derive (6.10) for the discrete setting. We assume that voxel i is fed by upstream, adjacent neighbors $j \in J$ with absolute flow p_j [ml/s] from voxel j to voxel i and that for each voxel it holds that $\sum_{j \in J} p_j = 0$ Also

assume that the contrast agent flow is proportional to the upstream concentration, which is in-line with the upwind discretization used in our experiments. In this case it follows that

$$\begin{aligned}(\phi_i c_i)'(t) &= \frac{1}{\text{Vol}(\Omega_i)}\left(\sum_{j\in J} p_j c_j(t) - \left(\sum_{j\in J} p_j\right)c_i(t)\right)\\ &= \sum_{j\in J} P_j c_j(t) - P c_i(t)\\ &= P(\tfrac{1}{P}\sum_{j\in J} P_j c_j(t) - c_i(t))\\ &= P(c_{\text{in}}(t) - c_i(t))\end{aligned}$$

Here $P_j := p_j/\text{Vol}(\Omega_i)$ the corresponding perfusion for feeding voxel j and $P := \sum_{j\in J} P_j$ the total perfusion for the voxel and

$$c_{\text{in}}(t) := \tfrac{1}{P}\sum_{j\in J} P_j c_j(t)$$

is the local arterial input for voxel i.

Numerical Verification

To verify relationship between the discretized continuous and the traditional models numerically, we simulated a tissue curve $C_i(t)$ using both the continuous PDE model as well as analytical recursive convolution by (6.11). We refer to the latter approach as *local convolution*. The two curves have an almost perfect match, as seen in Figure 6.4 (left).

As a direct consequence, it follows by recursion that the concentration at a voxel i can be written as a convolution of the (global) arterial input function with a weighted average of all upstream impulse response functions. Deconvolving a tissue concentration C_i with the global AIF will yield an impulse response function which depends not only on the local flow and porosity, but on flow and porosity of all upstream voxels. This relationship was also confirmed experimentally: Figure 6.4 (right) shows the impulse response function determined by analytical recursive convolution and the numerically achieved impulse response function obtained from deconvolving a tissue curve of the continuous model with the global arterial input function. The simulation was performed at location (1,20) of the software phantom. The two curves coincide almost perfectly, highlighting the validity of the established theory.

These results show that the PDE model and the convolution model are equal in terms of local, voxelwise flow estimates if the convolution model is applied with the local
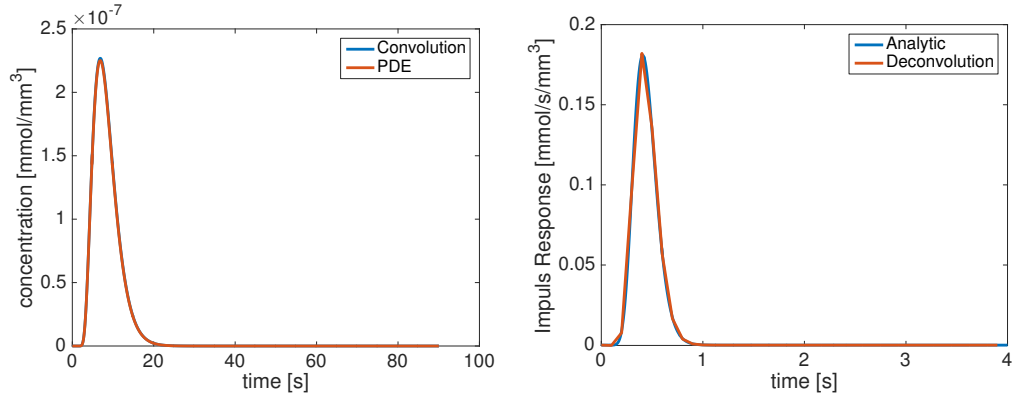
Figure 6.4: Left: Red curve shows the tissue concentrations (C) of the continuous PDE
model at location [32,35]. Blue curve shows recursive convolution by (6.11)
with experimental value of $P_v = 5328\,\text{ml/s/100ml}$ at the given location
and c_{in} taken locally from upstream voxels around the simulated voxel. The
two curves have an almost perfect overlap. Note that the numbers used for
the perfusion is unrealisticely high since normalization is performed with
respect to the volume of only one voxel. Right: Red curve shows the com-
puted impulse response functions (IR) at location [1,20] using the global
arterial input function. Blue curve shows the analytic impulse response
function given by a convolution over all upstream flow. The two curves
have an almost perfect overlap. These numerical experiments support that
the computed impulse response function by traditional methods is not the
directly feeding impulse response function, but rather a recursive impulse
response function depending on all upstream voxels.

arterial input. Also, the impulse response function obtained by convolution of the
global arterial input function is identical to an analytical recursive convolution along
all upstream voxels. This clearly demonstrates that the perfusion which is recovered by
traditional models will depend on all upstream flow. However, for meaningful interpre-
tation of the perfusion the entire streamline length within the capillary system needs
to be taken into consideration, where the blood is providing the tissue with nutrients
and oxygen.

6.2.2 Converting Flow to Perfusion

The model described in (6.4) uniquely determines the flux field $q(x)$. However, in
pharmacokinetic modeling the parameter of interest is usually the CBF, which we will
denote by $P(x)$ as the voxel wise field of perfusion. The surface flux and perfusion are
physically distinct, and there are at least two differences between $q(x)$ and $P(x)$. First,
the flux is a vector field and the perfusion is a scalar field. Second, the flux is normalized

to a surface area and the perfusion is normalized to a volume. Thus, the surface flux and the perfusion are strictly, mathematically different but still conceptually related. In the following we describe a method for converting flux into perfusion, motivated by the need to compare the ground-truth flux field to the scalar valued perfusion field obtained by traditional methods.

The common understanding of perfusion or volume flux $P(x)$ is the amount of blood feeding a tissue volume per unit time, with units $[\mathrm{mm}^3/(\mathrm{s}\,\mathrm{mm}^3)]$. One obvious approach for converting flux into perfusion could be to estimate the perfusion as the total inflow (or outflow) of fluid (e.g. arterial blood) into a control region per unit time, and then normalizing with the control region volume. This is a valid approach only if the control regions are not feeding each other, and is therefore well-founded for the entire organ. Such understanding of perfusion is in line with the theoretical foundation of traditional compartment models for perfusion where a control region has its own source of feeding arterial blood, independent of the neighbor regions.

On the other hand, if the control region is a single voxel or a sub-division of an organ with sequentially feeding arterial blood, the traditional model assumptions are violated since every control region will feed its neighbours, thus becoming a coupled system of flow. Simply summing the total inflow into a voxel and dividing by the voxel volume will strongly over-estimate the perfusion since the normalization would refer to the wrong volume. This phenomenon is demonstrated in Figure 6.5 where the volume on the left has the true perfusion of $P_1 = F_0/(2V)$ for an incoming flow F_0 $[\mathrm{mm}^3/(\mathrm{s}\,\mathrm{mm}^3)]$ and distribution volume $2V$ $[\mathrm{mm}^3]$. However, for another discretization as shown in the middle, the perfusion within each of these sub-volumes becomes $P_2 = F_0/V = 2P_1$. Taking the average across the two sub-volumes, it is clear that the perfusion is over-estimated with a factor of two. A discretization dependent perfusion estimate is not recommendable, and the perfusion estimate of P_2 is clearly wrong, see also [72].

In the following paragraph will introduce a meaningful notion of perfusion for the continuous model. To do this, we will consider distribution volumes which are following the streamlines. For each point of a streamline we will select a small perpendicular disk with radius chosen in such a way that the total flow over each disk is constant along the streamline. These disks will form a small tube with varying radii around the streamline with constant flow over each cross-section. Since the total flow over each disk is constant, we can define the perfusion for each voxel by the traditional model, obtaining a perfusion value which is constant along the streamline. To obtain a truly local perfusion, we will let the radii of the disks go to zero.

More precisely, let us consider an arbitrary streamline $S \subseteq \Omega \subseteq \mathbb{R}^3$ of length $l > 0$, parametrization $s : [0, l] \to \Omega$ and let us additionally assume that $Q = 0$. We start by calculating the total flow over a small 2-D disk which is perpendicular to the streamline. Let $y \in S$ be an arbitrary location along the streamline. The total flow over a 2-D disk
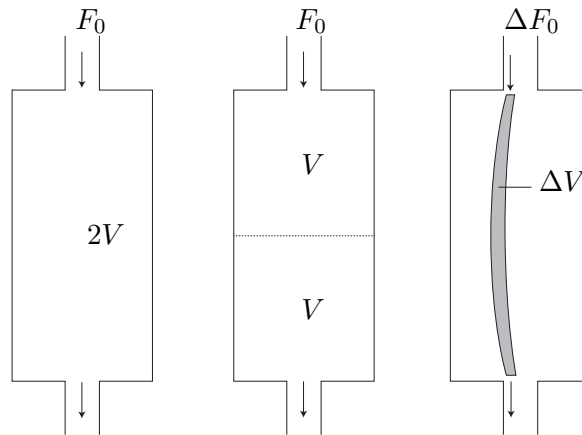
Figure 6.5: Perfusion within a small volume. Left: A compartment with volume $2V$ is exposed to a flow F_0 [mm^3/(s mm^3)] of fluid. By definition, the perfusion within this compartment becomes $P_1 = F_0/(2V)$. Middle: The same volume is divided into two compartments (e.g. voxels), and the perfusion for each of the compartments becomes $P_2 = F_0/V = 2P_1$. The discrepancy between the two discretizations occurs because the flow is counted twice as it is fed from one voxel to the other. Right: As a solution to the described problem we rather pick out a true distribution volume ΔV (area in this 2D sketch), which is a small area around a given streamline along the centre line of the grey area. This is the true distribution volume (area in this 2D sketch) which is fed with arterial blood from the incoming fractional flow ΔF_0. The correct perfusion within ΔV is therefore $\Delta F_0/\Delta V$. The entire compartment can further be divided into similar infinitesimal distribution volumes, thus providing locally correct perfusion estimates.

$B_r(y)$ perpendicular to the flow-field $q(y)$ is given by

$$F(y, r) = \int_{B_r(y)} q(x) \cdot \nu \, \mathrm{d}x \text{ where } \nu := q(y)/|q(y)|.$$

In order to calculate the perfusion, we need to establish the volume of a small tube around the streamline. We will not consider a tube with constant radius, but one with spatially varying radii $r : S \to \mathbb{R}^+$. The total volume of such a tube surrounding the streamline is given by

$$V(r) = \int_0^l r(s(u))^2 \pi \, \mathrm{d}u$$

We define the perfusion at an arbitrary point $y \in S$ by

$$P(y) := \lim_{\varepsilon \to 0} \frac{F(y, \varepsilon r(y))}{V(\varepsilon r)} \text{ for } r(y) := 1/\sqrt{|q(y)|}.$$

Note that the radii r are chosen in such a way that in the limit $\varepsilon \to 0$ the total flow is constant along the streamline. To see this, let us assume that q is differentiable with Jacobian J. Using the Taylor expansion of $q(x) = q(y) + J(\xi)(x-y)$ for some $\xi \in (0, x)$ as well as a change of coordinates $z = (x - y)/(\varepsilon u)$ yields for $u := r(y)$

$$\begin{aligned}
F(y, \varepsilon u) &= \int_{B_{\varepsilon u}(y)} \nu^\top q(x) \, \mathrm{d}x \\
&= \int_{B_{\varepsilon u}(y)} \nu^\top q(y) + \nu^\top J(\xi)(x - y) \, \mathrm{d}x \\
&= \varepsilon^2 \pi + \int_{B_1(0)} \nu^\top J(\zeta) z(u\varepsilon)^3 \, \mathrm{d}x \\
&= \varepsilon^2 \left(\pi + \varepsilon u^3 \int_{B_1(0)} \nu^\top J(\zeta) \, \mathrm{d}x \right)
\end{aligned}$$

where $\xi \in (0, x)$, $\zeta := (\xi - y)/(\varepsilon u)$ and simplifications are due to $u = 1/\sqrt{|q(y)|}$ and $\nu := q(y)/|q(y)|$. Note that since $V(\varepsilon r) = \varepsilon^2 V(r)$ it follows that

$$P(y) = \frac{1}{\int_0^l r(s(v))^2 \, \mathrm{d}v} \tag{6.12}$$

Equation (6.12) is independent of the spatial location y and an explicit formula for converting flux into perfusion.

We have simulated comparison of local perfusion (6.9) and perfusion along the streamlines (6.12) is performed in Figure 6.6. Streamlines were recovered from the flow-field **q** using a simple explicit tracking, which is also used in tractography for diffusion tensor imaging [106]. It can be observed that local perfusion varies with discretization and thus yields unrealistically high values. As (6.11) shows, this can nevertheless be regarded a valid definition of perfusion since it models the feeding of arterial blood in the voxel.
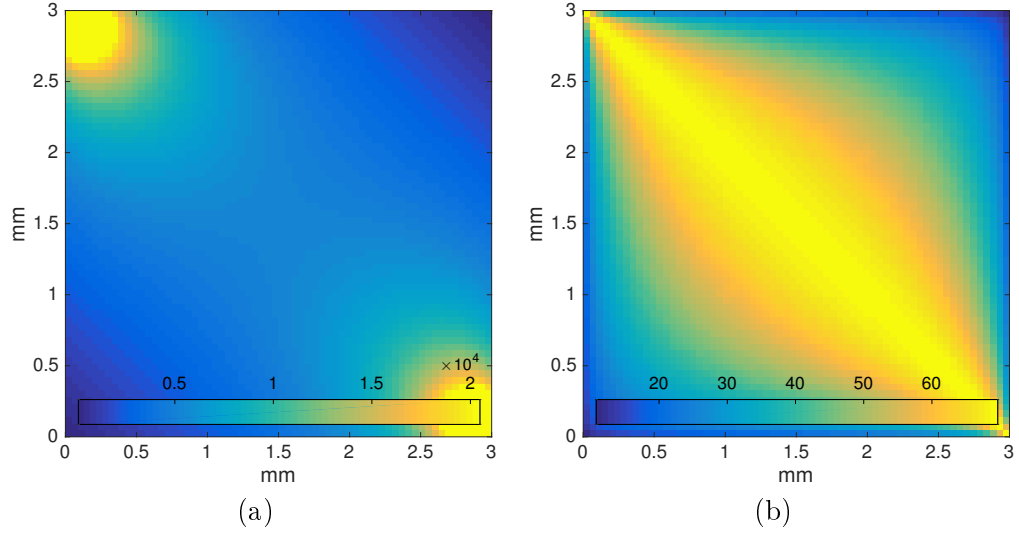
Figure 6.6: Comparison of different concepts of perfusion for the synthetic model. (a) Local (voxel-wise) perfusion according to (6.9). (b) Perfusion along the streamlines according to (6.12). It can be seen that the voxel-wise definition yields unrealistically high values since normalization is performed with respect to the wrong distribution volume.

