Image Reconstruction for Low-Field MRI

de Leeuw den Bouter, M.L.

DOI
10.4233/uuid:f3c4431d-368c-4a17-aacf-2d1283688a1a

Publication date
2022

Document Version
Final published version

Citation (APA)
de Leeuw den Bouter, M. L. (2022). *Image Reconstruction for Low-Field MRI*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:f3c4431d-368c-4a17-aacf-2d1283688a1a

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

IMAGE RECONSTRUCTION FOR LOW-FIELD MRI

IMAGE RECONSTRUCTION FOR LOW-FIELD MRI

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
maandag 2 mei 2022 om 10:00 uur

door

Merel Lisanne DE LEEUW DEN BOUTER

Wiskundig ingenieur, Technische Universiteit Delft, Delft, Nederland,
geboren te Capelle aan den IJssel, Nederland.

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. ir. M.B. van Gijzen,	Technische Universiteit Delft, *promotor*
Dr. ir. R.F. Remis,	Technische Universiteit Delft, *promotor*
Onafhankelijke leden:	
Prof. dr. ir. C. Vuik	Technische Universiteit Delft
Prof. dr. ir. A.W. Heemink	Technische Universiteit Delft
Prof. dr. J. Dankelman	Technische Universiteit Delft
Prof. dr. A.G. Webb,	Universiteit Leiden en Technische Universiteit Delft
Dr. J. Obungoloch,	Mbarara University of Science and Technology, Uganda

Printed by: ProefschriftMaken

Cover design by: Michelle Bettman

This research was supported by NWO-WOTRO (Netherlands Organization for Scientific Research) under grant W07.303.101 and by the TU Delft | Global Initiative, a program of the Delft University of Technology to boost Science and Technology for Global Development.

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

CONTENTS

SUMMARY

Each year, hundreds of thousands of infants develop hydrocephalus ("water on the brain"). This is a disease that, if untreated, leads to brain damage and ultimately death. The prevalence of hydrocephalus is relatively high in children living in the Global South (in sub-Saharan countries, for example), but access to advanced imaging technology is usually limited in countries belonging to the Global South. This is especially problematic for hydrocephalus, since magnetic resonance imaging often is the diagnostic tool of choice for this disease, but MRI scanners are essentially out of reach due to their cost, size, and stringent infrastructure demands. Therefore, the introduction of an inexpensive, portable, low-field MRI scanner is clinically relevant. An interdisciplinary team of researchers at the Leiden University Medical Center, Pennsylvania State University, Mbarara University of Science and Technology and Delft University of Technology has been working on the development of such low-field MRI scanners, with the first goal being to aid in the diagnosis of hydrocephalus in infants in sub-Saharan Africa. Within this project, several prototypes and various dedicated image reconstruction techniques have been developed. This dissertation focuses on the latter.

High-field MRI scanners have very strong and homogeneous static magnetic background fields, due to the superconducting magnets they are equipped with. To significantly reduce production costs, the low-field scanners considered in this work use permanent magnets to realize their static background fields. Obviously, such background fields are much weaker than in a high-field MRI scanner, leading to measured signals with a significantly lower signal-to-noise ratio, since this ratio scales with the magnitude of the background field. For spatial encoding (i.e., to distinguish what part of the signal originates from what part of the body or object inside the scanner), high-field scanners depend on gradient coils which superimpose a linearly varying magnetic field on the background field. The first prototype we consider does not have any gradient coils. Instead, spatial encoding is carried out by making use of the inhomogeneities in the static magnetic background field. Due to the nonbijective nature of the field, a single measurement does not yield enough information for a reconstruction. However, by carrying out several measurements and rotating the field between subsequent measurements, image reconstruction should be possible. The second prototype follows the design of high-field scanners more closely: it was designed such that the static magnetic field is as homogeneous as possible and the scanner is equipped with three gradient coils to allow for spatial encoding in three directions. In this case, the relationship between signal and image can be described by a Fourier Transform.

Initially, we consider the first prototype, and we use our knowledge of the inhomogeneities in the static magnetic background field, combined with a number of different field rotations, to obtain a linear system of equations describing the relationship between signal and image. This leads to an ill-posed inverse problem, which can be mitigated by posing it as a least-squares minimization problem, to which we add a regular-

ization term. Besides making the problem more well-posed, choosing an appropriate regularization operator helps reduce the effect of noise on the image, which is especially relevant for low-field MRI. We reformulate this minimization problem such that the Conjugate Gradient Minimal Error method can be used to solve it for nontrivial regularization matrices. We compare the performance of this method to the state-of-the-art Conjugate Gradient Least Squares algorithm for different regularization terms and show that the Generalized Conjugate Gradient Minimal Error method exhibits much faster convergence for a number of relevant regularization terms.

Subsequently, we use the same algorithm for image reconstruction on data acquired using a prototype of a single-sided portable MRI sensor which is being developed as a spin-off of the original project. Such a scanner could be used to image subcutaneous tissues and the spine, for example. Again, we use the inhomogeneities in the static magnetic field for spatial encoding, but instead of using different rotations, we now translate the object between subsequent measurements. We present some preliminary imaging results obtained in this manner using data acquired with the one-sided MRI sensor.

Using regularization in an additive way introduces an artificial regularization parameter which needs to be selected such that an optimal balance between the data fidelity term and the regularization term is attained. Usually, this parameter tuning is done in a heuristic manner, through extensive numerical experimentation, which can be time-consuming. We introduce multiplicative regularization for low-field imaging. By multiplying the data fidelity term with a regularization term, which is a total variation functional in our case, the problem of having to select such a regularization parameter is eliminated. The resulting minimization problem is solved using a nonlinear Conjugate Gradient algorithm. We apply this method to both simulated data and data acquired using the second low-field MRI scanner and demonstrate that this method can be relied upon to yield good results, both when it is applied in an image reconstruction setting and when it is used for denoising.

Additionally, we consider a scenario in which we are dealing with incomplete data. In MRI, scan times tend to be long and contemporary MRI research has a strong focus on techniques that can help reduce scan duration. In MRI, the acquired data is stored in k-space, which is the spatial frequency domain. By only partially filling k-space, which is equivalent to reducing the number of data points being sampled, scan times can be reduced. Different techniques exist that can overcome the reduction in image quality that is caused by the missing data points. We introduce a straightforward approach that incorporates spatial support information into the problem statement, i.e., we use our knowledge of the (approximate) location of the object to help invert the inverse problem. This problem can be solved using a Conjugate Gradient algorithm. We demonstrate that our relatively simple approach yields results that are of a quality comparable to images obtained using a significantly more complex compressed sensing framework.

Finally, we focus on deep learning approaches to tackle two different imaging challenges. The first one is the correction of image distortions which can occur due to inhomogeneities in the static magnetic background field and nonlinearities in the magnetic fields generated by the gradient coils. We train a neural network on a simulated dataset consisting of undistorted images and their noisy, distorted counterparts, with each input-output pair corresponding to different perturbations in the magnetic fields.

The resulting network is then applied to a number of distorted images acquired using the second low-field MRI scanner.

Instead of sampling only part of k-space to reduce scan time, one can also consider acquiring an image of a lower resolution. Therefore, as our second challenge, we consider a single image of a relatively low number of pixels, whose resolution we aim to increase using a neural network. For the training stage, we rely on a high-field dataset of brain images. By artificially downsampling these images and subsequently adding noise in the k-space domain, we simulate the acquisition of low-resolution low-field MR images. After training the network on these input-output pairs, we apply it to low-field brain images acquired using the second MRI scanner and obtain good quality super-resolution images.

SAMENVATTING

Op jaarbasis ontwikkelen honderdduizenden kinderen hydrocephalus ("waterhoofd"), een ziekte die, indien deze onbehandeld blijft, leidt tot hersenschade en uiteindelijk de dood tot gevolg heeft. De prevalentie van hydrocephalus is relatief hoog bij kinderen in het Mondiale Zuiden (bijvoorbeeld in sub-Sahara Afrika), maar de toegang tot geavanceerde beeldvormingstechnologie is daar meestal beperkt. Dit is problematisch in het geval van hydrocephalus, omdat magnetic resonance imaging als diagnostisch hulpmiddel meestal de voorkeur geniet, maar MRI scanners in wezen onbereikbaar zijn vanwege hun kosten, grootte en de strenge eisen waaraan de infrastructuur moet voldoen. Dit maakt de introductie van een betaalbare, mobiele MRI scanner met een lage veldsterkte klinisch relevant. Een interdisciplinair team van onderzoekers van het Leiden Universitair Medisch Centrum, Pennsylvania State University, Mbarara University of Science and Technology en de Technische Universiteit Delft reeds een aantal jaren werkt aan het ontwikkelen van dergelijke MRI scanners met een lage veldsterkte, waarbij het eerste doel is om te kunnen assisteren bij het diagnosticeren van hydrocephalus bij kinderen in landen in sub-Sahara Afrika. Binnen dit project zijn een aantal prototypes en verschillende toegespitste beeldreconstructietechnieken ontwikkeld. Deze dissertatie richt zich op het laatstgenoemde.

Hoge veldsterkte MRI scanners hebben zeer sterke en homogene statische magnetische achtergrondvelden, vanwege de supergeleidende magneten waarmee ze zijn uitgerust. Om de kosten significant te reduceren, gebruiken de lage veldsterkte MRI scanners die we in dit werk beschouwen permanente magneten om hun statische achtergrondvelden te realiseren. Uiteraard zijn dergelijke achtergrondvelden veel zwakker dan in een MRI scanner met een hoge veldsterkte. Dit leidt ertoe dat de signalen die gemeten worden een significant lagere signaal-ruisverhouding hebben, omdat deze verhouding schaalt met de sterkte van het magnetisch veld.

Voor spatiële codering (dat wil zeggen, om te kunnen onderscheiden welk deel van het signaal door welk deel van het lichaam of object in de scanner wordt opgewekt) gebruiken hoge veldsterkte MRI scanners gradiëntspoelen die een lineair variërend veld bovenop het achtergrondveld genereren. Het eerste prototype dat we beschouwen heeft geen gradiëntspoelen. In plaats daarvan vindt spatiële codering plaats door gebruik te maken van inhomogeniteiten in het statische magnetische achtergrondveld. Doordat het veld niet bijectief is, levert een enkele meting niet genoeg informatie op voor een reconstructie. Echter, door verschillende metingen uit te voeren en het veld te roteren tussen opeenvolgende metingen, zou het mogelijk moeten zijn om een beeld te reconstrueren. Het tweede prototype is meer conform het ontwerp van een hoge veldsterkte MRI scanner: het is zo ontworpen dat het achtergrondveld zo homogeen mogelijk is en het is uitgerust met drie gradiëntspoelen om spatiële codering in drie dimensies mogelijk te maken. In dit geval kan een Fourier Transform worden gebruikt om de relatie tussen signaal en afbeelding te beschrijven.

Eerst beschouwen we het eerstbeschreven prototype, en we gebruiken onze kennis van de inhomogeniteiten in het statische magnetische achtergrondveld, gecombineerd met een aantal verschillende veldrotaties, om een lineair stelsel van vergelijkingen te verkrijgen dat de relatie tussen signaal en afbeelding beschrijft. Dit leidt tot een slechtgesteld invers probleem, wat kan worden verholpen door het te poneren als een minimimalisatieprobleem in de vorm van een kleinste-kwadratenterm, waarbij we een regularisatieterm optellen. Behalve dat dit het probleem meer welgesteld maakt, zorgt het kiezen van de juiste regularisatieoperator ervoor dat het effect van ruis op de afbeelding beperkt wordt, wat bijzonder relevant is voor lage veldsterkte MRI. We herformuleren het minimalisatieprobleem dusdanig dat de Conjugate Gradient Minimal Error methode gebruikt kan worden om het op te lossen voor niet-triviale regularisatiematrices. We vergelijken de prestaties van deze methode met het gangbare Conjugate Gradient Least Squares Algoritme voor verschillende regularisatietermen en laten zien dat de Generalized Conjugate Gradient Minimal Error methode veel snellere convergentie vertoont voor een aantal relevante regularisatietermen.

Vervolgens gebruiken we hetzelfde algoritme om afbeeldingen te reconstrueren met data verkregen met een unilaterale mobiele MRI sensor die ontwikkeld wordt als een spin-off van het originele project. Een dergelijke scanner zou bijvoorbeeld gebruikt kunnen worden om onderhuidse weefsels en de ruggengraat in beeld te brengen. Wederom gebruiken we de inhomogeniteiten in het statische magnetische veld voor spatiële codering, maar in plaats van verschillende rotaties te gebruiken, transleren we het object nu tussen de metingen in. We presenteren enkele voorlopige resultaten in de vorm van afbeeldingen die we hebben gereconstrueerd met data uit de unilaterale MRI sensor.

Als regularisatie op een additieve manier wordt uitgevoerd, introduceert dat een artificiële regularisatieparameter die geselecteerd moet worden zodanig dat een optimale balans tussen de dataterm en de regularisatieterm wordt bewerkstelligd. Meestal wordt deze parameter afgestemd op een heuristische manier, door middel van uitgebreide numerieke experimentatie, wat tijdrovend kan zijn. Wij introduceren multiplicatieve regularisatie voor beeldvorming in de context van lage veldsterkte MRI. Door de dataterm met een regularisatieterm te vermenigvuldigen, wat een total variation functionaal is in ons geval, wordt het probleem van het selecteren van een regularisatieparameter geëlimineerd. We lossen het voortvloeiende minimalisatieprobleem op met een nonlineair Conjugate Gradient algoritme. We passen deze methode toe op gesimuleerde data en op data verkregen met het tweede lage veldsterkte MRI prototype en laten zien dat deze methode gebruikt kan worden om goede resultaten te bereiken, zowel wanneer deze gebruikt wordt in een reconstructieprobleem als om een afbeelding te ontdoen van ruis.

Vervolgens beschouwen we een scenario waarin we te maken hebben met incomplete data. MRI scans duren meestal lang en hedendaags MRI onderzoek heeft een sterke focus op technieken die kunnen helpen om de scanduur te reduceren. In MRI wordt de vergaarde data opgeslagen in de k-ruimte, oftewel het spatiële frequentiedomein. Door de k-ruimte slechts gedeeltelijk te vullen, wat equivalent is aan het reduceren van het aantal datapunten dat wordt bemonsterd, kan de scanduur worden verminderd. Er bestaat een verscheidenheid aan technieken die gebruikt kunnen worden om de reductie in beeldkwaliteit, die gepaard gaat met een kleiner aantal datapunten, te compenseren. We introduceren een eenvoudige aanpak waarbij we het inverse probleem oplossen door

onze kennis over waar het object zich (bij benadering) bevindt te verwerken in de probleembeschrijving. Dit probleem lossen we wederom op met een Conjugate Gradient algoritme. We tonen aan dat onze relatief eenvoudige aanpak resultaten oplevert die van een kwaliteit zijn die vergelijkbaar is met afbeeldingen verkregen met een aanpak gestoeld op compressed sensing, die significant complexer is.

Tot slot richten we ons op deep learning technieken om twee verschillende beeldvormingsproblemen aan te pakken. Het eerste is het corrigeren van vervormingen in afbeeldingen die kunnen optreden door inhomogeniteiten in het achtergrondveld en door nonlineariteiten in de magnetische velden die worden opgewekt door de gradiëntspoelen. We trainen een neuraal netwerk op een gesimuleerde dataset bestaande uit niet-vervormde afbeeldingen en hun ruizige vervormde tegenhangers, waarbij elk invoer-uitvoerpaar overeenkomt met andere verstoringen in de magnetische velden. Het voortvloeiende netwerk wordt vervolgens toegepast op een aantal vervormde afbeeldingen die we verkregen hebben uit data van de tweede lage veldsterkte MRI scanner.

In plaats van het slechts deel bemonsteren van de k-ruimte om de scanduur te verminderen, kan men ook overwegen om een afbeelding van een lagere resolutie te vergaren. Als onze tweede uitdaging beschouwen we daarom een afbeelding met een relatief klein aantal pixels, waarvan we de resolutie willen vergroten met een neuraal netwerk. Voor de trainingsfase beroepen we ons op een dataset van breinafbeeldingen verkregen met hoge veldsterkte MRI scanners. Door de resolutie van deze afbeeldingen kunstmatig te verkleinen en ruis toe te voegen aan de overeenkomstige data in de k-ruimte, simuleren we de acquisitie van lage veldsterkte MRI afbeeldingen met een lage resolutie. Na het netwerk getraind te hebben op deze invoer-uitvoerparen, passen we het toe op breinafbeeldingen verkregen met de tweede MRI scanner. We concluderen dat de superresolutie resultaten van goede kwaliteit zijn.

1

INTRODUCTION

1.1. BACKGROUND

Magnetic Resonance Imaging (MRI) scanners use strong magnetic fields to visualize the internal structure of the human body in a non-invasive manner. Unfortunately, access to MRI technology is largely confined to countries belonging in the upper-middle and high-income brackets. In a 2016 paper [1], it was calculated that there was 1 MRI scanner per 25 million people in SSA and 25 MRI scanners per 1 million people in the northern hemisphere. According to a 2019 paper [2], 11 countries in Africa, the largest of which has a population of 65 million, have no access to MRI technology at all. For a detailed overview on global MR accessibility, the reader is referred to [2].

There are several reasons for the lack of MRI scanners in SSA and low- and middle-income countries in general. Commercial MRI scanners, which usually have field strengths of 1.5 or 3 T, cost about 1 million euros per tesla [3]. Additionally, annual service contracts cost hundreds of thousands of euros [4]. Finally, in order to operate and repair MRI scanners, a high level of expertise is required.

Therefore, the introduction of low-cost, portable low-field MRI scanners is of great clinical relevance for low- and middle-income countries, although it should be noted that the impact of such scanners could potentially reach even further than that. An interdisciplinary team of researchers and engineers from the Leiden University Medical Center (LUMC, The Netherlands), Pennsylvania State University (PSU, United States of America), Mbarara University of Science and Technology (MUST, Uganda) and Delft University of Technology (TUD, The Netherlands) has been developing a low-cost, portable head-only low-field MRI scanner since 2017. We aim to keep the costs of this scanner below 50,000 euros and to make it straightforward to install, operate and maintain. To truly make MRI technology available to everyone, all software developed within this project is open source and by making the complete technical documentation available, the hardware is as well. This project is part of the Open Source Imaging Initiative (https://www.opensourceimaging.org) [5], which aims to make modern healthcare instruments available to a wider audience.

We consider resolving hydrocephalus the first milestone such a scanner should be able to achieve, since visualizing infant hydrocephalus is arguably one of the simplest challenges in MRI. Hydrocephalus is one of the most prevalent childhood neurological disorders [6]. As stated in [7]: "Although a precise definition is controversial, hydrocephalus generally refers to a disorder of cerebrospinal fluid (CSF) physiology resulting in abnormal expansion of the cerebral ventricles, typically associated with increased intracranial pressure." If untreated, hydrocephalus can lead to macrocephaly, cognitive dysfunction, and death [8]. In [8], it is estimated that each year, approximately 180,000 infants develop hydrocephalus in Africa, and 380,000 worldwide. It stands to reason that most of these 180,000 cases occur in sub-Saharan Africa (SSA), as the vast majority of the African population lives there. The elevated prevalence of infant hydrocephalus in SSA is due to the high birth rate as well as the very high incidence of hydrocephalus arising from infections such as ventriculitis and meningitis [9–11].

Magnetic resonance imaging is generally the diagnostic tool of choice for the treatment planning and monitoring of hydrocephalus, because unlike computed tomography (CT), which is another viable option, it does not require potentially dangerous ionizing radiation [12] to visualize the internal structure of the human body. Even for infants with an open fontanel, who can be screened using cranial ultrasonography, an MRI scan typically provides more information about cause and anatomy [7].

The vital step in visualizing hydrocephalus is the segmentation of the image into brain and CSF [13]. This is a straightforward task, since CSF and white/grey matter have very different T_1 times [14]. Additionally, to locate fluid accumulations in the brain, a relatively low spatial resolution is sufficient. In [15] it is reported that values of 2-3 mm in the in-plane direction, and 5 mm in the through-plane direction are required.

1.2. MAGNETIC RESONANCE IMAGING

MRI scanners consist of three main components: a main magnet, gradient coils and a radiofrequency (RF) system [16]. In conventional MRI, superconducting magnets are generally employed to generate a homogeneous static magnetic background field with a strength of several tesla. Such a strong field must be generated at a temperature of 4.2 K, which means that cryogenic cooling is required [17]. The magnet and cryostat make up a significant part of the total cost of an MRI scanner [18]. Secondly, there are three gradient coils, which are designed to generate time-dependent, spatially nonuniform magnetic fields. Usually, each of the gradient coils produces a magnetic field that varies linearly in one direction. The gradient coils allow for spatial encoding of the signal. Finally, the RF system consists of a transmitter coil that is responsible for the excitation of the spins inside the scanner, and a receiver coil which picks up the corresponding signal [19].

Roughly speaking, MRI works as follows. All physical objects consist of atoms, which in turn consist of nuclei and the electrons orbiting them. Nuclei of atoms with an odd atomic weight, such as the hydrogen atom, have an angular momentum called spin, which generates its own microscopic magnetic field (or magnetic moment). In the absence of an external magnetic field, the orientation of each of these is random. However, when an external magnetic field is applied, the magnetic moments align themselves in such a way that a net magnetic moment is generated. In a conventional MRI scanner, this external magnetic field is generated by the superconducting magnets. The precessional

frequency of the spins is linearly dependent on the strength of the magnetic field. By applying a linearly varying field on top of the strong homogeneous field, this frequency is made to vary linearly as well, which allows for spatial encoding. Applying such a gradient during readout enables frequency encoding, while the presence of a gradient before readout allows for phase encoding. Just before readout, a radiofrequency (RF) pulse is applied to change the orientation of the magnetic moments. After this, the magnetic moments gradually relax back to their equilibrium positions. This relaxation process generates a signal in the receiver coil. By carrying out this procedure repeatedly with different gradient strengths, the so-called k-space (or spatial Fourier domain) is filled with the different measurements. The image can subsequently be obtained by applying an inverse Fourier transform to this k-space matrix [19].

1.3. LOW-FIELD MRI

The stronger the magnetic field, the higher the signal-to-noise ratio (SNR). Eliminating the superconducting magnet from the MRI system, while significantly reducing the cost, inevitably leads to a weaker magnetic field and a several hundredfold reduction in SNR [20]. However, as Sarracanie and Salameh state in [21]: "One of the main misconceptions is that low-field MRI translates into poor image resolution, often associated with poor image quality. It is important, as scientists, to state that this concept is purely and simply wrong. Magnetic field strength has by no means ever been a limit to an achievable image resolution." In the same paper, the authors go on to explain that over the last 4 decades, the difference in images acquired using 1.5 T scanners shows that field strength is certainly not the only factor determining image quality, but instead, modern electronics, high performance acquisition techniques, high sensitivity detectors and image processing all play an important role. We believe that by using these and other advancements, low-field MR image quality can be improved to a sufficient level for diagnostic purposes as well.

Using low-field MRI to image the human body is no new concept [20]. The usefulness of low-field MRI was identified from the onset of MRI in the 1980s [13]. For example, Sepponen et al. [22] used static low-field MRI with a main field strength of 20 mT to study cerebral lesions in 1985. In 1992, do Nascimento et al. [23] demonstrated the feasibilty of human brain imaging using a home-built low-field MRI scanner operating at 16 mT. In 1993, Macovski and Connolly [24] showed that human hand and wrist imaging is possible in a homogeneous, weak magnetic field, after the application of a pre-polarizing pulsed magnetic field. This technique allows for signal acquisition at very low frequencies, while benefiting from the higher SNR provided by the pre-polarizing field.

Furthermore, a notable group of detectors is formed by the superconducting quantum devices (SQUIDs), see for example [25–27], that work in the microtesla range. More recently, Sarracanie et al. [28] and Lother et al. [15] produced Helmholtz coil-based systems designed for brain imaging operating at 6.5 and 23 mT respectively. The former team acquired high quality images of the human brain. However, the size and weight of these scanners are an impediment to making them truly portable.

The past few years have seen the advent of low-field MRI scanners based on Halbach cylinders (or Halbach arrays). The Halbach cylinder, which was introduced in 1980 by Klaus Halbach [29], is a configuration of permanent magnets that, combined, yield a ho-

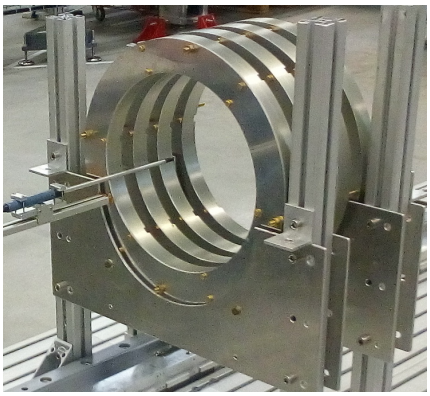
mogeneous static magnetic field, while, ideally, keeping the magnetic field strength outside the bore equal to zero. In practice, due to the discrete nature of the magnets and the finite length of the bore, the generated field will never be perfectly homogeneous. Cooley et al. [30, 31] showed that two-dimensional MR image reconstruction using a low-field MRI scanner based on a Halbach cylinder is feasible. The magnetic field generated by such a configuration of permanent magnets is oriented in the transverse direction, unlike in high-field MRI scanners. The field was mapped accurately and they incorporated this in the model describing the relationship between signal and image. To eliminate the difficulties caused by the non-bijective mapping from field to image, they carried out their imaging experiments by rotating the Halbach array between acquisitions. The Cooley group also introduced a low-field MRI scanner which, besides a built-in gradient in the static background field, uses two coil-based gradients for phase-encoding [32]. Ren et al. [33, 34] designed a portable low-field scanner for 2D head imaging based on a Halbach array in which the magnets are arranged in an inward-outward pair of rings, generating a magnetic field which is oriented along the bore, similar to high-field MRI scanners.

1.3.1. SCANNERS DEVELOPED WITHIN THIS PROJECT

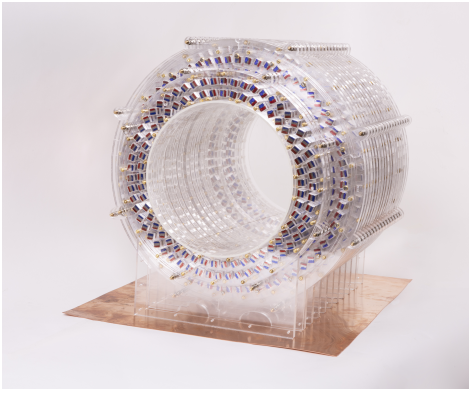
The PSU team developed a low-field MRI scanner using a pre-polarizing magnetic field, as described in [13], see Figure 1.1a. This scanner, which was designed such that it can be assembled manually using off-the-shelf electronics and materials that can be obtained easily. It has a Field of View (FoV) with a 22 cm diameter of spherical volume, a background magnetic field of 2.66 mT and a pre-polarizing magnetic field with a strength of 27 mT and is equipped with 3 gradient coils.

As a joint venture, the LUMC and TUD teams developed a low-field MRI scanner based on a Halbach cylinder. The original design, which can be seen in Figure 1.1b, consisted of 4 rings of 24 N52 neodymium magnets of $25 \times 25 \times 25$ mm^3 that generated a static magnetic field of about 60 mT. At a later stage, an additional ring of magnets designed specifically for shimming was added, as well as a ring to superimpose a gradient on top of the background field. A thorough description of this scanner can be found in [35]. The idea was, as Cooley et al. did in [30], to use an accurate map of the magnetic field for spatial encoding and to carry out several measurements of the same object, while the field (or object) was rotated between measurements. The rotations were meant to eliminate any problems arising because of the nonbijectivity of the magnetic field. Many insights were gained from this design but acquiring images of good quality proved too challenging. The idea of using different rotated versions of the same magnetic field to obtain sufficient information for image reconstruction was abandoned. Instead, the LUMC team designed a new scanner, see Figure 1.1c, with three coil-based gradients, as described in [36], to be used for frequency- and phase-encoding, which make 3D image reconstruction possible. Additionally, the configuration of the Halbach array was altered significantly: the second scanner, with a bore diameter of 27 cm and a length of 50 cm, consists of 2948 smaller N48 neodymium magnets of $12 \times 12 \times 12$ mm^3 distributed over 23 rings, which yield a more homogeneous field of approximately 50 mT. At a later stage, additional shimming rings were added to increase the homogeneity even more. A thorough description of the scanner can be found in [20] and [37].

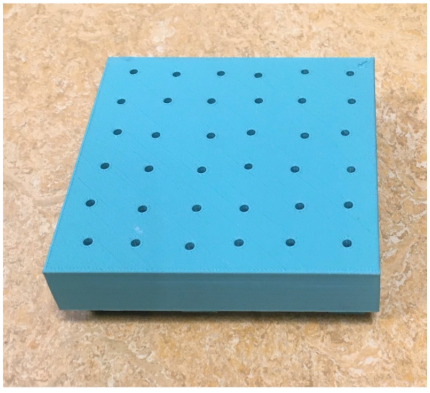
(a) PSU prototype



(b) LUMC/TUD prototype



(c) LUMC prototype



(d) LUMC single-sided MRI sensor

Figure 1.1: MRI scanners developed within this project

As a spin-off of the original project, a single-sided low-field MRI sensor was developed at the LUMC. The prototype, which is shown in Figure 1.1d, consists of a configuration of 36 neodymium magnets of $12 \times 12 \times 12$ mm^3 arranged in a square grid and has a size of $10 \times 10 \times 2$ cm^3. The configuration of magnets generates a magnetic field which is oriented mainly in the direction perpendicular to the surface of the magnet. It is approximately linear around a height of 2 mm above the magnet, which allows for slice selection. Over the course of the last few decades, several groups have developed single-sided MRI (or NMR) sensors, see for example [38–43], and, to the best of our knowledge, 2D and 3D images have only ever been acquired by mounting gradient coils on the sensor, to allow for spatial encoding in two or three dimensions. Our approach is different. The magnetic field generated by the magnets is not perfectly linear, but if we map the inhomogeneities accurately, we can use those for spatial encoding. Note that this is equivalent to the approach we aimed to take with the first LUMC/TUD prototype. However, instead of using

rotations to acquire a sufficient amount of information for a reconstruction, we translate the object in between measurements. Using a single-sided sensor places no restrictions on the size of the object being imaged. Additionally, this prototype is lightweight enough to be portable. Therefore, the introduction of such a sensor could potentially increase MRI accessibility. Possible applications include the imaging of subcutaneous structures and the spine.

1.4. IMAGE RECONSTRUCTION IN (LOW-FIELD) MRI

There is a vast literature on the subject of image reconstruction in MRI, which in itself covers a myriad of different aspects. Therefore, it is impossible to give a comprehensive overview of this topic here. We will focus on the subtopics that have been relevant within the low-field MRI project.

1.4.1. MODEL-BASED IMAGE RECONSTRUCTION

In conventional MRI, the relationship between signal (or k-space) and image is governed by a Fourier Transform. However, for this relationship to be valid, we need the assumptions of a homogeneous static background field and linear gradients to hold. If they do not, as was the case for the first low-field MRI prototype that was developed within the scope of our project, it is necessary to take these "imperfections" in the field into account. We chose to do that via model-based image reconstruction [44], where the accurate knowledge of the magnetic fields is used to build a matrix that relates the signal to the image in a linear fashion. If the static background field is homogeneous and the gradients perfectly linear, this matrix is equal to the FFT operator, in which case calculations with this operator and its inverse can be carried out very efficiently. However, in general it is not straightforward to find a solution in a direct manner. Therefore, iterative methods are usually applied to solve the inverse problem of converting signal to image.

1.4.2. REGULARIZATION

To find a(n approximate) solution to the linear system of equations, we can minimize the corresponding least-squares (data fidelity) term. Such a minimization problem is sometimes called a variational problem. It can be beneficial to incorporate a regularization term/penalty in the minimization problem, which enforces prior information that may be available about the solution. For instance, in case the measured signal is corrupted by a lot of noise, such a regularization term can help suppress the effects of noise on the resulting image. In this work, we will see several instances where a total variation (TV) penalty is incorporated in the minimization problem. TV regularization is a popular regularization method that penalizes jumps between neighboring pixels. The rationale behind using such a regularizer in MRI is based on the observation that neighboring pixels are likely to represent the same tissue and hence have the same intensity.

Regularization is usually carried out in an additive manner: the sum of a least-squares term and a regularization term constitutes the minimization problem, with a regularization parameter determining how much each term should contribute. In a noisy setting where TV is used to regularize the problem, the solution will be either overly smooth or noisy if the parameter is not tuned correctly. While methods exist that select this param-

eter automatically, see for example [45], this is usually simply done through extensive numerical experimentation. In this work, we look at regularization in an additive manner, but we also investigate multiplicative regularization for low-field MR imaging.

1.4.3. DEEP LEARNING

One of the drawbacks of solving a nonlinear image reconstruction problem, as is the case when TV regularization is used, is the long reconstruction time. Secondly, TV or any other type of sparsifying regularization can lead to reconstructions that appear overly blocky or unnatural, reducing acceptance by clinicians. Moreover, the regularization parameter needs to be selected, as well as (hyper)parameters of the algorithm that is used to solve the minimization problem, which can result in reconstructions of poor quality. One way to circumvent these problems is by using deep learning, which means that an artificial neural network is trained on a (large) dataset of available images and their corresponding k-space data. By training the network on appropriate data, it "learns" to reconstruct images and this process should, in general, teach the network to generalize to unseen data. Of course, the network needs to be trained, which can take a long time, and (hyper)parameters of the network should be chosen in advance, but after training, the problem of parameter selection is eliminated and reconstructions can be carried out rapidly.