6.2.3 A Method to Estimate Local Porosity

Porosity and CBV have the same definition, and we can therefore state that $\phi \equiv \text{CBV}$. It is known from literature on traditional models [137] for perfusion that CBV for the entire compartment can be expressed as

$$\phi = \frac{\int_0^\infty C(s)\,\mathrm{d}s}{\int_0^\infty c_\mathrm{a}(s)\,\mathrm{d}s}. \tag{6.13}$$

It is not obvious that (6.13) is valid also for a one-compartment field model where the voxels are feeding each other. We will now show that this is indeed the case.

Let us switch to a discrete setting let the following assumptions hold (see Section 6.2.1 for details):

(A1) The concentration in any voxel can be written as

$$\int_0^\infty c_i(s)\,\mathrm{d}s = \int_0^\infty (J_i * c_{\mathrm{in},i})(s)\,\mathrm{d}s. \tag{6.14}$$

and $J_i, c_{\mathrm{in},i} : [0,\infty) \to \mathbb{R}$ is specified in (A2).

(A2) In (A1) $J_i : [0,\infty) \to \mathbb{R}$ fulfills $\int_0^\infty J_i(s)\,\mathrm{d}s = 1$ (straight forward calculation) and the local arterial input $c_{\mathrm{in},i}(t)$ is a weighted average of adjacent, upstream

voxels $c_j(t)$, where J is the set of adjcacent, upstream voxels and $c_a(t)$, if the voxel lies in the source field:

$$c_{in,i}(t) = \tfrac{1}{P}\Big(p_0 c_a(t) + \sum_{j\in J} p_j c_j(t)\Big), \tag{6.15}$$

where $p = p_0 + \sum_{j\in J} p_j$ and $p_0 \neq 0$ if voxel i lies within the source/sink field. Consistent with (6.8) p_j is absolute flow.

(A3) **q** is a uni-directional flow field across each voxel face and there is no flow interaction between voxels in the same layer.

Lemma 21. Assume (A1)-(A3). Then $\int_0^\infty C_i(s)\,\mathrm{d}s = \phi_i \int_0^\infty c_a(s)\,\mathrm{d}s$ for all voxels.

PROOF. We use induction to show that $\int_0^\infty c_i(s)\,\mathrm{d}s = \int_0^\infty c_a(s)\,\mathrm{d}s$, then the claim follows. Let I_k denote the set of voxels which have k layers of upstream voxels. E.g. I_0 is the set of all voxels, which have no upstream voxels, I_1 is the set of voxels which are fed by I_0 and so on. Induction will be carried out over k.
Induction Basis:
Let $k = 0$ and let $i \in I_0$ be arbitrary. Following (A1) and (A2) it holds that $c_i(t) = (J_i * c_a)(t)$. Since the area under the convolution of two functions equals the product of the area of its factors, $\int_0^\infty c_i(s)\,\mathrm{d}s = \int_0^\infty c_a(s)\,\mathrm{d}s$ and the claim follows.
Induction Assumption: The claim holds for arbitrary $k \in \mathbb{N}$.
Induction Step:
For any voxel at location $i \in I_{k+1}$ which has the voxels $J \subseteq I_k$ as their upstream neighbors, we find the following expression:

$$
\begin{aligned}
\int_0^\infty c_i(s)\,\mathrm{d}s &\stackrel{(A1)}{=} \int_0^\infty (J_i * c_{in,i})(s)\,\mathrm{d}s \\
&= \int_0^\infty c_{in,i}(s)\,\mathrm{d}s \\
&\stackrel{(A2)}{=} \tfrac{1}{P}\int_0^\infty \Big(p_0 c_a(s) + \sum_{j\in J} p_j c_j(s)\Big)\,\mathrm{d}s, \\
&= \tfrac{1}{P}\Big(p_0 \int_0^\infty c_a(s)\,\mathrm{d}s + \sum_{j\in J} p_j \int_0^\infty c_j(s)\,\mathrm{d}s\Big), \\
&\stackrel{\substack{\text{I.A.,}\\(A3)}}{=} \tfrac{1}{P}\Big(p_0 + \sum_{j\in 1} p_j\Big) \int_0^\infty c_a(s)\,\mathrm{d}s, \\
&= \int_0^\infty c_a(s)\,\mathrm{d}s.
\end{aligned}
$$

Hence the claim (6.13) follows. $\qquad\square$

6.2.4 Discussion

In this section we have related traditional one compartment models with the novel synthetic model for contrast agent propagation in a tissue patch with highly developed capillary system. Equation (6.11) shows, that for each voxel Ω_i the contrast agent concentration can be expressed as

$$C_i(t) = (I_i * c_{\text{in}})(t) \text{ for } I_i(t) = P_{\text{in}} e^{-P_{\text{in}}/\phi_i t}.$$

In this equation the local arterial input function $c_{\text{in}}(t)$ is a weighted average of the adjacent, upstream voxels and (6.9) shows that local perfusion is given by

$$P_{\text{in}} = \frac{p}{\text{Vol}(\Omega_i)}.$$

where p [ml/s] describes the absolute inflow into Ω_i. These two relationships can be interpreted as follows: First, they show that if C_i is deconvolved with the (correct) local arterial input function, we expect to recover perfusion P_{in}, which is highly scale-dependent. Second, a recursive application of the first relationship shows, that $c_{\text{in}} = (J_1 * \cdots * J_l) * c_{\text{a}}$, where J_i are impulse-response functions of all upstream voxels. This means that we expect dispersion of the arterial input, which can cause underestimation of the local perfusion P_{in} [26].

We have additionally argued that the scale dependency is caused by the normalization with the wrong distribution volume. To cope with this problem, we have proposed an alternative definition of perfusion for coupled system which normalizes flow with respect to the streamline. However, note that such a definition is impractical for most clinical applications, since a directional flow field is often unknown.

In the next section we will experimentally validate, how the two effects –namely under- and overestimation of the true perfusion– relate.

6.3 Results: Reconstruction of Perfusion on Synthetic Data

In this section we will demonstrate experimentally that local application of traditional models for perfusion will cause a scale dependent overestimation of perfusion. We start by outlining the detailed setup of the software phantom we used to carry out the experiments.

6.3.1 Experiments on Phantom Data

We tested the convolution based traditional model (bSVD) (2.2) as well as maximum-slope (MS) model (2.8) for their capability to recover perfusion for the continuous phantom described in Section 6.1. Prior to reconstruction, the contrast agent concentration map $C(x_i, t_j)$ was downsampled to a time-resolution of 0.1 s. In order to simulate different spatial resolutions of the scanning process, the data was averaged using different block-sizes ranging from $(1, 1)$ pixel (i.e. same resolution as the PDE model) to $(64, 64)$ pixels (i.e. entire domain). Success of restoration was measured in terms of averaged relative error of the recovered perfusion with respect to the ground truth perfusion, $RE := (P_{\text{rec}} - P_{\text{true}})/P_{\text{true}} \cdot 100\%$. The recovered perfusion P_{rec} were compared against the two perfusion maps $P_{\text{true}} = \{P_{\text{v}}, P_{\text{s}}\}$ depicted in Figure 6.6.

The local perfusion map P_{v} was set up according to (6.9). Since normalization is performed with respect to voxel size, the values are unrealistically high and will vary with the discretization. As (6.11) shows, this can nevertheless be regarded a valid definition of perfusion since it models the feeding of arterial blood to a control region.

The global perfusion map P_{s} was set up using the definition along the streamlines (6.12). This definition most accurately reflects the physical perfusion at a given location and shows plausible perfusion values, cf. Figure 6.6. As an internal control, the average of P_{v} was found to be 49.59 ml/s/100ml, for all practical means identical to the global input perfusion of 50 ml/s/100ml. However, we do not expect the traditional models to be able to recover these values either. To quantify the errors occurring by traditional methods, the global arterial input function was used for the deconvolution, as measured in the source.

Results from deconvolution by traditional methods are displayed in Table 6.1. For the complete domain (i.e. block size 64×64), both the MS method as well as the convolution method were able to restore the ground truth perfusion of 50 ml/s/100ml accurately with errors of $< 1\%$ and $< 4\%$ respectively. However, the errors are increasing as methods are applied to smaller blocks of the system. If compared to P_{v}, one can see that results are improving with increasing block size. Note that the block size of $(0.5, 0.5)$mm is within the range of resolution available on clinical scanners today, and is therefore clinically interesting. Also a clear advantage of the bSVD method as compared to MS can be observed for larger block sizes.

Results from reconstructing the porosity ϕ (i.e. CBV) according to (6.13) are also shown in Table 6.1. The errors are low, independent of block size.

6.3.2 Discussion

We have studied the accuracy of traditional one compartment models for perfusion reconstruction in a coupled system of blood flow in the capillary system. To establish

Table 6.1: Mean relative error RE and standard deviation (both in %) of reconstructed perfusion compared to the ground truth values P_v, CBV, and P_s from the digital phantom. Both reconstruction models MS and bSVD are able to restore the perfusion for the entire domain, but fail when dividing the domain into smaller block sizes. For larger block sizes the bSVD model restores the perfusion more accurately than the MS model. However, the blood volume ϕ is recovered accurately, independent of block size. P_s is only defined within a coupled system having streamlines and can therefore not be compared with restored perfusion for the entire domain.

		block size (mm)			
		(0.05,0.05)	(0.23,0.23)	(0.47,0.47)	all
P_v	bSVD	-93%±4%	-67%±16%	-50%±23%	4%
	MS	-98%±2%	-88%±6%	-79%±11%	<1%
CBV		<1%	<1%	<1%	<1%
P_s	bSVD	753% ± 926%	650% ± 757%	476% ± 507%	
	MS	124% ± 79%	114% ± 66%	103% ± 51%	

ground truth values, we developed a PDE based digital phantom to simulate blood flow as porous media flow within a slab of capillary tissue.

Our results strongly support the usage of traditional models for entire regions which are exclusively fed by the measured arterial input. However, they also show that if traditional models are applied only to parts of the system, they tend to overestimate the actual perfusion. Although there is awareness in the community of such effects [72], studies reporting voxel wise perfusion maps without discussing their possible limitations are continuously published [104, 79]. Thus, a major motivation for our study is to stimulate the awareness around this topic and to push the development of more appropriate models for future applications.

There are at least two issues related to the overestimation of perfusion. The first issue is that blood passing through a voxel without being locally delivered to the capillary tissue will contribute to artificially high perfusion values. The second issue is thoroughly studied in this work, and relates to estimation of the correct distribution volume used for computing the perfusion. As soon as there is dependency of capillary flow between adjacent voxels, the correct distribution volume used for normalizing the absolute flow into perfusion (i.e. ml/s/100ml) is not known and over-estimation of perfusion will occur. Using local arterial input functions is no remedy for this problem, since the resulting perfusion will depend heavily on the voxel size and overestimate the actual flow, cf. Figure 6.5 and (6.11).

The results from the digital phantom are supported by real data experiments, where we showed local overestimation of perfusion for small voxel-sizes as compared to an averaging of concentrations for the entire volume of interest. Regarding the CBV estimates, one can observe from Table 6.1 that estimation of blood volume is far more stable, and even various block sizes had little impact on the results. These results

are in well agreement with the analytical considerations in Section 6.2.3, stating that (6.13) is valid for entire organs as well as for single voxels. Thus, these results support the usage of (6.13) for computing the CBV with high accuracy for any type of block size, including single voxels.

Furthermore we have introduced two theoretical definitions of voxelwise perfusion. The perfusion P_s models perfusion along the streamlines and most accurately reflects the physical notion of volume flow within the correct distribution volume according to mathematical definitions. We showed that P_s is a global quantity along the streamline, and scales with streamline length and geometry. Theory and experiments show that the traditional models cannot recover this perfusion. The usage of P_s for reconstruction of perfusion in real data might as well be challenging as the entire geometry and microscopical flow patterns would have to be known to track the streamlines. However, for our purpose, the concept of P_s was useful to clarify the definition of perfusion as a flow that must be normalized along its entire capillary length, where the blood undergoes a transition from arterial to venous blood. For future developments of field models, multi compartment models as suggested in [140] might be more applicable, where the perfusion was suggested as the non-zero divergence of the arterial flux.

Perfusion P_v was set up based on the interpretation of a coupled system between adjacent voxels. Theory and examples show that this definition of perfusion does not comply with the physical understanding of perfusion since it depends heavily on the discretization. However, we have shown that traditional models would restore this local value if the local arterial input function was selected. We have additionally analyzed, both analytically and experimentally, the impact of selecting a further upstream arterial input function. Specifically, we have justified that traditional perfusion measurements based on convolution will identify the recursive impulse response function for all upstream voxels (see Section 6.2.1). Locally, the correct distribution volume is not accounted for and the obtained perfusion will be overestimated compared to the actual perfusion. The coupling between the continuous model and the convolution model in Section 6.2.1 demonstrates that the two approaches physically provide the same results, and there is no contradiction between them. The problematic issue of the traditional models is related to physical interpretation and normalization with respect to correct distribution volume.

7 Real Data Results

Whereas evaluation of the proposed methods on software-phantom data has been performed in the respective sections, in this chapter we will present real data evaluation. As a brief overview, we will present results for the following methods:

- Optimal linear weights for T_1 estimation (Section 3.5).

- Coupled T_1 estimation with vector-field coupling (Section 4.3).

- COPTER: Coupled Model Based Parameter Estimation and Registration (Section 5.1).

- Blind Parameter Estimation using the Patlak Model (Section 5.3).

- Limitations of Traditional Models for Perfusion (Section 6).