Over the past few years, an abundance of deep learning techniques has been applied to MRI data and images. For a review on the application of deep learning to all aspects of MRI, the reader is referred to [46]. In [47], an overview on MR image reconstruction using deep learning can be found. In that paper, two different types of artifical neural networks are considered: those that are based on unrolled algorithms and those that are not. Neural networks of the first type typically attempt to find a solution to a variational problem which takes the mathematical model relating k-space data and image into account. Then, the network is based on an iterative reconstruction algorithm with a predetermined number of steps where each of the steps is explicitly incorporated into the network ("unrolling") and the regularization parameters and regularizing functions are learned through training. This effectively eliminates the choosing of hyperparameters for each new reconstruction problem, while still staying true to the structure of the variational model. Examples of unrolled neural networks are ISTA-net [48] (based on Iterative Shrinkage Thresholding Algorithm (ISTA)), ADMM-net [49] (based on the Alternating Direction Method of Multipliers (ADMM)), the Learned Primal-Dual algorithm [50] and the Variational Network [51]. The networks that are not based on unrolled algorithms typically use more standard network architectures. A neural network of this type that has received a lot of attention in the MRI community is the automated transform by manifold approximation (AUTOMAP) network by Zhu et al. [52], which directly learns a mapping from k-space data to image.

1.4.4. COMPRESSED SENSING

In MRI, scan times are typically long and a significant amount of contemporary research focuses on speeding up the acquisition process while maintaining image quality, with the focus usually being on compressed sensing (CS) techniques. Typically, acquisition speedup is accomplished by sampling only a subset of k-space. Traditionally, sampling

k-space is done by following Cartesian trajectories, but more complex trajectories, like radial and spiral ones, can be used as well. If only a subset of k-space is acquired, some of the pre-defined trajectories are skipped. This can be done in a random or a non-random fashion.

1.4.5. SUPER-RESOLUTION IMAGE RECONSTRUCTION

Super-resolution (SR) image reconstruction describes the process of using a number of low-resolution (LR) images to obtain a high-resolution (HR) image, see for example [53]. We can differentiate between 2 cases: single image super-resolution (SISR), in case only one LR image is available, and multiple image super-resolution (MISR), when the information contained in several different images of the same object is combined to yield one HR image. The super-resolution problem is an ill-posed one. The algorithms used to tackle this problem can be grouped into three categories: interpolation-based methods, reconstruction-based methods and learning-based algorithms [49]. Interpolation-based methods, like bicubic interpolation and k-space zero-padding, are generally efficient but lack accuracy in their ability to reconstruct high-frequency components. Reconstruction-based methods solve a minimization problem that incorporates prior knowledge about the solution. These methods usually perform better than interpolation-based methods, but they are computationally expensive and require extensive hyperparameter tuning. Learning-based methods, which learn relationships between HR and LR image pairs from data, tend to outperform the other methods. Of the learning-based category, deep learning-based methods yield the best results. For MRI, super-resolution techniques are interesting in case we have only LR images are available, which can occur when it is imperative to have short scanning times, in the case of subject motion or because of SNR considerations [54]. The first two considerations are especially relevant here because the low-field MRI scanner is meant to be used on infants, for whom lying still could prove challenging.

1.4.6. CORRECTING IMAGE DISTORTIONS

In low-field MRI, and to a lesser extent in high-field MRI, inhomogeneities are present in the static background field and gradients can deviate from the ideal linear scenario. When these imperfections are not taken into account and an inverse Fourier Transform is applied to the k-space data, distortions occur in the resulting images. It can be troublesome to obtain accurate maps of the magnetic fields. We present a deep learning approach that corrects simple images without explicit maps of the magnetic fields.

1.5. CONTRIBUTIONS AND OUTLINE

In this thesis, we focus on different aspects of image reconstruction and image processing for low-field MRI. The main contributions and outline of this thesis are as follows:

1. In Chapter 2, we first present a low-field MRI signal model that can be employed to reconstruct images from signals acquired using a low-field MRI scanner without gradient coils. We use this model to simulate the signal generation in a low-field MRI scanner. In low-field MRI, we generally need to include regularization to counteract the effects of a low SNR. To solve the resulting minimization prob-

lem, we introduce a new generalization of the conjugate gradient method for regularized and weighted least-squares problems: Generalized Conjugate Gradient Minimal Error (GCGME). We compare the performance of GCGME to the standard Generalized Conjugate Gradient Least Squares (GCGLS) method and conclude that GCGME outperforms GCGLS in terms of convergence speed in case a total variation function is used as a regularizer.

2. In Chapter 3, we reconstruct images from data acquired with a single-sided handheld low-field MRI sensor. Such a device is being developed as a spin-off of the original project. We show preliminary results that were obtained using data acquired with this sensor.

3. We discuss multiplicative regularization for low-field MRI in Chapter 4. Usually, regularization is carried out in an additive manner, which necessitates the inclusion of an artificial regularization parameter in the minimization problem. Tuning such a parameter typically requires extensive numerical experimentation which is not desirable in a clinical setting. By multiplying the least-squares term by a regularizing functional (instead of adding), the need for such a parameter is eliminated. We apply this technique to simulated data and measured data acquired using the LUMC low-field MRI scanner with gradient coils and demonstrate that this type of regularization has good denoising capabilities.

4. Chapter 5 introduces an imaging method that, given undersampled k-space data, uses (approximate) support information of the object being scanned to improve image quality. It hinges on a straightforward equation that can be solved using a conjugate gradient algorithm. Usually, compressed sensing techniques (in which support information is generally not taken into account) are used to carry out image reconstruction in the case of undersampled k-space data. However, these techniques tend to be computationally expensive and require the tuning of regularization parameters. We show that, for simulated and measured low-field MRI data sets acquired with the LUMC low-field MRI scanner, our method yields results which are of a similar quality as images obtained using a compressed sensing approach. Additionally, our method is much simpler to implement.

5. In Chapter 6, we discuss the occurrence of image distortions because of gradient nonlinearities and inhomogeneities in the static magnetic background field. Our aim is to use a deep learning-based approach to correct for such distortions in the acquired images. We use numerical simulations to generate a large dataset of undistorted images and their distorted, noisy counterparts. Each individual input-output sample in the training set corresponds to a different set of perturbed magnetic fields, because we strive to make the network applicable to images obtained using different low-field MRI scanners. Subsequently, we train a neural network to transform the distorted images back to their undistorted versions. The trained network is then used to correct images acquired using different versions of the LUMC low-field MRI scanner. We show preliminary results obtained using this deep learning approach.

6. Chapter 7 describes a deep learning-based super-resolution technique. We focus on single image super-resolution for low-field MR brain images. We take high-resolution images from a database of high-field MR images and generate their noisy, low-resolution counterparts. Subsequently, we train a neural network on these input-output pairs. After training, we apply the network to different low-field brain images. The super-resolution images produced by the network are of a similar quality compared to the high-resolution images that were acquired using the LUMC low-field scanner. However, the scan time needed to obtain these low-resolution images is four times shorter.

2

CG VARIANTS FOR ℓ_p-REGULARIZED IMAGE RECONSTRUCTION

2.1. INTRODUCTION

In this chapter, we consider the low-field MRI scanner shown in Figure 2.1a. Its design, which is described in more detail in [35], is based on a Halbach cylinder, i.e., a configuration of permanent magnets that generates a magnetic field inside the cylinder and a very weak, or in the ideal case, no magnetic field outside of it. Imaging can be done by making use of the variations in the magnetic field. However, due to the nonlinearity of the magnetic field inside the Halbach cylinder, the resulting reconstruction problem is very ill-posed. This field leads to non-bijective mappings and potentially gives rise to aliasing artifacts in the solution. Additionally, in the center of the cylinder, there is very little variation in the field, limiting the spatial resolution in that area. Another complication we face is low signal-to-noise ratios. Nevertheless, in a similar project, Cooley et al. [30] have shown that it is possible to reconstruct magnetic resonance images given signals obtained with a device based on a Halbach cylinder, using a simplified signal model in which similar assumptions are made as in high-field MRI. In this chapter, we revisit the underlying physics and formulate the general signal model for MRI without making these assumptions.

This chapter is based on the article:

M.L. de Leeuw den Bouter, M.B. van Gijzen, and R.F. Remis, *Conjugate gradient variants for ℓ_p-regularized image reconstruction in low-field MRI*, SN Applied Sciences **1**, 1736 (2019).

Regularization is required to limit the influence of noise on the solution of the image reconstruction problem as much as possible. In this chapter, we reformulate the weighted and regularized least-squares problem such that the Conjugate Gradient Minimal Error (CGME) method (see for example [55]) can be used to solve it for nontrivial covariance and regularization matrices, filling a gap in existing literature as far as we know. We do this by deriving the Schur complement equation for the residual. A similar approach is taken by Orban and Arioli [56] to derive generalizations of the Golub-Kahan algorithm. Using these algorithms, they formulate generalizations of LSQR, Craig's method and LSMR (see [55]) for the general regularization problem. We explain in which cases Generalized CGME (GCGME) may have an advantage over Generalized Conjugate Gradient least-squares (GCGLS). Additionally, we apply GCGME to MRI data with different types of regularization.

This chapter results from our efforts to address the challenges of low-field MRI using advanced image processing. It is interdisciplinary in nature, with an emphasis on image reconstruction techniques. The contributions of this chapter include a signal model for low-field MRI that does not rely on any field assumptions as encountered in high-field MRI. Also, a new generalization of the conjugate gradient method is presented for the weighted and regularized least-squares problem, including an analysis of when this generalization is expected to perform best. Although we focus on a low-field MRI setting, this algorithm is generally applicable to ℓ_p-regularized least-squares problems.

2.1.1. LOW-FIELD MRI

In MRI, the internal structure of the body is made visible by measuring a voltage signal that is induced by time variations of the transverse magnetization within a body part of interest. Based on this measured signal, an image of the spin density of different tissue types may be obtained. The spin density is often denoted by ρ but we will use the symbol X.

To be specific, first the body part of interest is placed in a static magnetic field $\vec{B} = B_0(\vec{r})\vec{i}_x$ that is oriented in the x-direction in our Halbach measurement setup (see Figure 2.1a) with a position-dependent x-component $B_0 = B_0(\vec{r})$. A net magnetization

$$\vec{M}_{\text{eq}} = M_0(\vec{r})\vec{i}_x \quad \text{with} \quad M_0(\vec{r}) = \frac{\gamma^2\hbar^2}{4k_{\text{B}}T}X(\vec{r})B_0(\vec{r}) \tag{2.1}$$

will be induced that is oriented in the same direction as the static magnetic field. In the above expression, $\gamma = 267 \cdot 10^6$ rad s^{-1} T^{-1} is the proton gyromagnetic ratio, $\hbar = 1.055 \cdot 10^{-34}$ m^2 kg s^{-1} is Planck's constant divided by 2π, $k_{\text{B}} = 1.381 \cdot 10^{-23}$ m^2 kg s^{-2} K^{-1} is Boltzmann's constant, and T is the temperature in kelvin.

Subsequently, a radiofrequency pulse is emitted to tip the magnetization towards the transverse yz-plane. After this pulse has been switched off (in our model this happens at $t = 0$), the magnetization rotates about the static magnetic field with a precessional frequency ω (also known as the Larmor frequency) given by

$$\omega(\vec{r}) = \gamma B_0(\vec{r}) \tag{2.2}$$

and will relax back to its equilibrium given by Eq. (2.1). During this process, an electromagnetic field is generated that can be locally measured outside the body using a

receiver coil. This measured signal is amplified, demodulated, and low-pass filtered, which means that the received signal, which we will denote by b, can be described by [19]:

$$b(t) = \int_{\vec{r}\in\mathbb{D}} c(\vec{r})\omega(\vec{r})e^{-t/T_2(\vec{r})}M_\perp(\vec{r},0)e^{-\mathrm{i}\Delta\omega t}\,\mathrm{d}\vec{r}, \tag{2.3}$$

where $\mathbb{D}$ is the domain occupied by the body part of interest, $T_2(\vec{r})$ is the transverse relaxation time, $c(\vec{r})$ is the so-called coil sensitivity with amplification included, $M_\perp(\vec{r},0)$ is the transverse magnetization at $t = 0$, and $\Delta\omega$ is the difference between the Larmor frequency and the demodulation frequency that is used. We take this demodulation frequency to be the frequency that corresponds to the static magnetic field at the center of our imaging domain.

Furthermore, using Eq. (2.2) in the expression for M_0, we have

$$M_0(\vec{r}) = \frac{\gamma\hbar^2}{4k_\mathrm{B}T}X(\vec{r})\omega(\vec{r}) \tag{2.4}$$

and since the initial transverse magnetization $M_\perp(\vec{r},0)$ is proportional to $M_0(\vec{r})$, we can also express our signal model as

$$b(t) = \int_{\vec{r}\in\mathbb{D}} c(\vec{r})\omega^2(\vec{r})e^{-t/T_2(\vec{r})}X(\vec{r})e^{-\mathrm{i}\Delta\omega t}\,\mathrm{d}\vec{r}, \tag{2.5}$$

where it is understood that all remaining proportionality constants have been incorporated in the coil sensitivity $c(\vec{r})$. Conventionally, the spatial dependence of ω is ignored. Therefore, the ω^2 term usually does not appear in MRI literature. However, we incorporate it into our model because of the relatively large inhomogeneities in the magnetic field we are considering. We remark that Eq. (2.5) is a general MRI signal model, but it is more suitable for low-field MRI because the assumptions made for high-field MRI (namely, a very strong and homogeneous magnetic field) do not hold for low-field. Ignoring T_2 relaxation, the final signal model becomes

$$b(t) = \int_{\vec{r}\in\mathbb{D}} c(\vec{r})\omega^2(\vec{r})X(\vec{r})e^{-\mathrm{i}\Delta\omega(\vec{r}t}\,\mathrm{d}\vec{r}. \tag{2.6}$$

We note that, in practice, the signal $b(t)$ is contaminated by noise.

MODEL-BASED IMAGE RECONSTRUCTION

In high-field MRI, the magnetic field is manipulated in such a way that Eq. (2.6) constitutes a Fourier Transform. The resulting linear problem is well-posed, and the image can be efficiently obtained using an inverse FFT. However, in low-field MRI, the magnetic field can be strongly inhomogeneous, which prevents us from using standard FFT routines. Model-based image reconstruction can be applied instead [44].

We divide the domain into pixels, where each pixel has side lengths of Δx and Δy in the x- and y-direction, respectively, so it has an area of $\Delta x\Delta y$. We denote the slice thickness by Δz. Now, we can discretize Eq. (2.6) using a "center-of-pixel" approximation to obtain a system of equations:

$$\mathbf{b}^{(1)} = \mathbf{A}^{(1)}\mathbf{x} + \mathbf{e}^{(1)}, \tag{2.7}$$

where $\mathbf{b}^{(1)}$ is the vector form of the noisy signal $b(t)$:

$$\mathbf{b} = \left[b(t_1), b(t_2), ..., b(t_Q)\right]^T, \tag{2.8}$$

acquired at discrete time instances t_q, $q = 1, ..., Q$, $\mathbf{x}$ is the discrete vector representation of the image $X(\vec{r}$ we aim to reconstruct, $\mathbf{e}^{(1)}$ is a noise vector and the elements of the model matrix $\mathbf{A}^{(1)}$ are described by

$$a_{jk} = c(\vec{r}_k)\omega^2(\vec{r}_k)e^{-i\Delta\omega(\vec{r}_k)t_j}\Delta x\Delta y\Delta z. \tag{2.9}$$

As can be seen in Figure 2.2, the field has a high degree of symmetry. The precessional frequency depends linearly on the magnitude of the field, which means that several pixels will correspond to the same frequency. Therefore, it is impossible to determine the contribution of each pixel to the signal if we use only one measured signal. By rotating the object to be imaged and hence obtaining a multitude of different signals corresponding to different rotations of the same object, we aim to mitigate this problem. The same approach was taken by Cooley et al. [30]. More formally, denoting the measurement number by l and disregarding the noise vector, we can write

$$\mathbf{b}^{(l)} = \mathbf{A}^{(l)}\mathbf{x}, \tag{2.10}$$

where $\mathbf{A}^{(l)}$ is recalculated according to the new position of the phantom. Then, all L measurements can be combined into one system of equations

$$\begin{pmatrix} \mathbf{b}^{(1)} \\ \mathbf{b}^{(2)} \\ \vdots \\ \mathbf{b}^{(L-1)} \\ \mathbf{b}^{(L)} \end{pmatrix} = \begin{pmatrix} \mathbf{A}^{(1)} \\ \mathbf{A}^{(2)} \\ \vdots \\ \mathbf{A}^{(L-1)} \\ \mathbf{A}^{(L)} \end{pmatrix} \mathbf{x}, \tag{2.11}$$

which we will write as

$$\mathbf{b} = \mathbf{A}\mathbf{x}. \tag{2.12}$$

2.2. METHODOLOGY

The model that is used to reconstruct X is given by the linear system of Eqs. (2.7). We can attempt to solve for $\mathbf{x}$ by finding a solution to the least-squares problem

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2. \tag{2.13}$$

This can be done by applying the conjugate gradient method introduced by Hestenes and Stiefel in 1952 [57] to the normal equations

$$\mathbf{A}^H\mathbf{A}\mathbf{x} = \mathbf{A}^H\mathbf{b}, \tag{2.14}$$

with $\mathbf{A}^H$ denoting the Hermitian transpose of $\mathbf{A}$.

The conjugate gradient method tailored to Eq. (2.14) was proposed in [57] and is usually denoted by CGLS (Conjugate Gradient for least-squares). The difference with the standard conjugate gradient method lies in the increased stability of the CGLS method. A review of the literature reveals that this method is known by other names as well. In [58], Saad calls it CGNR (Conjugate Gradient Normal Residual), while Hanke [59] and Engl [60] use the term CGNE (Conjugate Gradient for the Normal Equations).

On the other hand, the second normal equations

$$\mathbf{A}\mathbf{A}^H\mathbf{y} = \mathbf{b},\ \mathbf{x} = \mathbf{A}^H\mathbf{y} \tag{2.15}$$

can be solved using the conjugate gradient method as well. In the literature, this is usually called CGME (Conjugate Gradient Minimal Error). However, in [61] it is called CGNE (Conjugate Gradient Normal Error), while [62] uses the term Craig's method. It was introduced by Craig in 1955 [63]. CGLS and CGME are discussed by Björck in [55], Hanke in [59] and Saad in [58]. While CGLS minimizes the residual $\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x}$ in the ℓ_2 norm over the Krylov subspace $\mathbf{x}_0 + \mathscr{K}_k(\mathbf{A}^H\mathbf{A}, \mathbf{A}^H\mathbf{b} - \mathbf{A}^H\mathbf{A}\mathbf{x}_0)$, CGME minimizes the error (over the same subspace). The main drawback of this latter method is that, in theory, it only works for consistent problems for which $\mathbf{b} \in \mathbf{R}(\mathbf{A})$. This means that the method is of limited use for most problems in practice, because the presence of noise renders the system inconsistent. In [64], this problem is circumvented by defining an operator $\mathbf{Q}$ that projects $\mathbf{b}$ onto the column space of $\mathbf{A}$. Subsequently, $\mathbf{A}\mathbf{x} = \mathbf{Q}\mathbf{b}$ can be solved using CGME. The obvious disadvantage of this method is that $\mathbf{Q}\mathbf{b}$ has to be calculated and stored.

2.2.1. REGULARIZATION OF THE PROBLEM

The system of Eqs. (2.7) is ill-posed. Regularization of an ill-posed problem aims to make the problem less sensitive to noise by taking into account additional information, i.e., it aims at turning an ill-posed problem into a well-posed one. Like many iterative methods, both CGLS and CGME have a regularizing effect if the iterating procedure is stopped early: keeping the number of iterations low prevents the noise from corrupting the result too much. If a large number of iterations is used, noise can have a very strong effect on the solution. The regularizing properties of CGLS were established by Nemirovskii in [65] and are discussed in [55], [60] and [59], among others. CGME's regularizing effect was demonstrated by Hanke in [66]. However, we are interested in what Hansen [45] calls general-form Tikhonov regularization, i.e. adding a regularization term to minimization problem (2.13), leading to

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_{\mathbf{W}}^2 + \frac{1}{2}\tau\|\mathbf{x}\|_{\mathbf{R}}^2 \tag{2.16}$$

where $\mathbf{W}$ is a weighting matrix, and $\mathbf{R}$ is a Hermitian positive definite matrix. Using a CG algorithm to solve Eq. (2.16) is a natural choice [44]. The CG method is often used to solve image reconstruction problems in MRI when a conventional Fourier model is insufficient (see for example [67], [68] and [69]). Additionally, it is used as a building block for other algorithms used in MRI by Pruessman [70], Ramani and Fessler [71] and Ye et al. [72], among others. It is straightforward to generalize CGLS to regularized and weighted least-squares problems of the form of Eq. (2.16). In this case, because of the well-posedness of the resulting minimization problem, the noise does not influence the

solution as much as when Eq. (2.13) is considered and increasing the number of iterations does not lead to a noisier solution. In this work we will use $\mathbf{W} = \mathbf{C}^{-1}$, where $\mathbf{C}$ is the covariance matrix of the noise:

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2_{\mathbf{C}^{-1}} + \frac{1}{2}\tau\|\mathbf{x}\|^2_{\mathbf{R}} \tag{2.17}$$

For our application, the noise can be considered to be white, which means that $\mathbf{C} = \mathbf{I}$. However, for completeness, we consider the general case. In case $\mathbf{R} = \mathbf{I}$, Eq. (2.17) reduces to a minimization problem with standard Tikhonov regularization [73]. The optimal value of the regularization parameter τ is usually unknown. An approach that is often used to find a suitable value is the L-curve method [45]. By taking the gradient and setting it equal to $\mathbf{0}$, the normal equations are obtained:

$$(\mathbf{A}^H\mathbf{C}^{-1}\mathbf{A} + \tau\mathbf{R})\mathbf{x} = \mathbf{A}^H\mathbf{C}^{-1}\mathbf{b}. \tag{2.18}$$

Again, the conjugate gradient method can be used to solve Eq. (2.18). We will use the term GCGLS (Generalized CGLS) to refer to the conjugate gradient method applied to the normal Eqs. (2.18).

Saunders [62] extended Craig's method, which is mathematically equivalent to CGME, to the regularized least-squares problem with $\mathbf{C} = \mathbf{I}$ and $\mathbf{R} = \mathbf{I}$. He introduces an additional variable $\mathbf{s}$ and considers the constrained minimization problem

$$\min_{\mathbf{x},\mathbf{s}} \frac{1}{2}\left\|\begin{pmatrix}\mathbf{x}\\ \mathbf{s}\end{pmatrix}\right\|^2 \quad \text{subject to } \begin{pmatrix}\mathbf{A} & \sqrt{\tau}\mathbf{I}\end{pmatrix}\begin{pmatrix}\mathbf{x}\\ \mathbf{s}\end{pmatrix} = \mathbf{b}. \tag{2.19}$$

By defining $\tilde{\mathbf{r}} = \sqrt{\tau}\mathbf{s} = \mathbf{b} - \mathbf{A}\mathbf{x}$, he shows that this constrained minimization problem is equivalent to

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2 + \frac{1}{2}\tau\|\mathbf{x}\|^2. \tag{2.20}$$

For every $\tau > 0$, $\begin{pmatrix}\mathbf{A} & \sqrt{\tau}\mathbf{I}\end{pmatrix}\begin{pmatrix}\mathbf{x}\\ \mathbf{s}\end{pmatrix} = \mathbf{b}$ is consistent and hence, Eq. (2.20) can be solved using CGME. Unfortunately, no advantages to using CGME were found. Note that such a reformulization is necessary because the standard way of including the regularization matrix $\mathbf{R} = \mathbf{I}$, by simply solving the so-called damped least-squares problem

$$\begin{pmatrix}\mathbf{A}\\ \sqrt{\tau}\mathbf{I}\end{pmatrix}\mathbf{x} = \begin{pmatrix}\mathbf{b}\\ \mathbf{0}\end{pmatrix} \tag{2.21}$$

using CGME, is not possible, due to the inconsistency of the system. Reformulation of CGME for general-form regularization can be achieved using a Schur-complement approach as will be shown below.

Again, we consider Eq. (2.17). We introduce the variable $\mathbf{r} = \mathbf{C}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x})$ and we note that $||\mathbf{A}\mathbf{x}-\mathbf{b}||^2_{\mathbf{C}^{-1}} = ||\mathbf{r}||^2_{\mathbf{C}}$. Then, minimization problem (2.17) can be formulated as a constrained minimization problem:

$$\min_{\mathbf{r},\mathbf{x}} \frac{1}{2}||\mathbf{r}||^2_{\mathbf{C}} + \frac{1}{2}\tau||\mathbf{x}||^2_{\mathbf{R}} \tag{2.22}$$

$$\text{s.t. } \mathbf{r} = \mathbf{C}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x})$$

and using the technique of Lagrange multipliers, we find that

$$\mathbf{r} = \mathbf{C}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}) \text{ and } \tau\mathbf{R}\mathbf{x} = \mathbf{A}^H\mathbf{r}. \tag{2.23}$$

If we eliminate $\mathbf{r}$ from Eq. (2.23), the original normal Eqs. (2.18) are obtained, whereas if we assume $\tau\mathbf{R}$ is invertible and we subsequently eliminate $\mathbf{x}$, we end up with a different set of equations. As mentioned before, the first option leads to the GCGLS method. The latter approach leads to the GCGME method.

2.2.2. GCGLS

By applying the conjugate gradient method to Eq. (2.18) and making some adjustments to increase stability (see [55] for details), the GCGLS algorithm is obtained, see Algorithm 1. Here, M is the total number of data points measured and N is the number of pixels in the image. The residual of the normal Eqs. (2.18) is denoted by $\mathbf{s}_k$. We remark that the vectors on the left side can be overwritten by the vectors on the right. Only 8 vectors have to be stored, namely $\mathbf{x}$, $\mathbf{r}$, $\mathbf{s}$, $\mathbf{p}$, $\mathbf{q}$, $\mathbf{Rx}$, $\mathbf{Rp}$ and $\mathbf{C}^{-1}\mathbf{q}$. Note that the recursion for $\mathbf{Rx}_{k+1}$ is included to avoid an extra multiplication with $\mathbf{R}$. It can be ignored in case $\mathbf{R} = \mathbf{I}$. In this algorithm, only three matrix-vector multiplications are carried out per iteration: $\mathbf{Ap}_{k+1}$, $\mathbf{A}^H\mathbf{r}_{k+1}$ and $\mathbf{Rp}_k$. Additionally, one system with $\mathbf{C}$ has to be solved (if $\mathbf{C} \neq \mathbf{I}$). A slightly different formulation of the GCGLS algorithm can be found in [69].

Algorithm 1 GCGLS

Require: $\mathbf{A} \in \mathbb{C}^{M\times N}, \mathbf{C} \in \mathbb{C}^{M\times M}, \mathbf{R} \in \mathbb{C}^{N\times N}, \mathbf{x}_0, \in \mathbb{C}^N, \mathbf{b} \in \mathbb{C}^M, \tau \in \mathbb{R}_{\geq 0}$;
Ensure: Approximate solution $\mathbf{x}_k$ such that $\|\mathbf{A}^H\mathbf{r}_k - \tau\mathbf{R}\mathbf{x}_k\| \leqslant TOL$.

$\mathbf{r}_0 = \mathbf{C}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}_0)$, $\mathbf{s}_0 = \mathbf{A}^H\mathbf{r}_0 - \tau\mathbf{R}\mathbf{x}_0$, $\mathbf{p}_0 = \mathbf{s}_0$; $\mathbf{q}_0 = \mathbf{A}^H\mathbf{p}_0$, $\gamma_0 = \mathbf{s}_0^H\mathbf{s}_0$, $k = 0$
while $\sqrt{\gamma_k} > TOL$ **and** $k < k_{max}$ **do**
  $\xi_k = \mathbf{q}_k^H\mathbf{C}^{-1}\mathbf{q}_k + \tau\mathbf{p}_k^H\mathbf{R}\mathbf{p}_k$
  $\alpha_k = \frac{\gamma_k}{\xi_k}$
  $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\mathbf{p}_k$
  $\mathbf{Rx}_{k+1} = \mathbf{Rx}_k + \alpha_k\mathbf{Rp}_k$
  $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k\mathbf{C}^{-1}\mathbf{q}_k$
  $\mathbf{s}_{k+1} = \mathbf{A}^H\mathbf{r}_{k+1} - \tau\mathbf{R}\mathbf{x}_{k+1}$
  $\gamma_{k+1} = \mathbf{s}_{k+1}^H\mathbf{s}_{k+1}$
  $\beta_k = \frac{\gamma_{k+1}}{\gamma_k}$
  $\mathbf{p}_{k+1} = \mathbf{s}_{k+1} + \beta_k\mathbf{p}_k$
  $\mathbf{q}_{k+1} = \mathbf{A}\mathbf{p}_{k+1}$
  $k = k + 1$
end while

Algorithm 2 GCGME

Require: $\mathbf{A} \in \mathbb{C}^{M\times N}, \mathbf{C} \in \mathbb{C}^{M\times M}, \mathbf{R} \in \mathbb{C}^{N\times N}, \mathbf{r}_0 \in \mathbb{C}^M, \mathbf{b} \in \mathbb{C}^M, \tau \in \mathbb{R}_{>0}$;
Ensure: Approximate solution $\mathbf{x}_k$ such that $\|\mathbf{b}-\mathbf{A}\mathbf{x}_k - \mathbf{C}\mathbf{r}_k\| \leqslant TOL$.

$\mathbf{x}_0 = \frac{1}{\tau}\mathbf{R}^{-1}\mathbf{A}^H\mathbf{r}_0$
$\mathbf{s}_0 = \mathbf{b}-\mathbf{A}\mathbf{x}_0 - \mathbf{C}\mathbf{r}_0$, $\mathbf{p}_0 = \mathbf{s}_0$, $\mathbf{q}_0 = \mathbf{A}^H\mathbf{p}_0$, $\gamma_0 = \mathbf{s}_0^H\mathbf{s}_0$, $k = 0$
while $\sqrt{\gamma_k} > TOL$ **and** $k < k_{max}$ **do**
  $\xi_k = \frac{1}{\tau}\mathbf{q}_k^H\mathbf{R}^{-1}\mathbf{q}_k + \mathbf{p}_k^H\mathbf{C}\mathbf{p}_k$
  $\alpha_k = \frac{\gamma_k}{\xi_k}$
  $\mathbf{r}_{k+1} = \mathbf{r}_k + \alpha_k\mathbf{p}_k$
  $\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{\alpha_k}{\tau}\mathbf{R}^{-1}\mathbf{q}_k$
  $\mathbf{s}_{k+1} = \mathbf{s}_k - \alpha_k(\frac{1}{\tau}\mathbf{A}\mathbf{R}^{-1}\mathbf{q}_k + \mathbf{C}\mathbf{p}_k)$
  $\gamma_{k+1} = \mathbf{s}_{k+1}^H\mathbf{s}_{k+1}$
  $\beta_k = \frac{\gamma_{k+1}}{\gamma_k}$
  $\mathbf{p}_{k+1} = \mathbf{s}_{k+1} + \beta_k\mathbf{p}_k$
  $\mathbf{q}_{k+1} = \mathbf{A}^H\mathbf{p}_{k+1}$
  $k = k+1$
end while

2.2.3. GCGME

If $\tau\mathbf{R}$ is invertible, $\mathbf{x}$ can be eliminated from Eq. (2.23), yielding

$$\left(\frac{1}{\tau}\mathbf{A}\mathbf{R}^{-1}\mathbf{A}^H + \mathbf{C}\right)\mathbf{r} = \mathbf{b}. \tag{2.24}$$

Subsequently, $\mathbf{x}$ can be obtained from $\mathbf{r}$ as:

$$\mathbf{x} = \frac{1}{\tau}\mathbf{R}^{-1}\mathbf{A}^H\mathbf{r}. \tag{2.25}$$

In [56], Arioli and Orban derive a generalization of Craig's method [63] based on Schur complement (2.24). In Algorithm 2, we formulate a similar generalization of the CGME method applied to this system. We are not aware this generalization of CGME has been formulated elsewhere. Here, $\mathbf{s}_k$ is the residual of the normal Eqs. (2.24). Note that the original CGME algorithm can be recovered from the generalized CGME algorithm given above by taking $\frac{1}{\tau}\mathbf{R} = \mathbf{I}$ and $\mathbf{C} = \mathbf{O}$, the zero matrix. Only 7 vectors have to be stored, namely $\mathbf{x}$, $\mathbf{r}$, $\mathbf{s}$, $\mathbf{p}$, $\mathbf{q}$, $\mathbf{R}^{-1}\mathbf{q}$ and $\mathbf{C}\mathbf{p}$. GCGME needs four matrix operations per iteration: $\mathbf{C}\mathbf{p}_k$, $\mathbf{A}^H\mathbf{p}_{k+1}$, $\mathbf{R}^{-1}\mathbf{q}_k$ and $\mathbf{A}\mathbf{R}^{-1}\mathbf{q}_k$.

We remark that there is an essential difference between GCGLS and GCGME. GCGLS iterates for the solution vector $\mathbf{x}$ and the equality $\mathbf{r}_k = \mathbf{C}^{-1}(\mathbf{b}-\mathbf{A}\mathbf{x}_k)$ is explicitly imposed. The equality $\mathbf{x}_k = \frac{1}{\tau}\mathbf{R}^{-1}\mathbf{A}^H\mathbf{r}_k$ is not enforced, and is only (approximately) satisfied after convergence. GCGME, on the other hand, iterates for $\mathbf{r}_k$. The equality $\mathbf{x}_k = \frac{1}{\tau}\mathbf{R}^{-1}\mathbf{A}^H\mathbf{r}_k$ is enforced, while $\mathbf{r}_k = \mathbf{C}^{-1}(\mathbf{b}-\mathbf{A}\mathbf{x}_k)$ is only satisfied approximately after convergence.