We start in Section 7.1 by presenting results which demonstrate, that the novel optimized weights are yielding similar results as the nonlinear weights on real data. In Section 7.2 we will show how the coupling methods for vector fields which have been introduced in Section 4.3 are improving T_1 estimation from variable flip angle data. In Section 7.3 we will show how joint motion correction and registration as proposed in Section 5.1 are improving both registration and parameter estimation. First Real-Data results for the joint AIF and parameter estimation as described in Section 5.3 can be found in Section 7.4. We conclude this chapter with Section 7.5, where we will demonstrate that established methods for blood flow estimation will overestimate blood flow in coupled systems, see Chapter 6.

Note that evaluation of parameter estimation methods on real data is delicate, since there is a general lack of ground truth: Most dynamic parameters cannot be measured directly, since they depend on the tissue type and on individual patient characteristics [8, 139]. For real data, ground-truth parameter maps are hence generally not available. To cope with this issue, we compared our results with literature values, unless stated otherwise. However, note that even literature values only provide a rough estimate of the true parameters, as parameters might vary between patients considerably. Unless stated otherwise, results will be given in terms of an averaged absolute relative error, which was defined as

$$\mathrm{RE} := \tfrac{1}{\mathrm{Vol}(I)} \sum_{i \in I} \frac{|p_i - p_i^{\mathrm{ref}}|}{|p_i^{\mathrm{ref}}|}.$$

where I is set of voxels of interest (e.g. voxels in the kidney cortex) and $\mathrm{Vol}(I)$ its cardinality.

For evaluation, we relied on two datasets: Real kidney data came from a study on the function of the human kidney [74] and was made available by courtesy of Jarle Rørvik from the Haukeland University Hospital in Bergen, Norway. The brain perfusion data which was used to demonstrate limitations of traditional models for perfusion was made available by courtesy of Rashindra Manniesing from the DIAG group of the Radboud University Medical Center in Nijmegen, the Netherlands.

7.1 Linear Approximations for T_1 Estimation

We now present results for the linear approximations in T_1 estimation, which were introduced in Section 3.5. Since we are lacking ground-truth values, we will compare the (novel) optimized weights and the (traditional) linear weights with the nonlinear weights, which are generally considered most stable [154].

Data came from a study on the function of the human kidney [74]. Three healthy human volunteers were scanned with different flip angles, but otherwise constant acquisition parameters. Image acquisition was performed on an Avanto 1.5 T using TWIST with flip-angles $\alpha = (5°, 8°, 15°, 25°)$, TR $= 2.51$ ms, TE $= 0.89$ ms isotropic voxel-size $(0.5\,\text{mm})^3$ and matrix size $[256, 192] \times 52$. For each flip angle a total of 3 images was obtained during a single breathhold per flip-angle. To cope with motion artifacts, the data was pre-aligned using COPTER with a diffusive stabilizer, as described in Section 5.1. Note that this approach thus incorporated some mild regularity constraints for the parameter maps.

Results

For each flip-angle one of the three available datasets was used for reconstruction, which was performed with the nonlinear, linear and optimized weights. Results are depicted in Figure 7.1, where sample reconstructions and relative errors, given in percent, with respect to the nonlinear method are shown. Evaluation was performed on a 2D slice of each dataset, where kidney cortex and medulla were segmented by a medical expert. It can be seen that also on real data the optimized weights perform similar to the nonlinear weights. Whereas the traditional linear weights are yielding errors of up to 10%, it can be observed that the linear weights are yielding similar estimates, even in the medulla region, where high T_1 are making the recovery difficult.

The real data experiments thus confirm that the results obtained with optimized weights are indeed high quality approximations of the ones obtained with the nonlinear weights even for real data. However, a visual inspection shows that in the lung region estimated parameters from the optimized and the nonlinear weights still differ. This is since parameter estimation in this area is highly ill-posed due to the low signal intensity, see Chapter 3.3 for details. The approximations are thus expected to work
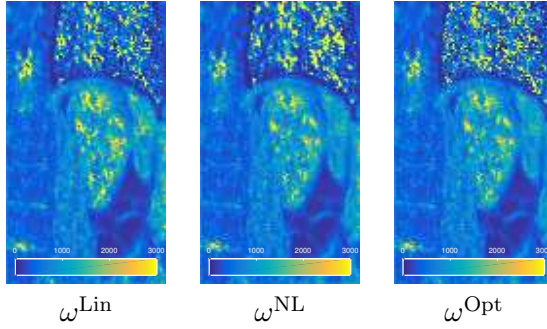
		Cortex		Medulla	
		Lin	Opt	Lin	Opt
D1	left	9.60	0.99	23.70	2.21
	right	9.61	1.00	13.66	1.76
D2	left	2.62	0.19	10.32	0.39
	right	8.97	0.38	15.17	1.30
D3	left	2.14	0.19	05.46	0.52
	right	3.29	0.15	08.97	0.44

ω^{Lin} $\qquad$ ω^{NL} $\qquad$ ω^{Opt}

Figure 7.1: Figure displaying results of the nonlinear, linear and the optimized reconstruction for real data. Given is the mean relative error of T_1 in percent as compared to the results of the nonlinear reconstruction. Weights were obtained by solving (3.9) with $T_1 = (1412, 966)$, $M_0 = (3000, 3000)$ and $f = (0.5, 0.5)$. The images depict the reconstruction of T_1 for D_1, left kidney. Note that T_1 cannot be estimated reliably in regions with zero signal intensity, as it is the case for the lung, see Section 3.3.2. Visual inspection confirms the close agreement of the nonlinear and the optimized approach.

best for the area, where the weights were fine-tuned, see Section 3.5 for details on the recovery method.

7.2 Spatial Coupling for T_1 Estimation

We will now present real data results which confirm, that coupled recovery is capable to improve T_1 estimation for low quality real data or sequences, where the flip angles were not optimized for the expected range or T_1.

We begin by demonstrating that coupling of (T_1, M_0) is superior to coupling of only T_1 [155] or simple denoising. We then proceed to demonstrate that usage of tailored vector-field coupling terms has further positive impact on the recovery. Whereas Section 4.4 has shown that this effect is clear for phantom data, real data improvements will be less obvious. Note, however, that the lack of gold standard makes the evaluation difficult. Whereas numbers show only small improvements, visual inspection of the parameter maps will show a tighter coupling of the recovered parameters, thus indicating overall improved estimation.

Real kidney data was acquired with the same parameters as outlined in Section 7.1.

7.2.1 Superiority of Fully Coupled Estimation

To highlight the positive impact of fully coupled methods, we compared uncoupled nonlinear recovery with various forms of coupled recovery. Specifically, we chose a Total Variation term and compared coupling in (T_1, M_0) with only coupling in T_1 and a standard denoising approach. Optimal parameters were determined with respect to visual plausibility of the parameter maps for all methods separately.

Visual inspection clearly shows that compared to the purely nonlinear approach, all coupled methods yield more homogeneous results. Advantages of the two parameter approach can be observed mainly in reconstruction in the medulla, where T_1 is high as compared to the cortex. Numerical evaluation was done with respect of the mean relative error on cortex on medulla. Since we were lacking a ground-truth, the relative error was calculated voxel-wise with respect to the mean of an unstabilized approach in the respective region of interest. Also in numerical evaluation, we found positive impact of the fully coupled approach as compared to the denoising approach. We found the largest impact of coupling for high T_1, see Figure 7.2. This is expected as it is difficult to estimate high T_1 stably, see Section 3.3.2.

7.2.2 Impacts of Vector Field Coupling

Additionally, we tested the coupling terms which were tailored for vector fields on real data. Note that this time recovery was performed not on 2D slices but on the complete 3D datasets, where cortex and medulla were again segmented by a medical expert. To highlight the merits of this work, comparison was performed with respect coupling methods described in [114, 155]. Specifically we considered $\mathcal{C}_{\mathrm{TP}}(p) := \int_\Omega |\nabla T_1|^2 \, \mathrm{d}x + \int_\Omega |\nabla M_0|^2 \, \mathrm{d}x$ and $\mathcal{C}_{\mathrm{TV}}(p) := \int_\Omega |\nabla T_1| \, \mathrm{d}x$ respectively. Note that results from the previous section already indicate that the latter approach might be inferior, since only one parameter is considered in the coupling.

Results are shown in Figure 7.3. Given is the mean relative error with respect to T_1 from literature (Cortex $\approx 966 \, \mathrm{ms}$, Medulla $\approx 1412 \, \mathrm{ms}$). Again in all cases spatial coupling yields significantly improved results as compared to no coupling. However, the benefit of vector field coupling techniques is on real data not as clear as on phantom data. Here only slight advantages of the vector-field coupling techniques can be observed. To further illustrate our results, RGB maps were created by using T_1 in the red, and M_0 in the green channel. This visualization highlights joint properties of T_1 and M_0 and allows to visually inspect the amount of coupling of T_1 and M_0. Here again clear advantages of methods which couple both (T_1, M_0) can be observed. However, improvements of the tailored vector-field terms $\mathcal{C}_{\mathrm{Nuc}}$ and $\mathcal{C}_{\mathrm{Fro}}$ as compared to the decoupling $\mathcal{C}_{\mathrm{TP}}$ are minor, but nevertheless present.

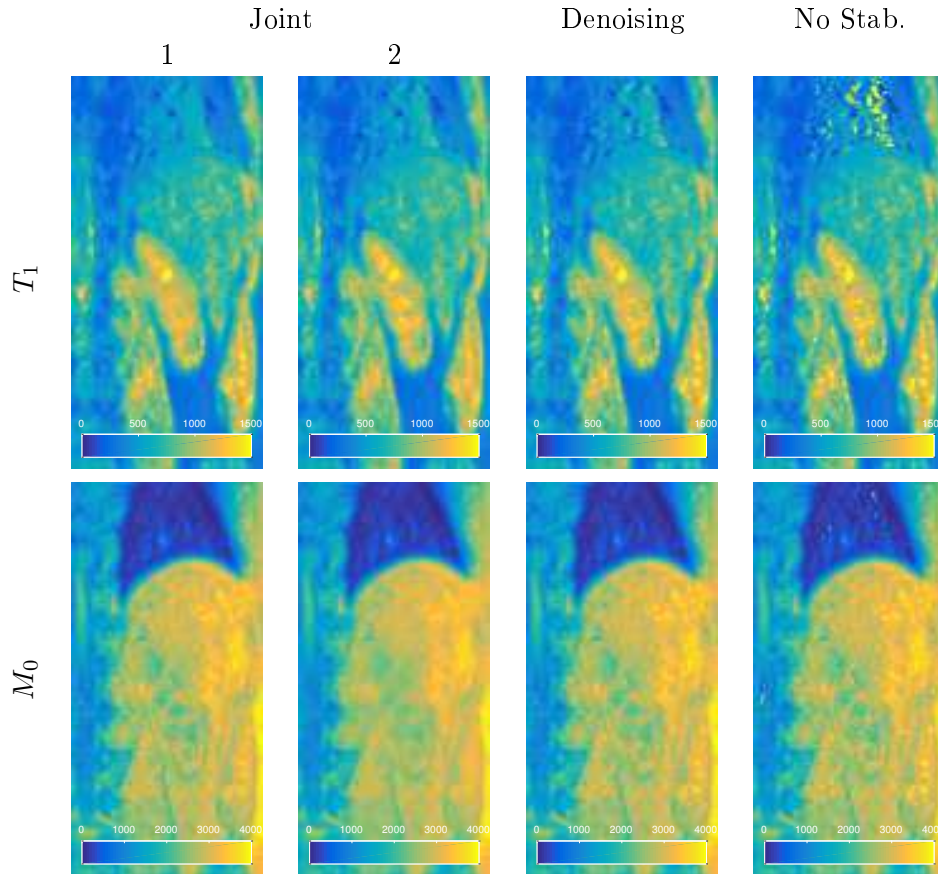| # Parameters		Cortex				Medulla			
		Simul.		Denoising	No Stab.	Simul.		Denoising	No Stab.
		1	2	1	1	1	2	1	1
D1	left	24.17	22.76	25.74	29.02	19.15	17.92	33.47	36.73
	right	21.97	21.21	27.10	30.62	26.44	23.23	42.38	45.53
D2	left	31.76	28.02	32.18	33.48	12.80	12.40	18.12	19.06
	right	18.33	15.13	20.38	21.15	13.84	10.90	18.01	18.84
D3	left	16.03	15.19	15.67	16.31	12.51	11.64	13.25	14.62
	right	10.52	9.71	10.61	11.15	11.24	6.59	13.80	14.12

Figure 7.2: Figure displaying the impact of spatial coupling on real data. The error was computed voxel-wise with respect to the mean of the nonlinear reconstruction over the area of interest and is given in percent. Comparison was performed with respect to: Uncoupled nonlinear recovery, joint recovery with coupling in (T_1, M_0), joint approach with coupling in T_1 and denoising approach. An example of the reconstruction is given for dataset D3, right kidney.

		Cortex					Medulla				
		No Coupl.	$\mathcal{C}_{\text{TP}}$	$\mathcal{C}_{\text{TV}}$	$\mathcal{C}_{\text{Fro}}$	$\mathcal{C}_{\text{Nuc}}$	No Coupl.	$\mathcal{C}_{\text{TP}}$	$\mathcal{C}_{\text{TV}}$	$\mathcal{C}_{\text{Fro}}$	$\mathcal{C}_{\text{Nuc}}$
D1	left	19.2	15.6	16.3	15.1	14.9	37.8	10.8	12.8	10.6	10.4
	right	27.0	21.7	22.4	21.2	21.0	37.0	19.6	21.1	19.0	18.6
D2	left	20.4	19.2	19.1	19.7	19.2	24.6	10.7	12.7	14.0	10.2
	right	29.7	14.9	21.7	16.0	15.4	49.1	7.8	9.8	8.0	7.3
D3	left	35.2	22.2	19.0	23.3	22.8	149.5	11.9	28.8	14.0	12.9
	right	71.0	22.2	21.5	15.5	14.3	108.8	10.3	16.4	10.4	9.7

No Coupl.　　　$\mathcal{C}_{\text{TP}}$　　　$\mathcal{C}_{\text{TV}}$　　　$\mathcal{C}_{\text{Fro}}$　　　$\mathcal{C}_{\text{Nuc}}$
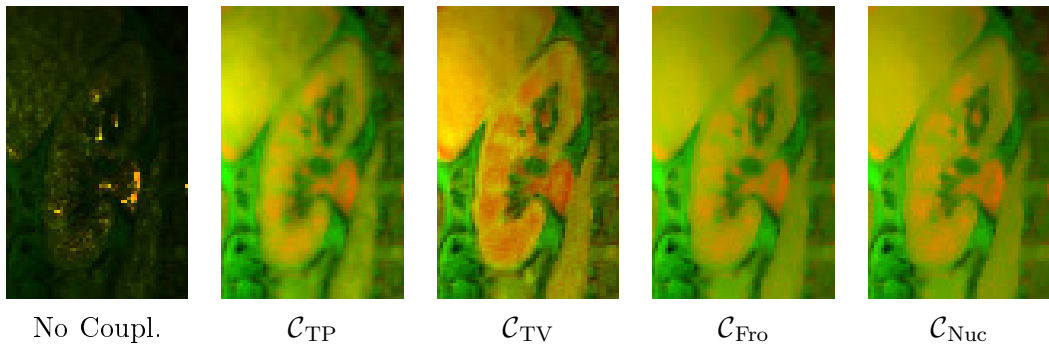
Figure 7.3: Parameter estimation results for vector field coupling on real data. Given is the mean relative error in percent over cortex and medulla with respect to T_1 from literature (Cortex $\approx 966\,\text{ms}$, Medulla $\approx 1412\,\text{ms}$). Also depicted are color images for dataset D3 (right), where T_1 was in the red and M_0 and the green channel.

7.2.3 Discussion

Both real data results demonstrate that coupling of both (T_1, M_0) has clear advantages as compared to methods which use spatial coupling only for T_1. However, whereas results of Section 4.4 have shown clear advantages of vector-field coupling methods for software phantom data, only minor benefits could be observed for real data. A reason for this might be the lack of ground-truth parameters for T_1. Also, manual segmentation of cortex and medulla is challenging for poor data quality. Nevertheless as expected a visual comparison shows clearer contours for the vector-field coupling as compared to standard coupling.

Based on these findings, we advocate spatial coupling of both variables (T_1, M_0) in T_1 estimation in the following cases: if data quality is poor, if only limited flip-angles are available or if flip-angles were not chosen optimally for the expected range of T_1. If computational aspects are no issue, tailored vector-field coupling terms might yield additional improvements. Nevertheless our results indicate that standard Tikhonov-Philips coupling can already lead to major improvements as compared to uncoupled approaches. The main parameter for choosing the coupling term is the expected structure of the tissue: For data where large partial volume effects are expected, $\mathcal{C}_{\mathrm{TP}}$ is a suitable choice. However, in modern scanners resolution is constantly increasing and sharper edges which separate anatomical structures are expected in the images. In such cases $\mathcal{C}_{\mathrm{Nuc}}$ or the computationally less demanding $\mathcal{C}_{\mathrm{Fro}}$ could be better choices. However, we expect coupling to have most impact for poor quality data: For well-designed sequences with good SNR we expect only minor improvements using spatial coupling.

7.3 Joint T_1 Estimation and Registration

We now present preliminary real data results which highlight the merits of joint registration and parameter estimation for nonlinear transformations. As in Section 5.2.2 we compared the uncoupled approach of minimizing J_{Ref} with the coupled approach of minimizing J_{COPTER}.

Real kidney data was acquired with the same parameters as outlined in Section 7.1. For each flip-angle we chose the first two of the three available datasets, yielding eight 3D datasets for one T_1 recovery. Note that since the data was acquired with one breath-hold per flip-angle, motion artifacts could be observed mainly between different flip-angles.

As proposed in [69], for the reference method the data was pre-processed with gaussian-smoothing. The standard deviation was experimentally determined to be half voxel-width. To avoid outliers, the reconstructed T_1 was cut-off at 2000 ms. This clipping was only necessary for the uncoupled approach, the coupled approach yielded reasonable T_1 throughout the recovery. For the coupled approach the parameters γ were determined

Table 7.1: Estimated T_1 [ms] for the different registration methods. Given are mean and standard deviation over cortex and medulla. Reference values from [8] were $966\,\mathrm{ms}$ (cortex) and $1412\,\mathrm{ms}$ (medulla) [8].

Dataset		Cortex		Medulla	
		Ref	COPTER	Ref	COPTER
D1	left	1231 ± 421	891 ± 170	1563 ± 405	1323 ± 194
	right	995 ± 403	855 ± 243	1214 ± 424	1230 ± 278
D2	left	919 ± 324	802 ± 153	1298 ± 427	1228 ± 140
	right	1338 ± 471	1025 ± 206	1660 ± 379	1465 ± 133
D3	left	1006 ± 469	834 ± 233	1273 ± 439	1266 ± 194
	right	1515 ± 478	1183 ± 257	1751 ± 316	1628 ± 194

experimentally: For D1 we chose $\gamma = 1 \times 10^{-6}$, for dataset D2 $\gamma = 5 \times 10^{-6}$ and for dataset D3 $\gamma = 1 \times 10^{-6}$. Again we decided to use the alternating optimization strategy, since it was computationally more efficient and it performed similar to the joint strategy on phantom data, see Section 5.2.2 for details. For the evaluation, cortex and medulla were manually segmented in the dataset acquired with flip-angle 8°, which had best contrast with respect to these structures. We calculated the mean and the standard deviation for both structures. The mean was compared to reference values given in [8].

Results

Our preliminary findings, where we compared the reference approach (minimize J_{Ref}) with the proposed coupled approach (COPTER) are shown in Table 7.1 and Figure 7.4. The table shows clearly that the coupled approach yields a significantly reduced standard deviation of T_1. However, as expected it can also be observed that it comes in parts with significantly reduced mean. Whereas on the kidney cortex in half of the experiments the standard method was slightly superior, on the medulla the novel method yields improved results in four out of six cases. This indicates that the stabilized approach is capable to improve T_1 estimation for large T_1 values, which are generally more difficult to determine. This notion is supported by the visual evaluation of the recovered T_1 maps in Figure 7.4. It can readily be observed, that at many points constraints are active in the uncoupled method and large motion artifacts can be observed. The results from the coupled methods on the other hand show more structure: Medulla and cortex can be distinguished in the reconstruction. However, improved results can also be observed in the motion field. Here the checker-board evaluation shows that both methods can account for small nonlinear motion, as in the left side of D2. Differences can be observed for larger motion, as in the case of the right side of D2. Here it can be observed better motion correction is achieved for the coupled method. In the unstabilized method, motion artifacts above the liver can be observed both in
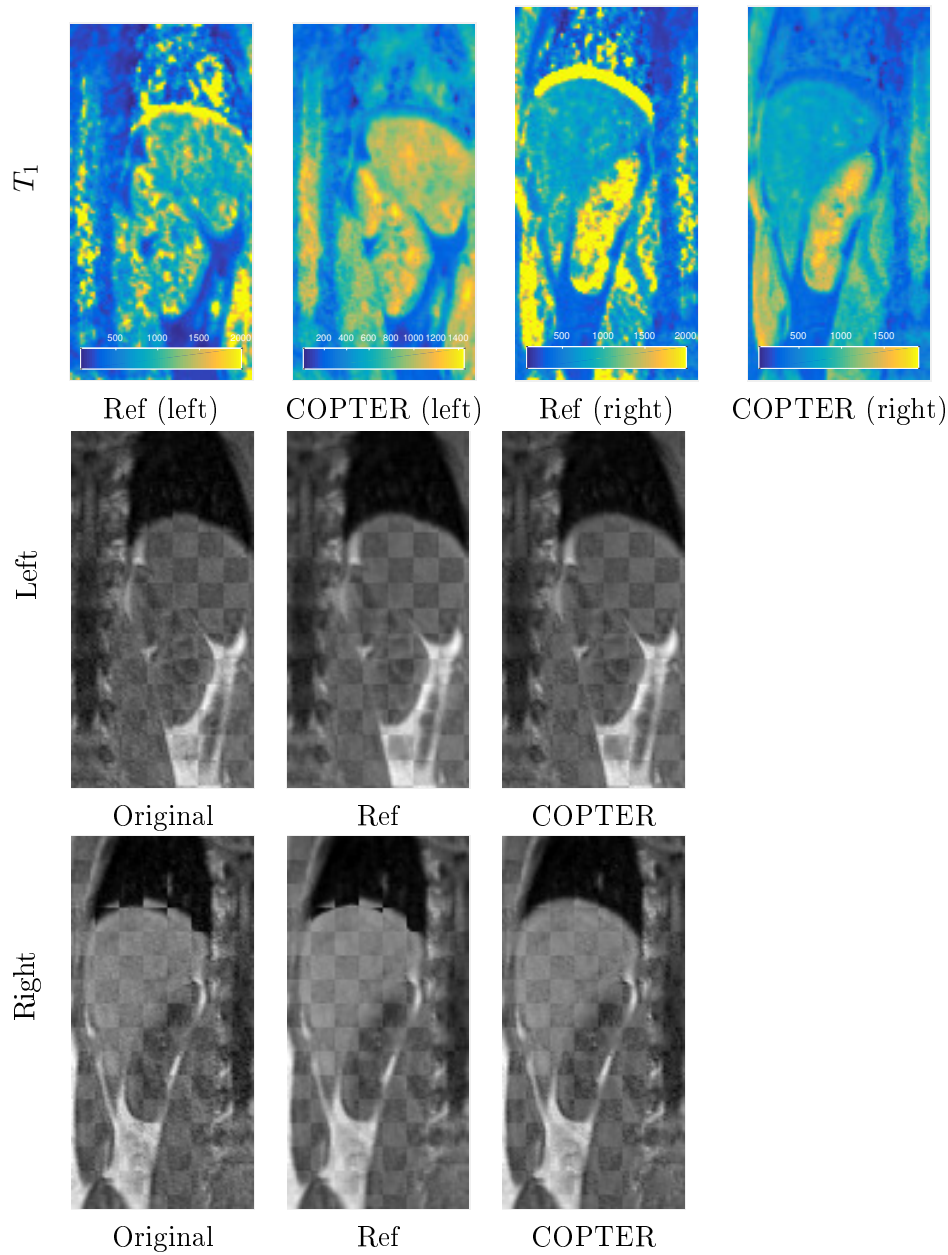
Figure 7.4: Estimated T_1 maps and checkerboard evaluation of the registration. Displayed slice 34 of dataset D2. Checker-board evaluation is with respect to flip-angles 5° and 8°. Compared to the uncoupled reference method, COPTER yielded smoother T_1 maps and thus improved recovery. For the uncoupled approach the checker-board evaluation shows residual motion artifacts above the liver both in T_1 and in the transformed data, which was prevented by COPTER.

the transformed data and in the recovered T_1. This means that T_1 is reconstructed in such a way that motion is simulated by the model. However, the spatial coupling prevents larger jumps and recovers more suitable transformations.

Discussion

It can be observed, that spatial coupling has the potential to improve parameter estimation in the joint setting. For the inclusion of coupling a balancing between speed and accuracy is necessary: Coupling in the (E_1, N) coordinate system as proposed in Section 5.2.1 allows to exploit the linear structure of the objective function. In this case the optimization can be sped up by using the variable projection method described in Section 5.1. However, as the results show, coupling in (E_1, N) is difficult and lead for our data only to minor improvements in T_1 estimation, see also Section 4.4.

In order to obtain more accurate results for T_1 estimation, the obtained results indicated that coupling of (T_1, M_0) might be preferable. Whereas phantom data experiments indicated that the coupled joint approach yields improved results, numerical evaluation of the joint approach showed only minor improvements for real data. However, again the lack of gold-standard parameters complicates the evaluation. Our experiments show that spatial coupling is capable to improve both registration and parameter estimation: Whereas in the uncoupled approach model parameters might be selected which simulate movement, this unwanted effect seems to be suppressed in the coupled version.

As described in Section 5.2.2, the coupled approaches are expected to yield most improvement for low-quality data and more than two flip-angles: Other, not documented, experiments indicated that for only few flip-angles and high-quality data, reconstruction using the joint approach without stabilization yields comparable results.

7.4 Blind Parameter Estimation for the Patlak–Rutland Model

In this section we will present results which show how joint methods can improve parameter estimation, if no or only a low-quality AIF is available. Specifically we will give first real data results of the method which was proposed in Section 5.3.

Data again came from a study on the human kidney [74]. However, this time we used the DCE-MRI sequence to test our algorithm. One healthy human volunteer was scanned on a 1.5T MRI system (Siemens Avanto, Siemens, Erlangen, Germany) using a TWIST sequence after injection of 4ml MultiHance injected as a bolus at a rate of $3\,\mathrm{ml/s}$. A time sequence of 49 images with variable spacing between 2.5s during first-pass and 60s in the late-phase was acquired. The DCE-MRI images were registered and contrast agent concentrations were calculated from the MRI data. Afterwards,
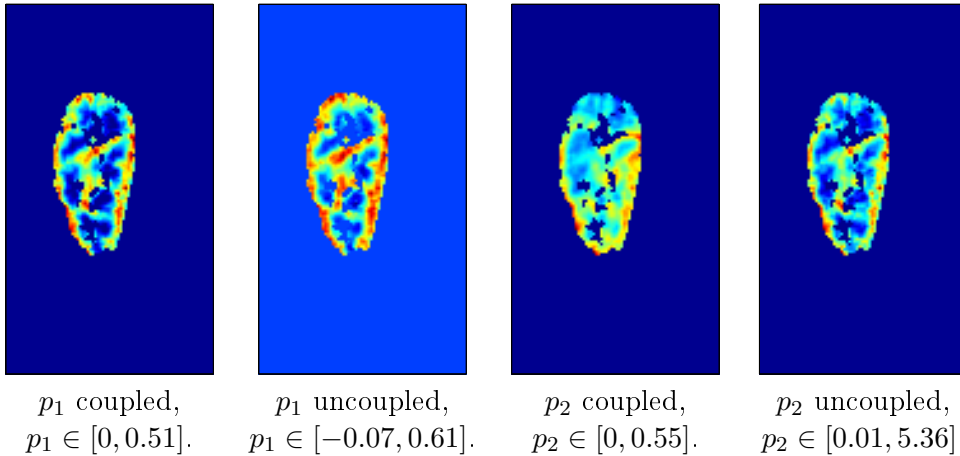
p_1 coupled,	p_1 uncoupled,	p_2 coupled,	p_2 uncoupled,
$p_1 \in [0, 0.51]$.	$p_1 \in [-0.07, 0.61]$.	$p_2 \in [0, 0.55]$.	$p_2 \in [0.01, 5.36]$.