2.2.4. CONVERGENCE OF GCGLS AND GCGME

The convergence of the conjugate gradient method depends on the condition number of the system matrix. Again, suppose that CG is used to solve the system $\mathbf{L}\mathbf{u} = \mathbf{f}$ for the

unknown vector $\mathbf{u}$, where $\mathbf{L}$ is a Hermitian positive definite (HPD) matrix and $\mathbf{f}$ is a known vector. Then the following classical convergence bound holds [55]:

$$\|\mathbf{u}-\mathbf{u}_k\|_{\mathbf{L}} \leq 2\left(\frac{\sqrt{\kappa_2(\mathbf{L})}-1}{\sqrt{\kappa_2(\mathbf{L})}+1}\right)^k \|\mathbf{u}-\mathbf{u}_0\|_{\mathbf{L}}, \tag{2.26}$$

where $\kappa_2(\mathbf{L})$ is the ℓ_2-norm condition number of $\mathbf{L}$, which, for HPD matrices, is equal to

$$\kappa_2(\mathbf{L}) = \frac{\lambda_{\max}(\mathbf{L})}{\lambda_{\min}(\mathbf{L})} \tag{2.27}$$

in which $\lambda_{\max}(\mathbf{L})$ and $\lambda_{\min}(\mathbf{L})$ are the largest and smallest eigenvalue of $\mathbf{L}$, respectively. In this section we bound the condition numbers of the two Schur complement matrices in Eqs. (2.18) and (2.24) to gain insight into when GCGME can be expected to perform better than GCGLS, and vice versa. Given two HPD matrices $\mathbf{K}$ and $\mathbf{M}$, the following bound on the condition number holds:

$$\frac{\lambda_{\max}(\mathbf{K})+\lambda_{\min}(\mathbf{M})}{\lambda_{\min}(\mathbf{K})+\lambda_{\max}(\mathbf{M})} \leq \kappa_2(\mathbf{K}+\mathbf{M}) \leq \frac{\lambda_{\max}(\mathbf{K})+\lambda_{\max}(\mathbf{M})}{\lambda_{\min}(\mathbf{K})+\lambda_{\min}(\mathbf{M})}. \tag{2.28}$$

This inequality follows from Weyl's theorem [74], which states that for eigenvalues of Hermitian matrices $\mathbf{K}$ and $\mathbf{M}$, the following holds:

$$\lambda_i(\mathbf{K})+\lambda_{\min}(\mathbf{M}) \leq \lambda_i(\mathbf{K}+\mathbf{M}) \leq \lambda_i(\mathbf{K})+\lambda_{\max}(\mathbf{M}). \tag{2.29}$$

Here $\lambda_i(\mathbf{K})$ denotes any eigenvalue of the matrix $\mathbf{K}$. For GCGLS we have that

$$\mathbf{K} = \tau\mathbf{R} \ , \ \mathbf{M} = \mathbf{A}^H\mathbf{C}^{-1}\mathbf{A} \tag{2.30}$$

and, using the following inequalities

$$\lambda_{\max}(\mathbf{A}^H\mathbf{C}^{-1}\mathbf{A}) \leq \frac{\sigma_{\max}(\mathbf{A})^2}{\lambda_{\min}(\mathbf{C})} \ , \quad \lambda_{\min}(\mathbf{A}^H\mathbf{C}^{-1}\mathbf{A}) \geq 0, \tag{2.31}$$

with $\sigma_{\max}(\mathbf{A})$ the largest singular value of $\mathbf{A}$, we get that

$$\begin{aligned} \frac{\tau\lambda_{\max}(\mathbf{R})\lambda_{\min}(\mathbf{C})}{\tau\lambda_{\min}(\mathbf{R})\lambda_{\min}(\mathbf{C})+\sigma_{\max}(\mathbf{A})^2} &\leq \kappa_2(\mathbf{A}^H\mathbf{C}^{-1}\mathbf{A}+\tau\mathbf{R}) \\ &\leq \frac{\tau\lambda_{\max}(\mathbf{R})\lambda_{\min}(\mathbf{C})+\sigma_{\max}(\mathbf{A})^2}{\tau\lambda_{\min}(\mathbf{R})\lambda_{\min}(\mathbf{C})} \end{aligned} \tag{2.32}$$

Analogously, for CGME, we have

$$\mathbf{K} = \mathbf{C} \ , \ \mathbf{M} = \frac{1}{\tau}\mathbf{A}\mathbf{R}^{-1}\mathbf{A}^H \tag{2.33}$$

and using similar manipulations as above we obtain

$$\frac{\tau\lambda_{\min}(\mathbf{R})\lambda_{\max}(\mathbf{C})}{\tau\lambda_{\min}(\mathbf{R})\lambda_{\min}(\mathbf{C})+\sigma_{\max}(\mathbf{A})^2} \leq \kappa_2(\frac{1}{\tau}\mathbf{A}\mathbf{R}^{-1}\mathbf{A}^H+\mathbf{C}) \tag{2.34}$$

$$\leq \frac{\tau\lambda_{\min}(\mathbf{R})\lambda_{\max}(\mathbf{C}) + \sigma_{\max}(\mathbf{A})^2}{\tau\lambda_{\min}(\mathbf{R})\lambda_{\min}(\mathbf{C})}$$

These inequalities indicate that if

$$\lambda_{\max}(\mathbf{C})\lambda_{\min}(\mathbf{R}) \gg \lambda_{\max}(\mathbf{R})\lambda_{\min}(\mathbf{C}) \Leftrightarrow \kappa_2(\mathbf{C}) \gg \kappa_2(\mathbf{R}) \tag{2.35}$$

GCGLS can be expected to perform best, and that if

$$\kappa_2(\mathbf{R}) \gg \kappa_2(\mathbf{C}) \tag{2.36}$$

GCGME should be preferred. This latter situation may occur when the regularization term is minimized in the ℓ_p-norm with $p \in (0,1]$, as we will discuss in the next section. In Appendix B, we analyze a specific and simple case and determine the specific values of the regularization parameter for which we expect GCGME to converge faster than GCGLS and vice versa.

2.2.5. TYPES OF REGULARIZATION

Instead of an ℓ_2-penalty, we will consider the more general case of an ℓ_p-penalty with $p \in (0,2]$. Then, the minimization problem becomes

$$\min_{\mathbf{x}} \frac{1}{2}||\mathbf{Ax}-\mathbf{b}||_2^2 + \frac{1}{p}\tau||\mathbf{Fx}||_p^p. \tag{2.37}$$

A vast literature regarding this $\ell_2\ell_p$-minimization problem is available. In for example [75], [76], [77] and [78], this problem is solved using a majorization-minimization approach. In this work, we will focus on the classical approach using Iterative Reweighted Least-Squares (IRLS), also known as Iterative Reweighted Norm (IRN), see for example [55], for solving minimization problem (2.37), in which GCGLS and GCGME can be used as building blocks. Their performances will be compared. We choose the IRLS algorithm for three reasons: its simplicity, the fact that it is a well-known technique and that in this algorithm, the regularization matrix changes in each iteration, which makes it especially interesting for us, because we can test whether GCGME indeed performs better in case Eq. (2.36) holds. This work is not meant to evaluate the performance of IRLS as a solver for Eq. (2.37) and we do not compare it with other methods. For completeness, however, we do mention that we could also have chosen to evaluate both approaches as a building block of the Split-Bregman method [79] for the ℓ_1-regularized problem, for example. In [76], Chan and Liang use CG as a building block for their half-quadratic algorithm that solves Eq. (2.37) as well. A comparison between GCGLS and GCGME could be carried out in this context too.

IRLS is an iterative method that can solve an ℓ_p-regularized minimization problem by reducing it to a sequence of ℓ_2-regularized minimization problems. Note that for a vector $\mathbf{m}$ of length N,

$$||\mathbf{m}||_p = \left(\sum_{i=1}^{N} |m_i|^p\right)^{1/p} \tag{2.38}$$

so

$$||\mathbf{m}||_p^p = \sum_{i=1}^{N} |m_i|^p. \tag{2.39}$$

Furthermore, $\mathbf{F}$ is some regularizing matrix. We introduce

$$\min_{\mathbf{x}} \frac{1}{2}||\mathbf{Ax}-\mathbf{b}||_2^2 + \frac{1}{2}\tau||\mathbf{x}||_{\mathbf{F}^H\mathbf{DF}}^2, \tag{2.40}$$

where

$$\mathbf{D} := \text{diag}\left(\frac{1}{|\mathbf{Fx}|^{2-p}}\right), \tag{2.41}$$

and $|\mathbf{Fx}|$ is the element-wise modulus of $\mathbf{Fx}$. This is simply another instance of minimization problem (2.17), with $\mathbf{R} = \mathbf{F}^H\mathbf{DF}$. However, now $\mathbf{R}$ depends on $\mathbf{x}$. So, when the kth iterate $\mathbf{x}_k$ is known, $\mathbf{x}_{k+1}$ is found as follows:

$$\mathbf{x}_{k+1} = \arg\min_{\mathbf{x}} \frac{1}{2}||\mathbf{Ax}-\mathbf{b}||_2^2 + \frac{1}{2}\tau||\mathbf{x}||_{\mathbf{F}^H\mathbf{D}_k\mathbf{F}}^2, \tag{2.42}$$

where

$$\mathbf{D}_k = \text{diag}\left(\frac{1}{|\mathbf{Fx}_k|^{2-p}+\epsilon}\right). \tag{2.43}$$

This is repeated until convergence. Furthermore, in Eq. (2.43), ϵ is a small number that is added to the denumerator to prevent division by zero. We will use $\epsilon = 10^{-6}$. We observe that in each IRLS iteration, we simply encounter an instance of minimization problem (2.17) again with $\mathbf{R}_k = \mathbf{F}^H\mathbf{D}_k\mathbf{F}$, which can be solved using either GCGLS or GCGME. When carrying out calculations with $\mathbf{D}_k^{-1}$, we will use

$$\mathbf{D}_k^{-1} = \text{diag}\left(|\mathbf{Fx}_k|^{2-p}\right). \tag{2.44}$$

Due to the sparsity-inducing property of the ℓ_p penalty when $p \leq 1$ (see for example [80]), $\mathbf{D}_k^{-1}$ will contain an increasing number of entries nearly equal to zero. In cases where $\mathbf{F}$ is an invertible matrix, $\mathbf{R}_k^{-1} = \mathbf{F}^{-1}\mathbf{D}_k^{-1}(\mathbf{F}^H)^{-1}$. When GCGME is used, we can take advantage of this structure, instead of calculating $\mathbf{R}_k$ and working with its inverse. Moreover, when $\mathbf{F}$ is an orthogonal matrix, no additional computations are necessary to compute inverses.

The regularization matrix $\mathbf{R} = \mathbf{F}^H\mathbf{D}_k\mathbf{F}$ will become ill-conditioned when elements of $\mathbf{Fx}_k$ become small. Therefore, we expect that, when combined with IRLS, GCGME will perform better than GCGLS for $p \leq 1$. Numerical experiments are carried out to investigate this further.

DIFFERENT CHOICES FOR p

We will minimize the following ℓ_1-regularized least-squares problem and the $\ell_{1/2}$-regularized least-squares problem to obtain approximations to the optimal solution $\mathbf{x}$. For a general $\mathbf{F}$, this results in the following two minimization problems:

$$\min_{\mathbf{x}} \frac{1}{2}||\mathbf{Ax}-\mathbf{b}||_2^2 + \tau||\mathbf{Fx}||_1. \tag{2.45}$$

and

$$\min_{\mathbf{x}} \frac{1}{2}||\mathbf{Ax}-\mathbf{b}||_2^2 + 2\tau||\mathbf{Fx}||_{1/2}^{1/2}. \tag{2.46}$$

We note that in the latter case, the objective function is not convex which means that the obtained solution does not necessarily correspond to a global minimum, see for example [81]. For each of these two minimization problems, we will consider two different regularization operators.

REGULARIZING USING THE IDENTITY MATRIX

First, we set $\mathbf{F} = \mathbf{I}$. In case the ℓ_1 penalty is used, the minimization problem reduces to

$$\min_{\mathbf{x}} \frac{1}{2}||\mathbf{Ax}-\mathbf{b}||_2^2 + \tau||\mathbf{x}||_1. \tag{2.47}$$

This is known as LASSO (Least Absolute Shrinkage and Selection Operator) regularization which was first introduced by Tibshirani in [82]. If the regularization parameter is set to a sufficiently high value, the resulting solution will be sparse. The same holds for the $\ell_{1/2}$-regularized minimization problem:

$$\min_{\mathbf{x}} \frac{1}{2}||\mathbf{Ax}-\mathbf{b}||_2^2 + 2\tau||\mathbf{x}||_{1/2}^{1/2}. \tag{2.48}$$

The rationale behind choosing the regularization operator this way is the fact that the intensity of many pixels in MRI images is equal to 0. In both cases ($p = 1$ and $p = 1/2$), the regularization matrix reduces to $\mathbf{R}_k = \mathbf{D}_k = \text{diag}\left(\frac{1}{|\mathbf{x}_k|^{2-p}}\right)$ and its inverse is simply $\mathbf{R}_k^{-1} = \mathbf{D}_k^{-1} = \text{diag}\left(|\mathbf{x}_k|^{2-p}\right)$. This is especially useful for GCGME, because calculating the product of $\mathbf{R}^{-1}$ and a vector is trivial in this case.

REGULARIZING USING FIRST ORDER DIFFERENCES

Additionally, we consider the case where $\mathbf{F}$ is a first order difference matrix $\mathbf{T}$ that calculates the values of the jumps between each pair of neighboring pixels. We define the 2D first-order difference operators that calculate the value of the jumps between neighboring pixels in the x-direction and y-direction as $\mathbf{T}_x$ and $\mathbf{T}_y$ respectively. Then, the complete regularizing operator is

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_x \\ \mathbf{T}_y \end{pmatrix}. \tag{2.49}$$

This type of regularization is known as anisotropic total variation regularization. A reason for choosing $\mathbf{F} = \mathbf{T}$ is that neighboring pixels are very likely to have the same values in MR images. This is due to the fact that neighboring pixels tend to represent the same tissue. However, $\mathbf{T}$ is not a square matrix, which means that $\mathbf{R}_k$ has to be calculated explicitly and then inverted when GCGME is used. Although this makes regularization with first order differences in combination with GCGME less attractive than with GCGLS, we do include this technique to investigate the relative reconstruction quality of

this widely used regularization method. The resulting minimization problems are equal to Eqs. (2.45) and (2.46) with $\mathbf{F} = \mathbf{T}$:

$$\min_{\mathbf{x}} \frac{1}{2}||\mathbf{Ax}-\mathbf{b}||_2^2 + \tau||\mathbf{Tx}||_1. \tag{2.50}$$

and

$$\min_{\mathbf{x}} \frac{1}{2}||\mathbf{Ax}-\mathbf{b}||_2^2 + 2\tau||\mathbf{Tx}||_{1/2}^{1/2}. \tag{2.51}$$

FOUR DIFFERENT MINIMIZATION PROBLEMS

We will investigate all four minimization problems (2.47), (2.48), (2.50) and (2.51). Since the least-squares term is the same in all four minimization problems, the difference between them lies in the penalty term used, as summarized in Table 2.1. In each of the four cases, we will use both GCGLS and GCGME to compare their rate of convergence.

Table 2.1: Overview of the 4 different minimization problems considered in this work.

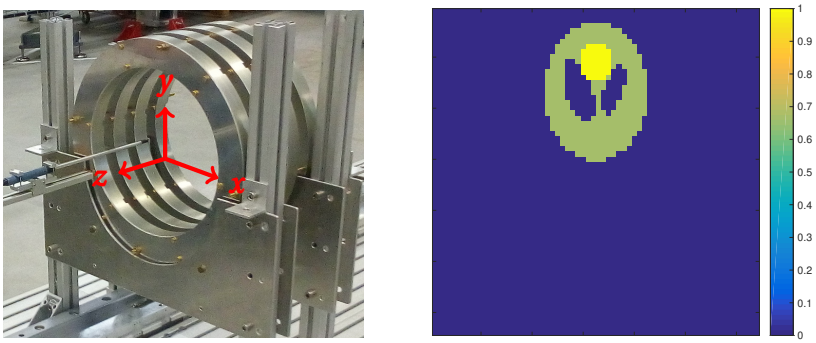
	Regularization matrix $\mathbf{I}$	Regularization matrix $\mathbf{F}$
ℓ_1-penalty	$\min_{\mathbf{x}} \frac{1}{2}\|\|\mathbf{Ax}-\mathbf{b}\|\|_2^2 + \tau\|\|\mathbf{x}\|\|_1$	$\min_{\mathbf{x}} \frac{1}{2}\|\|\mathbf{Ax}-\mathbf{b}\|\|_2^2 + \tau\|\|\mathbf{Tx}\|\|_1$
$\ell_{1/2}$-penalty	$\min_{\mathbf{x}} \frac{1}{2}\|\|\mathbf{Ax}-\mathbf{b}\|\|_2^2 + 2\tau\|\|\mathbf{x}\|\|_{1/2}^{1/2}$	$\min_{\mathbf{x}} \frac{1}{2}\|\|\mathbf{Ax}-\mathbf{b}\|\|_2^2 + 2\tau\|\|\mathbf{Tx}\|\|_{1/2}^{1/2}$

2.2.6. NUMERICAL SIMULATIONS

For our simulations, we use a simulated magnetic field as shown in Figure 2.1a. (We also have access to a measured field map, but it is measured on a very coarse grid, making it unsuitable for our purposes.) The magnetic field within the FoV of 14 cm by 14 cm is clearly inhomogeneous, as can be seen in Figure 2.2. The magnetic field has an approximately quadrupolar profile. This is because the Halbach cylinder is designed to generate a field that is as uniform as possible. However, due to practical limitations, such as the finite length of the cylinder, this uniformity cannot be attained, leading to a quadratic residual field profile, see for example [30] and [83]. We do not use a switched linear gradient coil, as is done in conventional MRI. Instead, the inhomogeneous background field is used for readout encoding. For a thorough exploration of the use of non-bijective encoding maps in MRI, we refer the reader to [83], [84], [85], [86] and [87].

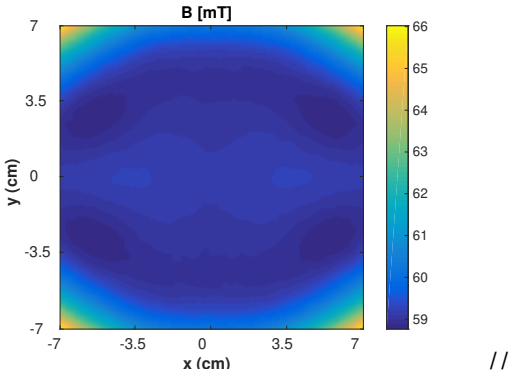
Performing slice selection in the presence of a nonhomogeneous background field is nontrivial, but this complication is ignored here. We assume that the entire measured signal originates from one slice. We simulate the signal generation inside the Halbach cylinder using Eqs. (2.9) and (2.7). The dwell time is set to $\Delta t = 5 \cdot 10^{-6}$, and the readout window is 0.5 ms, leading to 101 data points per measurement. Additionally, the field is rotated by 5° after each individual measurement, so in order to cover a full circle, 72 different angles are considered. We note that this is similar to a radial frequency-domain trajectory dataset in conventional MRI. In [83], quadrupolar fields are used to generate such a dataset. However, the field we are using is only approximately quadrupolar, so it is not a true radial frequency-domain trajectory experiment. The system consists of $72 * 101 = 7272$ equations. The numerical phantom of 64x64 pixels is shown in Figure 2.1b, resulting in a matrix $\mathbf{A}$ of size 7272×4096. We assume that the repetition time T_R is

long enough for the magnetization vector to relax back to its equilibrium. Also, the echo time is assumed to be so short as to make T_2-weighting negligible.

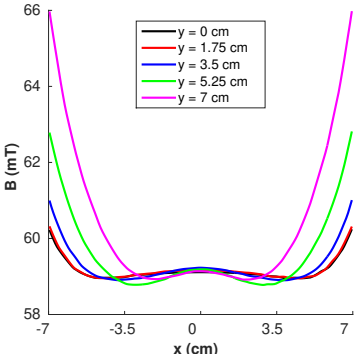


(a) The Halbach cylinder (b) Numerical phantom

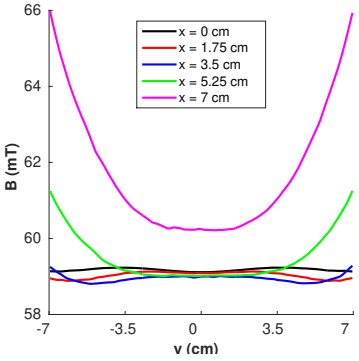
Figure 2.1: Low-field MRI prototype and numerical phantom.



(a) B_0



(b) 1D variations, y fixed



(c) 1D variations, x fixed

Figure 2.2: Magnetic field B_0 within the Field of View. (b) and (c) show the 1D variations in the FoV.

Since the background field is almost homogeneous in the center, as can be seen in Figure 2.2, we decided to place the object of interest in the numerical phantom off-center. Within a homogeneous region in the field, distinguishing between the different pixels is impossible. Another obstacle in the reconstruction process is the fact that the background field is almost symmetrical in both the x- and the y-axis, potentially leading to aliasing artifacts in the lower half of the image (because the object of interest is placed in the upper half of the image). We could reconstruct the phantom by leaving out all the columns in matrix $\mathbf{A}$ corresponding to the pixels in the lower half of the image. Another way of circumventing this problem is by using several receiver coils with different sensitivity maps to break the symmetry of the problem [83], [87]. However, we choose not to take these approaches so we can see how severe these artifacts are for the different objective functions.

The coil sensitivity c is assumed to be constant so it is left out of the calculations. White Gaussian noise is added so the covariance matrix $\mathbf{C}$ is simply the identity matrix. We assume an SNR of 20. The numerical experiments are carried out using MATLAB version 2015a. Often, CG is stopped once the residual is small enough. However, GCGLS and GCGME are solving different normal equations, so the residuals are different for both methods. Therefore, a comparison using such a stopping criterion would not be fair. Instead, a fixed number of CG iterations is used per IRLS iteration. The value of the regularization parameter τ is chosen heuristically. The number of IRLS iterations is set to 10. We consider both 10 and 1000 CG iterations per IRLS iteration. The initial guess $\mathbf{x}_0$ in GCGLS (and $\mathbf{r}_0$ in GCGME) is the zero vector. During the first IRLS iteration, we set $\mathbf{D} = \mathbf{I}$, which means that $\mathbf{R} = \mathbf{F}^*\mathbf{F}$. After the first IRLS iteration, we calculate the weight matrix $\mathbf{D}$ according to Eq. (2.43). We use warm starts, i.e., we use the final value of our iterate $\mathbf{x}_k$ (or $\mathbf{r}_k$ for GCGME) of the previous IRLS iteration as an initial guess for the next IRLS iteration.

2.3. RESULTS AND DISCUSSION

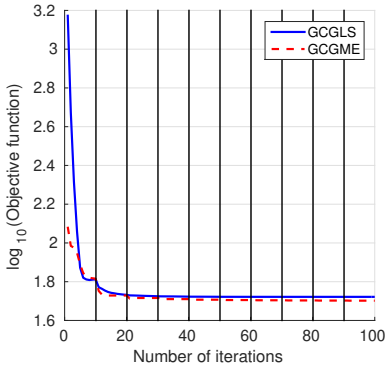
Table 2.2 shows the parameters that were chosen for all four different minimization problems. The regularization parameter was chosen heuristically in each case.

Table 2.2: Overview of the choice of parameters for the four different minimization problems considered in this work.

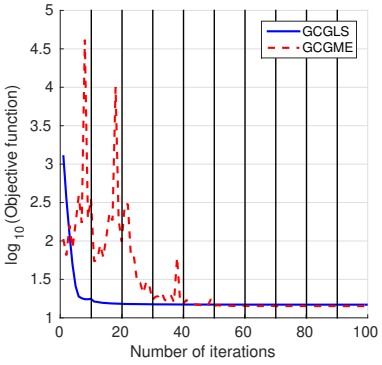
	Regularization matrix $\mathbf{I}$	Regularization matrix $\mathbf{F}$
ℓ_1-penalty	$\min_{\mathbf{x}} \frac{1}{2}\|\|\mathbf{Ax}-\mathbf{b}\|\|_2^2 + \tau\|\|\mathbf{x}\|\|_1$ 10 IRLS iterations 10 CG iterations $\tau = 1.5 \cdot 10^{-1}$	$\min_{\mathbf{x}} \frac{1}{2}\|\|\mathbf{Ax}-\mathbf{b}\|\|_2^2 + \tau\|\|\mathbf{Tx}\|\|_1$ 10 IRLS iterations 10 CG iterations $\tau = 1 \cdot 10^{-2}$
$\ell_{1/2}$-penalty	$\min_{\mathbf{x}} \frac{1}{2}\|\|\mathbf{Ax}-\mathbf{b}\|\|_2^2 + 2\tau\|\|\mathbf{x}\|\|_{1/2}^{1/2}$ 10 IRLS iterations 10 CG iterations $\tau = 5 \cdot 10^{-3}$	$\min_{\mathbf{x}} \frac{1}{2}\|\|\mathbf{Ax}-\mathbf{b}\|\|_2^2 + 2\tau\|\|\mathbf{Tx}\|\|_{1/2}^{1/2}$ 10 IRLS iterations 10 CG iterations $\tau = 2.5 \cdot 10^{-3}$

The objective function values are plotted as a function of the iteration number in

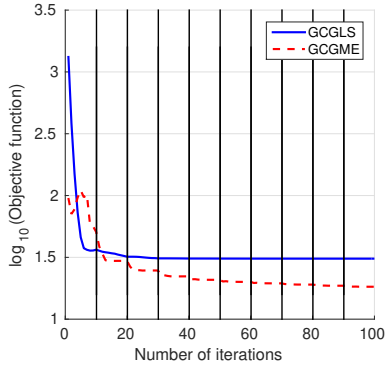
Figure 2.3 and all resulting images are shown in Figure 2.4. We note that in all cases (except perhaps the $\|\mathbf{x}\|_{1/2}^{1/2}$ one), GCGME yields a result that resembles the original more than GCGLS does. GCGLS tends to yield aliasing artifacts in the lower half of the image. This effect is less pronounced for the GCGME results, especially when $\|\mathbf{Tx}\|_{1/2}^{1/2}$ is used as the penalty term. We see that GCGME attains a lower objective function value in all cases. However, both methods should in theory converge to the same value for the $\|\mathbf{x}\|_1$- and $\|\mathbf{Tx}\|_1$-penalty terms. Evidently, GCGLS has not converged yet. If we increase the number of CG iterations to 1000, GCGLS and GCGME converge to the same result, as can be seen in Appendix C. The GCGME result is the same, whether 10 or 1000 CG iterations are carried out, from which we conclude that GCGME has already converged in the first case. However, GCGLS needs a significantly larger number of iterations to converge. In case $\mathbf{F} = \mathbf{I}$, GCGLS and GCGME both need 0.069 seconds per iteration. When $\mathbf{F} = \mathbf{T}$, GCGME needs slightly more time per iteration than GCGLS: 0.072 versus 0.069 seconds.



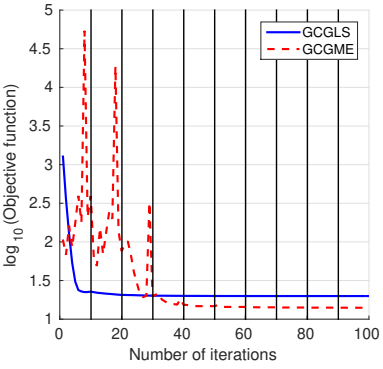
(a) Penalty term $\|\mathbf{x}\|_1$



(b) Penalty term $\|\mathbf{Tx}\|_1$

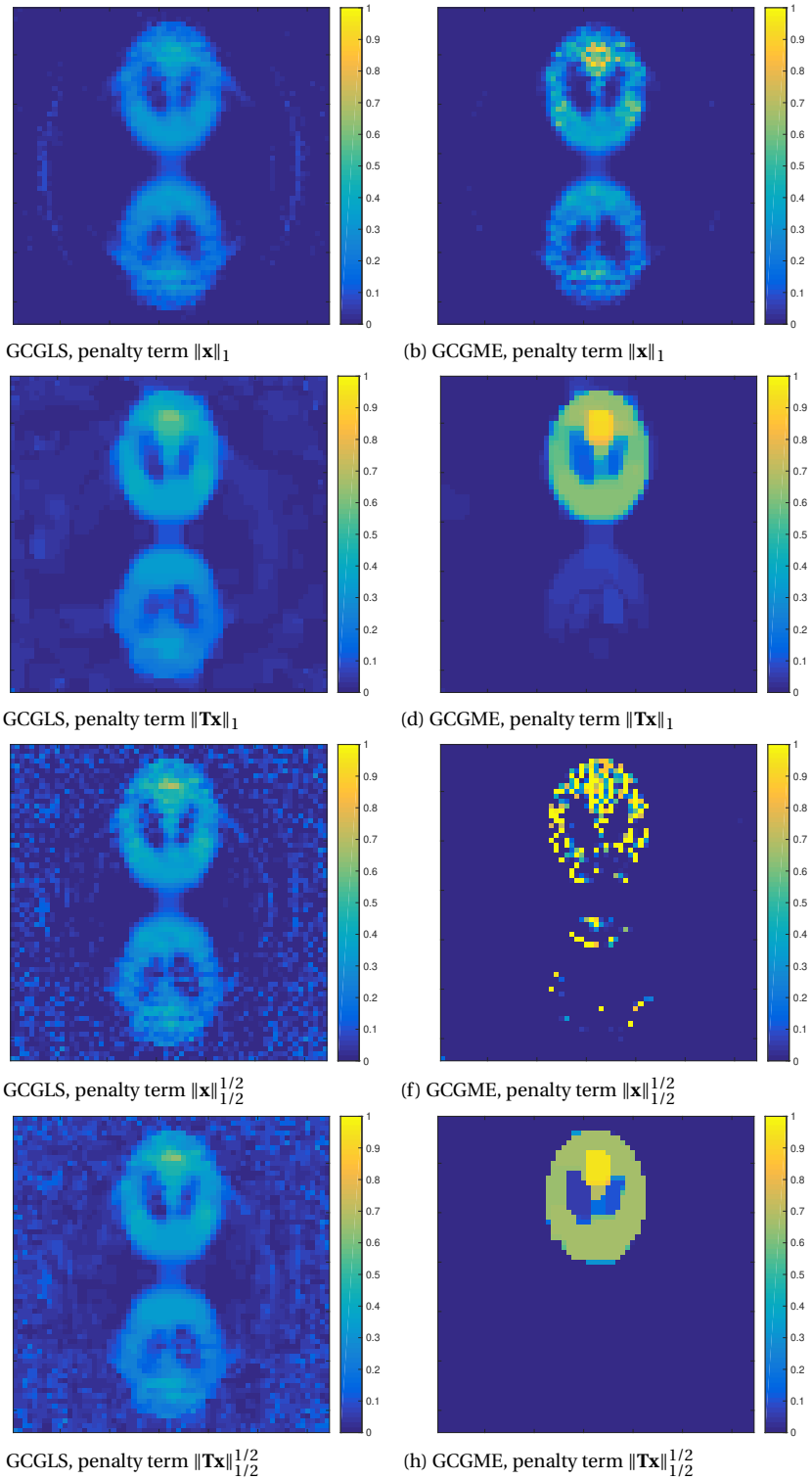


(c) Penalty term $\|\mathbf{x}\|_{1/2}^{1/2}$



(d) Penalty term $\|\mathbf{Tx}\|_{1/2}^{1/2}$

Figure 2.3: Objective function value as a function of the iteration number for the four different penalty terms. The vertical black lines indicate the start of a new IRLS iteration.



(a) GCGLS, penalty term $\|\mathbf{x}\|_1$

(b) GCGME, penalty term $\|\mathbf{x}\|_1$

(c) GCGLS, penalty term $\|\mathbf{Tx}\|_1$

(d) GCGME, penalty term $\|\mathbf{Tx}\|_1$

(e) GCGLS, penalty term $\|\mathbf{x}\|_{1/2}^{1/2}$

(f) GCGME, penalty term $\|\mathbf{x}\|_{1/2}^{1/2}$

(g) GCGLS, penalty term $\|\mathbf{Tx}\|_{1/2}^{1/2}$

(h) GCGME, penalty term $\|\mathbf{Tx}\|_{1/2}^{1/2}$

Figure 2.4: Reconstruction results for the four different penalty terms. In all four cases, the GCGLS and the GCGME result are shown.

2.3.1. DISCUSSION OF THE RESULTS

GCGLS needs a large number of CG iterations to converge, while for GCGME, this number is small (typically, 10 is sufficient). This can be explained by the observation that as we get closer to the solution, many elements of the vector $|\mathbf{F}\mathbf{x}_k|^{2-p}$ will converge to zero, due to the sparsity-enforcing properties of the ℓ_p penalty when $p \leq 1$. Therefore, $\mathbf{D}_k^{-1} = \text{diag}\left(|\mathbf{F}\mathbf{x}_k|^{2-p}\right)$ will contain an increasing number of very small entries, resulting in the matrix $\mathbf{R}_k = \mathbf{F}^H\mathbf{D}_k\mathbf{F}$ becoming more and more ill-conditioned as the number of IRLS iterations grows. That means that, after a few IRLS iterations, $\kappa_2(\mathbf{R}_k) \gg \kappa_2(\mathbf{I})$ will hold, in which case GCGME performs better than GCGLS, which is consistent with our results.

It is interesting to note that when the number of CG iterations for GCGLS is set to 10, GCGLS appears to have reached convergence after 4-5 IRLS iterations, yielding an image with aliasing artifacts in the form of an additional shape in the lower half of the image, as well as regions of intensity in the corners of the image. However, convergence is not actually attained yet, but it is simply very slow. The number of CG iterations needs to be increased to 1000 before convergence is reached.

We observe that the $\|\mathbf{T}\mathbf{x}\|_{1/2}^{1/2}$ penalty is best at repressing the aliasing artifacts in the lower half of the image.