Figure 7.5: Shown in slice 34 of the recovered p_1 and p_2 of the kidney data from the Haukeland clinic. The regularization and constraints help to reduce outlier and to enforce a narrower range of parameters. However, also a possible oversmoothing in p_2 can be observed.

the kidney medulla and the cortex were segmented. For details on the scan and the postprocessing pipeline see [74]. For validation, an arterial input function was measured in the descending aorta, slightly above the renal artery, see Figure 7.6. Since the Patlak model is only valid in the early uptake phase, recovery was performed for the time interval $t \in [1, 1.7]$min, see [68] for details. We employed our new regularized method to calculate the 3D parameter maps for p_1 and p_2 as well as the AIF.

Slice 34 of the 3D dataset is shown in Figure 7.5. It can be observed that combined coupling and constraint optimization help to reduce outliers and to enforce a narrower range of parameters. The most prominent effect can be observed in p_2, where maximum parameters were reduced from $5.36\,\mathrm{ml}^3/(\mathrm{s}\,\mathrm{ml}^3)$ to $0.55\,\mathrm{ml}^3/(\mathrm{s}\,\mathrm{ml}^3)$. However, since the coupling parameters were tailored for the phantom, also a possible over-smoothing in p_2 can be observed. Larger differences between the unstabilized and the stabilized method can be observed in the arterial input function, depicted in Figure 7.6. The integral was normalized to one for visualization. To calculate the GFR, the maximum peak was adjusted to fit the maximum peak of the measured arterial input function. It can be observed that the uncoupled method predicts negative concentrations, whereas these effects are removed due to the constrained optimization. Nevertheless both methods are giving a false prediction of 0 concentration at time point $t = 1\,\mathrm{min}$, possibly leading to an underestimated GFR of $21.6\,\mathrm{ml/min}$ for the coupled method.

Discussion

Our simulations show that whereas experimental results on phantom data in Section 5.4 were promising, transfer to real data is still difficult. Reasons for this behavior
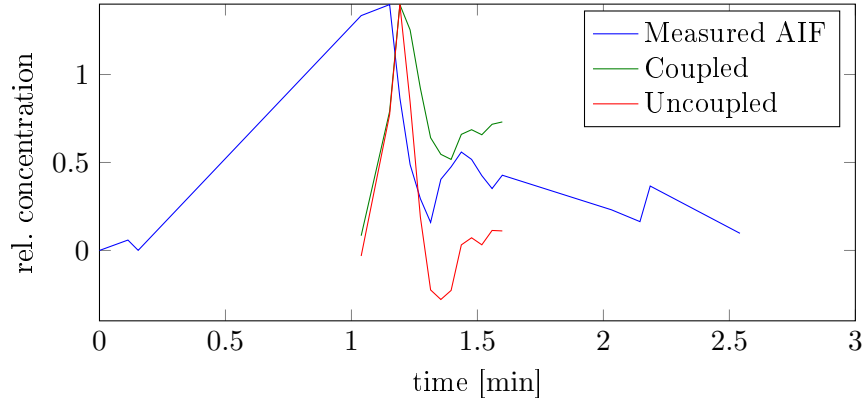
Figure 7.6: Comparison of a measured arterial input function with estimated arterial input functions, which obtained from coupled and uncoupled recovery. In the recovered arterial input functions, a false prediction of 0 at time point $t = 1\,\text{min}$ can be observed, leading to an underestimated GFR of $21.6\,\text{ml/min}$ for the uncoupled method (reference approximately $105\,\text{mL/min}$ for both kidneys using Iohexol clearance).

might be manifold. We suspect them to be linked mainly to the problem of scale: Due to the underlying model, our method cannot yet estimate parameters uniquely and thus a proper scaling approach has still to be identified. Note that this scale factor has a direct anti-proportional effect on the calculated GFR. This problem complicates especially the evaluation on clinical data, where the gold-standard GFR of $105\,\text{mL/min}$ was determined using Iohexol clearance and was thus given with respect to both kidneys [74]. Whereas the latter problem could be overcome using novel data from the Haukeland University Hospital, where donor patients with only one kidney were scanned, reasonable methods to determine scale still need to be identified, see Section 5.3 for a brief outline of state-of-the art methods to do this.

7.5 Limitations of Traditional Models for Perfusion

In this section we will present real data results which indicate that traditional models will overestimate perfusion if applied to coupled systems. In order to illustrate that this effect also may be observed on a complete dataset, we applied the deconvolution model to a clinically acquired human perfusion CT dataset of a 56 years old male admitted with suspicion of stroke to the Radboud University Medical Center in Nijmegen, the Netherlands. The perfusion scan was obtained using a Toshiba Aquilon ONE scanner, pixel-size $0.43\,\text{mm} \times 0.43\,\text{mm}$, slice thickness $0.5\,\text{mm}$, contrast agent $50\,\text{ml}$ Xentix 300, total scan-time $114\,\text{s}$, time resolution ranging from $2.1\,\text{s}$ in the early- to $30\,\text{s}$ in the late phase of contrast agent uptake. To cope with motion artifacts, the data was pre-aligned using rigid transformations. The arterial input function was manually selected
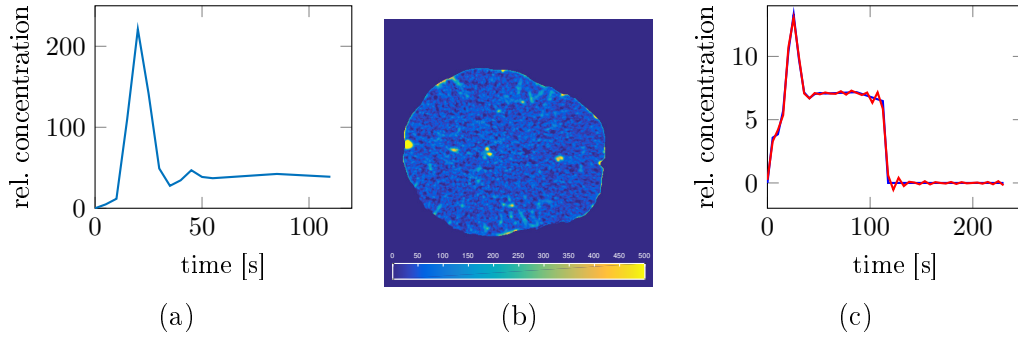
(a) (b) (c)

Figure 7.7: Results from real-data experiments. (a) AIF manually selected from the MCA. (b) One slice of the voxel-wise scaled CBF-reconstruction [ml/min/100mg] for a 3D volume of interest. (c) Mean concentration curve for the complete 3D volume of interest and the curve approximation by bSVD.

by a medical expert in the middle cerebral artery. Since we expected to see local over estimation effects mainly for small voxel sizes, the data was processed at full resolution $(512 \times 512 \times 320)$ voxels. However, in order to deal with noise it was necessary to apply a prior gaussian smoothing with standard deviation of 1 voxel. Relative concentrations were estimated from the CT signal assuming a spatially independent proportionality constant. The brain tissue was segmented automatically and used as ROI for the perfusion analysis.

CBF was then estimated voxelwise using an own Matlab implementation of bSVD, yielding an average scaled CBF of 64.357 ml/min/100mg with a global threshold of 4%. Furthermore, we estimated the perfusion for the whole volume of interest by averaging the concentration values first and then performing the bSVD using the same threshold, yielding a total scaled CBF of 24.791 ml/min/100mg. Voxelwise results are depicted in Figure 7.5.

8 Conclusion

Although traditionally dynamic models were designed to describe the response of a single system, such models are used increasingly to recover parameters from spatially structured data. In this work we have shown how such estimations can be improved and have narrowed down boundaries under which such estimation produces physically meaningful results.

First and foremost we have focused on improving parameter estimation using dynamic models from spatially arranged, low-quality 4D data. Although there are various publications which have used coupled dynamic models for parameter estimation [114, 155, 135], these approaches were often either using coupling of only one parameter or were simple sums of established coupling terms for scalar functions. As such, regularity was only introduced in space but not between different parameter maps. In this work we have presented a family of coupling terms for parameter estimation which introduce both coupling in space and between different parameters. For the example of T_1 estimation from high-resolution data we have shown that the proposed approaches are capable to improve parameter estimation from low-quality data considerably, see Section 4.4. Parts of these results are to appear in the *Journal of Algorithms and Computational Technology* [S3].

Next to standard parameter recovery, we have also shown that spatial coupling can improve joint methods, in which data redundancy is exploited to estimate additional parameters such as motion or control parameters of the pharmacokinetic model. Here the results indicate that especially joint motion correction and parameter estimation [23, 11, 69] can profit from spatial stabilization. As outlined in Section 5.2 and Section 7.3, we assume that this is the case since spatial regularity prevents the dynamic models to adjust in such a way that motion is emulated. For double-blind approaches [121] we have found that spatial coupling can improve results, but that results will be especially sensitive to the chosen stabilization parameter, see Section 5.3. These results have been published in the *Proceedings of the BVM* [H1, H2].

Finally we have presented boundaries under which one-dimensional dynamic models produce meaningful results if coupling of the models is expected, see Section 6.3 and Section 7.5. In the case of blood flow estimation we have shown that in pure capillary tissue one-dimensional dynamic models produce scale dependent results, see also [73]. More specifically, we have presented a novel PDE model which describes perfusion in pure capillary tissue and have connected this model to traditional one-dimensional dynamic models for perfusion estimation. For simple parameters we have further given

explicit expressions for the perfusion which is expected to be recovered from such data using dynamic models. We have also shown both analytically and experimentally that in this case the recovered flow will depend heavily on the discretization. Additionally, we have presented real-data results which indicate that such effects might be observed on coarse scale in real-life applications. These results are submitted to the *IEEE Transactions on Biomedical Engineering* and are currently under review [S4].

We will now start a more detailed discussions of the obtained results:

Linear Approximations of Nonlinear Weights

To improve T_1 recovery from noisy variable flip-angle data in MRI, we have introduced a unified framework for T_1 estimation from such data. The framework consists of a family of weighted parameter estimation techniques and includes the two most commonly used ones, namely the linear and the nonlinear approach, as special cases. We have established sensitivities of the parameter estimation for a gaussian noise model for all members of the proposed family. For the linear and the nonlinear estimation, our sensitivity results coincide with the ones obtained in [154, 43]. The obtained sensitivities were then used to determine a set of weights which allows for linear estimation of T_1 with lower relative errors than the original linear method. We have evaluated the proposed approach both on phantom and on real data. The results presented in Section 3.6 and Section 7.1 show that the optimized weights allow to reconstruct T_1 with similar errors as the nonlinear estimation, which is most robust [43, 33]. The optimized weights can hence be used for both fast and robust estimation of T_1, avoiding the computational overhead of nonlinear optimization. Results for synthetic as well as real data confirm improvements in T_1 estimation of up to 24% as compared to standard linear estimation. However, it was also found that the amount of improvement depends both on the sequence and the expected T_1: If the linear and the nonlinear method have similar sensitivities, only minor improvements could be observed. Matlab code to determine the optimized set of weights has been made publicly available under `http://www.mic.uni-luebeck.de/people/constantin-sandmann.html`. A code example can be found in Listing 8.1

Spatial Coupling for Parameter Estimation

We have found that T_1 reconstruction becomes increasingly ill-posed for low signal intensities, see Lemma 10. This makes estimation especially of high T_1 delicate, since for T_1 weighted sequences low signal is often the result of high T_1, see (2.19). To cope with such problems we have additionally analyzed the impact of spatial coupling, where information on the expected spatial structure of the T_1 maps is included in the recovery.

Results on phantom data presented in Section 4.4.1 indicate that fully coupled recovery might only be superior to standard denoising techniques if both variables T_1 and M_0 are stabilized. Note that some approaches for T_1 recovery apply coupling only with respect to T_1 [155]. Our results both on phantom and on real data suggest that improved parameter estimation is possible by using coupling of both variables. Additionally, we have investigated the impact of coupling in the (E_1, N) coordinate system, which is widely used if the linear method is used for T_1 estimation. Here it was found that coupling in (T_1, M_0) is superior, as the change of coordinates introduces systematic biases for large T_1.

We have found that best results in recovery can be achieved, if different coupling parameters are used for T_1 and M_0. Since this is both tedious and impractical for real data, we have additionally introduced tailored coupling methods for vector fields, which were inspired by RGB denoising techniques. The proposed coupling terms are based on l^p norms of the singular values of the Jacobian, the so-called Schatten-p-norms. Here we have found that best recovery results can be achieved if the low-rank inducing Schatten-1-norm is used, since common edges in recovered parameter maps are enforced. It was demonstrated on phantom data, that these terms have additional advantages with respect to errors in T_1 reconstruction, see Section 4.4. Note that the proposed terms have been studied for for RGB denoising extensively [61, 88, 105] and are currently increasingly used for joint image reconstruction [148, 81]. However, we have not found indications that these terms have yet been proposed for T_1 estimation or even parameter estimation in general.

As expected we have also found that spatial coupling can come with underestimation of the parameters for large stabilization parameters. Strategies to cope with this problems could be to use an L^1 related distance measure, see the outlook for details. Also, the choice of coupling is heavily dependent of the data: The proposed coupling terms are designed to recover sharp edges in the parameter maps. These are hence best suited for high resolution scans of anatomical structures which are expected to be clearly distinct. For other cases a Tikhonov-Philips coupling which recovers smooth solutions might be preferable. We also stress that the impact of spatial coupling naturally depends on the quality of the data: For high-quality data with good SNR we expect only minor improvement by spatial coupling. Also note that the proposed Schatten-norm based stabilizers require computationally more complex optimization methods, which

Listing 8.1: Matlab code demonstrating the calculation of optimized weights for
T_1 recovery. The documentation of the function is omitted. The code
is publicly available under `http://www.mic.uni-luebeck.de/people/`
`constantin-sandmann.html`.

```matlab
function wOpt = getOptimalWeights(T1Exp,flips,TR,varargin)
        %if no input is given: show help and run an minimal example
        if nargin == 0
                clc
                help(mfilename);
                runMinimalExample;
                return;
        end
        %default parameters
        L        = numel(T1Exp); %number of expected T1
        M0       = 3000*ones(L,1); %default M0
        r        = ones(L,1)/L;     %default weights in obFun
        tolJ     = 1e-4; tolY = 1e-4; tolG = 1e-4; %stopping criteria
        maxIter  = 1000; LSMaxIter = 20; %maximum iterations
        regH     = 1e-5; %parameter for Levenberg-Marquardt
        %overwrites default parameter
        for j=1:2:length(varargin),
                eval([varargin{j},'=varargin{',int2str(j+1),'};']);
        end
        %prepare variables
        T1Exp = T1Exp(:);
        M0    = M0(:);
        flips = flips(:);
        k     = numel(flips);
        %check input
        assert(numel(T1Exp)==numel(M0),'M0 and T1 must have an equal
            number of elements');
        assert(max(abs(flips))<pi  ,'Flip angles must be given in RAD.')
            ;
        %prepare optimization
        obFun    = @(w) obFunMultiT1(w,T1Exp,M0,flips,TR,r);
        wLin     = ones(k,1);
        paraOpt  = {'tolJ',tolJ,'tolY',tolY,'tolG',tolG,...
                    'maxIter',maxIter,'LSMaxIter',LSMaxIter,...
                    'regH',regH};
        %do the optimization
        wOpt     = GaussNewtonSimplex(obFun,wLin,paraOpt{:});

end
```

might be slower than optimization methods which can be used for Tikhonov-Philips coupling.

Joint Parameter Estimation and Motion Correction

Another part of this work has been concerned with joint methods, where dynamic parameters are to be estimated from both motion and noise corrupted data, see Section 5.1. The results which were obtained in Section 5.2 and Section 7.3 indicate that spatial coupling can also improve parameter estimation in this joint setting. Note that our results indicate that for simple data (affine deformations, clear structure of the data) joint approaches only lead to negligible improvements in parameter estimation and registration, see Section 5.2.1. From a practical perspective this is promising, as conventional techniques can be employed for recovery. However, for more complicated data results of [69] indicate that joint approaches indeed come with advantages. In this case our results from Section 5.2.2 show that spatial coupling can additionally improve both parameter estimation and registration: To demonstrate these improvements, we have compared an unstabilized approach with a stabilized approach both on phantom and on real data. We have found that in the unstabilized approach dynamic parameters are recovered which emulate motion. Since this often leads to parameter maps with large outliers, such effects are mitigated by spatial coupling. On real data we have observed both clearer parameter maps and improved motion correction by using spatial coupling. We have also found that spatial coupling is capable to improve parameter estimation especially for difficult to recover high T_1, see Lemma 10. However, both the coupled and the established uncoupled method failed to recover the parameters accurately even in the presence of no noise. Reasons for this behavior need to be investigated further and might be linked to the considerable deformations which were simulated. For the actual optimization, we have compared two optimization strategies: A coupled approach, where optimization is performed with respect to all variables at once, and an alternating approach, where subsequent registration and parameter estimation are performed. Here we have found no large differences between the coupled and the subsequent approach. Since the subsequent approach is easier to implement and has more stable convergence, we conclude that the subsequent approach is more suited for applications. Note that in our experiments we have used only Tikhonov-Philips coupling since we found our implementation of the Schatten-norm based stabilizers to be too slow for real 4D data. We nevertheless expect improved results for the more advanced coupling strategies. Again, as already described above, both the type and impact of the coupling terms will depend on the data. Also note that in the case of only two flip-angles we do not expect the joint approach to work, since any two data points completely determine the dynamic parameters. The impact of coupling in such setting yet needs to be determined.

Blind Parameter Estimation for the Patlak Rutland Model

Additionally to joint models which are used to recover dynamic parameters and motion, we have also analyzed joint models which are used to recover dynamic parameters and control parameters of the model, see Section 5.3. Here we have focused on joint estimation of the arterial input function and pharmacokinetic parameters and have again extended the established approach [121] by a spatial coupling term. Results from Section 5.4 indicate that also in this setting, spatial coupling can improve parameter estimation especially at low SNR. In the software phantom experiments we have measured improved results of up 45% percent for K^{trans}, which is the main parameter to recover the GFR. However, in this method we have found different challenges: First, the selection of stabilization parameters is difficult, since in a double-blind setting one variable determines the other: Both too high coupling and too low coupling will introduce errors in the arterial input function, which again will lead to errors in the parameters. This implied that we could not eliminate the dependency of the arterial function completely and still rely on estimate, which could be given by a low-quality measurement or a population-averaged arterial input function as in [113]. Additionally the problem of scale remains: Since in the double-blind setting both variables can only be estimated up to a scaling constant, external criteria need to be found to determine the scaling factor. Other approaches are using either reference values for the pharmacokinetic parameters [146, 162] or using the integral of nearby arterial concentration [53]. Note that if one is only interested in relative values to distinguish local functionality this normalization can be omitted.

Limitations of Traditional Models for Perfusion

Our results described in Section 6.3 and Section 7.5 show that if traditional methods for perfusion estimation are applied to a coupled system they perform well if applied to averaged snapshots of the system, but tend to fail when they are applied to only parts of the system. This effect was demonstrated both analytically and experimentally. We have illustrated this effect in detail in the case of high resolution phantom images where the voxel size reaches the spatial dimension of the capillary systems, but also showed the effect on coarse scale real data. The reason for this failure is not numerical instabilities in the deconvolution, but rather that perfusion becomes overestimated since traditional models will not account for the correct distribution volume. Whereas the the scaling effect for generic volumes has been described in [72], we have presented a real-life example and additionally performed an exact quantitative analysis for coupled systems. This scaling problem is expected to become more pronounced in future as imaging hardware is constantly improving in spatial resolution. We expect to find overestimation also in pathological tissue, where fluid is passing through multiple, adjacent voxels, and recommend to be take this effect into consideration in clinical evaluation of the data. To account for this problem, the development of new field

models for perfusion is therefore highly demanded, in line with approaches described in [140, 101].

Outlook

Although this work has clarified assets and drawbacks of one-dimensional dynamic models in coupled settings, there are still multiple open questions which can be addressed in future work:

Most prominently, the ideas of joint motion correction and parameter estimation could be extended to concentration recovery and motion correction for DCE-MRI. Multiple approaches have been taken to extend the joint approaches for motion estimation and dynamic modeling to such data [23, 11]. To the best of the author's knowledge, none of these approaches is currently used in everyday clinical practice. Possible reasons for this have been outlined in Section 5.1 and include the resulting, highly nonlinear joint model: Given a T_1 map, a straight-forward approach would consider the concatenated signal equation and dynamic model. Unknown parameters would thus be the deformation fields and the dynamic parameters for each time point. As a first step, one could seek to improve only the estimation of contrast agents without including the dynamic modeling. Results from [69] and recent evaluations from the University of Bergen have shown that better recovery of contrast agent concentrations from DCE-MRI images might be the key for more accurate estimation of pharmacokinetic parameters. Note that this step already requires registration, since each image needs to be related to the baseline. Our results indicate that spatial coupling might have important influence for accurate registration and concentration estimation. In a non-coupled approach, increasing the contrast agent concentration at a given point every signal intensity belonging to a shorter T_1 can be emulated, which might simulate motion. The work to extend such joint approaches to DCE-MR images is subject of a current grant application of the Institute of Mathematics and Image Computing with several partners.

Another extension of our results could concern the problem that regularization parameters in spatial coupling can lead to underestimation of the parameters. Here effects of an L^1 based distance could be investigated: In the current setting we have combined a squared distance with a non-squared stabilizer. Both terms thus scale significantly differently. As results from [30] indicate, in the denoising setting L^1 based distances allow for perfect recovery for a whole range of stabilization parameters without any loss of contrast. This might be tried for parameter estimation as well. Connected to this problem are different optimization schemes for the primal-dual objective functions: We have found that the PDHG yields both fast and accurate results. Nevertheless approaches such as iPiano [109], which are using ideas from the heavy-ball optimization method, might converge even more rapidly and help to avoid possible local minima.

Another idea is to use the vector-field stabilizers for CBF estimation with the Meier-Zierler model: By coupling each component of the residue functions, more reliable CBF

results might be obtained. A first implementation of this approach could be easily done, since it would lead to a linear data-term $D(R) := \|\tilde{L}R - C\|^2$ for $R \in R^{kn}$), $C \in \mathbb{R}^{kn}$, $\tilde{L} := \mathrm{kron}(\mathrm{id}_n, L)$ and $L \in \mathbb{R}^{k,k}$ as in Section 3.2.2. However, this does not address the question of how to include monotonicity of R.

Concerning CBF estimation, our experiments show that spatially coupled models for perfusion might be more suitable for higher scanner resolutions. A spatially coupled model for blood flow which accounts both for spatial propagation and multiple compartments has been proposed recently [140]. However, since the proposed model is highly complex, experimental evaluation is still pending. Further work on this topic is highly advised and subject of a recent grant application of the University of Bergen.

In this work we have proposed various strategies to improve parameter estimation from voxel-wise 4D data with a focus on the role of spatial coupling. We have transferred state of the art modeling from applied mathematics to such problems and have found that in some applications these methods are indeed capable to improve parameter results. Further, we have used the modeling to demonstrate and to test limits of established methods for parameter recovery. We thus conclude that especially for voxel-wise applications and one-dimensional dynamic models, coupling is an essential ingredient if stable results are to be estimated. Nevertheless, we also hope that with the development of continuous methods for contrast agent propagation as e.g. in [140], coupling will be an inherent component and will not need to be included externally.

Bibliography

[1]   J E Aarnes, T Gimse, and K-A Lie. *An introduction to the numerics of flow in porous media using Matlab*. Springer Verlag, 2007.

[2]   B Abels, E Klotz, B F Tomandl, et al. "Perfusion CT in acute ischemic stroke: a qualitative and quantitative comparison of deconvolution and maximum slope approach". In: *Am J Neuroradiol* 31.9 (2010), pp. 1690–1698.

[3]   M Abramowitz, I A Stegun, et al. "Handbook of mathematical functions". In: *Applied mathematics series* 55 (1966), p. 62.

[4]   D C Alsop, J A Detre, X Golay, et al. "Recommended implementation of arterial spin-labeled perfusion MRI for clinical applications: A consensus of the ISMRM perfusion study group and the European consortium for ASL in dementia". In: *Magn Reson Med* 73.1 (2015), pp. 102–116.

[H5]   N Ammann, A Derksen, and C Heck. "A novel magnetometer-accelerometer calibration based on a least squares approach". In: *Proc. ICUAS 2015*. IEEE. 2015, pp. 577–585.

[5]   L Annet, L Hermoye, F Peeters, et al. "Glomerular filtration rate: Assessment with dynamic contrast-enhanced MRI and a cortical-compartment model in the rabbit kidney". In: *J Magn Reson Imaging* 20.5 (2004), pp. 843–849.

[6]   R C Aster, B Borchers, and C H Thurber. *Parameter estimation and inverse problems*. 2nd. Waltham, MA: Academic Press, 2013.

[7]   D M Bates and D G Watts. "Relative curvature measures of nonlinearity". In: *J Roy Stat Soc B* 42.1 (1980), pp. 1–25.

[8]   Cedric MJ de Bazelaire et al. "MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0T: preliminary results 1". In: *Radiol* 230.3 (2004), pp. 652–659.

[9]   D P Bertsekas. "Projected Newton methods for optimization problems with simple constraints". In: *SIAM J Control Optim* 20.2 (1982), pp. 221–246.

[10]   J C Bezdek and R J Hathaway. "Some notes on alternating optimization". In: *AFSS International Conference on Fuzzy Systems*. Springer. 2002, pp. 288–300.

[11]   M Bhushan et al. "Motion correction and parameter estimation in dceMRI sequences: application to colorectal cancer". In: *Proc. MICCAI 2011*. Toronto: Springer, 2011, pp. 476–483.

[12]   A Bivard, C Levi, V Krishnamurthy, et al. "Defining acute ischemic stroke tissue pathophysiology with whole brain CT perfusion". In: *J Neuroradiology* 41.5 (2014), pp. 307–315.

[13]   J Blinn. "Consider the lowly 2x2 matrix". In: *IEEE Comput Graph Appl* 16.2 (1996), pp. 82–88.

[14]   F Bloch. "Nuclear Induction". In: *Physical Review* 70.7–8 (1946), pp. 460–474.

[15]   P Blomgren and T F Chan. "Color TV: total variation methods for restoration of vector-valued images". In: *IEEE T Image Process* 7.3 (1998), pp. 304–309.

[16]   L Bokacheva, H Rusinek, J L Zhang, et al. "Estimates of glomerular filtration rate from MR renography and tracer kinetic models". In: *J Magn Reson Imaging* 29.2 (2009), pp. 371–382.

[17]   T Boutelier et al. "Bayesian hemodynamic parameter estimation by bolus tracking perfusion weighted imaging". In: *IEEE T Med Imaging* 31.7 (2012), pp. 1381–1395.

[18]   William G Bradley. "History of medical imaging". In: *Proc Am Philos Soc* 152.3 (2008), pp. 349–361.

[19]   C Broit. "Optimal registration of deformed images". Ph.D. Thesis. University of Pennsylvania, 1981.

[20]   J A Brookes, T W Redpath, F J Gilbert, et al. "Accuracy of T1 measurement in dynamic contrast-enhanced breast MRI using two-and three-dimensional variable flip angle fast low-angle shot". In: *J Magn Reson Imaging* 9.2 (1999), pp. 163–171.

[21]   A Buades, B Coll, and J-M Morel. "A non-local algorithm for image denoising". In: *Proc. IEEE CVPR 2005*. Vol. 2. IEEE. San Diego, 2005, pp. 60–65.

[22]   D L Buckley, A E Shurrab, C M Cheung, et al. "Measurement of single kidney function using dynamic contrast-enhanced MRI: Comparison of two models in human subjects". In: *J Magn Reson Imaging* 24.5 (2006), pp. 1117–1123.

[23]   G A Buonaccorsi, J P B O'Connor, A Caunce, et al. "Tracer kinetic model–driven registration for dynamic contrast-enhanced MRI time-series data". In: *Magn Reson Med* 58.5 (2007), pp. 1010–1019.