2.4. CONCLUSION

We formulated a general MRI signal model describing the relationship between measured signal and image which is more suitable for low-field MRI because the assumptions that are usually made in high-field MRI do not necessarily hold in this case. The discretized version yields a linear system of equations that is ill-posed. Regularization is needed to obtain a reasonable solution. We considered the weighted and regularized least-squares problem. A second set of normal equations was derived, which allowed us to generalize the Conjugate Gradient Minimal Error (CGME) method to include nontrivial weighting and regularization matrices.

We compared our GCGME method to the classical GCGLS method by applying both to data simulated using our signal model. Different regularization operators were considered: the identity matrix and the anisotropic total variation operator that determines the size of the jumps between neighboring pixels. The regularization term was measured in the ℓ_1-norm and the $\ell_{\frac{1}{2}}$-norm and Iterative Reweighted least-squares (IRLS) was used to solve the resulting minimization problems. In each IRLS iteration, an ℓ_2-regularized minimization problem was solved using GCGLS or GCGME.

GCGME converges much faster than GCGLS, due to the regularization matrix becoming increasingly ill-conditioned as the number of IRLS iterations grows. This makes GCGME the preferred algorithm for our application.

3

TWO-DIMENSIONAL IMAGING USING A HAND-HELD SINGLE-SIDED MRI SENSOR: PRELIMINARY FINDINGS

3.1. INTRODUCTION

Several sensors based on single-sided nuclear magnetic resonance (NMR) have been developed over the course of the last few decades. While the principle is the same, such NMR sensors are called MRI sensors if they are used for medical purposes. NMR sensors do not place any restrictions on the size of the object being imaged as the sample is placed outside of the magnet, as opposed to conventional NMR (or MRI) scanners in which the object of interest is enclosed by the magnet. Hence, these sensors allow for near-surface scanning of arbitrarily large objects. A complete overview on single-sided (mobile) NMR can be found in [88] and [89]. Applications include well logging [90], food analysis [91, 92] and medical testing/imaging [40, 42]. If such a sensor is small and lightweight enough to be truly portable, it increases MRI accessibility.

An example of a one-sided NMR sensor is the NMR-MOUSE (Nuclear Magnetic Resonance MObile Universal Surface Explorer) as described in [38], which consists of two semicylindrical permanent magnets with antiparallel magnetization placed 13 mm apart. The magnets generate a static magnetic field parallel to the magnet surface. In the gap between the magnets, an RF coil is placed to excite the object and detect the NMR signal. In [39], another type of NMR-MOUSE was introduced, which is based on a simple bar magnet that generates a static magnetic field with a built-in, approximately constant gradient. The RF coil is placed on top of one of the poles. In [93], Casanova and Blümich mounted an appropriate gradient coil system on top of this NMR-MOUSE and demonstrated the feasibility of two-dimensional imaging, and in Perlo et al. [94], using

this same magnet, three-dimensional spatial resolution was achieved by combining a two-dimensional imaging sequence with the application of slice selection pulses.

Ever since the NMR-MOUSE was introduced over two decades ago, different research groups have developed their own single-sided NMR sensors. Examples focusing on hardware design can be found in [43, 95–100]. Other papers describe the hardware design and present 2D images as well. For example, in [101] 2D images were obtained using a single-sided NMR sensor based on an adjustable array of permanent magnets. He et al. [40] introduced a design based on a reduced Halbach magnet meant to assess the depth of burn injuries and showed promising preliminary results, including images of a simple burn model. Greer et al. [41] noted that, due to their commercial unavailability, single-sided NMR sensors are still relatively uncommon and introduced an NMR sensor that is easily reproducible, single-sided and portable and they presented a 2D image that was reconstructed from data acquired using this sensor.

The single-sided MRI sensor we consider, which is shown in Figure 1.1d, was developed at the Leiden University Medical Center. It consists of a configuration of 36 neodymium permanent magnets. Each magnet is of size $12\times12\times12$ mm^3. The orientation and location of each of the magnets was determined using an artificial intelligence (AI) model that was developed at the LUMC, which combines a deep neural network-based encoder and an analytical forward model [102]. This model aims to find a configuration of the 36 magnets that generates a static magnetic field with a built-in linear gradient and is as homogeneous as possible. The resulting magnetic field has a linearly increasing static magnetic field with a gradient of approximately -0.5 T/m at a height of 2 cm above the surface of the magnets, allowing for slice selection. The output of the AI model was used to 3D print a casing to hold the 36 magnets. The resulting device is $10\times10\times2$ cm^3 in size and weighs less than 1 kilogram. Possible applications of this scanner include the imaging of subcutaneous structures and the spine.

As far as we know, all 2D (and 3D) images presented in single-sided NMR literature have been reconstructed based on data acquired using a magnet equipped with gradient coils. Adding such gradient coils to a magnet enables 2D (or 3D) spatial encoding, but it also makes the sensor more complex. Therefore, in our design, we decided to forgo the gradient coils. The inhomogeneities in the static magnetic field will be used for spatial encoding instead, as is done in for example [103] and [104]. However, due to the nonbijective nature of the magnetic field generated by the magnet, as was the case in Chapter 2, a single measurement will not yield a sufficient amount of information for a reconstruction. Therefore, we will use an approach that is similar to the one we took in Chapter 2: several measurements are carried out, with the object of interest being translated within the field of view (FoV) between subsequent measurements.

3.2. IMAGE RECONSTRUCTION

The relationship between the measured signal $b(t)$ and the spin density $X(\vec{r})$ of the object of interest is given by

$$b(t) = \int_{\vec{r}\in\mathbb{D}} c(\vec{r})X(\vec{r})e^{-\mathrm{i}\Delta\omega(\vec{r})t}\,\mathrm{d}\vec{r}, \tag{3.1}$$

with $c(\vec{r})$ the receiver coil sensitivity and $\Delta\omega(\vec{r})$ defined as

$$\Delta\omega(\vec{r}) = \gamma B(\vec{r}) - \omega_{\text{mod}}, \tag{3.2}$$

where γ is the gyromagnetic ratio, B the static magnetic field strength and ω_{mod} the demodulation frequency. We can discretize Eq. (3.1), yielding the linear system of equations

$$\mathbf{b}^{(1)} = \mathbf{A}^{(1)}\mathbf{x}, \tag{3.3}$$

where $\mathbf{b}^{(1)}$ is the signal $b(t)$ measured at discrete time instances, $\mathbf{x}$ is the unknown image (whose values represent the spin density in each voxel) and $\mathbf{A}^{(1)}$ is the model matrix whose elements, ignoring constants, are described by

$$a_{jk} = c(\vec{r}_k)e^{-i\Delta\omega(\vec{r}_k)t_j}, \tag{3.4}$$

with j denoting the jth time sample and k denoting the voxel number, using lexicographic ordering. To acquire a sufficient amount of information for a reconstruction, we carry out a number of measurements, between which the sample is translated. Denoting the translation vector corresponding to measurement l by $\vec{r}^{(l)}$, we can write

$$\mathbf{b}^{(l)} = \mathbf{A}^{(l)}\mathbf{x}, \tag{3.5}$$

with

$$a_{jk}^{(l)} = c(\vec{r}_k - \vec{r}^{(l)})e^{-i\Delta\omega(\vec{r}_k - \vec{r}^{(l)})t_j}. \tag{3.6}$$

Then, we can combine all L measurements into one system of equations

$$\begin{pmatrix} \mathbf{b}^{(1)} \\ \mathbf{b}^{(2)} \\ \vdots \\ \mathbf{b}^{(L-1)} \\ \mathbf{b}^{(L)} \end{pmatrix} = \begin{pmatrix} \mathbf{A}^{(1)} \\ \mathbf{A}^{(2)} \\ \vdots \\ \mathbf{A}^{(L-1)} \\ \mathbf{A}^{(L)} \end{pmatrix} \mathbf{x}, \tag{3.7}$$

which we will write as

$$\mathbf{b} = \mathbf{A}\mathbf{x}. \tag{3.8}$$

Instead of solving Eq. (3.8) directly, we solve the more well-posed minimization problem

$$\arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{T}\mathbf{x}\|_1, \tag{3.9}$$

where $\mathbf{T}$ is defined as the 3D anisotropic total variation (TV) operator

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_x \\ \mathbf{T}_y \\ \mathbf{T}_z \end{pmatrix}, \tag{3.10}$$

with $\mathbf{T}_x$, $\mathbf{T}_y$ and $\mathbf{T}_z$ first-order differencing matrices that calculate the values of the jumps between neighboring pixels in the x-, y- and z-direction respectively.

3.3. EXPERIMENT

The static magnetic background field was mapped on a grid of 30x30x10 mm^3, between 22 mm and 32 mm above the magnet, with a step size of 1 mm, using an F71 Lakeshore Teslameter with a 3-axis Hall probe. We interpolate the resulting field map to a resolution of 0.25 mm in the z-direction (i.e., perpendicular to the magnet). This choice was motivated by the observation that the field varies significantly in the z-direction, much more so than in the other two directions, and it is likely that the original resolution is not sufficient for an accurate discretization of the field. We leave the resolution in the other two directions unaltered so as not to make the problem more underdetermined than necessary. The demodulation frequency corresponds to 49.866 mT, which approximately coincides with the center of the magnetic field "slice" at $z = 28.5$ mm.

The complete experimental setup is shown in Figure 3.1. In this setup, the configuration of magnets is positioned inside a Faraday cage, which was mounted on top of an xy-stage. This stage allows for the translation of the magnets inside the Faraday cage. The magnets are located just below the sample housing, which contains the phantom, as well as an RF coil circuit. The RF coil is a copper wire with a diameter of 0.8 mm that was wound around the phantom. The sample housing can be removed from the setup.

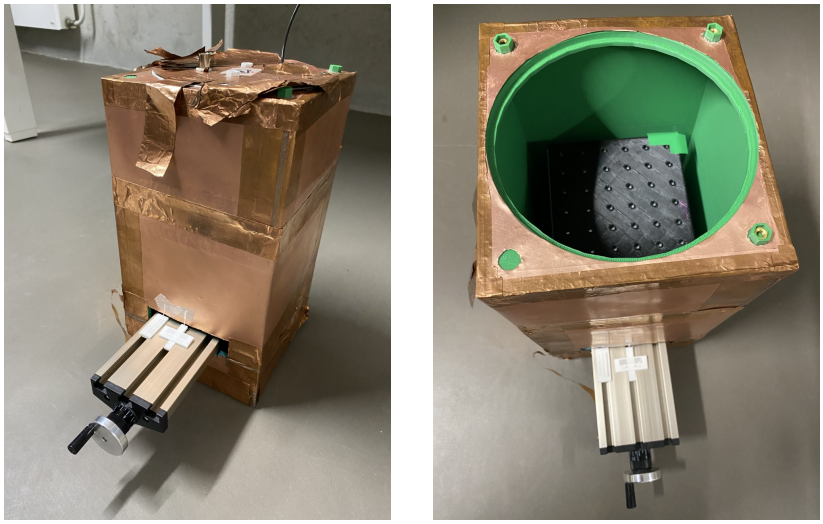
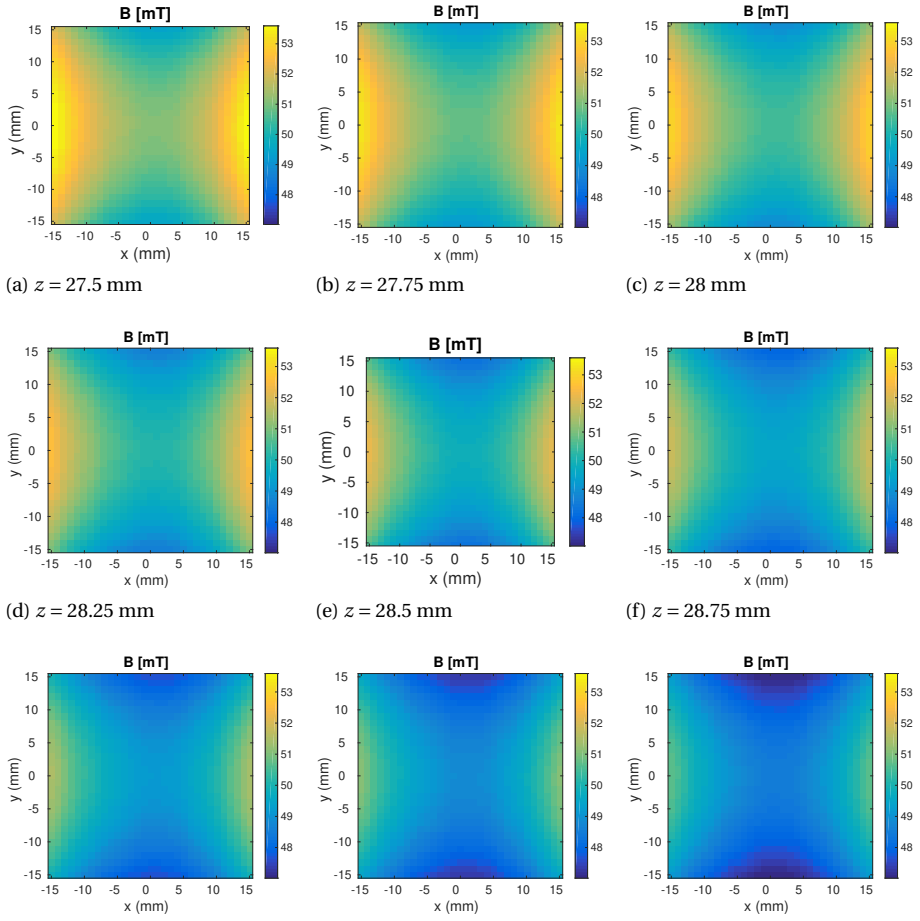


Figure 3.1: On the left, the experimental setup is shown. On the right, the lid has been removed, showing the magnet. The lid functions as a sample housing. It contains the phantom and the RF coil circuit.

Figure 3.2 shows the magnetic field map for nine different equidistant slices between $z = 27.5$ and $z = 29.5$ mm. The coil sensitivity maps for the same slices are shown in Figure 3.3. These maps were obtained by first determining the reflection coefficient S_{11} of the RF coil as a function of frequency and then mapping these values to the corresponding frequencies in the field. We note that the coil sensitivity plots exhibit contour lines which are of a somewhat unusual nature. Their appearance can be explained by

the fact that during the design phase, this specific configuration of magnets was selected because the transverse magnetic field components exhibit an almost linear decay. However, the z-component is not negligible, as is reflected in the coil sensitivity maps. We observe that in each slice, a relatively small part of the FoV is "seen" by the coil.

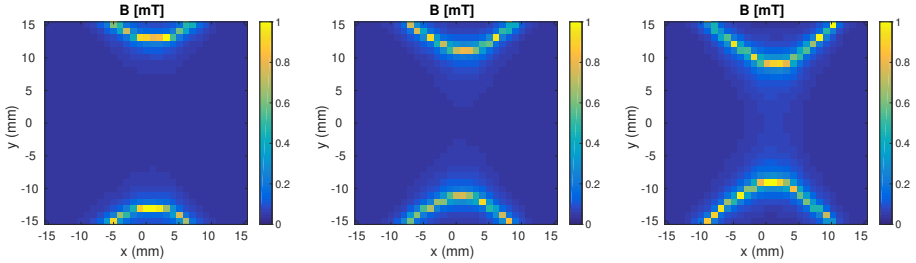


(a) $z = 27.5$ mm (b) $z = 27.75$ mm (c) $z = 28$ mm

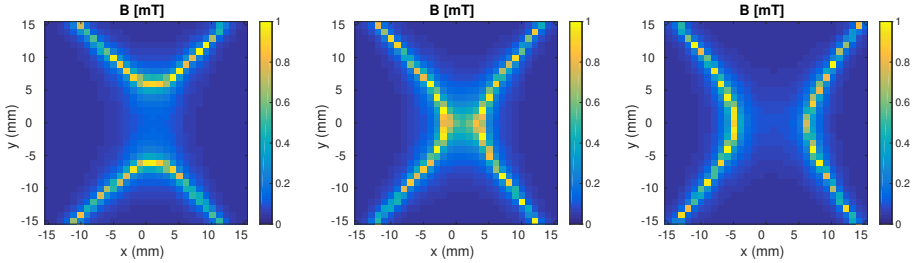
(d) $z = 28.25$ mm (e) $z = 28.5$ mm (f) $z = 28.75$ mm

(g) $z = 29$ mm (h) $z = 29.25$ mm (i) $z = 29.5$ mm

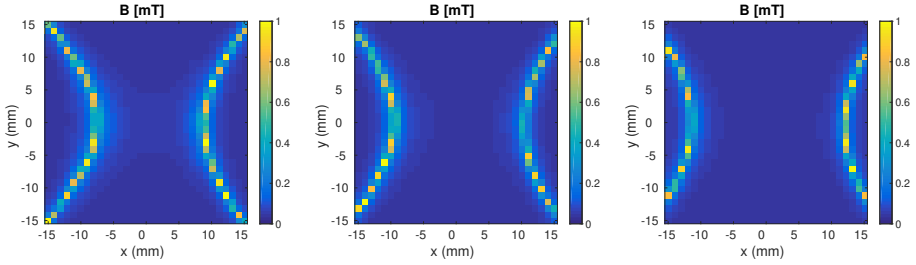
Figure 3.2: Magnetic field strength for different values of z.



(a) $z = 27.5$ mm (b) $z = 27.75$ mm (c) $z = 28$ mm



(d) $z = 28.25$ mm (e) $z = 28.5$ mm (f) $z = 28.75$ mm



(g) $z = 29$ mm (h) $z = 29.25$ mm (i) $z = 29.5$ mm

Figure 3.3: Coil sensitivity maps for different values of z.

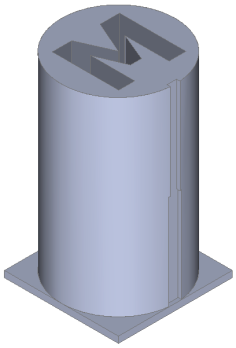


Figure 3.4: Schematic used for the 3D printing of the phantom. The tube is 30 mm in length and 18 mm in diameter. The depth of the M-shaped cutout is equal to half of the tube length.

For the first test of the setup and the image reconstruction method, we use a simple phantom in the shape of the letter M. The schematic used for the 3D printing of the mold is shown in Figure 3.4. The M-shaped cutout was filled with olive oil. We acquire a signal for seven different translations of this object, with the translations being 2 mm in the x- and/or y-direction. For each translation, the acquisition parameters were as follows: repetition time 200 ms, echo time 20 ms, dwell time 5 μs, 512 data points, 2000 averages.

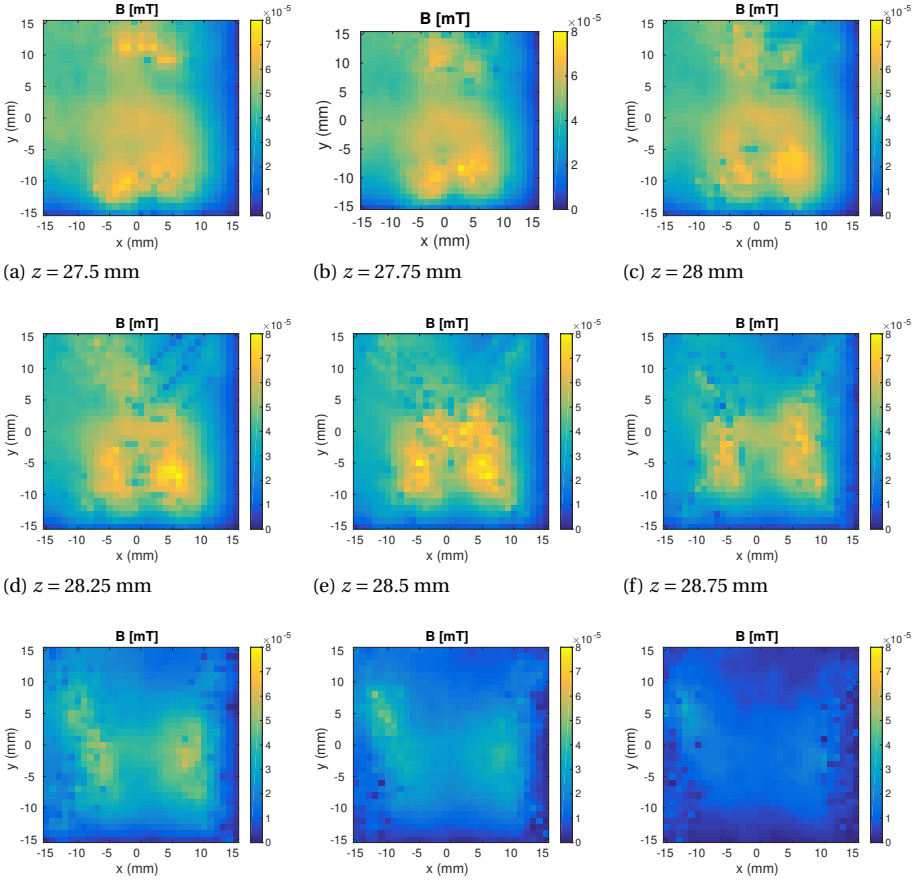
3.4. RESULTS

We solve minimization problem (3.9) for this experiment, using the Alternating Direction Method of Multipliers (ADMM) [105]. The resulting 3D image is shown in Figure 3.5. We see that, while the result is far from perfect, the shape of an M can be discerned.

3.5. DISCUSSION

While the result we obtained is promising, we are far from being able to make any definitive claims about the quality of our image. The M-shape is clearly visible in some slices. However, we noticed that the shape of the coil sensitivity map seems to influence the outcome, with bands of higher intensity where the coil sensitivity is at its highest. We cannot completely rule out the possibility that the coil sensitivity pattern could have influenced the outcome of the image reconstruction process in such a way that symmetric shapes with vertical or horizontal lines, like an M-shape, are favored, especially considering that we used a type of regularization that encourages piece-wise constant solutions. Other factors which are known to have an impact on the final solution include the accuracy of the field map, the value of the regularization parameter and the accuracy of the translations. Therefore, to ensure the consistency of our approach, we need to carry out more experiments with different kinds of phantoms.

To obtain a sufficient amount of information for a reconstruction, we carried out seven measurements, with the phantom being translated by 2 millimeters between them. These translations were executed by hand, which is a relatively inaccurate method for such short distances. It is not unlikely that this had a negative impact on the reconstruc-



(a) $z = 27.5$ mm (b) $z = 27.75$ mm (c) $z = 28$ mm

(d) $z = 28.25$ mm (e) $z = 28.5$ mm (f) $z = 28.75$ mm

(g) $z = 29$ mm (h) $z = 29.25$ mm (i) $z = 29.5$ mm

Figure 3.5: Slices of the reconstructed three-dimensional image.

tion. One of the next steps is the introduction of a stepper motor driver that can be used to move the stage on which the phantom is placed in an automated manner. This should eliminate the translations as a source of inaccuracy in the model.

In conventional MRI scanners, the static magnetic background field is oriented in the z-direction, i.e., along the bore. The magnetic field being oriented exclusively in one of the three Cartesian directions allows for a relatively simple model describing the relationship between signal and image. The main magnetic field component of the MRI sensor we considered points in the x-direction, by which we mean one of the two directions parallel to the surface of the magnet. However, the other two components are not negligible. In the center of each slice, these two components are very small, but moving 5 mm in each direction yields a z-component which is about 10% of the x-component, and if we move even further we encounter percentages of more than 30%. It is yet unclear to what extent these concomitant field components could invalidate the signal model. By starting out with the Bloch equation, it is possible to derive a model describing the relationship between signal and image for a general static magnetic background field, see Appendix D. However, the resulting expression depends on some quantities that are hard to determine, most notably the orientation of the magnetization. In our current setup, it is unclear what this orientation is. In theory, though, it is possible to determine the orientation of the magnetic moment in each voxel by numerically solving the Bloch equation describing the precession of the magnetization vector before readout. To do that, we need, besides a map of the B_1-field, all parameters describing the RF pulse, which is far from trivial. Due to the very high frequency of the B_1-pulse, the time step in the numerical integration scheme has to be very small. This makes carrying out this analysis very computationally expensive, especially for a 3D scenario. Additionally, we are working with a spin echo sequence. To accurately model the orientation of the magnetization at the time we are acquiring the signal, we would need to incorporate T_2 and T_2^* (and possibly T_1) relaxation. While this might still be feasible for a phantom with known relaxation times, it is not when attempting to image human tissues, which is the application we are interested in. Therefore, we decided to not pursue using the more general signal model described in Appendix D any further. Instead, we simply stuck to the original signal model of Eq. (3.1).

3.6. CONCLUSION

In this chapter, we described our first attempt at image reconstruction using data acquired with a single-sided portable MRI sensor. Instead of mounting gradient coils on top of the sensor for spatial encoding, we used an accurate map of the relatively inhomogeneous static magnetic field generated by the sensor's array of permanent magnets. Since a single measurement does not yield a sufficient amount of information for a reconstruction, we carried out a number of different measurements, with the sample being translated between them. We used a sample in the shape of the letter M for our experiment. Our reconstruction results showed an approximate outline of an M in a number of slices.

4

Low-Field Magnetic Resonance Imaging Using Multiplicative Regularization

4.1. Introduction

A direct consequence of having a weaker background field is that low-field MR scanners inevitably yield significantly noisier signals than their high-field counterparts. Moreover, depending on the type of low-field scanner that is used, the background field may not be (sufficiently) uniform throughout the object and the gradient fields that are used for imaging may not be linear or are only approximately linear within some subdomain of the object or body part that we want to image. Loosely speaking, a uniform background field and linear gradient fields ultimately lead to a Fourier transform representation of the measured MR signals and MR imaging essentially amounts to applying an inverse Fourier transform to the measured data. Deviations from these ideal background and gradient fields lead to imaging artifacts when an inverse Fourier transform is applied,

This chapter is based on the article:

M.L. de Leeuw den Bouter, M.B. van Gijzen, and R.F. Remis, *Low-field magnetic resonance imaging using multiplicative regularization*, Magnetic Resonance Imaging, **75**, 21-33 (2021)

and additional work.

as can be seen in [37]. Even when the background field and gradient fields can be considered constant and linear throughout the object, then still noisy reconstructions are obtained, since the acquired signals usually have a low signal-to-noise ratio (SNR).

There is a vast literature on image reconstruction in MRI. An excellent overview of many relevant techniques is given in [106]. In the case of nonlinear gradients, a model-based image reconstruction as described in [44] and [104], among others, can be employed instead of the standard inverse Fourier Transform. In this chapter, we will make use of this model-based image reconstruction as well. However, other methods exist that correct for nonlinear magnetic fields as a post-processing step, i.e. these methods are applied to the image obtained after having applied a standard technique like the inverse Fourier Transform to the k-space data. In [107] for example, spherical harmonic deconvolution methods are used to achieve this result.

In this chapter, we employ an MR imaging technique that, using model-based image reconstruction, addresses the aforementioned issues of imaging using nonlinear magnetic fields and the contamination of the signal by noise. Specifically, we pose our low-field imaging problem as an optimization problem that minimizes a (least-squares) data fidelity term in which nonuniform background fields and nonlinear gradients are taken into account through a generalized signal model. Furthermore, noise effects are suppressed by incorporating a weighted ℓ_2-norm total variation objective function into the optimization procedure. Total variation regularization penalizes jumps between neighboring pixels. The additive variant is often used in MR imaging to denoise images while still maintaining their edges, see for example Chapter 2 and [108–111], among many others.

It is customary to include such a regularization term in an additive manner into an optimization framework. However, one of the main drawbacks of such an additive scheme is that an artificial regularization parameter must be chosen that balances the data fidelity and regularization terms. While methods for choosing this parameter exist, see for example [112], they are often computationally expensive and do not allow for fast (real-time) imaging. Moreover, typically a regularizing parameter needs to be determined for each new available data set, which leads to even larger computation times in case multiple data sets need to be processed.

Inspired by the success of multiplicative regularization in wave field inversion, see [113–116], for instance, we include regularization in a multiplicative manner as well. In the resulting iterative imaging algorithm, the image is then updated using a Polak-Ribière type of conjugate gradient directions, see for example [117]. Two practical advantages of this scheme are that there is no effective regularization parameter that needs to be computed and reconstruction results can be monitored as the scheme progresses. A similar multiplicative regularization approach was applied to image deblurring problems in [118].

We apply our imaging method to low-field noise-corrupted simulated data using nonlinear gradient fields and to measured data obtained with the low-field scanner shown in Figure 1.1c.

We consider k-space data corresponding to an apple, a melon and the brain of a healthy volunteer, which were acquired using the MR scanner described in [20]. Moreover, the effects of a nonlinear frequency encoding gradient are also studied using noise-

corrupted simulated data. We resort to simulated data in this case, since we do not have measured background or gradient fields available. We do stress, however, that if such measured fields are available, then these can be easily incorporated into our imaging scheme. Imaging results will be presented for a regularizing objective function based on the minimization of a weighted total variation term. We demonstrate that the method indeed effectively suppresses noise for a given data set without any extra computations to determine a regularization parameter.

4.2. SIGNAL MODELING AND REGULARIZATION

The starting point of our imaging procedure is the voltage signal $b(t)$ that is measured by a receive coil of the low-field scanner. Assuming that the complete object is excited, we have for this signal the representation

$$b(t) = \int_{\mathbf{r}\in\mathbb{D}} e^{-\mathrm{i}\Delta\omega_0(\mathbf{r})\,t}\, e^{-\mathrm{i}2\pi\mathbf{k}(\mathbf{r},t)\cdot\mathbf{r}}\, X(\mathbf{r})\,\mathrm{d}V, \tag{4.1}$$

where $\Delta\omega_0(\mathbf{r}) = \gamma B_0(\mathbf{r}) - \omega_{\mathrm{mod}}$ is the difference between the local Larmor frequency $\gamma B_0(\mathbf{r})$ and the demodulation frequency ω_{mod}, with γ the proton gyromagnetic ratio. Furthermore, $X(\mathbf{r})$ is the image and $\mathbb{D}$ is the imaging domain or Field of View (FoV). The image X depends on the proton density, coil sensitivity, etc. and is expressed in $[\mathrm{Vm}^{-3}]$. Note that here, we used a slightly different version of Eq. (2.6). Finally,

$$\mathbf{k}(\mathbf{r},t) = \frac{\gamma}{2\pi}\int_{\tau=0}^{t} \mathbf{G}(\mathbf{r},\tau)\,\mathrm{d}\tau, \tag{4.2}$$

where $\mathbf{G}(\mathbf{r},\tau) = G_x(\mathbf{r},\tau)\mathbf{i}_x + G_y(\mathbf{r},\tau)\mathbf{i}_y + G_z(\mathbf{r},\tau)\mathbf{i}_z$ $[\mathrm{Tm}^{-1}]$ is the gradient vector corresponding to the application of the gradient coils of the scanner. This model takes background field inhomogeneities into account through $\Delta\omega_0(\mathbf{r})$ and nonuniformities in the gradient fields are modeled using the spatially dependent gradient vector $\mathbf{G}$. In case the background field can be considered uniform throughout the object, we have $B_0(\mathbf{r}) = B_0$, and the gradient vector does not depend on position, which means that $\mathbf{G}(\mathbf{r},\tau) = \mathbf{G}(\tau)$, our signal model simplifies to

$$b(t) = \tilde{X}[\mathbf{k}(t)] = \int_{\mathbf{r}\in\mathbb{D}} e^{-\mathrm{i}2\pi\mathbf{k}(t)\cdot\mathbf{r}}\, X(\mathbf{r})\,\mathrm{d}V, \tag{4.3}$$

where we have taken $\omega_{\mathrm{mod}} = \gamma B_0$ and

$$\mathbf{k}(t) = \frac{\gamma}{2\pi}\int_{\tau=0}^{t} \mathbf{G}(\tau)\,\mathrm{d}\tau. \tag{4.4}$$

In this case the signal $b(t)$ is a three-dimensional spatial Fourier transform $\tilde{X}[\mathbf{k}(t)]$ of the image $X(\mathbf{r})$.

The data used for imaging consists of time samples of the voltage signal. Introducing the time instances $t_n = (n-1)\Delta t$ for $n = 1,2,\ldots,N$ with $\Delta t > 0$ and taking into account that the image X has the domain $\mathbb{D}$ as its support, we may write

$$b(t_n) = \int_{\mathbf{r}\in\mathbb{D}} e^{-\mathrm{i}\Delta\omega_0(\mathbf{r})\,t_n}\, e^{-\mathrm{i}2\pi\mathbf{k}(\mathbf{r},t_n)\cdot\mathbf{r}}\, X(\mathbf{r})\,\mathrm{d}V, \tag{4.5}$$

for $n = 1, 2, ..., N$.

We discretize the unit cube by subdividing the domain into nonoverlapping voxels, where each voxel has a positive side length Δx, Δy, and Δz in the x-, y-, and z-direction, respectively. The volume of a voxel is denoted by $V_\Delta = \Delta x \Delta y \Delta z$. Then, we discretize Having introduced the voxelization of our imaging domain, we discretize the integral in (4.5) using the composite midpoint rule and obtain

$$\mathbf{b} = \mathbf{A}\mathbf{x}, \tag{4.6}$$

where

$$\mathbf{b} = [b(t_1), b(t_2), ..., b(t_N)]^T, \tag{4.7}$$

and $\mathbf{A}$ is the matrix representation of the integral operator in Eq. (4.5). In case we have a homogeneous background field and linear gradient fields and $N = PQR$, we use $\mathbf{A} = \mathscr{F}$, where $\mathscr{F}$ is the 3D unitary Discrete Fourier Transform (DFT) matrix, so $\mathbf{A}^H = \mathscr{F}^{-1}$.

In practice, the measured data contains noise. Now in high-field MRI, the noise level is much lower than in low-field MRI. Simply applying an inverse 3D Fourier transform to the measured signal generally yields images of excellent quality in high-field MRI. In a low-field setting, however, the background and gradient fields may not be hommogeneous and linear, respectively, within the complete object and, as mentioned earlier, the required signals typically have a much lower SNR. Consequently, even when the background and gradient fields can be considered uniform, a straightforward application of an inverse Fourier transform to the measured data will lead to very noisy reconstructions in general.

To address both of these issues, we pose the imaging reconstruction problem as an optimization problem and minimize an objective function that consists of a data fidelity term describing the mismatch between the observed and the modeled data, and a regularizing term, which suppresses the influence of noise on the reconstructions. In particular, in many regularized optimization methods the objective function that needs to be minimized is of the form

$$F(\mathbf{x}) = F^{\text{data}}(\mathbf{x}) + \lambda F^{\text{reg,add}}(\mathbf{x}), \tag{4.8}$$

where $F^{\text{data}}(\mathbf{x})$ is the data fidelity term, $F^{\text{reg,add}}(\mathbf{x})$ some regularizing function (often chosen to be a variant of the total variation (TV) operator), and λ is a regularization parameter. The main drawback of using objective functions of the form (4.8) is that to reconstruct an image, the artificial regularization parameter λ needs to be selected for each new data set that is acquired. Strategies for choosing this parameter exist, of course (such as the L-curve method [45]), but such approaches are computationally demanding and generally require extensive numerical experimentation for each new available data set.

Instead of using an additive approach, we follow [113, 114, 116, 118], for example, and set up an iterative reconstruction algorithm that is based on multiplicative regularization in which at each iteration the objective function is of the form

$$F(\mathbf{x}) = F^{\text{data}}(\mathbf{x}) F^{\text{reg,mult}}(\mathbf{x}). \tag{4.9}$$

Here, $F^{\text{reg,mult}}$ is the regularizing function, which is, in general, of a different form than the regularizer in Eq. (4.8). Specifically, for multiplicative regularization, it is usually

defined such that $F^{\text{reg,mult}}(x) = 1$ at an optimal point. A major advantage of such a multiplicative approach is that no extra computations are required to determine an effective regularization parameter. Here, we focus on the application of multiplicative regularization to invert low-field MR data. For theoretical properties of multiplicative regularization the reader is referred to [114, 119, 120], for example.

As a first step we introduce the standard 2-norm data misfit objective function as

$$F^{\text{data}}(\mathbf{x}) = \frac{\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2}{\|\mathbf{b}\|_2^2}. \tag{4.10}$$

Subsequently, we set up our iterative scheme and assume that at the kth iteration we have some approximation $\mathbf{x}_{k-1}$ of the image available. The next iterate $\mathbf{x}_k$ is now constructed by minimizing the objective function

$$F_{k-1}(\mathbf{x}) = F^{\text{data}}(\mathbf{x}) F^{\text{TV}}_{k-1}(\mathbf{x}), \tag{4.11}$$

where $F^{\text{TV}}_{k-1}(\mathbf{x})$ is the discretized counterpart of the weighted ℓ_2-norm total variation functional

$$\mathscr{F}^{\text{TV}}_{k-1}(X) = \int_{\mathbf{r}\in\mathbb{D}_u} \frac{|\nabla X|^2 + \delta^2_{k-1}}{|\nabla X_{k-1}|^2 + \delta^2_{k-1}} \, \mathrm{d}V. \tag{4.12}$$

Here, X and X_{k-1} are the continuous counterparts of $\mathbf{x}$ and $\mathbf{x}_{k-1}$, respectively, and δ^2_{k-1} is given by

$$\delta^2_{k-1} = F^{\text{data}}(\mathbf{x}_{k-1})^2 \int_{\mathbf{r}\in\mathbb{D}_u} |\nabla X_{k-1}|^2 \, \mathrm{d}V. \tag{4.13}$$

In our algorithm, we only work with this particular choice of δ^2_{k-1}, but other choices are possible as well, see [113] and [115]. In each iteration, Eq. (4.11) is minimized. This choice of $F^{\text{TV}}(\mathbf{x})$ promotes solutions that are piecewise constant. Note that if this process converges, we have $\lim_{k\to\infty} \mathbf{x}_{k-1} = \mathbf{x}$ and hence $\lim_{k\to\infty} \nabla\mathbf{x}_{k-1} = \nabla\mathbf{x}$, which means that $\lim_{k\to\infty} F^{\text{TV}}_{k-1}(\mathbf{x}) = 1$.

For ill-posed problems, as encountered in wavefield imaging and contrast source inversion (CSI), for example, the total objective function typically shows a convergence behavior as descibed in [113, 114]. Specifically, in CSI the data functional is typically large, at the start of the iterative process, giving a large weight to the regularizing functional. As the iterative process progresses, the regularizing function will start to approach a value of 1 and the focus will shift towards minimizing the data functional, as illustrated in [121]. However, here we consider imaging problems which are well-posed (Fourier transform) or at least less ill-posed than in CSI, which means that, given a reasonable initial guess, the value of the data functional will be small. Therefore, we cannot expect the same behavior, except that if the process converges, the regularizing functional will converge to 1.

4.3. NUMERICAL DISCRETIZATION

To arrive at the discretized counterpart $F_{k-1}(\mathbf{x})$ of $\mathscr{F}_{k-1}(X)$, we use the weighting function

$$w_{k-1}(\mathbf{r}) = \frac{1}{|\nabla X_{k-1}|^2 + \delta^2_{k-1}} \tag{4.14}$$

and write Eq. (4.12) as

$$\mathscr{F}_{k-1}^{\mathrm{TV}}(X) = \delta_{k-1}^2 \int_{\mathbf{r}\in\mathbb{D}_{\mathrm{u}}} w_{k-1}(\mathbf{r})\,\mathrm{d}V + \int_{\mathbf{r}\in\mathbb{D}_{\mathrm{u}}} w_{k-1}(\mathbf{r})|\nabla X(\mathbf{r})|^2\,\mathrm{d}V,. \tag{4.15}$$

where $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$ is the nabla operator. Discretizing the integrals in (4.15) using the composite midpoint rule, we obtain

$$\begin{aligned}\mathscr{F}_{k-1}^{\mathrm{TV}}(X) \approx \delta_{k-1}^2 V_\Delta \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R} w_{k-1}(x_p, y_q, z_r) \\ + V_\Delta \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R} w_{k-1}(x_p, y_q, z_r)|\nabla X|^2(x_p, y_q, z_r),\end{aligned} \tag{4.16}$$

with P, Q and R denoting the number of voxels in the x-, y- and z-direction, respectively. The partial derivatives in the gradient operator on the right-hand side of Eq. (4.16) are approximated by two-point forward or backward difference formulas with a Dirichlet boundary condition. We can use matrices $\mathbf{D}_{x;f}$, $\mathbf{D}_{y;f}$ and $\mathbf{D}_{z;f}$ ($\mathbf{D}_{x;b}$, $\mathbf{D}_{y;b}$ and $\mathbf{D}_{z;b}$) to carry out forward (or backward) differencing in the x-, y- and z-direction. Denoting the 3D image array by $\mathbf{X}$, we introduce its vectorized counterpart as $\mathbf{x} = \mathrm{vec}(\mathbf{X})$, where we use lexicographical ordering. Similarly, $\mathbf{w}_{k-1}$ represents the vector form of Eq. (4.14).

Having introduced these vectors and matrices, we can write Eq. (4.16) more compactly as

$$\begin{aligned}\mathscr{F}_{k-1}^{\mathrm{TV}}(X) &\approx \delta_{k-1}^2 V_\Delta \mathbf{e}^T \mathscr{W}_{k-1}\mathbf{e} + V_\Delta \left(\mathbf{g}_x^H \mathscr{W}_{k-1}\mathbf{g}_x + \mathbf{g}_y^H \mathscr{W}_{k-1}\mathbf{g}_y + \mathbf{g}_z^H \mathscr{W}_{k-1}\mathbf{g}_z\right) \\ &= \delta_{k-1}^2 V_\Delta \mathbf{e}^T \mathscr{W}_{k-1}\mathbf{e} + V_\Delta \sum_{i=x,y,z} \mathbf{g}_i^H \mathscr{W}_{k-1}\mathbf{g}_i,\end{aligned} \tag{4.17}$$

where $\mathbf{e}$ is the $PQR \times 1$ vector of ones, $\mathscr{W}_{k-1} = \mathrm{diag}(\mathbf{w}_{k-1})$, and

$$\mathbf{g}_x = \mathbf{D}_x\mathbf{x}, \quad \mathbf{g}_y = \mathbf{D}_y\mathbf{x}, \quad \text{and} \quad \mathbf{g}_z = \mathbf{D}_z\mathbf{x} \tag{4.18}$$

with $\mathbf{D}_\xi$ a forward or backward differencing matrix. Finally, substituting the gradient vectors of Eq. (4.18) into Eq. (4.17) gives

$$\mathscr{F}_{k-1}^{\mathrm{TV}}(X) \approx F_{k-1}^{\mathrm{TV}}(\mathbf{x}) := \delta_{k-1}^2 V_\Delta \mathbf{e}^T \mathscr{W}_{k-1}\mathbf{e} + V_\Delta \mathbf{x}^H \mathbf{L}_w \mathbf{x}, \tag{4.19}$$

where $\mathbf{L}_w$ is a three-dimensional weighted approximate Laplacian given by

$$\mathbf{L}_w = \mathbf{D}_x^T \mathscr{W}_{k-1}\mathbf{D}_x + \mathbf{D}_y^T \mathscr{W}_{k-1}\mathbf{D}_y + \mathbf{D}_z^T \mathscr{W}_{k-1}\mathbf{D}_z. \tag{4.20}$$

In practice, one can use forward finite differences or backward finite differences to implement the total variation functional. Another option is to use mixed finite differences, which combines these two.

4.3.1. MIXED FINITE DIFFERENCE APPROACH

As mentioned before, the partial derivatives in Eq. (4.16) are approximated by forward or backward two-point finite difference formulas and the differentiation matrices $\mathbf{D}_\xi$ in

the above formulas are either all forward differentiation matrices $\mathbf{D}_{\xi;\mathrm{f}}$ or backward differentiation matrices $\mathbf{D}_{\xi;\mathrm{b}}$. Another option is to mix the forward and backward differencing operators. Specifically, introducing the forward and backward x-coordinate gradient vectors as

$$\mathbf{g}_{x;\mathrm{f}} = \mathbf{D}_{x;\mathrm{f}}\mathbf{x} \quad \text{and} \quad \mathbf{g}_{x;\mathrm{b}} = \mathbf{D}_{x;\mathrm{b}}\mathbf{x} \tag{4.21}$$

with similar definitions for the forward and backward y- and z-coordinate gradient vectors, we approximate $\mathscr{F}^{\mathrm{TV}}_{k-1}(X)$ as

$$\mathscr{F}^{\mathrm{TV}}_{k-1}(X) \approx \delta^2_{k-1} V_\Delta \, \mathbf{e}^T \mathscr{W}_{k-1}\mathbf{e} + \frac{V_\Delta}{2} \sum_{i=x,y,z} \mathbf{g}^H_{i;\mathrm{b}} \mathscr{W}_{k-1}\mathbf{g}_{i;\mathrm{b}} + \mathbf{g}^H_{i;\mathrm{f}} \mathscr{W}_{k-1}\mathbf{g}_{i;\mathrm{f}} \tag{4.22}$$

which is equivalent to Eq. (4.19), but with the mixed finite difference Laplacian

$$\begin{aligned} \mathbf{L}_w = &\frac{1}{2}\left[\mathbf{D}^T_{x;\mathrm{b}}\mathscr{W}_{k-1}\mathbf{D}_{x;\mathrm{b}} + \mathbf{D}^T_{x;\mathrm{f}}\mathscr{W}_{k-1}\mathbf{D}_{x;\mathrm{f}}\right] \\ &+ \frac{1}{2}\left[\mathbf{D}^T_{y;\mathrm{b}}\mathscr{W}_{k-1}\mathbf{D}_{y;\mathrm{b}} + \mathbf{D}^T_{y;\mathrm{f}}\mathscr{W}_{k-1}\mathbf{D}_{y;\mathrm{f}}\right] \\ &+ \frac{1}{2}\left[\mathbf{D}^T_{z;\mathrm{b}}\mathscr{W}_{k-1}\mathbf{D}_{z;\mathrm{b}} + \mathbf{D}^T_{z;\mathrm{f}}\mathscr{W}_{k-1}\mathbf{D}_{z;\mathrm{f}}\right]. \end{aligned} \tag{4.23}$$

Extensive numerical testing has shown that the mixed finite-difference approach leads to faster convergence than implementations that use forward or backward difference operators only. Therefore, we use this mixed finite-difference approach in our implementation of multiplicative regularization.

4.4. MR IMAGING USING MULTIPLICATIVE REGULARIZATION

As mentioned above, at the kth step of the algorithm we assume that we have an approximation $\mathbf{x}_{k-1}$ available. We update the image according to the update formula

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \beta_k \mathbf{d}_k, \tag{4.24}$$

where $\mathbf{d}_k$ is the Polak-Ribière update direction given by [113]

$$\mathbf{d}_k = \mathbf{g}_k + \frac{\mathrm{Re}[\mathbf{g}^H_k(\mathbf{g}_k - \mathbf{g}_{k-1})]}{\|\mathbf{g}_{k-1}\|^2_2}\mathbf{d}_{k-1} \tag{4.25}$$

with $\mathbf{d}_0 = \mathbf{0}$ and $\mathbf{g}_k$ the gradient of the objective function

$$F_{k-1}(\mathbf{x}) = F^{\mathrm{data}}(\mathbf{x})F^{\mathrm{TV}}_{k-1}(\mathbf{x}) \tag{4.26}$$

evaluated at $\mathbf{x} = \mathbf{x}_{k-1}$. Using the product rule, this gradient is given by

$$\mathbf{g}_k = \mathbf{g}^{\mathrm{data}}_k F^{\mathrm{TV}}_{k-1}(\mathbf{x}_{k-1}) + F^{\mathrm{data}}(\mathbf{x}_{k-1})\mathbf{g}^{\mathrm{TV}}_k = \mathbf{g}^{\mathrm{data}}_k + F^{\mathrm{data}}(\mathbf{x}_{k-1})\mathbf{g}^{\mathrm{TV}}_k, \tag{4.27}$$

where we have used $F^{\mathrm{TV}}_{k-1}(\mathbf{x}_{k-1}) = 1$ and $\mathbf{g}^{\mathrm{data}}_k$ is the gradient of F^{data} at $\mathbf{x} = \mathbf{x}_{k-1}$ given by

$$\mathbf{g}^{\mathrm{data}}_k = -2\|\mathbf{b}\|^{-2}_2 \mathbf{A}^H \mathbf{r}_{k-1} \tag{4.28}$$

with $\mathbf{r}_{k-1} = \mathbf{b} - \mathbf{A}\mathbf{x}_{k-1}$ the data residual. Furthermore, $\mathbf{g}_k^{\mathrm{TV}}$ is the gradient of F_{k-1}^{TV} at $\mathbf{x} = \mathbf{x}_{k-1}$ and using Eq. (4.19) is easily obtained as

$$\mathbf{g}_k^{\mathrm{TV}} = 2V_\Delta \mathbf{L}_w \mathbf{x}_{k-1}. \tag{4.29}$$

The gradient of F_{k-1} at $\mathbf{x} = \mathbf{x}_{k-1}$ now follows as

$$\mathbf{g}_k = 2\left[-\|\mathbf{b}\|_2^{-2}\mathbf{A}^H\mathbf{r}_{k-1} + V_\Delta F^{\mathrm{data}}(\mathbf{x}_{k-1})\mathbf{L}_w\mathbf{x}_{k-1}\right]. \tag{4.30}$$

Note that for position independent background and linear gradient fields we have $\mathbf{A}^H = \mathscr{F}^{-1}$, i.e. $\mathbf{A}^H$ is an inverse Fourier Transform.

Finally, to obtain the update coefficient β_k, we substitute $\mathbf{x} = \mathbf{x}_{k-1} + \beta\mathbf{d}_k$ in the objective function $F_{k-1}(\mathbf{x}) = F^{\mathrm{data}}(\mathbf{x})F_{k-1}^{\mathrm{TV}}(\mathbf{x})$ and determine the update coefficient by solving the equation

$$\left.\frac{\partial F_{k-1}(\mathbf{x}_{k-1} + \beta\mathbf{d}_k)}{\partial\beta}\right|_{\beta=\beta_k} = 0. \tag{4.31}$$

Explicitly, $F_{k-1}(\mathbf{x}_{k-1} + \beta\mathbf{d}_k) = F^{\mathrm{data}}(\mathbf{x}_{k-1} + \beta\mathbf{d}_k)F_{k-1}^{\mathrm{TV}}(\mathbf{x}_{k-1} + \beta\mathbf{d}_k)$, where

$$F^{\mathrm{data}}(\mathbf{x}_{k-1} + \beta\mathbf{d}_k) = a_0 + a_1\beta + a_2\beta^2 \quad \text{and} \quad F_{k-1}^{\mathrm{TV}}(\mathbf{x}_{k-1} + \beta\mathbf{d}_k) = 1 + b_1\beta + b_2\beta^2 \tag{4.32}$$

with

$$a_0 = F^{\mathrm{data}}(\mathbf{x}_{k-1}), \quad a_1 = -2\frac{\mathrm{Re}\,(\mathbf{r}_{k-1}^H\mathbf{A}\mathbf{d}_k)}{\|\mathbf{b}\|_2^2}, \quad \text{and} \quad a_2 = \frac{\|\mathbf{A}\mathbf{d}_k\|_2^2}{\|\mathbf{b}\|_2^2} \tag{4.33}$$

and

$$b_1 = 2V_\Delta\,\mathrm{Re}\,(\mathbf{x}_{k-1}^H\mathbf{L}_w\mathbf{d}_k) \quad \text{and} \quad b_2 = V_\Delta\,\mathbf{d}_k^H\mathbf{L}_w\mathbf{d}_k. \tag{4.34}$$

Note that a_2 and b_2 are always positive. With these results, the update coefficient follows from Eq. (4.31) as the root for which the polynomial $a_0b_1 + a_1 + 2(a_0b_2 + a_1b_1 + a_2)\beta + 3(a_1b_2 + a_2b_1)\beta^2 + 4a_2b_2\beta^3$ is minimized. The roots can be found analytically, or using a built-in polynomial root-finding algorithm. To summarize, the overall algorithm is as follows:

LOW-FIELD MULTIPLICATIVELY REGULARIZED MR IMAGING

1. Given an initial guess of the low-field MR image $\mathbf{x}_0$
2. For $k = 1, 2, \ldots$,
   (a) Compute the gradient vector $\mathbf{g}_k$ as given by Eq. (4.30);
   (b) Compute the Polak-Ribière update direction $\mathbf{d}_k$;
   (c) Compute the update coefficient β_k;
   (d) Update the low-field MR image: $\mathbf{x}_k = \mathbf{x}_{k-1} + \beta_k\mathbf{d}_k$.

Remark: Consider the case of position independent background and linear gradient fields. We then have $\mathbf{A}^H = \mathscr{F}^{-1}$ and the gradient vector becomes

$$\mathbf{g}_k(\mathbf{x}_{k-1}) = 2\left[-\|\mathbf{b}\|_2^{-2}\mathscr{F}^{-1}\mathbf{r}_{k-1} + V_\Delta F^{\text{data}}(\mathbf{x}_{k-1})\mathbf{L}_w\mathbf{x}_{k-1}\right]. \tag{4.35}$$

As an initial guess, let us take $\mathbf{x}_0 = \mathscr{F}^{-1}\mathbf{b}$. This initial guess is very noisy, but the data is exactly matched and $F^{\text{data}}(\mathbf{x}_0) = 0$. Consequently, $\mathbf{g}(\mathbf{x}_0) = \mathbf{0}$ and the algorithm stops after one iteration with $\mathbf{x}_1 = \mathbf{x}_0$. In this case, where we have a homogeneous background field and perfectly linear gradient fields, which is often assumed in practice, we use our algorithm as a denoising algorithm by setting the gradient at the kth iteration to $\mathbf{g}(\mathbf{x}_{k-1}) = \mathbf{L}_w\mathbf{x}_{k-1}$. We now update in directions determined by the gradient of the TV-term only. Starting with a masked initial guess $\mathbf{x}_0 = \mathbf{M}\mathscr{F}^{-1}\mathbf{b}$, where the mask $\mathbf{M}$ zeroes out any components of $\mathscr{F}^{-1}\mathbf{b}$ for which it is a priori known that the image at the corresponding location vanishes, $F^{\text{data}}(\mathbf{x}_k)$ will generally increase as k increases and if $\lim_{k\to\infty} F^{\text{TV}}_{k-1}(\mathbf{x}_k) = 1$ then the objective function $F_{k-1}(\mathbf{x})$ will asymptote to $F^{\text{data}}(\mathbf{x}^*)$ as $k \to \infty$, where $\mathbf{x}^* = \lim_{k\to\infty}\mathbf{x}_{k-1} = \lim_{k\to\infty}\mathbf{x}_k$ is the converged image in which noise variations have been minimized (suppressed).

4.5. RESULTS

In this section we illustrate the performance of our multiplicatively regularized imaging and denoising algorithm. In Section 4.5.1 we use the algorithm to reconstruct the well-known 2D Shepp-Logan phantom from simulated and noise-corrupted low-field data with a frequency encoding gradient that is nonlinear. The data matrix $\mathbf{A}$ is not equal to a standard DFT matrix in this case and simply applying an inverse DFT to the data would lead to highly distorted reconstructions. We therefore use the algorithm as an iterative reconstruction algorithm to recover the Shepp-Logan phantom from the data.

Subsequently, in Section 4.5.2, we apply our algorithm to data acquired using the low-field scanner described in [20]. This scanner has been constructed such that the background and gradient fields are approximately constant and linear throughout the imaging domain. We take this into account in our data model and the data matrix $\mathbf{A}$ is assumed to be equal to a unitary three-dimensional DFT matrix in this case. The algorithm is now used as a denoising algorithm and images of an apple, a melon and a brain will be presented to demonstrate the performance of the method on measured data for which a Fourier signal representation is assumed to be applicable. In Section 4.5.3, we compare our multiplicatively regularized 2D and 3D results with images obtained by solving the additively regularized minimization problem

$$\mathbf{x}^{\text{add}} = \arg\min\left\{\|\mathbf{b}-\mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{T}\mathbf{x}\|_1\right\}, \tag{4.36}$$

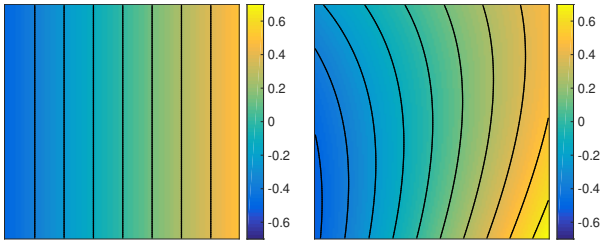
where the second term is the total variation term, with $\mathbf{T}$ the anisotropic total variation operator as defined in Eqs. (2.49) and (3.10) for 2D and 3D, respectively. We solve this minimization problem using the Alternating Direction Method of Multipliers (ADMM), see for example [105].

The experiments were carried out in MATLAB R2015a on a desktop PC with an Intel(R) Xeon(R) W-2123 CPU (3.60GHz).

4.5.1. Two-Dimensional Imaging of Simulated Noise-Corrupted Low-Field MR Data

In low-field MRI, inhomogeneities may be present in the background field and gradients may not be perfectly linear. To investigate the performance of the algorithm as an image reconstruction algorithm in such cases, we consider a two-dimensional low-field reconstruction problem in which field perturbations are taken into account. Many different field perturbations can be considered, of course, but in all cases signals are obtained for which the relationship between the signal and the object is no longer governed by a DFT. To fix the idea, we therefore focus on signal generation in an MR scanner in which the frequency encoding gradient is not perfectly linear. Specifically, we introduce a perturbation of the gradient field that is relative to the gradient strength. To this end, we scale the frequency encoding gradient such that we can describe its strength to be in the range $[-.5,.5)$ and we use $\tilde{x}$ and $\tilde{y}$ to denote the location in the FoV such that $\tilde{x} = -.5$ corresponds to the left and $\tilde{x} = .5$ to the right boundary of the FoV, and $\tilde{y} = -.5$ and $\tilde{y} = .5$ to the lower and upper boundary, respectively. The gradient field is now perturbed by the function $.5\tilde{x}\tilde{y} + .05\tilde{x}^2 + .35\tilde{y}^2$. The unperturbed gradient field is shown in Figure 4.1a, while the perturbed gradient field is shown in Figure 4.1b.

Since the background and (nonlinear) gradient profiles are known, we can use the discretized version of Eq. (4.5) to obtain the model matrix $\mathbf{A}$. As an image or model solution, we take the Shepp-Logan phantom of Figure 4.2a and we use matrix $\mathbf{A}$ to generate the data. Subsequently, white Gaussian noise with an SNR of 20 is added to the data in the spatial frequency domain. We note that not taking the distorted gradient into account leads to the image shown in Figure 4.2b, which is obviously a deformed image of the true object.



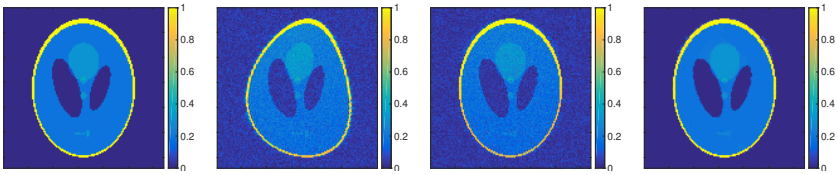
(a) Ideal gradient (b) Perturbed gradient

Figure 4.1: In an ideal scenario, the frequency encoding gradient is linear (a). In these simulations, however, the gradient is perturbed and nonlinear (b). The black lines are contour lines.

Having the noise-corrupted data available, we apply our multiplicative regularization scheme to this data in an attempt to reconstruct the Shepp-Logan model solution. Mixed finite differences and the 2D version of the Laplacian matrix given in Eq. (4.23) are used to implement the total variation functional. As an initial guess, we use $\mathbf{x}_0 = \alpha \mathbf{A}^H \mathbf{b}$, where α is chosen such that $F^{\text{data}}(\mathbf{x}_0)$ is minimized. This initial guess is depicted in Figure 4.2c. Subsequently, we use this $\mathbf{x}_0$ to define $F_0^{\text{TV}}(\mathbf{x})$, and start iterating. The reconstruction that we obtain after 50 iterations is shown in Figure 4.2d. The algorithm's

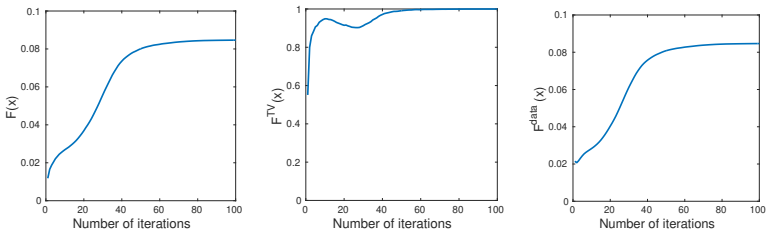
progress is shown in Appendix E, where we can see that the algorithm manages to progressively denoise the image, while maintaining the edges and structures in the phantom.

Figure 4.3 shows the value of the objective function $F(\mathbf{x})$ as a function of the iteration number, for the two-dimensional Shepp-Logan image. The values of $F^{\text{TV}}(\mathbf{x})$ and $F^{\text{data}}(\mathbf{x})$ are plotted as well. We note that, as expected, F^{TV} converges to 1, while the other two steadily grow to a larger value that enables this convergence. The increasing value of $F^{\text{data}}(\mathbf{x})$ can be explained by the observation that our matrix $\mathbf{A}$ does not deviate all that much from a Fourier Transform, which means that the term $\|\mathbf{b} - \mathbf{A}\mathbf{x}_0\|^2$ is close to 0. However, since $\mathbf{x}_0$ is too noisy, this solution does not meet the smoothness requirements of the total variation functional. Instead we iterate towards a solution that minimizes the least-squares term under the constraint that $F^{\text{TV}}(\mathbf{x})$ is equal to 1. One iteration takes 1.79 seconds, which is relatively long but this can be explained by the fact that we need to explicitly calculate the matrix $\mathbf{A}$ and its transpose, since we cannot rely on the Fast Fourier Transform (FFT) due to the nonlinear gradient.



(a) Model solution (b) IFFT (c) Initial guess (d) Algorithm result

Figure 4.2: Reconstruction results.

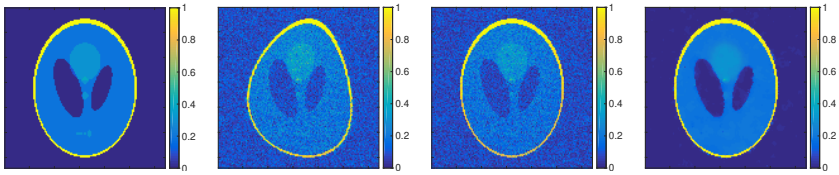


(a) $F(\mathbf{x})$ (b) $F^{TV}(\mathbf{x})$ (c) $F^{\text{data}}(\mathbf{x})$

Figure 4.3: Plots of the objective function value $F(x)$ (left), which is equal to the product of F^{TV} (middle) and F^{data} (right), for the two-dimensional Shepp-Logan image.

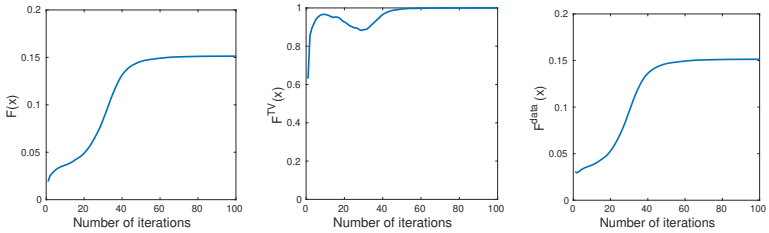
Additionally, we consider the same Shepp-Logan phantom, but with a decreased SNR of 5. This is more realistic in a low-field MRI setting. The inverse FFT reconstruction, initial guess and the multiplicatively regularized reconstruction are shown in Figs. 4.4b, 4.4c and 4.4d, respectively, and for different iteration numbers, the reconstructions are shown in Appendix E. This result was obtained after 50 iterations. We see that the algorithm manages to denoise the image, but some structures are lost. This is not due to the

method itself, but because the SNR is simply too low. As will be shown later, an additively regularized method is not able to recover these structures either. The convergence plots are shown in Figure 4.5, which showcase the same behavior as in the higher SNR case.



(a) Model solution (b) IFFT (c) Initial guess (d) Algorithm result

Figure 4.4: Reconstruction results for an SNR of 5.



(a) $F(\mathbf{x})$ (b) $F^{TV}(\mathbf{x})$ (c) $F^{\text{data}}(\mathbf{x})$

Figure 4.5: Plots of the objective function value $F(x)$ (left), which is equal to the product of F^{TV} (middle) and F^{data} (right), for the two-dimensional Shepp-Logan image. In this case, the SNR was lowered to 5.

4.5.2. THREE-DIMENSIONAL IMAGING OF MEASURED DATA

This section contains reconstructed images of an apple, a melon and a brain that were scanned using the low-field scanner of [20].

Table 4.1: Parameter settings for the different imaging experiments

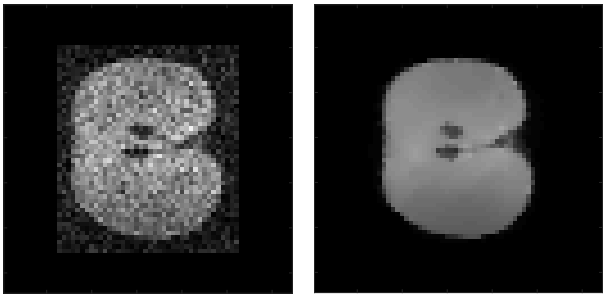
	Apple	**Melon**	**Brain**
Repetition time (TR)	3000 ms	2000 ms	500 ms
Echo time (TE)	30 ms	30 ms	20 ms
Number of voxels	$64 \times 64 \times 64$	$128 \times 128 \times 128$	$128 \times 128 \times 50$
Imaging domain/FoV	$128 \times 128 \times 128$ mm^3	$200 \times 200 \times 200$ mm^3	$224 \times 224 \times 175$ mm^3
Pulse duration	100 μs	100 μs	100 μs
Acquisition bandwidth	10 kHz	20 kHz	20 kHz
Echo train length	Not applicable	Not applicable	4

APPLE EXPERIMENT

In our first experiment, an apple is imaged using a spin-echo sequence whose parameters are given in the second column of Table 4.1. As an initial guess, we take a masked

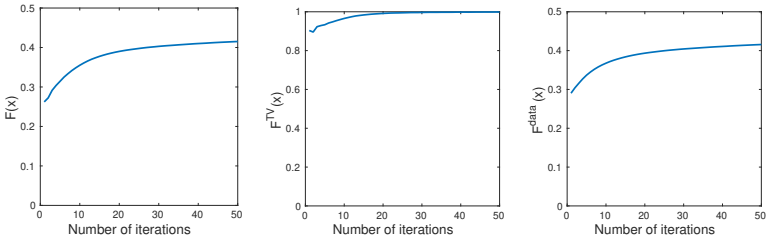
version of the three-dimensional inverse DFT of the k-space obtained during the apple experiment. The 35th slice of this initial guess is shown in Figure 4.6a. Clearly, the initial guess is contaminated by noise and the mask that is used is visible. We use this initial guess to determine $F_0^{\mathrm{TV}}(\mathbf{x})$, after which we start the iterative process.

To remove the noise from the initial image, we use the algorithm as a denoising algorithm and use a mixed difference approach with the Laplacian matrix of Eq. (4.23) in the discretized total variation functional. One iteration takes approximately 0.3 seconds. This is very fast compared to the Shepp-Logan problem, for which the problem is smaller (64×64 pixels instead of $64 \times 64 \times 64$ pixels), which can be explained by the fact that here, the FFT can be employed. We observe that the amount of noise in the image decreases as the iteration process continues, as can be seen in Figure E.3 in the Appendix. After 30 iterations an image of good quality is obtained, with the apple's shape and seeds being clearly visible and the noise having been eliminated. Figure 4.7 shows the convergence plots for our algorithm. Again, we see that F^{TV} converges to 1.