[24]   M Burger, J Modersitzki, and L Ruthotto. "A hyperelastic regularization energy for image registration". In: *SIAM J Sci Comput* 35.1 (2013), pp. 132–148.

[25]   R B Buxton. "Quantifying CBF with arterial spin labeling". In: *J Magn Reson Imaging* 22.6 (2005), pp. 723–726.

[26]   F Calamante, P J Yim, and J R Cebral. "Estimation of bolus dispersion effects in perfusion MRI using image-based computational fluid dynamics". In: *Neuroimage* 19.2 (2003), pp. 341–353.

[27]   G Casella and R L Berger. *Statistical inference*. 2nd. Vol. 2. Duxbury Press, 2002.

[28] A Chambolle and T Pock. "A first-order primal-dual algorithm for convex problems with applications to imaging". In: *J Math Imaging Vis* 40.1 (2011), pp. 120–145.

[29] A Chambolle, V Caselles, D Cremers, et al. "An introduction to total variation for image analysis". In: *Theoretical Foundations and Numerical Methods for Sparse Recovery.* Ed. by M Fornasier. Radon Series on Computational and Applied Mathematics 9. Walter de Gruyter, 2010, pp. 263–340.

[30] T F Chan and S Esedoglu. "Aspects of total variation regularized L1 function approximation". In: *SIAM J Appl Math* 65.5 (2005), pp. 1817–1837.

[31] Y Chen, B T Mallison, and L J Durlofsky. "Nonlinear two-point flux approximation for modeling full-tensor effects in subsurface flow simulations". In: *Computat Geosci* 12.3 (2008), pp. 317–335.

[32] H-L M Cheng. "T1 measurement of flowing blood and arterial input function determination for quantitative 3D T1-weighted DCE-MRI". In: *J Magn Reson Imaging* 25.5 (2007), pp. 1073–1078.

[33] H-L M Cheng and G A Wright. "Rapid high-resolution T1 mapping by variable flip angles: Accurate and precise measurements in the presence of radiofrequency field inhomogeneity". In: *Magn Reson Med* 55.3 (2006), pp. 566–574.

[34] Y-I Cho and D J Cho. "Hemorheology and microvascular disorders". In: *Korean Circ J* 41.6 (2011), pp. 287–295.

[35] K A Christensen et al. "Optimal determination of relaxation times of Fourier transform nuclear magnetic resonance. Determination of spin-lattice relaxation times in chemically polarized species." In: *J Phys Chem-US* 78.19 (1974), pp. 1971–1977.

[36] J Chung, E Haber, and J Nagy. "Numerical methods for coupled super-resolution". In: *Inverse Problems* 22.4 (2006), p. 1261.

[37] P Colella. "Multidimensional upwind methods for hyperbolic conservation laws". In: *J Comput Phys* 87.1 (1990), pp. 171–200.

[38] M Cutajar et al. "The importance of AIF ROI selection in DCE-MRI renography: reproducibility and variability of renal perfusion and filtration". In: *Eur J Radiol* 74.3 (2010), e154–e160.

[39] B Dacorogna. *Introduction to the Calculus of Variations.* London: Imperial Press, 2004.

[40] H Darcy. "Les Fontaines Publiques de la Ville de Dijon". In: *Victor Dalmont* (1856), p. 647.

[41] A C Davison. *Statistical models.* Vol. 11. Cambridge University Press, 2003.

[42] P Deetjen, E-J Speckmann, and J Hescheler. *Physiologie.* 4th ed. München, Jena: Urban & Fischer, 2005.

[43]   S CL Deoni, T M Peters, and B K Rutt. "Determination of optimal angles for variable nutation proton magnetic spin-lattice, T1, and spin-spin, T2, relaxation times measurement". In: *Magn Reson Med* 51.1 (2004), pp. 194–199.

[44]   N G Dowell and P S Tofts. "Fast, accurate, and precise mapping of the RF field in vivo using the 180 signal null". In: *Magn Reson Med* 58.3 (2007), pp. 622–630.

[45]   M Droske and M Rumpf. "A variational approach to nonrigid morphological image registration". In: *SIAM J Appl Math* 64.2 (2004), pp. 668–687.

[46]   E Durand. "Comparison of magnetic resonance imaging with radionuclide methods of evaluating the kidney". In: *Semin Nucl Med.* Vol. 44. 2. Elsevier. 2014, pp. 82–92.

[47]   R Ellinger, C Kremser, M F H Schocke, et al. "The impact of peak saturation of the arterial input function on quantitative evaluation of dynamic susceptibility contrast-enhanced MR studies". In: *J Comput Assist Tomogr* 24.6 (2000), pp. 942–948.

[48]   A Elmoataz, M Toutain, and D Tenbrinck. "On the p-Laplacian and ∞-Laplacian on Graphs with Applications in Image and Data Processing". In: *SIAM J Imaging Sci* 8.4 (2015), pp. 2412–2451.

[49]   L Evans. *Partial differential equations.* American Mathematical Society, 1998.

[50]   R Eymard, T Gallouët, and R Herbin. "Finite volume methods". In: *Handbook of numerical analysis* 7 (2000), pp. 713–1018.

[51]   J A Fessler. "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): Applications to tomography". In: *IEEE Trans Image Process* 5.3 (1996), pp. 493–506.

[52]   A E Fick. "Über die Messung des Blutquantums in den Herzventrikeln". In: *Sitzungsberichte der Physikalisch-Medizinischen Gesesllschaft zu Würzburg.* 1870, p. 16.

[53]   J U Fluckiger, M C Schabel, and E V R DiBella. "Model-based blind estimation of kinetic parameters in dynamic contrast enhanced (DCE)-MRI". In: *Magn Reson Med* 62.6 (2009), pp. 1477–1486.

[54]   J Frahm, A Haase, and D Matthaei. "Rapid NMR imaging of dynamic processes using the FLASH technique". In: *Magn Reson Med* 3.2 (1986), pp. 321–327.

[55]   E K Fram et al. "Rapid calculation of T1 using variable flip angle gradient refocused imaging". In: *Magn Reson Imag* 5.3 (1987), pp. 201–208.

[56]   M Furlan, G Marchal, J-M Derlon, et al. "Spontaneous neurological recovery after stroke and the fate of the ischemic penumbra". In: *Ann Neurol* 40.2 (1996), pp. 216–226.

[57]   R Fusco et al. "Influence of parametrization on tracer kinetic modeling in DCE-MRI". In: *Journal of Medical and Biological Engineering* 34.2 (2014), pp. 157–163.

[58] P E Gill, W Murray, and M H Wright. *Practical optimization.* San Diego: Academic press, 1981.

[59] G T Gobbel and J R Fike. "A deconvolution method for evaluating indicator-dilution curves". In: *Phys Med Biol* 39.11 (1994), p. 1833.

[60] S K Godunov. "A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics". In: *Sb Math* 89.3 (1959), pp. 271–306.

[61] B Goldluecke, Strekalovskiy E, and D Cremers. "The Natural Vectorial Total Variation Which Arises from Geometric Measure Theory". In: *Siam J Imaging Sciences* 5.2 (2012), pp. 537–563.

[62] G H Golub and V Pereyra. "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate". In: *SIAM J Numer Anal* 10.2 (1973), pp. 413–432.

[63] G H Golub and C F Van Loan. *Matrix computations.* 4th. J. Hopkins Uni. Press, 2012.

[64] P Gravel, G Beaudoin, and J A De Guise. "A method for modeling noise in medical images". In: *IEEE Trans Med Imaging* 23.10 (2004), pp. 1221–1232.

[65] H Gudbjartsson and S Patz. "The Rician distribution of noisy MRI data". In: *Magn Reson Med* 34.6 (1995), pp. 910–914.

[66] R K Gupta. "A New Look at the Method of Variable Nutation Angle for the Measurement of Spin-Lattice Relaxation Times Using Fourier Transform NMR". In: *J Magn Reson* 25 (1977), pp. 231–235.

[67] E Haber and J Modersitzki. "A multilevel method for image registration". In: *SIAM J Sci Comput* 27.5 (2006), pp. 1594–1607.

[68] N Hackstein, J Heckrodt, and W S Rau. "Measurement of single-kidney glomerular filtration rate using a contrast-enhanced dynamic gradient-echo sequence and the Rutland-Patlak plot technique". In: *J Magn Reson Imaging* 18.6 (2003), pp. 714–725.

[69] A Hallack, M A Chappell, M J Gooding, et al. "A new similarity metric for groupwise registration of variable flip angle sequences for improved T10 estimation in DCE-MRI". In: *Biomedical Image Registration.* Ed. by S Ourselin and M Modat. Cham: Springer International Publishing, 2014, pp. 154–163.

[70] P C Hansen. "Analysis of discrete ill-posed problems by means of the L-curve". In: *SIAM review* 34.4 (1992), pp. 561–580.

[71] S Hartmann. "The world as a process". In: *Modelling and simulation in the social sciences from the philosophy of science point of view.* Springer, 1996, pp. 77–100.

[H2]   C Heck, M Benning, and M Modersitzki. "Joint Registration and Parameter Estimation of T1 Relaxation Times Using Variable Flip Angles". In: *Bildverarbeitung für die Medizin 2015*. Ed. by H Handels, T M Deserno, T-P Meinzer, et al. Heidelberg: Springer Berlin Heidelberg, 2015, pp. 215–220.

[H1]   C Heck et al. "Model-Based Parameterestimation in DCE-MRI Without an Arterial Input Function". In: *Bildverarbeitung für die Medizin 2014*. Ed. by H Handels, T M Deserno, T-P Meinzer, et al. Heidelberg: Springer Berlin Heidelberg, 2015, pp. 246–251.

[72]   R M Henkelman. "Does IVIM measure classical perfusion?" In: *Magn Reson Med* 16.3 (1990), pp. 470–75.

[73]   R M Henkelman. "Measurement of signal intensities in the presence of noise in MR images". In: *Med Phys* 12.2 (1985), pp. 232–233.

[74]   E Hodneland et al. "In vivo estimation of glomerular filtration in the kidney using DCE-MRI". In: *Proc. ISPA*. IEEE. Dubrovnik, 2011, pp. 755–761.

[75]   J Hom et al. "Optimal duration of acquisition for dynamic perfusion CT assessment of blood-brain barrier permeability using the Patlak model". In: *American Journal of Neuroradiology* 30.7 (2009), pp. 1366–1370.

[76]   G-H Jahng et al. "Perfusion magnetic resonance imaging: a comprehensive update on principles and techniques". In: *Korean J Radiol* 15.5 (2014), pp. 554–577.

[77]   Daniel E Jones. "Fourier transform nuclear magnetic resonance III. Spin-lattice relaxation times". In: *J Magn Reson* 6.2 (1972), pp. 191–196.

[78]   *K/DOQI Clinical Practice Guidelines for Chronic Kidney Disease: Evaluation, Classification and Stratification*. Am J Kidney Dis 39: S1-S266 (suppl 1). 2002.

[79]   P Kickingereder, A Radbruch, S Burth, et al. "MR Perfusion–derived Hemodynamic Parametric Response Mapping of Bevacizumab Efficacy in Recurrent Glioblastoma". In: *Radiology* (2015), p. 151172.

[80]   E Klotz and M König. "Perfusion measurements of the brain: using dynamic CT for the quantitative assessment of cerebral ischemia in acute stroke". In: *Eur J Radiol* 30.3 (1999), pp. 170–184.

[81]   F Knoll, M Holler, R Koesters, et al. "Joint MR-PET reconstruction using a multi-channel image regularizer". In: *IEEE Trans Med Imaging* (2016).

[82]   T Knopp and T M Buzug. "How Magnetic Particle Imaging Works". In: *Magnetic Particle Imaging*. Springer, 2012, pp. 11–70.

[83]   K Königsberger. *Analysis 2*. 2nd. Berlin: Springer Verlag, 1997.

[84]   A A Konstas, G V Goldmakher, T-Y Lee, et al. "Theoretic basis and technical implementations of CT perfusion in acute ischemic stroke, part 1: theoretic basis". In: *Am J Neuroradiol* 30.4 (2009), pp. 662–668.

[85] D A Kovar, M Lewis, and G S Karczmar. "A new method for imaging perfusion and contrast extraction fraction: input functions derived from reference tissues". In: *J Magn Reson Imaging* 8.5 (1998), pp. 1126–1134.

[86] K Kudo et al. "Differences in CT Perfusion Maps Generated by Different Commercial Software: Quantitative Analysis by Using Identical Source Data of Acute Stroke Patients 1". In: *Radiology* 254.1 (2010), pp. 200–209.

[87] C L Lawson and R J Hanson. *Solving Least Squares Problems*. Originally published 1974 by Prentice-Hall. Philadelphia: SIAM's Classics in Applied Mathematics, 1995.

[88] S Lefkimmiatis et al. "Convex generalizations of total variation based on the structure tensor with applications to inverse problems". In: *Proc. SSVM 2013*. Graz: Springer, 2013, pp. 48–60.

[89] R J LeVeque. *Finite volume methods for hyperbolic problems*. Vol. 31. Cambridge university press, 2002.

[90] X Li, L R Arlinghaus, G D Ayers, et al. "DCE-MRI analysis methods for predicting the response of breast cancer to neoadjuvant chemotherapy: Pilot study findings". In: *Magn Reson Med* 71.4 (2014), pp. 1592–1602.

[91] Shuaiqi Liu et al. "Medical image fusion based on nuclear norm minimization". In: *International Journal of Imaging Systems and Technology* 25.4 (2015), pp. 310–316.

[92] L Ljung and T Glad. "On global identifiability for arbitrary model parametrizations". In: *Automatica* 30.2 (1994), pp. 265–276.

[93] D C Look and D R Locker. "Time saving in measurement of NMR and EPR relaxation times". In: *Rev Sci Instrum* 41.2 (1970), pp. 250–251.

[94] F Maggi. *Sets of finite perimeter and geometric variational problems: an introduction to geometric measure theory*. Vol. 135. Cambridge University Press, 2012.

[95] F Man, J T Patrie, W Xin, et al. "Delay-sensitive and delay-insensitive deconvolution perfusion-CT: similar ischemic core and penumbra volumes if appropriate threshold selected for each". In: *Neuroradiology* 57.6 (2015), pp. 573–581.