(a) Initial guess (b) Algorithm result

Figure 4.6: Reconstruction result of the 35th slice of the apple.



(a) $F(\mathbf{x})$ (b) $F^{TV}(\mathbf{x})$ (c) $F^{\mathrm{data}}(\mathbf{x})$

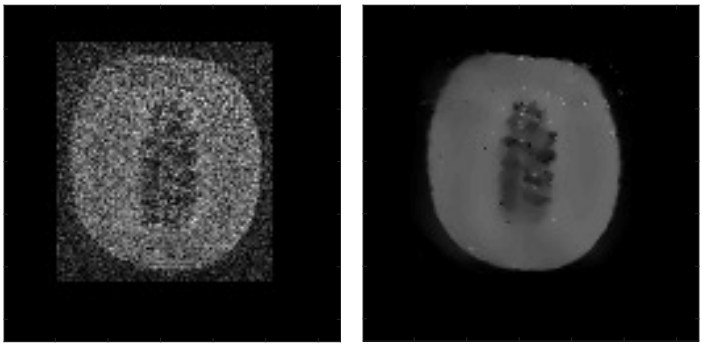
Figure 4.7: Plots of the objective function value $F(x)$ (left), which is equal to the product of F^{TV} (middle) and F^{data} (right), for the three-dimensional apple image.

Iterating further leads to a result that is somewhat oversmoothed, which is similar to the results obtained in the simulated 2D setting with a low SNR. Therefore, it is advisable to stop the iterative process early (semi-convergence). Finally, we note that the

reconstructed images are obtained essentially in real-time and the reconstructions can be monitored as the iterative scheme progresses, which is of great importance in practice when a scanner is used for diagnostic purposes.

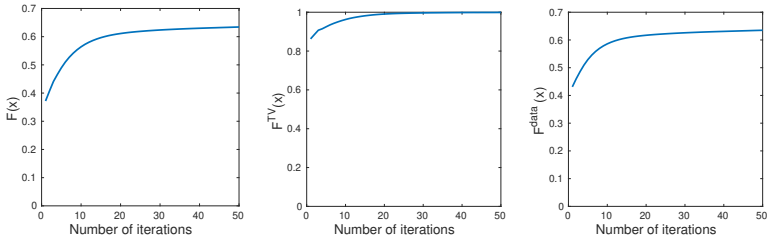
MELON EXPERIMENT

In our second experiment, a melon is imaged using a spin-echo sequence with the parameter settings given in the third column of Table 4.1. Here, too, we construct an initial guess by masking the image that is obtained by applying an inverse three-dimensional DFT to the data. Note that this data set has twice as many data points in each Cartesian direction as the data set in the apple experiment. The 64th slice of the initial guess is shown in Figure 4.8a. The mask that is used is clearly visible and again a very noisy initial image is obtained. To remove this noise, we use mixed finite differences to implement the total variation functional and the resulting reconstruction of the melon is shown in Figure 4.8b.



(a) Initial guess (b) Algorithm result

Figure 4.8: Reconstruction result of the 64th slice of the melon.



(a) $F(\mathbf{x})$ (b) $F^{TV}(\mathbf{x})$ (c) $F^{\text{data}}(\mathbf{x})$

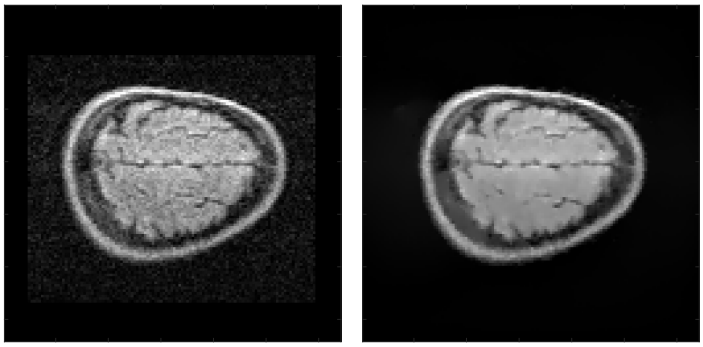
Figure 4.9: Plots of the objective function value $F(x)$ (left), which is equal to the product of F^{TV} (middle) and F^{data} (right), for the three-dimensional melon image.

As can be seen in Figure E.4 in Appendix E, the noise level in the images decreases with the iteration number. We note that some individual pixels appear overly bright or

dark, which may be removed by iterating further but this leads to oversmoothing, since the SNR is very low. We therefore terminate the iterative process after 20 iterations. Finally, even though the complex data set of the melon is eight times larger than the apple data set, the method still produces images essentially in real time that can be monitored as the scheme progresses. Each iteration takes approximately 2 seconds. In Figure 4.9, the convergence plots are shown, which follow the same pattern as before.

IN VIVO BRAIN EXPERIMENT

In the final experiment, an in vivo 3D data set was acquired from the brain of a healthy volunteer, using a turbo spin echo sequence, with the parameters as given in Table 4.1. To speed up acquisition time, a cylindrical k-space coverage pattern was used. The left column of Figure 4.10 shows two slices of the initial guess, which is a masked version of the image obtained after having applied an inverse DFT to the k-space data. We terminate the algorithm after 40 iterations.



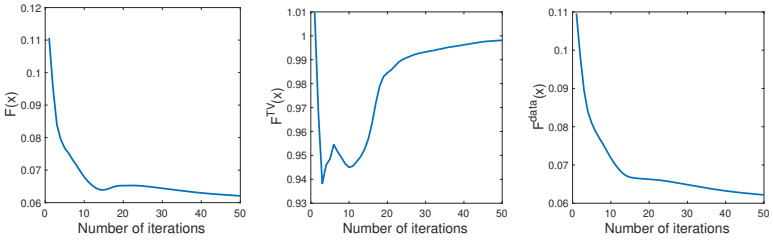
(a) Initial guess (slice 24) (b) Algorithm result (slice 24)

(c) Initial guess (slice 34) (d) Algorithm result (slice 34)

Figure 4.10: Two different slices of the reconstructed brain image.

Two slices of the final image are shown in the right column of Figure 4.10. Figure E.5 in Appendix E shows the algorithm's output for different iteration numbers, again

demonstrating that the longer the algorithm is allowed to iterate, the more it denoises the image.



(a) $F(\mathbf{x})$ (b) $F^{TV}(\mathbf{x})$ (c) $F^{\text{data}}(\mathbf{x})$

Figure 4.11: Plots of the objective function value $F(x)$ (left), which is equal to the product of F^{TV} (middle) and F^{data} (right), for the three-dimensional brain image.

4.5.3. ADDITIVE AND MULTIPLICATIVE REGULARIZATION

Finally, in Figure 4.12, our multiplicatively regularized reconstructions are shown next to additively regularized reconstructions. We tuned the regularization parameter λ in a heuristic manner such that the background noise disappears, while the different structures are still visible. For all experiments, we used 10 ADMM iterations and within each ADMM iteration, we used 10 iterations of the Conjugate Gradient (CG) algorithm to solve the first minimization problem. We observe that, in terms of image quality, additive and multiplicative regularization yield comparable results, but no parameter tuning is required in the multiplicative case. This is reflected in the peak signal-to-noise ratios (PSNR), shown in Table 4.2, which we were only able to calculate for the simulated case because no ground truth was available for the measured data. We see that multiplicative regularization achieves a somewhat higher PSNR for low SNR and a marginally lower PSNR for high SNR. Furthermore, in Table 4.3, the computing times for the 2 different algorithms (our algorithm and the ADMM algorithm solving the additively regularized problem) are shown. We note that it might be possible to lower the computation time for the additively regularized problem by decreasing the number of either ADMM or CG iterations. This table shows that, for our test problems, the computation times for multiplicative regularization are quite competitive compared to additive regularization. Additionally, the computation time shown for additive regularization does not include the time spent searching for an appropriate regularization parameter (requiring repeated execution of the algorithm in general), which is a problem that is completely eliminated when using a multiplicatively regularized approach.

Table 4.2: PSNR for different reconstructions of the Shepp-Logan phantom.

	Multiplicative regularization	Additive regularization
SNR 20	37.30	**37.34**
SNR 5	**33.83**	30.41

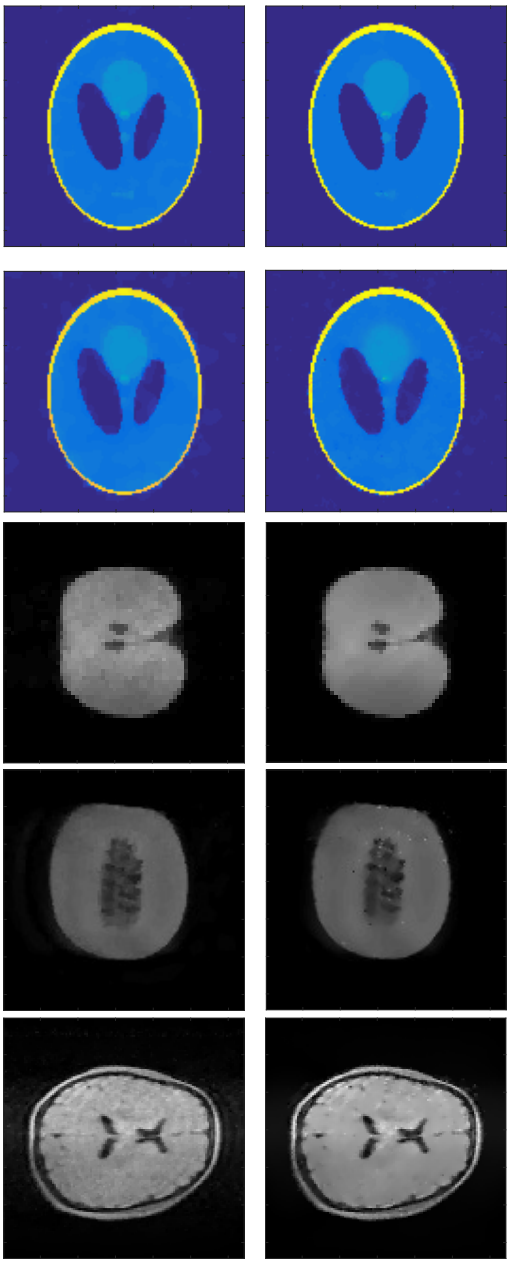


Figure 4.12: Images of the Shepp-Logan phantom using simulated data with an SNR of 20 (top row) and an SNR of 5 (second row), an apple (third row), a melon (fourth row) and a brain (bottom row) based on measured data, using additive regularization (left column) and multiplicative regularization (right column).

Table 4.3: Computing time (and total number of iterations) for the different reconstructions in seconds.

	Multiplicative regularization	Additive regularization
Shepp-Logan (SNR 20)	94.19 (50 iterations)	194.81 (100 iterations)
Shepp-Logan (SNR 5)	89.87 (50 iterations)	193.29 (100 iterations)
Apple	5.42 (30 iterations)	6.3 (100 iterations)
Melon	36.50 (20 iterations)	81.20 (100 iterations)
Brain	24.49 (40 iterations)	28.34 (100 iterations)

4.6. CONCLUSION AND DISCUSSION

In this chapter we applied a multiplicative regularization approach to low-field MR imaging. By multiplying a least-squares data fidelity function by a regularizing total variation function that is differentiable, we avoid the problem of having to determine and compute a regularization parameter as is required for additive regularization. The resulting multiplicative regularization problem is nonlinear and is solved by using a nonlinear conjugate gradient scheme with Polak-Ribière update directions. Furthermore, we showed that the algorithm can be used as an image reconstruction and denoising algorithm by applying the method to two-dimensional simulated noise-corrupted MR data obtained with a nonlinear gradient field and to three-dimensional measured data obtained with a low-field Halbach scanner. We demonstrated that multiplicative regularization yields very promising results, converging within a few iterations, whether we are dealing with two-dimensional noisy data for which a Fourier signal representation is no longer valid, or with three-dimensional measured data for which a Fourier transform relationship between signal and image can be assumed. Moreover, accurate reconstructions are obtained essentially in real time in case a Fourier signal model is applicable. We observed that in case of a low SNR, it is better to stop the iterative process somewhat early (semi-convergence), to maintain the edges in the image. For a high SNR, edges are preserved while noise is eliminated, even for a large number of iterations.

In this work, we focused on a low-field MRI setting with one single receiver coil. However, multiplicative regularization may also be applied to data acquired in a high-field MRI scanner equipped with several coils, of course, and we intend to test its performance against more standard high-field reconstruction techniques used today.

Moreover, in future work we will also focus on incorporating measured background and gradient fields of practical low-field MR scanners into the data model that is used in our multiplicative regularization scheme. Obviously, this is particularly important in case the background and gradient fields are such that a standard Fourier transform signal representation is no longer valid, since otherwise significant distortions in the image are obtained as we demonstrated for two-dimensional simulated MR data. Standard FFTs can no longer be used in this case and accurate images may only be obtained by solving the image reconstruction problem as an optimization problem. In our multiplicative scheme, forming matrix-vector products with the model matrix $\mathbf{A}$ is then the main computational bottleneck, since computing the action of the Laplacian of the total variation functional on a vector involves sparse finite difference operators only. Consequently, to reduce the reconstruction time of the method, efficient routines to compute matrix-vector products with the data matrix $\mathbf{A}$ have to be developed, possibly involv-

ing nonuniform FFTs. Finally, we intend to include compressed sensing techniques into our multiplicative regularization framework as well, since this may lead to reduced scan times and a reduction of motion artifacts, for example.

5

INVERSION OF INCOMPLETE K-SPACE DATA USING SUPPORT INFORMATION

5.1. INTRODUCTION

In many application areas, ranging from geophysics to MRI, one is confronted with the problem of reconstructing an object, a function, or an image from incomplete Fourier spectral data (see, e.g. [122–124]). This is an ill-posed problem in general and very difficult or impossible to solve without any additional information (support or sparsity information, for example). However, by taking *a priori* information about the object into account, it may be possible to successfully reconstruct the object of interest based on incomplete Fourier data. In compressed sensing (CS), for example, we take into account that the object or image has a sparse representation in some basis and accurate reconstructions are possible provided that the undersampling artefacts are incoherent [125]. As an illustration, in MRI the prototype CS problem consists of minimizing the objective function

$$F_{\mathrm{CS}}(\mathbf{x}) = \|\mathbf{S}_k(\mathbf{d} - \mathscr{F}\mathbf{x})\|_2^2 + \lambda\|\Psi\mathbf{x}\|_1, \tag{5.1}$$

where Ψ is a sparsifying transform, $\mathscr{F}$ the (unitary) discrete Fourier transform (DFT) matrix, $\mathbf{d}$ the data vector, and $\mathbf{S}_k$ a diagonal matrix with ones and zeros on the diagonal

This chapter is based on the article:

M.L. de Leeuw den Bouter, P.M. van den Berg, and R.F. Remis, *Inversion of incomplete spectral data using support information with an application to magnetic resonance imaging*, Journal of Physics Communications, **5**, 055006 (2021).

representing incoherent k-space measurements, where a diagonal entry equal to one corresponds to a point in k-space for which data is available. Often, an additional total variation functional is added to the above objective function as well and the bases that are used for the sparsifying transform Ψ are typically global bases (wavelets, noiselets, etc.) defined over the complete field of view (FoV). With CS, no information about the support of the object is required to successfully image the object of interest. However, optimization algorithms that minimize objective functions that consist of a least-squares objective function characterizing the data fidelity and an ℓ_1 objective function that enforces sparsity are generally more complex than algorithms that minimize only a least-squares objective function. In addition, in CS a regularization parameter needs to be determined to balance the two terms in the total objective function.

In this chapter, we consider the problem of reconstructing a bounded object from incomplete spectral data in case the support of the object is known. Our motivation comes from MRI where, at least in principle, the support of the object or body part of interest can be determined before an actual clinical scan takes place. During the so-called prescan, for example, an image is produced based on not-fully-sampled k-space data. From this image, the FoV is determined and this scan can also be used to estimate the support of the object and its location within the FoV. In general, this approximate support is not exact, of course, but for single body parts (a head, an arm, or a leg, for example) the support can be determined in a fairly straightforward manner using binary maps. Determining approximate supports in case multiple (possibly disjoint) objects or body parts are present within the FoV may be more challenging, but the overall idea remains the same. Here, we essentially follow this approach and determine the approximate support of an object (an apple) based on incomplete data that was obtained with a low-field MR scanner [20, 37, 126]. We show that image quality improves when this (approximate) support information is included in the reconstruction algorithm. Further examples in which different undersampling patterns are used are presented as well to demonstrate that directly including support information in the reconstruction algorithm generally improves image quality.

Having support information available, our objective is to reconstruct the image within the FoV from incomplete spectral data. In this case, we can still use CS techniques, of course. In [127], for example, a CS technique is introduced that takes partially known support information into account, while in [128] and [129] the CS problem is formulated as an optimization problem that aims at finding the solution to the data equation which has the smallest number of components outside of the support.

In this work, however, we propose a much simpler reconstruction algorithm in case support information is known. Specifically, we build upon the work presented in [122] and propose an image reconstruction algorithm, which is essentially the conjugate-gradient least-squares (CGLS) algorithm applied to the normal equation that corresponds to a space- and frequency-restricted Fourier transform equation. In [122], this approach was proposed for one-dimensional space- and band-limited Fourier transform problems, while here we generalize this approach to two- and three-dimensional imaging problems. In particular, we consider two- and three-dimensional (random) undersampling patterns in the spectral domain that are typically used in MRI to speed up data acquisition and study the performance of the method. We stress that for the two- and

three-dimensional spatial support functions and spectral domain undersampling patterns considered here, we have to resort to numerical methods, since analytical results concerning the eigenfunctions and singular functions of the corresponding truncated Fourier operators are not available (as in Slepian-Pollak theory [124, 130]). Finally, we also compare our reconstructions to reconstructions obtained using CS techniques and possible extensions of the method (parallel imaging and total variation regularization) are briefly discussed as well.

5.2. BASIC EQUATIONS

The starting point of our analysis is the discrete Fourier transform equation

$$\mathbf{d} = \mathscr{F}\mathbf{x}, \tag{5.2}$$

where $\mathbf{x}$ is a discrete (vectorized) image function defined on the FoV, $\mathscr{F}$ is the unitary discrete Fourier transform (DFT) in one, two, or three dimensions, and $\mathbf{d}$ is a vector containing Fourier transform data. As we have seen before, Eq. (5.2) follows from discretizing Eq. (4.3)).

We assume that the support of an object located within the FoV is known and introduce its indicator or support matrix as a diagonal matrix $\mathbf{S}_x$, where a diagonal element is equal to the indicator function of the object applied to the pixel (2D) or voxel (3D) that corresponds to this diagonal element. The support matrix $\mathbf{S}_x$ is obviously idempotent, that is, it satisfies $\mathbf{S}_x^2 = \mathbf{S}_x$ and the image vector $\mathbf{x}$ satisfies $\mathbf{x} = \mathbf{S}_x\mathbf{x}$. Furthermore, when spectral data is only available within a subdomain of Fourier space, we introduce a support matrix $\mathbf{S}_k$ for this subdomain in k-space as well, where the diagonal elements of $\mathbf{S}_k$ indicate for which points in k-space spectral data is available. The support matrix $\mathbf{S}_k$ is idempotent and available k-space data is given by the vector $\mathbf{S}_k\mathbf{d}$.

Having introduced the support matrices $\mathbf{S}_x$ and $\mathbf{S}_k$, we can now formulate our reconstruction problem, which consists of retrieving the image vector $\mathbf{x} = \mathbf{S}_x\mathbf{x}$ from the equation

$$\mathbf{S}_k\mathscr{F}\mathbf{S}_x\mathbf{x} = \mathbf{S}_k\mathbf{d} \qquad \text{or} \qquad \mathbf{A}\mathbf{S}_x\mathbf{x} = \mathbf{S}_k\mathbf{d}, \tag{5.3}$$

where we have introduced the space- and band-limited discrete Fourier transform as $\mathbf{A} = \mathbf{S}_k\mathscr{F}\mathbf{S}_x$. As pointed out in e.g. [122], in a continuous setting such a reconstruction problem can be formulated in terms of an integral equation of the second kind for the image function restricted to its support. A similar approach can be followed in the general discrete case considered here. In particular, we start with the full DFT of the image and write

$$\mathscr{F}\mathbf{x} = \mathscr{F}\mathbf{S}_x\mathbf{x} = \mathbf{S}_k\mathscr{F}\mathbf{S}_x\mathbf{x} + (\mathbf{I}-\mathbf{S}_k)\mathscr{F}\mathbf{S}_x\mathbf{x} = \mathbf{S}_k\mathbf{d} + (\mathbf{I}-\mathbf{S}_k)\mathscr{F}\mathbf{S}_x\mathbf{x}. \tag{5.4}$$

Applying the inverse Fourier transform $\mathscr{F}^{-1} = \mathscr{F}^H$ to this equation, we get

$$\mathbf{x} = \mathscr{F}^H\mathbf{S}_k\mathbf{d} + \mathscr{F}^H(\mathbf{I}-\mathbf{S}_k)\mathscr{F}\mathbf{S}_x\mathbf{x} \tag{5.5}$$

and by restricting this equation to the object domain by multiplying the above equation on the left by the support matrix $\mathbf{S}_x$, we arrive at the equation

$$(\mathbf{I}-\mathbf{K})\mathbf{S}_x\mathbf{x} = \mathbf{h}, \tag{5.6}$$

where $\mathbf{h} = \mathbf{S}_x\mathscr{F}^H\mathbf{S}_k\mathbf{d}$ and $\mathbf{K} = \mathbf{S}_x\mathscr{F}^H(\mathbf{I}-\mathbf{S}_k)\mathscr{F}\mathbf{S}_x$. Equation (5.6) is the discrete counterpart of the integral equation presented in [122] for one-dimensional problems. If this equation is solved using the Neumann series starting with a vanishing initial guess one arrives at the Papoulis-Gerchberg algorithm or alternating orthogonal projection method [123].

However, as pointed out in [122] for the one-dimensional continuous case, the one-dimensional continuous counterpart of (5.6) can also be written as a normal equation involving the truncated Fourier transform operator. This approach can be extended to the general case considered here as well. Specifically, given the definition of the space- and band-limited DFT matrix $\mathbf{A}$, it is easily verified that

$$(\mathbf{I}-\mathbf{K})\mathbf{S}_x = \mathbf{A}^H\mathbf{A} \tag{5.7}$$

and since $\mathbf{h} = \mathbf{S}_x\mathscr{F}^H\mathbf{S}_k\mathbf{d} = \mathbf{A}^H\mathbf{d}$, we can write (5.6) as

$$\mathbf{A}^H\mathbf{A}\mathbf{x} = \mathbf{A}^H\mathbf{d}, \tag{5.8}$$

5

showing that solving (5.6) is equivalent to solving the normal equation (5.8). As is well known, any solution that minimizes the least-squares objective function

$$F(\mathbf{x}) = \|\mathbf{S}_k\mathbf{d}-\mathbf{A}\mathbf{x}\|_2^2 \tag{5.9}$$

satisfies the normal equation (5.8) as well.

From Slepian-Pollak theory [130] we know that in the one-dimensional continuous case, the left and right singular functions of the Fourier transform that is band- and space-limited to the intervals $[-k_0,k_0]$ and $[-a_0,a_0]$ in k- and x-space, respectively, have these intervals as their support. Moreover, the left and right singular functions are complete on $L^2[-k_0,k_0]$ and $L^2[-a_0,a_0]$, respectively, and all singular values of the truncated Fourier transform belong to the interval $[0,1]$ with clustering occurring around zero and one. Some of these one-dimensional spectral results have been extended to higher dimensions [124] (in two dimensions to functions with circular support in ordinary and k-space, for example), but no such theory exists for the truncated two- or three-dimensional discrete Fourier transforms considered here with general x- or k-space indicator functions.

Therefore, let us first consider computing the singular value decomposition (SVD) of matrix $\mathbf{A}$ given by $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^H$, where the columns $\mathbf{u}_i$ of $\mathbf{U}$ and the columns $\mathbf{v}_i$ of $\mathbf{V}$ are the left and right singular vectors of matrix $\mathbf{A}$, respectively, and Σ is a diagonal matrix with the nonnegative singular values σ_i of $\mathbf{A}$ on its diagonal arranged in decreasing order. With the SVD of matrix $\mathbf{A}$ at our disposal, let $\mathscr{V}_k = \text{span}\{\mathbf{v}_1,\mathbf{v}_2,...,\mathbf{v}_k\}$ be the space spanned by the first k right singular vectors that correspond to the first k largest singular values $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_k > 0$. We can then take the kth truncated SVD solution

$$\mathbf{x}_k = \text{argmin}_{\mathbf{x}\in\mathscr{V}_k} F(\mathbf{x}) \tag{5.10}$$

as an approximate solution to (5.8). Care must be exercised when selecting k, however, since a poorly chosen k may lead to an overly smooth reconstruction or the approximate solution $\mathbf{x}_k$ is heavily affected by noise that is present in the data.

For reconstruction problems encountered in practice, however, computing the SVD of matrix $\mathbf{A}$ comes at prohibitively high computational costs. Computing matrix-vector products with matrix $\mathbf{A}$, on the other hand, can be carried out at "FFT speed" and we therefore resort to iterative methods that solve the reconstruction problem. Specifically, instead of the approximate solutions of (5.10), we take

$$\mathbf{x}_k = \text{argmin}_{\mathbf{x}\in\mathcal{K}_k} F(\mathbf{x}), \tag{5.11}$$

as approximate solutions, where $\mathcal{K}_k = \text{span}\{\mathbf{A}^H\mathbf{d}, (\mathbf{A}^H\mathbf{A})\mathbf{A}^H\mathbf{d}, ..., (\mathbf{A}^H\mathbf{A})^{k-1}\mathbf{A}^H\mathbf{d}\}$ is the kth Krylov subspace generated by $\mathbf{A}^H\mathbf{A}$ and vector $\mathbf{A}^H\mathbf{d}$. The CGLS algorithm [55] that starts with a vanishing initial guess produces approximations $\mathbf{x}_k$ that satisfy (5.11) and we therefore use this algorithm to solve the reconstruction problem. The reason for taking the $\mathbf{x}_k$ of (5.11) as approximate solutions comes from the observation that the first Krylov subspace vector (right-hand side vector)

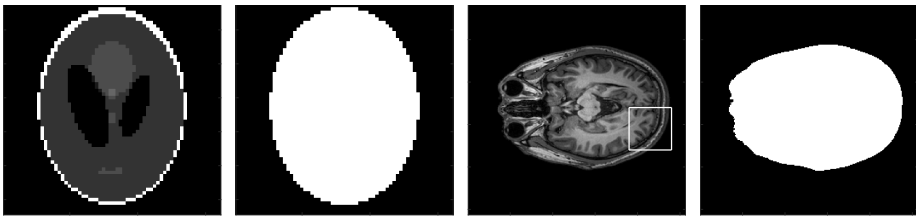
$$\mathbf{A}^H\mathbf{d} = \mathbf{V}\Sigma\mathbf{U}^H\mathbf{d} = \sum_{i=1}^{N} \sigma_i(\mathbf{u}_i^H\mathbf{d})\mathbf{v}_i \tag{5.12}$$

predominantly contains contributions from right singular vectors that correspond to the largest singular values. Furthermore, matrix $\mathbf{A}$ has many zero singular values and since in CGLS the largest singular values are typically approximated first, we expect that $\mathbf{x}_k$ mainly contains information about the right singular vectors that correspond to the first k singular values. In fact, in case of noisy data, the regularizing effects of CGLS may be used to obtain stable reconstructions by terminating the iterative process after a certain number of iterations (semiconvergence, see e.g. [45]). For completeness, we mention that the approximations $\mathbf{x}_k$ of (5.11) may also be generated by the LSQR algorithm [55].

Finally, we note that if we add a sparsifying regularizer to the objective function of (5.9), we end up with the prototype objective function of CS (5.1). If the image has a sparse representation in some basis and the imaging artefacts are incoherent, CS may accurately reconstruct the image without any support information. However, when this support information is available then CS techniques may not be necessary and solving the normal equation (5.8) with CGLS and a vanishing initial guess may already give satisfactory reconstruction results. In the following section we present reconstruction results for k-space undersampling patterns typically used in MRI, which illustrate that solving the normal equation with CGLS indeed may be sufficient in these cases and no CS techniques are necessary.

5.3. IMAGE RECONSTRUCTION

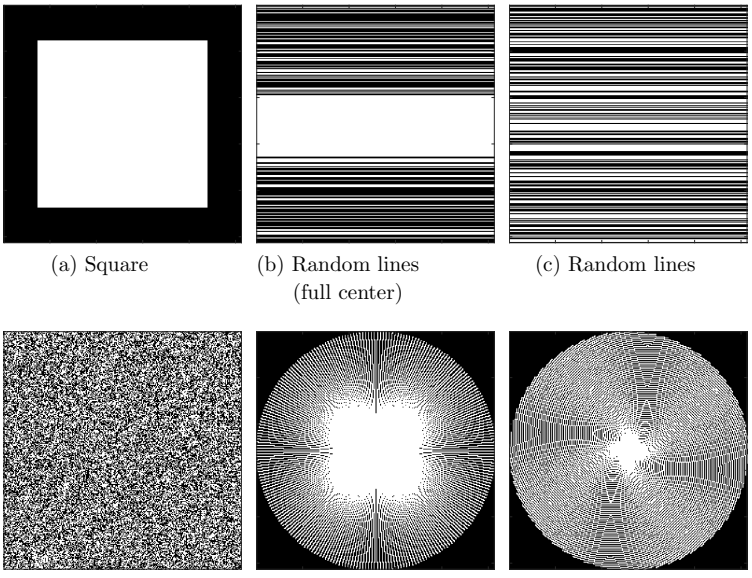
To illustrate the performance of the CGLS imaging algorithm in case support information is available, we consider reconstruction problems for noise-free and noisy 2D and 3D data sets. In our 2D experiments we use a Shepp-Logan phantom with 64×64 pixels [131] as the model solution and a brain image from the Kirby 21 dataset [132] with 256 pixels in each spatial direction. These phantoms and their supports are shown in Figure 5.1. The support of the 2D Shepp-Logan phantom consists of 1988 out of a total of 4096 pixels ($\approx 49\%$), and the support of the 2D brain image has a size of 23892 out of a total of 65536 pixels ($\approx 36\%$).



(a) Shepp-Logan phantom (b) Shepp-Logan support function (c) Brain image (d) Brain support function

Figure 5.1: Model solutions (a) and (c) and their support (b) and (d) shown in white.

We consider six different undersampling patterns with Fourier or k-space support functions illustrated in Figure 5.2. The first support function (Figure 5.2a) is a square undersampling pattern that leaves out the edges of k-space. The second support function (Figure 5.2b) is a pattern of random lines with all center lines being sampled, while the third pattern (Figure 5.2c) consists of completely random lines (so not all center lines are sampled). An undersampling pattern consisting of random points (Figure 5.2d) is also considered, along with a radial pattern and a spiral pattern (Figures 5.2e and 5.2f, respectively).



(a) Square (b) Random lines (full center) (c) Random lines

(d) Random points (e) Radial (f) Spiral

Figure 5.2: Tested undersampling patterns. White represents k-space points that are taken into account during reconstruction, black points are not used during reconstruction. The exact undersampling patterns that are used for the brain image experiment with 256×256 k-space points are shown. For the 64×64 Shepp-Logan case, the undersampling patterns look similar.

In all cases, the undersampling factor is (approximately) equal to two. We note that in MRI, some of these patterns are more difficult to implement in practice than others due to hardware limitations. Additionally, we remark that in reality, when using a radial or spiral pattern, the sampled points do not necessarily fall on a Cartesian grid. To make the application of an inverse Fourier transform to k-space possible, an iterative process called gridding [133, 134] may be used to transform the measured data into a Cartesian format. In this work, however, we do not take these limitations into consideration and simply assume that we have access to undersampled k-space data in Cartesian format.

Figure 5.3 shows the reconstruction results for the Shepp-Logan phantom and the six different undersampling patterns of Figure 5.2. Specifically, the model solution is shown in the left column, while the reconstructions obtained by simply applying an inverse FFT (IFFT) to incomplete spectral data are shown in the second column. The reconstructions that are obtained using CGLS starting with a vanishing initial guess are shown in the fourth column of Figure 5.3. These results were obtained after 1000 CGLS iterations or when the residual dropped below a tolerance of 10^{-10}. We observe that for all cases, CGLS with known support information yields images of improved quality compared to images obtained by simply applying an IFFT. This is also reflected in the peak signal-to-noise ratio (PSNR) values presented in Table 5.1.

For the brain image, the reconstruction results are shown in Figure 5.4 and detailed views of the reconstructions within the domain indicated by the white square in Figure 5.1c are shown in Figure 5.5. Comparing the reconstruction results obtained using an IFFT with the reconstructions obtained using CGLS, we again observe that including support information significantly improves the quality of the reconstructions. To quantify this statement, we compute the PSNR values of the various reconstructions shown in Figure 5.4 and the results are presented in Table 5.2. We observe that for all undersampling patterns the PSNR values of the reconstructions obtained with CGLS and support information included are higher than the PSNR values of the reconstructions obtained with an IFFT without support information. We note that the CGLS algorithm performs particularly well for a random points undersampling pattern (The norm of the error of the reconstruction result is in the order of 10^{-8} for the brain image). Unfortunately, such a sampling pattern is very difficult or impossible to realize during practical MR scans. Furthermore, in the case of a spiral or radial undersampling pattern, our algorithm using support information significantly increases the PSNR of the reconstruction.

Table 5.1: PSNRs of the Shepp-Logan reconstructions obtained using an IFFT, CS, and the CGLS algorithm with support information included for the six undersampling patterns of Figure 5.2.

Undersampling pattern	PSNR (IFFT)	PSNR (CS)	PSNR (CGLS)
(a) Square	21.40	26.46	38.49
(b) Random lines (full center)	19.97	34.56	30.52
(c) Random lines	18.05	24.25	22.16
(d) Random points	17.21	24.13	72.31
(e) Radial	20.71	47.61	39.82
(f) Spiral	19.85	51.51	45.12

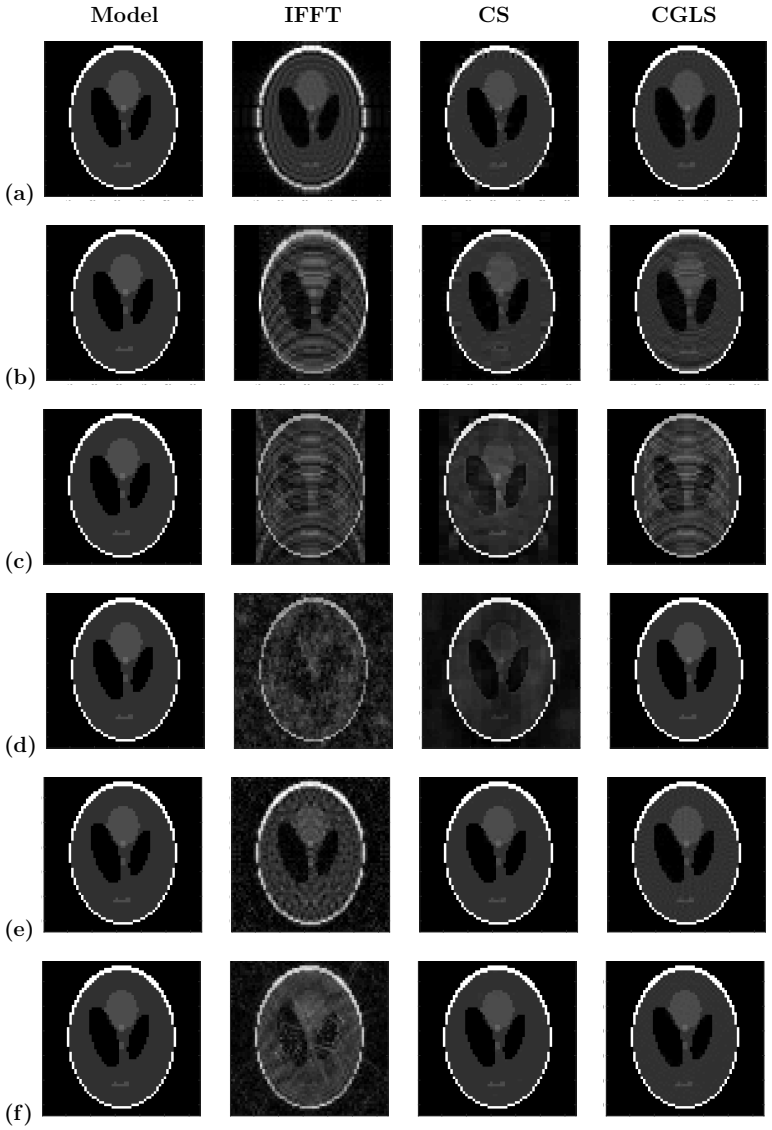


Figure 5.3: Reconstructed Shepp-Logan images: (a) square, (b) random lines (with center), (c) random lines, (d) random points, (e) radial, (f) spiral undersampling pattern. First column: model solution, second column: IFFT reconstruction, third column: CS reconstruction, fourth column: CGLS reconstruction with support information.

Table 5.2: PSNRs of the Kirby reconstructions obtained using an IFFT, CS, and the CGLS algorithm with support information included for the six undersampling patterns of Figure 5.2.

Undersampling pattern	PSNR (IFFT)	PSNR (CS)	PSNR (CGLS)
(a) Square	36.37	30.76	38.38
(b) Random lines (full center)	30.37	33.97	37.73
(c) Random lines	16.54	17.26	32.00
(d) Random points	15.61	16.97	209.45
(e) Radial	33.80	35.99	45.25
(f) Spiral	29.80	38.35	47.55

5.3.1. COMPRESSED SENSING

The image reconstruction problem may also be solved with CS techniques, of course. No support information is required in this case, but CS imaging algorithms are generally more complex than the CGLS approach considered in this paper.

As an illustration, let us consider the CS problem that consists of finding the image function $\mathbf{x}$ that minimizes the objective function

$$F_{\text{CS}}(\mathbf{x}) = \|\mathbf{S}_k(\mathbf{d} - \mathscr{F}\mathbf{x})\|_2^2 + \lambda\|\Psi\mathbf{x}\|_1, \tag{5.13}$$

for the undersampling patterns of Figure 5.2. Here, Ψ is the Daubechies wavelet operator [135] and λ is a regularization parameter, which is determined in a heuristic manner through numerical experimentation. No support information is included and the minimization problem is solved using the Alternating Direction Method of Multipliers (ADMM). For all experiments, we used 100 ADMM iterations and within each ADMM iteration, we used 10 iterations of the Conjugate Gradient (CG) algorithm to solve the first minimization problem. The reconstruction results for the Shepp-Logan phantom and the Kirby model are shown in the third column of Figures 5.3, 5.4, and 5.5. The PSNR values of the corresponding CS reconstructions are presented in Tables 5.1 and 5.2.

From Figure 5.3 and Table 5.1 we observe that for the Shepp-Logan phantom, CGLS with support information included generally produces reconstructions of a similar quality as the reconstructions obtained with CS. In most cases, the PSNR values of the CS reconstructions are (somewhat) larger, except for the square and random points undersampling patterns. In these cases, CGLS outperforms CS and, as mentioned above, especially the CGLS reconstruction for a random points undersampling pattern is of a very high quality.

From Figures 5.4 and 5.5 and Table 5.2 we observe that for the Kirby data set, CGLS in general again produces reconstructions of a similar quality as the reconstructions obtained with CS, except that here the PSNRs of all CGLS reconstructions are larger than the PSNRs of the corresponding CS reconstructions. A random points undersampling pattern, in particular, produces a CGLS reconstruction that is clearly superior to the CS reconstruction. Moreover, for a random lines undersampling pattern, CGLS produces a reconstruction that is significantly better than the reconstruction obtained with CS as well with a PSNR that is almost twice as large as the PSNR of the CS reconstruction.

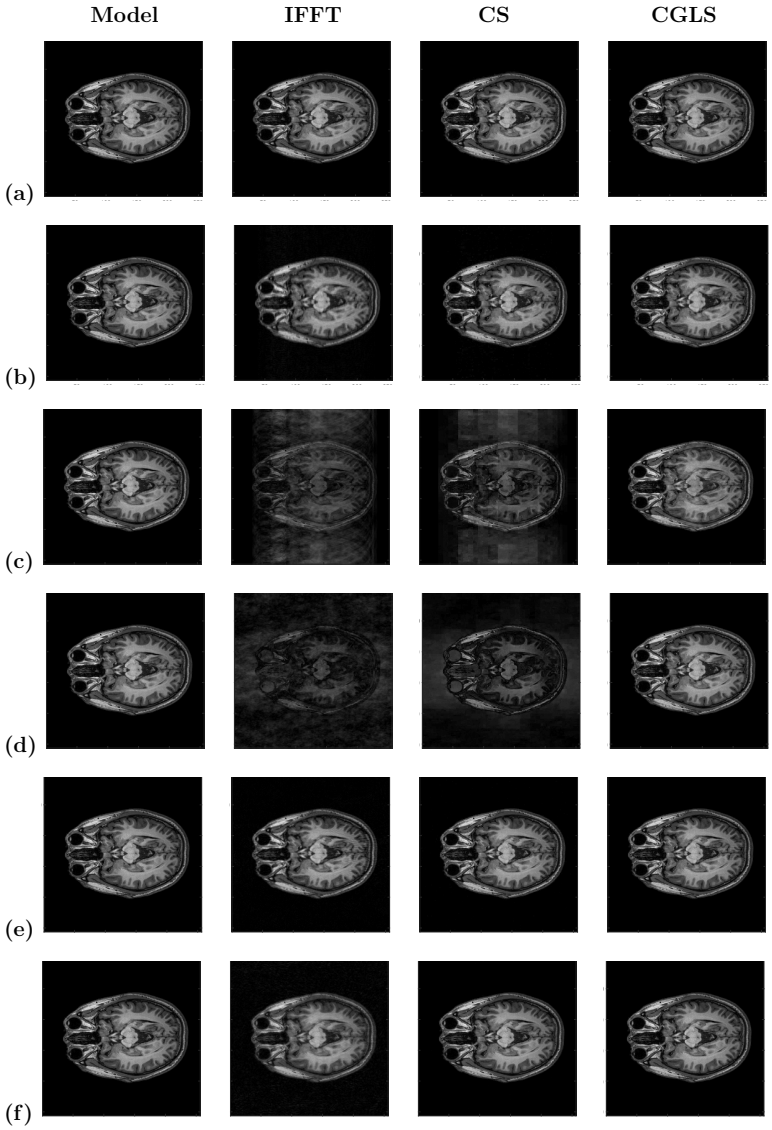


Figure 5.4: Reconstructed brain images: (a) square, (b) random lines (with center), (c) random lines, (d) random points, (e) radial, (f) spiral undersampling pattern. First column: model solution, second column: IFFT reconstruction, third column: CS reconstruction, fourth column: CGLS reconstruction with support information.

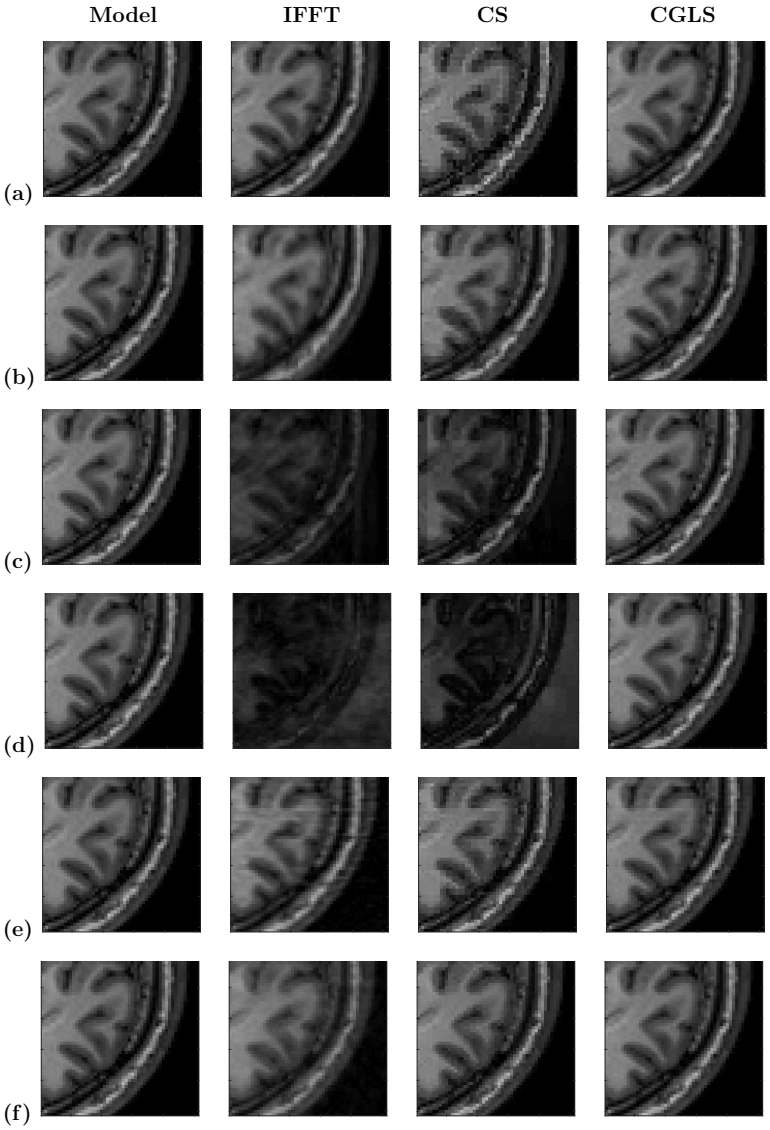


Figure 5.5: Patch of each of the reconstructed brain images: (a) square, (b) random lines (with center), (c) random lines, (d) random points, (e) radial, (f) spiral undersampling pattern. First column: model solution, second column: IFFT reconstruction, third column: CS reconstruction, fourth column: CGLS reconstruction with support information.

5.3.2. NOISY MEASUREMENTS

To demonstrate the performance of the CGLS reconstruction method in the case of noisy data, we again consider reconstructing the Shepp-Logan and Kirby phantoms from their corresponding incomplete k-space data, but now the data is contaminated by noise with an SNR of 50. For both phantoms, we restrict ourselves to a spiral undersampling pat-

tern, since the effects of noise are similar for the other undersampling patterns and noisy data is also considered in Section 5.3.3, where we apply the proposed CGLS algorithm to measured data.

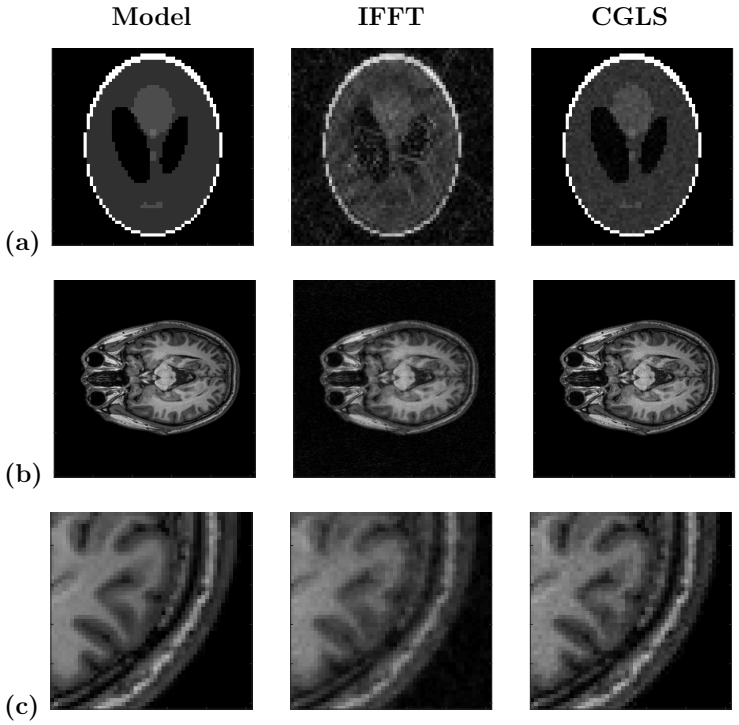


Figure 5.6: Reconstructed images based on noisy spiral data with an SNR of 50: (a) Shepp-Logan phantom, (b) brain image, (c) patch of the same brain image. Left column: Model solution, center column: IFFT reconstruction, right column: CGLS reconstruction with support information.

In Figure 5.6, the model solutions (first column), the reconstructions obtained with an inverse FFT (second column), and the reconstructions obtained with the CGLS algorithm with support information included (third column) are shown for the Shepp-Logan phantom (first row), the Kirby head model (second row), and for the region indicated by the white square in Figure 5.1c (third row). These results clearly show that the CGLS algorithm, with support information included, yields reconstructions that resemble the model solution much more accurately than the reconstruction obtained by simply applying the inverse Fourier transform to the undersampled data. This is also confirmed by the PSNR values shown in Table 5.3.

Table 5.3: PSNRs of the reconstructions obtained by applying a simple IFFT to noisy spiral data and using CGLS with support information.

	PSNR (IFFT)	PSNR (CGLS)
Shepp-Logan	19.75	34.82
Brain	29.34	37.82

5.3.3. RECONSTRUCTIONS BASED ON EXPERIMENTAL LOW-FIELD MRI DATA

In this section, we apply the proposed reconstruction method to a 3D data set, which was acquired using a spin-echo sequence on the low-field MRI scanner described in [20]. The scanning parameters of the measurement are reported in Table 5.4. An apple was placed inside the scanner and all imaging was carried out using a single receive coil. We note that the signals that were obtained have an SNR that is much lower than the SNR of signals measured in typical commercial MR scanners (in this case, the SNR is about 6).

Table 5.4: Scanning parameters used in the apple imaging experiment.

Parameter	Value
Repetition time (TR)	3 s
Echo time (TE)	30 ms
Imaging domain/FoV	$128 \times 128 \times 128$ mm^3
Pulse duration	100 μs
Acquisition bandwidth	10 kHz
Number of averages	1

Applying a 3D inverse Fourier transform to fully sampled k-space data results in a 3D reconstruction of the apple and some slices of this reconstruction are shown in the left column of Figure 5.7. These reconstructions serve as model solutions, since we obviously do not have a perfect or high SNR model solution available in this case.

To illustrate the performance of CGLS, we again consider a k-space undersampling pattern of random lines with the center of k-space completely sampled. The resulting undersampling factor is equal to two and the reconstruction results obtained by simply applying an inverse FFT to the incomplete data set and after 100 iterations of the CGLS algorithm are shown in the second and third column of Figure 5.7, respectively. We observe that an image is obtained which visually resembles the original full k-space solution. The PSNR values in Table 5.5 also show that using support information increases the PSNR from 24.02 (IFFT reconstruction) to 61.92 (CGLS reconstruction).

Table 5.5: PSNR of the 3D apple reconstructions obtained using a simple inverse Fourier transform and using CGLS with support information included.

	PSNR (IFFT)	PSNR (Support)
Apple image	24.02	61.92

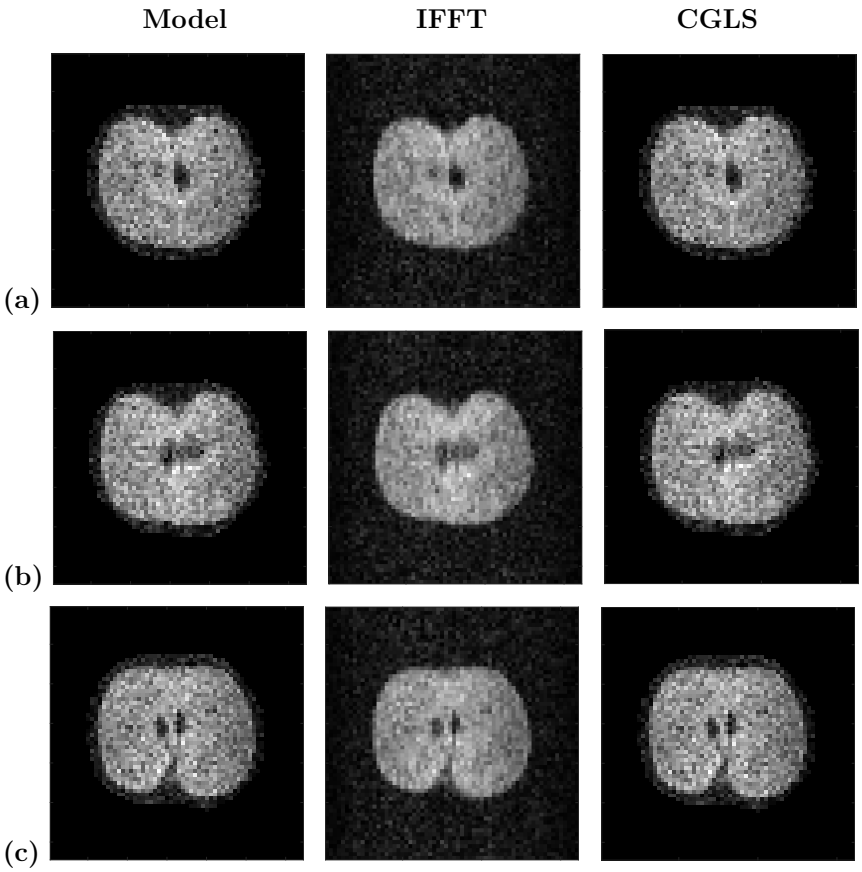


Figure 5.7: Some slices of the reconstructed 3D apple for a stack of random lines undersampling patterns. Left: Model solution, center: IFFT reconstruction, right: CGLS reconstruction with support information included.

5.4. CONCLUSION AND DISCUSSION

In this chapter we discussed an imaging method that can be applied if the support of an object within a certain FoV is known and its spatial Fourier transform is only known on a certain k-space undersampling pattern. We demonstrated that the CGLS algorithm applied to the corresponding truncated Fourier transform equation produces reconstructions that are essentially of a similar quality as reconstructions obtained by solving a standard CS problem in which support information is not taken into account. In particular, for the Shepp-Logan phantom and a range of two-dimensional k-space undersampling patterns, the CGLS algorithm produces reconstructions with PSNR values that are generally slightly smaller than the PSNR values of the corresponding CS reconstructions. However, for a realistic head model the PSNR values of the CGLS reconstructions are typically larger than the PSNR values of the CS reconstructions. Specifically, in the case of a random points undersampling pattern, the CGLS reconstructions of the Shepp-

Logan phantom and the head model have significantly larger PSNR values than their CS counterparts. For both phantom models, the PSNR values of the CGLS reconstructions are also larger than the PSNR values of the reconstructions obtained by simply applying an IFFT to the available incomplete data sets.

In 3D, an improvement in the reconstructions was also observed when support information is taken into account. The CGLS algorithm was applied to a measured data set obtained with a low-field MR scanner. For a random lines undersampling pattern the method was able to provide reconstructions with sufficient detail and PSNR values that are significantly larger than the PSNR values of the reconstructions obtained via a standard IFFT. In conclusion, when support information about the object is available, a straightforward application of the CGLS algorithm to a truncated Fourier transform equation definitely improves simple IFFT-based reconstructions and, at least for the undersampling patterns considered here, generally provides reconstructions with a quality similar to reconstructions obtained via standard CS techniques.

Carrying out a Fourier reconstruction using a reduced number of data points invariably leads to imaging artifacts. In this work, we considered basic CGLS and CS reconstruction algorithms to address this problem. More advanced CS techniques can be used as well, of course, (see [136–145], for example), but the same holds true for the CGLS algorithm. For example, in an MR setting, the use of multiple receive coils can be included in the reconstruction process and parallel imaging techniques analogous to SENSE [146] and GRAPPA [147] may be developed. Furthermore, total variation regularization may be included in the reconstruction process as well, either in an additive or multiplicative manner [148].

6

DEEP LEARNING FOR CORRECTION OF IMAGE DISTORTIONS CAUSED BY FIELD INHOMOGENEITIES AND GRADIENT NONLINEARITIES

6.1. INTRODUCTION

Because of the weaker background field, MR signals generated in low-field MRI scanners suffer from a lower signal-to-noise ratio (SNR) than those generated in high-field scanners. Additionally, as is the case in high-field MR scanners as well, the static background field may not be sufficiently homogeneous throughout the Field of View (FoV) and the magnetic fields generated by the gradient coils may not be linear or are only approximately linear, depending on the type of scanner used. In case the static background field is homogeneous and the gradient fields are linear, the signal model describing the relationship between image and measured signal reduces to a Fourier transform. If the static background field and/or the gradient fields deviate from these assumptions, artifacts occur in the image when a simple inverse Fourier transform is applied to the measured data. Many approaches exist that correct for image distortions caused by gradient nonlinearities and by inhomogeneities in the static background field. Model-based image

This chapter is based on the article:

M.L. de Leeuw den Bouter, R.F. Remis, and M.B. van Gijzen, *Deep learning for correction of image distortions caused by background field inhomogeneities and gradient nonlinearities in low-field MRI*, to be submitted to SCEE 2022 proceedings, July 2022.

reconstruction can be used for example [103, 104]. Another approach is given in [149], where Janke et al. use spherical deconvolution methods. For these methods to work, the inhomogeneities and nonlinearities need to be known explicitly. If accurate field maps are not available, a method like conjugate phase reconstruction, as introduced in [150], can be used. In [151], field and image are estimated simultaneously. However, the first method requires a set of images of the same object with different echo times, and the second one uses multiple receiver coils to acquire a sufficient amount of information to achieve its goal. In this chapter, we focus on a scenario in which we have one image, acquired using a single receiver coil. We assume we do not have access to exact maps of the static magnetic background field nor the fields generated by the gradient coils. Our goal is to correct images that have been distorted because of nonlinear gradient fields and/or inhomogeneities in the static background field, without having to incorporate any exact field maps. Three examples of distorted low-field MR images are shown in Figures 6.1, 6.2 and 6.3. These images originated from measurements carried out using different versions of the low-field scanner described in [20] and [37]. The phantom that was used to acquire the first image was a small real-life version of the Shepp-Logan phantom [131] shown in Figure 6.1. The second image was obtained by imaging a melon. For the final experiment, an additional Shepp-Logan phantom was created. While the first one was less than 10 cm in diameter, this later version was comparable in size to a human head. Clearly, the resulting images have been distorted by gradient nonlinearities and/or static background field inhomogeneities. The small Shepp-Logan phantom was imaged using the earliest version of the scanner and the resulting image is more severely distorted than the other two images, which is consistent with the fact that the homogeneity of the scanner's background field and the linearity of the gradient fields were improved in each updated version of the scanner. However, the melon's shape still deviates from the round shape we would expect to see and so does the second Shepp-Logan image. Our goal is to correct for these distortions. To accomplish this goal, we train a convolutional neural network on a large dataset generated using simulations, in which the inhomogeneities and nonlinearities causing the distortions vary from sample to sample. Over the last couple of years, neural networks have been applied successfully to different problems in MR imaging, see for example [46, 152–154]. By using a dataset with distortions caused by variable static background fields and gradient nonlinearities, we aim to realize a trained network that is capable of correcting image distortions for low-field MRI scanners in general, without having to retrain it for every single scanner with a different measured dataset. Therefore, testing the network on experimental data originating from three different versions of the low-field MRI scanner, with a different static background field and different magnetic fields generated by the gradient coils, is warranted.

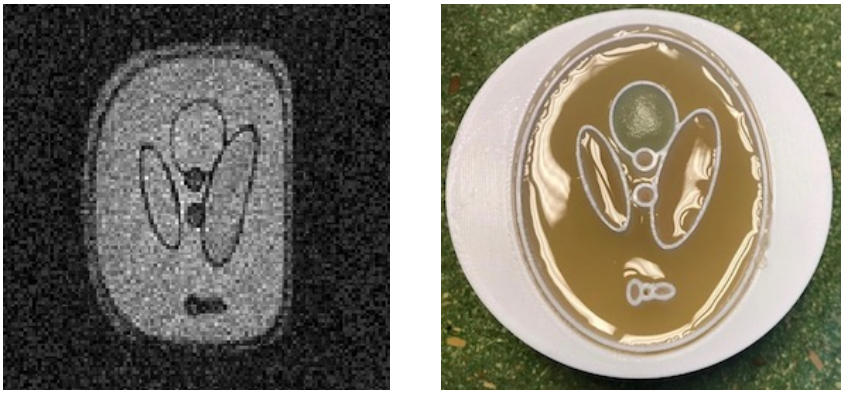


Figure 6.1: Initial reconstruction (left) of a small Shepp-Logan phantom (right) obtained by applying a 2D inverse Fast Fourier Transform to the measured data.

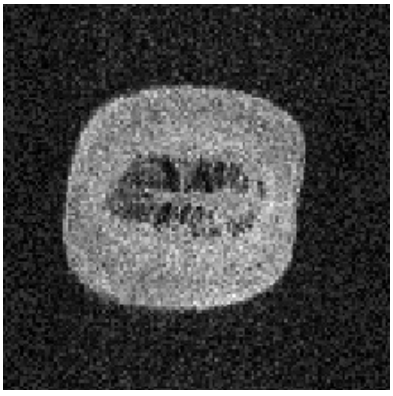


Figure 6.2: Initial distorted reconstruction of a melon.

Figure 6.3: Initial reconstruction (left) of a head-sized Shepp-Logan phantom (right) obtained by applying a 2D inverse Fast Fourier Transform to the measured data.

6.2. SIGNAL MODEL

Using $\mathbf{r}$ to denote position and assuming a 2D scenario, the relationship between signal $b(t)$ and image/spin density $X(\mathbf{r})$, ignoring relaxation effects, is described by the well-known equation [19]

$$b(t) = \int_{\mathbf{r}\in \mathrm{FoV}} X(\mathbf{r}) e^{-i\Delta\omega_0(\mathbf{r})t} e^{-i2\pi\mathbf{k}(t)\cdot\mathbf{r}} \,\mathrm{d}\mathbf{r}, \tag{6.1}$$

where γ denotes the gyromagnetic ratio, $\Delta\omega_0(\mathbf{r}) = \gamma B_0(\mathbf{r}) - \omega_{\mathrm{mod}}$ is the difference between the static background field B_0 and the demodulation frequency ω_{mod}, and $\mathbf{k}$ can be described by

$$\mathbf{k}(t) = \frac{\gamma}{2\pi}\int_0^t \mathbf{G}(\tau)\,\mathrm{d}\tau. \tag{6.2}$$

Here, $\mathbf{G}(\mathbf{r},\tau) = \left[G_x(\mathbf{r},\tau),\ G_y(\mathbf{r},\tau)\right]^T$ is the vector corresponding to the gradients generated by the gradient coils. For a spin-echo experiment, Eq. (6.1) can also be written as [19]

$$b(t) = \iint_{\mathrm{object}} X(x,y) e^{-i\gamma(B_f(x,y)t + n\Delta B_p(x,y)T_{\mathrm{pulse}})} \,\mathrm{d}x\,\mathrm{d}y, \tag{6.3}$$

where n is used to denote the nth excitation cycle, T_{pulse} the duration of the phase-encoding pulse, $B_f(x,y)$ is the magnetic field present during readout (the subscript f stands for frequency encoding) and $\Delta B_p(x,y)$ is the difference in phase encoding gradient from one excitation cycle to the next, such that the total magnetic field generated by the phase encoding gradient is described by $B_p(x,y) = n\Delta B_p(x,y)$ (where the subscript p is used to denote phase encoding)[1]. Note that in Eq. (6.3), the static background field and the field generated by the frequency encoding gradient have been absorbed into the same term $B_f(x,y)$. The gradient vector $\mathbf{G}$ in Eq. (6.2) is then given by:

$$\mathbf{G} = \begin{cases} \nabla B_f, & \text{for } t > T_{\mathrm{pulse}} \text{ (frequency encoding)} \\ \nabla B_p, & \text{for } t < T_{\mathrm{pulse}} \text{ (phase encoding).} \end{cases} \tag{6.4}$$

In Eq. (6.3), we use $B_f(x,y) = \tilde{B}_f(x,y) - \frac{1}{\gamma}\omega_{\mathrm{mod}}$ and $B_p(x,y) = \tilde{B}_p(x,y) - \frac{1}{\gamma}\omega_{\mathrm{mod}}$, where $\tilde{B}_f(x,y)$ and $\tilde{B}_p(x,y)$ are the actual magnetic field strengths (of around 50 mT) present during frequency and phase encoding, respectively. As mentioned before, $B_f(x,y)$ is used to signify the sum of the background field and the additional magnetic field that is generated by a frequency-encoding gradient coil. In the term $B_y(x,y)$, we only consider the nonlinearities in the magnetic field generated by the phase-encoding coil, not the background field inhomogeneities. The reason is that since the length of the phase-encoding pulse is constant, the background field inhomogeneities yield a constant phase difference whose only effect is an additional phase in the resulting image.

[1] The magnitude of the field $B_p(x,y)$ generated during the phase encoding pulse is determined by the amount of current being sent through the phase encoding gradient coil. Increasing (or decreasing) the current by a factor will simply lead to an increase (or decrease) of the magnitude of the magnetic field by the same factor. This means we can also include nonlinearities present in $B_p(x,y)$, since the nonlinearities are simply rescaled from one excitation cycle to the next. Hence, the term $\Delta B_p(x,y)$ is equal to the magnetic field generated by the phase encoding gradient coil corresponding to the excitation cycle denoted by $n = 1$ and, in general, for each excitation cycle, we can simply write $B_p(x,y) = n\Delta B_p(x,y)$.

6.3. METHODS

6.3.1. EXPERIMENTAL SETUP

SMALL SHEPP-LOGAN PHANTOM

The small Shepp-Logan phantom shown in Figure 6.1b was imaged using a very early version of the low-field MRI scanner described in [20] and [37] in March of 2019. A spin-echo pulse sequence was used with the parameter settings in Table 6.1.

Table 6.1: Parameter settings for the small Shepp-Logan imaging experiment.

Parameter	**Value**
Repetition time (TR)	1 s
Echo time (TE)	10 ms
Number of complex data points	128 × 128
Imaging domain/Field of view (FoV)	120 × 120 mm^2
Pulse duration	100 μs
Acquisition bandwidth	50 kHz
Number of averages	8

MELON

To assess whether the method generalizes to other objects, we will also include the image shown in Figure 6.2. This image is not as clearly distorted as the Shepp-Logan image, but its shape does deviate from the shape we would expect a melon to have. The image shown is the 64th slice of a 128 × 128 × 128 pixel 3D image. For this experiment, a turbo spin-echo sequence was used with the parameters as shown in Table 6.2. This experiment was carried out on a more advanced version of the same low-field MRI scanner in March of 2020.

Table 6.2: Parameter settings for the melon imaging experiment.

Parameter	**Value**
Repetition time (TR)	2 s
Echo time (TE)	30 ms
Number of complex data points	128 × 128 × 128
Imaging domain/Field of view (FoV)	200 × 200 × 200 mm^3
Pulse duration	100 μs
Acquisition bandwidth	20 kHz
Echo train length	16

HEAD-SIZED SHEPP-LOGAN PHANTOM

The final experiment was carried out using a larger Shepp-Logan phantom in November of 2020. While the earlier Shepp-Logan phantom had a diameter of less than 10 cm, this one was comparable in size to a human head. The parameters that were used here were comparable to the ones used for the first Shepp-Logan experiment, except that the FoV was significantly larger.