[96] P Mansfield and P G Morris. *NMR Imaging in Biomedicine*. New York: Academic Press, 1982.

[97] MATLAB. *Version 8.6.0 (R2015b)*. Natick, Massachusetts: The MathWorks Inc., 2015.

[98] P Meier and K L Zierler. "On the Theory of the Indicator-Dilution Method for measurement of Blood Flow and Volume". In: *J Appl Physiol* 6.12 (1954), pp. 731–744.

[99]   I A Mendichovszky, M Cutajar, and I Gordon. "Reproducibility of the aortic input function (AIF) derived from dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) of the kidneys in a volunteer study". In: *Eur J Radiol* 71.3 (2009), pp. 576–581.

[100] D R Messroghli et al. "Modified Look-Locker inversion recovery (MOLLI) for high-resolution T1 mapping of the heart". In: *Magn Reson Med* 52.1 (2004), pp. 141–146.

[101] C Michler et al. "A computationally efficient framework for the simulation of cardiac perfusion using a multi-compartment Darcy porous-media flow model". In: *Int J Numer Method Biomed Eng* 29.2 (2013), pp. 217–232.

[102] J Modersitzki. *FAIR: Flexible Algorithms for Image Registration*. Ed. by N J Higham. Vol. 6. Fundamentals of Algorithms. Philadelphia: SIAM, 2009.

[103] J Modersitzki. *Numerical methods for image registration*. Oxford: Oxford University Press, 2003.

[104] M Mokin, C C Ciambella, MW Masud, et al. "Whole-Brain Computed Tomographic Perfusion Imaging in Acute Cerebral Venous Sinus Thrombosis". In: *Interv Neurol* 4.3-4 (2016), pp. 104–112.

[105] T Möllenhoff et al. "Low rank priors for color image regularization". In: *Proc. EMMCVPR 2015*. Springer. Hong Kong, 2015, pp. 126–140.

[106] Susumu Mori, Barbara J Crain, and Peter C Van Zijl. "3D brain fiber reconstruction from diffusion MRI". In: *Neuroimage* 7.4 (1998).

[107] Jorge Nocedal and Stephen Wright. *Numerical optimization*. Second. New York: Springer Science & Business Media, 2006.

[108] W D Obrist et al. "Cerebral blood flow and metabolism in comatose patients with acute head injury: relationship to intracranial hypertension". In: *J Neurosurg* 61.2 (1984), pp. 241–253.

[109] P Ochs et al. "iPiano: Inertial proximal algorithm for nonconvex optimization". In: *SIAM J Imaging Sci* 7.2 (2014), pp. 1388–1419.

[110] A V Oppenheim, R W Schafer, J R Buck, et al. *Discrete-time signal processing*. Vol. 2. Englewood Cliffs, NJ: Prentice Hall, 1989.

[111] L Østergaard et al. "High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: Mathematical approach and statistical analysis". In: *Magn Reson Med* 36.5 (1996), pp. 715–725.

[112] J W Pan, D B Twieg, and H P Hetherington. "Quantitative spectroscopic imaging of the human brain". In: *Magn Reson Med* 40.3 (1998), pp. 363–369.

[113] G J M Parker, C Roberts, A Macdonald, et al. "Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI". In: *Magn Reson Med* 56.5 (2006), pp. 993–1000.

[114] F de Pasquale et al. "Bayesian estimation of relaxation times T1 in MR images of irradiated Fricke-agarose gels". In: *Magn Reson Imag* 18.6 (2000), pp. 721–731.

[115] C S Patlak, R G Blasberg, and J D Fenstermacher. "Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data". In: *J Cereb Blood Flow Metab* 3.1 (1983), pp. 1–7.

[116] C S Patlak, D A Goldstein, and J F Hoffman. "The flow of solute and solvent across a two-membrane system". In: *J Theor Biol* 5.3 (1963), pp. 426–442.

[117] M A Perazella. "Current status of gadolinium toxicity in patients with kidney disease". In: *Clinical Journal of the American Society of Nephrology* 4.2 (2009), pp. 461–469.

[118] W Perl, N A Lassen, and R M Effros. "Matrix proof of flow, volume and mean transit time theorems for regional and compartmental systems". In: *Bull Math Biol* 37 (1975), pp. 573–588.

[119] J R Petrella and J M Provenzale. "MR perfusion imaging of the brain: techniques and applications". In: *Am J Roentgenol* 175.1 (2000), pp. 207–219.

[120] B Recht, M Fazel, and P A Parrilo. "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization". In: *SIAM review* 52.3 (2010), pp. 471–501.

[121] D Y Riabkov and E V R Di Bella. "Estimation of kinetic parameters without input functions: analysis of three methods for multichannel blind identification". In: *IEEE Trans. Biomed. Eng.* 49.11 (2002), pp. 1318–1327.

[122] S O Rice. "Mathematical analysis of random noise". In: *Bell Syst Tech J* 23.3 (1944), pp. 282–332.

[123] R T Rockafellar. *Convex analysis*. Reprinted in 1997. Princeton, New Jersey: Princeton University Press, 1970.

[124] R Rosencranz and S A Bogen. "Clinical laboratory measurement of serum, plasma, and blood viscosity". In: *Am J Clin Pathol* 125 Suppl (June 2006), pp. 78–86.

[125] T J Rothenberg. "Identification in parametric models". In: *Econometrica* 39.3 (1971), pp. 577–591.

[126] L I Rudin, S Osher, and E Fatemi. "Nonlinear total variation based noise removal algorithms". In: *Physica D* 60.1 (1992), pp. 259–268.

[127] W Rudin. *Principles of mathematical analysis*. 3rd. Aukland: McGraw-Hill, 1987.

[128] L Ruthotto. "Hyperelastic Image Registration". Ph.D. Thesis. University of Münster, 2012.

[129] M D Rutland. "A single injection technique for subtraction of blood background in 131I-hippuran renograms". In: *Br J Radiol* 52.614 (1979), pp. 134–137.

[S3]   C Sandmann, J Modersitzki, and E Hodneland. "A Practical Guideline for T1 Reconstruction from Various Flip Angles in MRI". In: *J Algorithm Comput Technol* (2016). accepted.

[S4]   C Sandmann et al. "Limitations of Traditional Models for Perfusion". submitted to IEEE Trans Biomed Eng.

[130] G Sapiro and D L Ringach. "Anisotropic diffusion of multivalued images with applications to color filtering". In: *IEEE T Image Process* 5.11 (1996), pp. 1582–1586.

[131] M Sasaki et al. "Assessment of the accuracy of a Bayesian estimation algorithm for perfusion CT by using a digital phantom". In: *Neuroradiology* 55.10 (2013), pp. 1197–1203.

[132] Joachim Schmitt. "Bestimmung der glomerulären Filtrationsrate bei Kindern und Jugendlichen: Alternative Methoden zur Inulin-Clearance und der Einfluss der laborchemischen Untersuchungsmethoden". Ph.D. Thesis. Philipps-Universität Marburg, 2004.

[133] WP Segars et al. "4D XCAT phantom for multimodality imaging research". In: *Med Phys* 37.9 (2010), pp. 4902–4915.

[134] A M Smith et al. "Whole brain quantitative CBF, CBV, and MTT measurements using MRI bolus tracking: Implementation and application to data acquired from hyperacute stroke patients". In: *J Magn Reson Imaging* 12.3 (2000), pp. 400–410.

[135] J C Sommer. "Regularized estimation and model selection in compartment models". Ph.D. Thesis. LMU München, 2013.

[136] S P Sourbron. "Technical aspects of MR perfusion". In: *Eur J Radiol* 76.3 (2010), pp. 304–313.

[137] S P Sourbron and D L Buckley. "Classic models for dynamic contrast-enhanced MRI". In: *NMR Biomed* 26.8 (2013), pp. 1004–1027.

[138] S P Sourbron and DL Buckley. "On the scope and interpretation of the Tofts models for DCE-MRI". In: *Magn Reson Med* 66.3 (2011), pp. 735–745.

[139] S P Sourbron, H J Michaely, M F Reiser, et al. "MRI-measurement of perfusion and glomerular filtration in the human kidney with a separable compartment model". In: *Invest Radiol* 43.1 (2008), pp. 40–48.

[140] Steven Sourbron. "A tracer-kinetic field theory for medical imaging". In: *IEEE Trans Med Imaging* 33.4 (2014), pp. 935–946.

[141] *Sterbefälle (absolut, Sterbeziffer, Ränge, Anteile) für die 10/20/50/100 häufigsten Todesursachen (ab 1998)*. in www.gbe-bund.de (Startseite – Gesundheitliche Lage – Sterblichkeit – Mortalität und Todesursachen – Filter Tabellen), from 04.05.2016. Primärquelle: Todesursachenstatistik sowie Fortschreibung des Bevölkerungsstandes, Statistisches Bundesamt. Bonn.

[142] G N Stewart. "Researches on the Circulation Time in Organs and on the Influences which affect it". In: *J Physiol* 15.1-2 (1893), pp. 1–89.

[143] M Straka et al. "Real-time diffusion-perfusion mismatch analysis in acute stroke". In: *J Magn Reson Imaging* 32.5 (2010), pp. 1024–1037.

[144] M Straka, R D Newbould, G W Albers, et al. "What is a good sampling rate for DSC-MRI brain perfusion measurements?" In: *Proc. ISMRM*. Toronto, 2008, p. 1899.

[145] G J Strijkers et al. "MRI contrast agents: current status and future perspectives". In: *Anti-Cancer Agent Me* 7.3 (2007), pp. 291–305.

[146] T Taxt, R Jirik, C B Rygh, et al. "Single-channel blind estimation of arterial input function and tissue impulse response in DCE-MRI". In: *IEEE Trans Biomed Eng* 59.4 (2012), pp. 1012–1021.

[147] A N Tikhonov and V I Arsenin. *Solutions of ill-posed problems*. Winston, 1977.

[148] "Total generalized variation based joint multi-contrast, parallel imaging reconstruction of undersampled k-space data". In: *Proc. 23 Mag. Reson. Med.* 2015, pp. 80–.

[149] M I Townsley. "Structure and composition of pulmonary arteries, capillaries, and veins". In: *Compr Physiol* (2012).

[150] R Venkatesan, W Lin, and E M Haacke. "Accurate determination of spin-density and T1 in the presence of RF-field inhomogeneities and flip-angle miscalibration". In: *Magn Reson Med* 40.4 (1998), pp. 592–602.

[151] P H Vivier, P Storey, H Rusinek, et al. "Kidney function: glomerular filtration rate measurement with MR renography in patients with cirrhosis". In: *Radiology* 259.2 (2011), pp. 462–470.

[152] W Walter. *Gewöhnliche Differentialgleichungen*. 6th ed. Berlin Heidelberg New York: Springer, 2000.

[153] D Wang et al. "Concatenated and parallel optimization for the estimation of T1 maps in FLASH MRI with multiple flip angles". In: *Magn Reson Med* 63.5 (2010), pp. 1431–1436.

[154] H Z Wang, S J Riederer, and J N Lee. "Optimizing the precision in T1 relaxation estimation using limited flip angles". In: *Magn Reson Med* 5.5 (1987), pp. 399–416.

[155] H Wang and Y Cao. "Spatially regularized T1 estimation from variable flip angles MRI". In: *Med Phys* 39.7 (2012), pp. 4139–4148.

[156] H J Weinmann, M Laniado, and W Mützel. "Pharmacokinetics of GdDTPA / dimeglumine after intravenous injection into healthy volunteers." In: *Physiol Chem Phys Med NMR* 16.2 (1983), pp. 167–172.

[157] M Wintermark, W S Smith, N U Ko, et al. "Dynamic perfusion CT: optimizing the temporal resolution and contrast volume for calculation of perfusion CT parameters in stroke patients". In: *Am J Neuroradiol* 25.5 (2004), pp. 720–729.

[158] M Wintermark, M Reichhart, J-P Thiran, et al. "Prognostic accuracy of cerebral blood flow measurement by perfusion computed tomography, at the time of emergency room admission, in acute stroke patients". In: *Ann Neurol* 51.4 (2002), pp. 417–432.

[159] R Wirestam, L Andersson, L Ostergaard, et al. "Assessment of regional cerebral blood flow by dynamic susceptibility contrast MRI using different deconvolution techniques". In: *Magn Reson Med* 43.5 (2000), pp. 691–700.

[160] O Wu et al. "Tracer arrival timing-insensitive technique for estimating flow in MR perfusion-weighted imaging using singular value decomposition with a block-circulant deconvolution matrix". In: *Magn Reson Med* 50.1 (2003), pp. 164–174.

[161] O Wu, L Østergaard, and A G Sorensen. "Technical aspects of perfusion-weighted imaging". In: *Neuroimag Clin N Am* 15.3 (2005), pp. 623–637.

[162] T E Yankeelov, J J Luci, M Lepage, et al. "Quantitative pharmacokinetic analysis of DCE-MRI data without an arterial input function: a reference region model". In: *Magn Reson Imaging* 23.4 (2005), pp. 519–529.

[163] V L Yarnykh. "Actual flip-angle imaging in the pulsed steady state: a method for rapid three-dimensional mapping of the transmitted radiofrequency field". In: *Magn Reson Med* 57.1 (2007), pp. 192–200.

[164] G Zaharchuk et al. "Comparison of arterial spin labeling and bolus perfusion-weighted imaging for detecting mismatch in acute stroke". In: *Stroke* 43.7 (2012), pp. 1843–1848.

[165] T Zhang, J M Pauly, and I R Levesque. "Accelerating Parameter Mapping with a Locally Low Rank Constraint". In: *Magn Reson Med* 73.3 (2015), pp. 655–661.

[166] Y D Zhang, C J Wu, J Zhang, et al. "Feasibility study of high-resolution DCE-MRI for glomerular filtration rate (GFR) measurement in a routine clinical modal". In: *Magn Reson Imaging* 33.8 (2015), pp. 978–983.

[167] K L Zierler. "Indicator dilution methods for measuring blood flow, volume, and other properties of biological systems: a brief history and memoir". In: *Ann Biomed Eng* 28.8 (2000), pp. 836–848.

[168] Y Zur, S Stokar, and P Bendel. "An analysis of fast imaging sequences with steady-state transverse magnetization refocusing". In: *Magn Reson Med* 6.2 (1988), pp. 175–193.