6.3.2. DATASET GENERATION

We simulated a dataset of 100,000 2D model images of 128 × 128 pixels and their distorted counterparts. Each image consists of a set of superimposed ellipses of intensities varying between 0 and 1. The reasoning behind this is that the images we are trying to improve, namely, images of a Shepp-Logan phantom and a melon, consist of ellipses as well. This training set was inspired by [50], where a Shepp-Logan phantom was reconstructed using a convolutional neural network trained to do iterative image reconstruction on Computed Tomography (CT) data simulated using images consisting of ellipses.

In MRI, it is common to expand the magnetic fields in terms of spherical harmonics for correction and analysis [155–158]. Therefore, for each sample, $B_f(x, y)$, the magnetic field present during readout, and $B_p(x, y)$, the magnetic field which is present during the phase encoding step, are perturbed by adding spherical harmonics of the second and third order of random magnitude to the ideal magnetic fields whose magnitudes increase linearly in the x- and y-direction, respectively. When perturbing B_f, we actually perturb the total sum of the static background field and the magnetic field generated by a frequency encoding gradient and do not distinguish between the two. We stress again that by using a dataset with varying gradient nonlinearities and inhomogeneities, we aim to make our approach applicable to a variety of low-field MRI scanners using a gradient-based design.

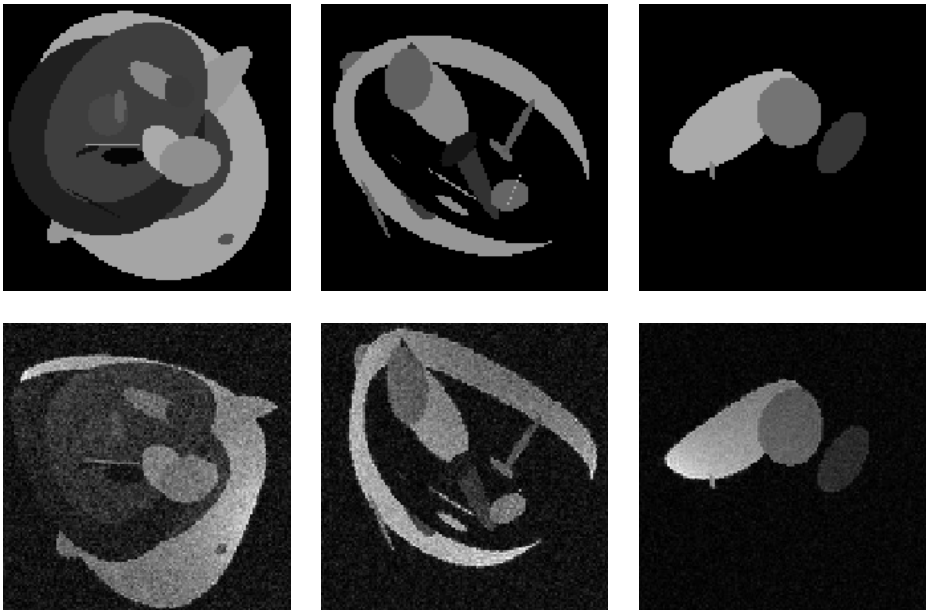


Figure 6.4: Three samples in the dataset. Top row: model images, bottom row: distorted images. Frequency encoding was done in the horizontal direction, phase encoding in the vertical direction.

Specifically, field perturbations were introduced by first mapping the "model" magnetic field values between -1 and 1. We restricted the magnitude of the perturbations

by allowing the maximum deviation in the magnetic fields to be in the range (-0.6,0.6). Additionally, images which were distorted to the point that ellipses crossed an image boundary were eliminated from the dataset.

Using these distorted magnetic field maps, we simulated the signal generation using Eq. (6.3), leading to distorted images with intensity variations. We added complex Gaussian white noise to all signals, with variable signal-to-noise ratios (SNRs). Applying an IFFT to these noisy signals yields a complex-valued, noisy, distorted image. We took the magnitude of these images as the input to the network. Subsequently, the intensities of both distorted and undistorted images were rescaled such that the maximum pixel value in both types of images is equal to 1. Some samples in the dataset are shown in Figure 6.4.

To ensure rotational invariance as much as possible, each sample in the dataset was included 4 times: the original images (distorted and undistorted) and the same images rotated by 90°, 180° and 270°, leading to a dataset of 400,000 samples.

6.3.3. CONVOLUTIONAL NEURAL NETWORK

A convolutional neural network of the U-Net architecture was chosen for the purpose of correcting the image distortions caused by gradient nonlinearities. U-Net was introduced in 2015 by Ronneberger et al. [159] for biomedical image segmentation. Variations of it have been used for medical image reconstruction, see for example [160].

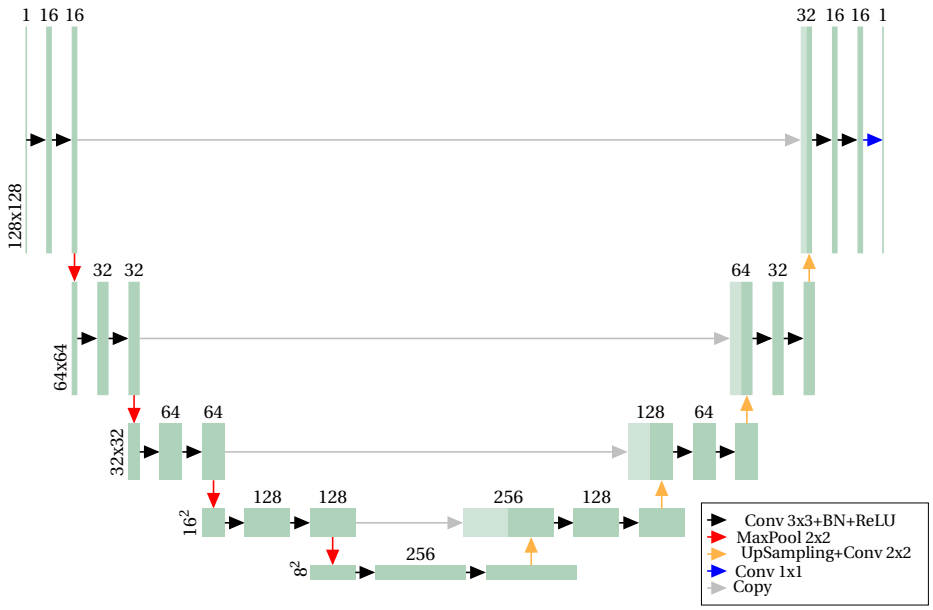


Figure 6.5: The U-Net convolutional neural network.

U-Net consists of a contracting path and an expanding path. The contracting path is built up of blocks consisting of 3x3 convolutional layers, batch normalization layers, rectified linear unit (ReLU) activation functions and max pooling layers. The max pooling layers make sure that the feature maps are downsampled, and with each max pooling layer, the number of channels is doubled. The blocks in the expanding path consist of 3×3 convolutional layers, batch normalization layers, ReLU activation functions and upsampling layers. After each upsampling layer, the number of channels is halved. The output of each level of the encoder is used as part of the input of the corresponding decoder level (skip connections). The exact architecture is shown in Figure 6.5.

We used 90% of the samples in the simulated dataset for training and 10% for validation. Each input-output pair consisted of the noisy distorted image and the corresponding undistorted image. We chose L1-loss as our loss function. The batch size was set to 20, and we used the Adam optimizer [161] with a learning rate of 10^{-3}, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$, for 20 epochs.

6.4. RESULTS

6.4.1. SMALL SHEPP-LOGAN PHANTOM

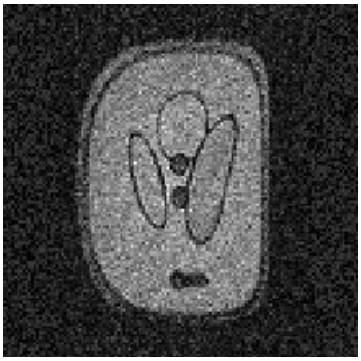
After training the network, we feed the distorted original image shown in Figure 6.6a to the network, which was obtained by simply applying an inverse Fourier transform to the measured k-space data. The resulting U-Net reconstruction is shown in Figure 6.6b. We see that the network succeeds at correcting most of the distortions caused by the nonlinearities in the magnetic fields of the MRI scanner. The ellipses in the phantom have recovered their shape. The outermost ellipse is only partially reconstructed, but this can be explained by the observation that this particular ellipse corresponds to a thin line, especially at the bottom and the sides of the physical phantom, which is difficult to fill, therefore containing very little fluid, see Figure 6.1b. This is confirmed by the original reconstruction, in which some segments of the outermost ellipse are barely visible, as can be seen in Figure 6.6a. Moreover, we note that the network manages to successfully denoise the entire image.

6.4.2. MELON

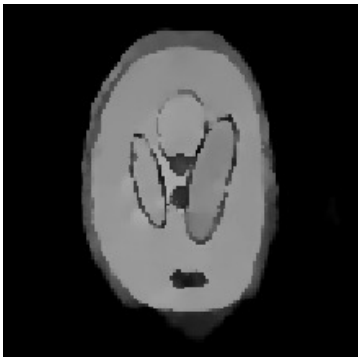
Additionally, we feed the image shown in Figure 6.7a to the neural network, and we obtain the result in Figure 6.7b. The network output looks rounder and is much more reminiscent of a melon's shape. Additionally, all noise has been removed from the image.

6.4.3. HEAD-SIZED SHEPP-LOGAN PHANTOM

Finally, we provide the network with the distorted Shepp-Logan image shown in Figure 6.8a. The resulting output is shown in Figure 6.8b. While the resulting image has been denoised and the very top part of the object contour looks rounder and arguably better than before, some distortions have not been corrected at all and the shape of the uppermost circle inside the phantom has even deteriorated.

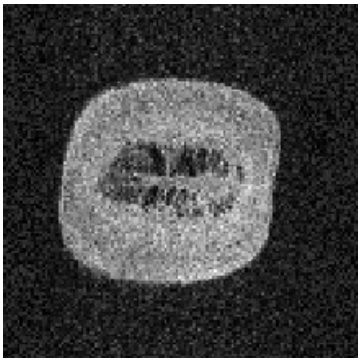


(a) Distorted image

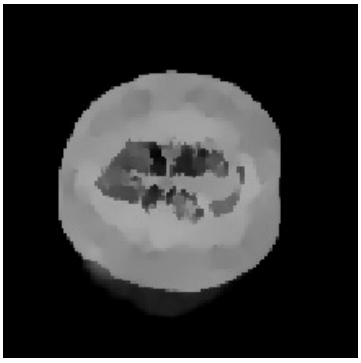


(b) Network result

Figure 6.6: Small Shepp-Logan image: input and output of the trained convolutional neural network.



(a) Distorted image



(b) Network result

Figure 6.7: Melon image: input and output of the trained convolutional neural network.

(a) Distorted image

(b) Network result

Figure 6.8: Head-sized Shepp-Logan image: input and output of the trained convolutional neural network.

6.5. DISCUSSION

We have seen that the network's performance is not always robust when it comes to different types of input. It is difficult to pinpoint what is causing this, due to the lack of interpretability exhibited by neural networks.

We are using simulated data to train the network, while we want to apply it to real data. A possible explanation of the network's inability to fully correct the distortions in the head-sized Shepp-Logan reconstruction is that it might have been subjected to magnetic fields whose nonlinearities are incongruent with the types of nonlinearities present in the training set. In the training set, we only included magnetic field nonlinearities in the form of spherical harmonics of the second and third order. We might not be able to correct for all types of artifacts in this way, but often an approximation of a low order can already be sufficient. It is possible that this is the case for the first two experiments, but not the final one.

Besides the simulated training data possibly not being able to generalize to the real-world scenario, the problem might lie in the specific architecture we used. While the U-Net is omnipresent in the field of image segmentation and has been used for some image reconstruction problems, we are attempting to tackle a different and very specific problem. We experimented with different architectures and the U-Net yielded the best results. However, in the rapidly expanding field of deep learning, different and more advanced neural networks are being developed on a daily basis. It is very likely that some of these would outperform the U-Net in this specific context.

Another assumption we made in this work is that the images consist of only ellipses. We made this choice because we know that our target images, the Shepp-Logan phantoms and the melon, can also be described as such. We will work on extending this approach to different images that do not necessarily consist of ellipses.

The current network architecture does not yield any output from which to effectively gauge the size and shape of the nonlinearities present in the magnetic fields, causing the distortions. We attempted to remedy this by, among other approaches, adding a spatial transformer or a diffeomorphic layer as introduced by Jadenberg et al. in [162] and Detlefsen et al. in [163], respectively, to the network. These are layers that deform the image grid. Since gradient nonlinearities and background field inhomogeneities also effectively deform the grid, we hoped that such an approach would give us the exact grid deformation caused by the nonlinearities and inhomogeneities. However, these did not prove to be effective for our dataset and further research is required.

6.6. CONCLUSION

In this work, we presented our first attempts at correcting distortions caused by background field inhomogeneities and gradient nonlinearities in low-field MR images (which were acquired using a single receiver coil) using a convolutional neural network. We trained a convolutional neural network of the U-Net architecture on a simulated dataset consisting of distorted, noisy images and their undistorted counterparts. By training the network on a simulated dataset with varying static background field inhomogeneities and gradient nonlinearities, we aimed to make it applicable to a variety of low-field MRI scanners using a gradient-based design. To test our approach, we used images of two

real-world Shepp-Logan phantoms and a melon and obtained varying results. While the most recent experiment, carried out using a head-sized Shepp-Logan phantom, did not lead to an improvement in image quality, the network performed well when provided with the original reconstructions of the first small Shepp-Logan phantom and the melon: the resulting images were denoised and the distortions were eliminated to a large extent.

7

DEEP LEARNING-BASED SINGLE IMAGE SUPER-RESOLUTION

7.1. INTRODUCTION

The low-field scanner described in [37] was designed to accommodate infants' heads. In general, scan times are long in MRI and a significant amount of contemporary MRI research focuses on accelerating the acquisition process. Clearly, staying still for dozens of minutes on end is even more challenging for infants than for adults, and hence efforts should be taken to make the scan duration as short as possible, while maintaining good image quality.

Assuming a full sweep of the spatial frequency domain, i.e., k-space, a shorter scan duration corresponds to an image of a lower resolution. Super-resolution image reconstruction techniques attempt to reconstruct a high-resolution (HR) image from one low-resolution (LR) image or several LR images, see for example [53]. We will focus on single image super-resolution (SISR), in which case we only have one LR image at our disposal. SISR is an ill-posed problem, because one LR image can correspond to several different HR images, and hence it does not have a unique solution.

We note that, to decrease the scan time and still reconstruct an image of good quality, one could also resort to compressed sensing (CS) techniques [106] by sampling only part of k-space and assuming that the image is sparse in some transform domain. However,

This chapter is based on the article:

M.L. de Leeuw den Bouter, G. Ippolito, T.P.A. O'Reilly, R.F. Remis, M.B. van Gijzen, and A.G. Webb, *Deep learning-based single image super-resolution for low-field MR brain images*, submitted to Nature Scientific Reports, December 2021.

we will focus on super-resolution (SR) image reconstruction instead. This decision was based on the observation that the SR problem is more straightforward than the CS problem. We will use a deep learning technique to improve image quality and, after being trained on LR-HR image pairs, our neural network should be applicable to all low-field MR brain images of the resolution the network was trained on. In the CS problem, on the other hand, there is an additional variable: the k-space sampling pattern. If we were to train a network on data acquired using a specific k-space sampling pattern, it is highly unlikely that the network would generalize to images acquired using a different pattern.

One class of methods that can be used for SISR relies on interpolation, like bicubic interpolation [164], Lanczos interpolation [165] and zero-padding the k-space data, the latter of which is common for MRI images. These methods are efficient and straightforward but the results they provide are generally lacking in accuracy when it comes to recovering high-frequency components [49]. Reconstruction-based SR methods generally pose the problem as a minimization problem in which prior knowledge about the solution is incorporated, thereby restricting the solution space. Examples can be found in [166], [167], [168], [169] and [170]. Drawbacks of reconstruction-based methods are that they are generally computationally expensive and that they require the tuning of hyperparameters. Learning-based (or example-based) methods are the third class of methods that can be used to tackle the problem of SISR. In general, these methods use machine learning to extract relationships between LR images and their HR counterparts. They are computationally fast and usually perform well [49]. Examples include methods based on Markov random fields [171], sparse representations [172, 173] and neighbor embeddings [97]. Of the learning-based methods, methods using deep learning (DL) generally outperform all other reconstruction-based and learning-based methods [49]. Examples of convolutional neural networks that have been used for SISR are SRCNN [174], SRDenseNet [175] and SRGAN [176], among many others. For a more complete overview of the DL methods that have been used for super-resolution, the reader is referred to [177].

Deep learning has been applied to the problem of SISR in MRI before. Pham et al. [178] applied a 3D version of the SRCNN network to MR brain images. Similarly, Chen et al. [81] developed DCSRN, a densely connected convolutional network, which is closely related to a 3D version of SRDenseNet, that was trained to carry out super-resolution on 3D MR brain images as well. Masutani et al. [179] applied modified versions of the SRCNN architecture and the U-Net [159] architecture to cardiac MRI images. Chen et al. [180] trained a multi-level densely connected network using a generative adversarial network (GAN)-based loss function.

In this work, we employ a convolutional neural network of the SRDenseNet architecture [175] to carry out SISR on MR brain images. The main contribution and the novelty of our work is in the application: we focus on low-field MR brain images. Low-field MRI is a rapidly growing field with a lot of potential, but it also comes with challenges. This paper demonstrates the potential of deep learning-based methods in addressing the challenges in low-field MR imaging.

7.2. METHODS

7.2.1. BACKGROUND

Given a low-resolution 2D image $\mathbf{y}$, our aim is to acquire its high-resolution counterpart $\mathbf{x}$. The relationship between $\mathbf{x}$ and $\mathbf{y}$ can be modeled as follows:

$$\mathbf{y} = \mathscr{F}_{LR}^{-1}\mathbf{D}\mathscr{F}_{HR}\mathbf{x} + \mathbf{n}, \tag{7.1}$$

where $\mathscr{F}_{LR}^{-1}$ is the inverse FFT operator applied in the LR domain, $\mathbf{D}$ is an operator that selects only the low-frequency components in k-space, which, in our case, is of size 64x64, $\mathscr{F}_{HR}$ is the FFT operator for the HR regime (128x128) and $\mathbf{n}$ is an (unknown) noise vector. The goal of super-resolution is to find an approximate inverse of the operator $\mathscr{F}_{LR}^{-1}\mathbf{D}\mathscr{F}_{HR}$. We note that the standard super-resolution problem is generally posed differently, i.e., the HR image is assumed to undergo blurring and downsampling, culminating in an LR image. However, using Equation (7.1) follows the low-resolution MRI acquisition process more accurately.

7.2.2. CONVOLUTIONAL NEURAL NETWORK

We chose a convolutional neural network of the SRDenseNet architecture for our application. Our choice was motivated by SRDenseNet's good performance, combined with its manageable number of parameters. We note that, as there is a vast literature on deep learning-based methods for super-resolution, other networks may be applicable as well [174, 176, 177]. SRDenseNet was introduced by Tong et al. in [175]. It consists of blocks of densely connected convolutional layers ("dense blocks"). In every dense block, which is consistent with the DenseNet architecture [181], each convolutional layer receives as input the concatenated outputs of all preceding convolutional layers, as shown in Figure 7.1.

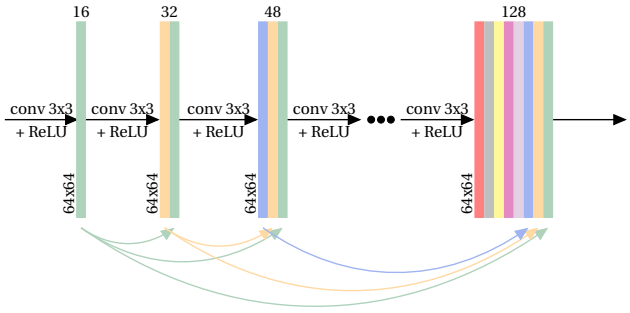


Figure 7.1: A dense block, which is a fundamental component of the SRDenseNet architecture, contains eight convolutional layers that receive the outputs of all preceding layers as input.

By reusing feature maps in this way the learning of redundant features is avoided. Instead, the current layer is forced to learn supplemental information. As in the original paper, we will use 8 dense blocks of 8 convolutional layers each, where each convolutional layer produces 16 feature maps, which means that each dense block yields 128

feature maps. In each convolutional layer, the kernel size is 3x3. After the final dense block, a bottleneck layer with convolutional kernels of size 1x1 is used to reduce the number of feature maps to 256, followed by a transpose convolutional layer (which is often called a deconvolution layer) which upsamples the image to HR space. Note that in this work, the upsampling factor is equal to 2 and hence we use only one single transpose convolutional layer with a stride of 2, as opposed to the 2 transpose convolutional layers in the original SRDenseNet which was used for an upsampling factor of 4. Finally, another convolutional layer with a 3x3 kernel is applied to reduce the output to a single channel.

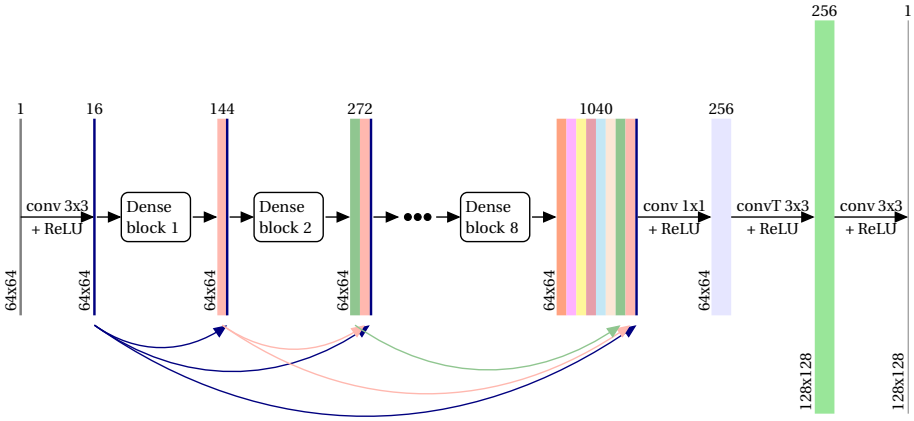


Figure 7.2: The SRDenseNet convolutional neural network [175] that is used to carry out single image super-resolution on low-field MR images.

All layers except for the final convolutional layer use a nonlinear ReLU (Rectified Linear Unit) activation function. Additionally, skip connections are employed to feed the output of each dense block to each of the subsequent dense blocks, as is consistent with the SRDenseNet_All architecture showcased in the original paper [175]. The complete architecture, which has 1,910,689 trainable parameters, is shown in Figure 7.2.

7.2.3. DATASET AND TRAINING

In this work, we focused on 2D images, but it should be noted that this approach can be extended to 3D. We generated a training and validation set using 2D images obtained from the NYU fastMRI Initiative database (`fastmri.med.nyu.edu`) [182, 183]. As such, NYU fastMRI investigators provided data but did not participate in analysis or writing of this manuscript. A listing of NYU fastMRI investigators, subject to updates, can be found at the aforementioned website. The primary goal of fastMRI is to test whether machine learning can aid in the reconstruction of medical images. The database consists of slices of T1-weighted, T2-weighted and FLAIR (fluid-attenuated inversion recovery) images, acquired using 1.5 T and 3 T MRI scanners. By training on such a variety of MR brain images, the resulting network should be applicable to images acquired using different sequences as well, without the need to re-train the network whenever the parameter

settings change. We note that, even if we were planning on applying the trained network to, for example, T1-weighted low-field MR images only, it would still make sense to train the network on high-field MR images acquired using different kinds of sequences, making it adaptable to different kinds of input. The reason for this is that the relaxation times vary with field strength and hence, a T1-weighted image acquired using a low-field scanner might look different from one acquired using a high-field scanner. One parameter to be careful with, though, is the image size. All input images and output images in the training set are of size 64×64 and 128×128, respectively. Because of the purely convolutional nature of the network, it is possible to use images of a different size as input. The network should be able to accommodate small deviations in size. However, it is unlikely that it would generalize to images that deviate significantly in size from the images in the training set.

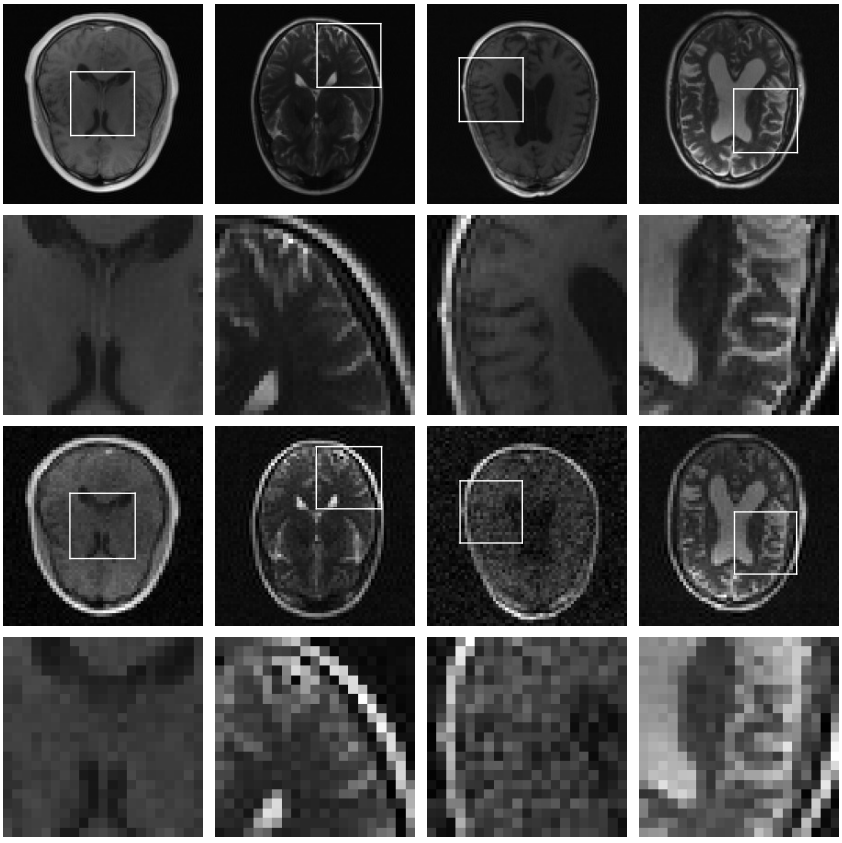


Figure 7.3: Examples of HR-LR image pairs in the training set. The first row contains four different HR images, with the white squares denoting patches whose zoomed-in versions are shown in the second row. In the third row, the corresponding (noisy) LR images are shown, with the fourth row containing LR versions of the patches in the second row.

The images in the database have different sizes. As we are interested in HR images of 128×128 pixels, all images were resized to 128×128 pixels. This was done by using an FFT to convert the images to k-space data, selecting the central part of k-space and subsequently applying an inverse Fast Fourier Transform (FFT), as in Eq. (7.1). After that, we downsample these HR images to LR images of 64×64 pixels, by, again, using Eq. (7.1), i.e., we use an FFT to convert the image to k-space, select the central part of k-space (of size 64×64) and apply an inverse FFT to obtain an LR image. To obtain noisy LR images, we add complex Gaussian noise in k-space, with the noise level varying from image to image. We used a range of noise levels consistent with the low-field MR images we have seen in practice. This step is necessary to make the convolutional neural network generalize to images acquired using a low-field MRI scanner, which, due to the weaker magnetic field, yields signals with a relatively low SNR [21]. In this way, 29,059 and 17,292 image pairs were obtained from the training and validation sets that are provided in the dataset, respectively. We assigned 10,000 of the 17,292 image pairs in the validation set to our own validation set, and the other 7,292 to our test set. Some examples of image pairs present in the training set are shown in Figure 7.3. We note that the data was split at the patient level, and hence, no data leakage occurred.

Since SRDenseNet is a purely convolutional neural network, it is possible to train on patches instead of complete images, which requires less memory during training. Furthermore, using patches allows us to generate more data. Therefore, we used the HR-LR image pairs to create 190,000 pairs of patches to train the network on, and 10,000 pairs of patches for validation, the HR and their corresponding LR patches having a size of 32×32 pixels and 16×16 pixels, respectively.

The convolutional neural network was implemented in TensorFlow [184]. The Adam optimizer [161] with a learning rate of 10^{-3} was used to minimize the mean-squared error loss between the network output and the model HR image patches. Additionally, we investigated two different loss functions: ℓ_1-loss and HFEN (High-Frequency Error Norm) loss [185]. However, after visual inspection of the resulting images, we found that the mean-squared error loss outperformed the others. We used a batch size of 20 and a total number of epochs of 74 because this corresponded to the smallest value of the validation loss. The training was carried out on a Titan X Geforce GPU (12GB) and took about 5 hours.

LOW-FIELD MR IMAGE ACQUISITION

Two three-dimensional in vivo scans of the brains of two healthy volunteers were acquired using the low-field MRI scanner described in [20, 37]. We will use different (2D) slices of the resulting 3D images as our network input. Both experiments were carried out using a turbo spin echo sequence. For the first experiment, the following parameters were used: FoV (field of view) $224 \times 224 \times 175$ mm^3, $128 \times 128 \times 50$ voxels, T_R / T_E (repetition time/echo time) = 500 ms/20 ms, echo train length 4, acquisition bandwidth 20 kHz, no signal averaging, cylindrical k-space coverage. The second experiment was carried out with a different set of parameters: FoV $180 \times 180 \times 240$ mm^3, $160 \times 120 \times 60$ voxels, T_R / T_E = 400 ms/20 ms, echo train length 5, acquisition bandwidth 20 kHz, no signal averaging.

7.3. RESULTS

To get some idea of what image quality we can expect when we apply SRDenseNet to noisy LR images, we first focus on the output of the network when we present it with high-field MR images that were artificially down-sampled and contaminated with noise. To this end, we use images from the test set (so the network has not seen these images before). In the first column of Figure 7.4, HR images (of 128x128 pixels) of three different brains are shown. The corresponding LR images (of 64x64 pixels), which we obtained by eliminating the high-frequency components in k-space, are shown in the second column. By zero-padding the k-space data corresponding to the LR image to a size of 128x128, the images in the third column were obtained. Additionally, the LR images were fed into the trained convolutional neural network, resulting in the images in the fourth column. We observe that using a convolutional neural network, instead of a simple zero-padding operation, can improve the quality of SR images. The images produced by the network are sharp and most of the structures that are present in the HR images are recovered. The peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) values of the SR images are shown in Table 7.1.

Table 7.1: PSNR and SSIM values of the reconstructions obtained by zero-padding the k-space data and by applying SRDenseNet to the three different MR brain images shown in Figure 7.4.

	Zero-padding (PSNR/SSIM)	SRDenseNet (PSNR/SSIM)
First brain	35.39/0.9506	**41.79/0.9875**
Second brain	37.14/0.9706	**45.47/0.9953**
Third brain	36.54/0.9265	**42.69/0.9876**

Three slices, of 128x128 pixels, of the first 3D low-field image acquired are shown in the first column of Figure 7.5. (We note that, since cylindrical k-space coverage was employed, the HR image could have been of a slightly better quality. However, we believe it suffices for our purposes.) The second column shows the corresponding 64x64 pixel LR images, which were obtained after eliminating the high-frequency components from k-space and adding noise. These images are used as the network input. The third and fourth column show the images obtained after zero-padding the k-space data (corresponding to the LR image) and the results obtained using SRDenseNet. We observe that for these low-field LR images, a convolutional neural network can help improve image quality, compared to a simple zero-padding result. The edges are sharp and high-frequency components seem to be recovered to a large extent.

Figure 7.6 shows the 128x128 pixel HR images reconstructed using the data acquired during the second experiment (note that these HR images were obtained from the original 160x120 slices by eliminating the outermost pixels in one direction and zero-padding the image in the other), the corresponding 64x64 pixel LR images which are used as the input to the network, the image obtained after carrying out zero-padding in k-space and the super-resolution images generated by the network. Here, too, we see that the network yields images that are sharp and contain significantly more details than the input LR images.

7.4. DISCUSSION AND CONCLUSION

Scan times are long in MRI, and it is desirable to decrease scan duration as much as possible, while maintaining sufficient image quality. Super-resolution image reconstruction aims to increase the resolution of one or several LR images. In this work, we used a convolutional neural network of the SRDenseNet architecture to carry out single-image super-resolution for brain images acquired using a low-field MRI scanner. We pre-processed images from a publicly available dataset to obtain pairs of noisy LR images (which were meant to emulate low-field MR images as much as possible) and noise-free HR images, on which we subsequently trained the network. The network was shown to yield sharp super-resolution images upon being presented with noisy low-resolution images acquired using a low-field MRI scanner. We note that, by training on noisy LR images and their noise-free HR counterparts, the network has strong denoising capabilities as well. This is especially relevant for low-field MRI, compared to high-field MRI, as one of the key differences between the two is the much lower SNR we have to deal with in the former. Additionally, this approach is fast: it can be carried out in only a few seconds.

The low-field HR images that are available are somewhat contaminated by noise, which makes them different from objective ground truth images. Therefore, gauging the quality of our SR reconstructions is challenging. However, based on visual inspection, the convolutional neural network seems to capture most of the structures and details in the SR image that are present in the HR image.These results indicate that it might be possible to decrease the scan duration by a factor of 4 and still reconstruct images of sufficient quality, using a deep learning-based method.

We do note that it is of the utmost importance to be cautious when it comes to the output of neural networks. While they are very powerful, it is well known that convolutional neural network are not infallible: at times they can produce artifacts or they can fail to produce significant details. We will apply this network to more low-field MR brain images, to gauge how often such errors occur. Additionally, it is vital that radiologists are included in the evaluation of the performance of our approach (or any deep learning-based medical imaging approach, for that matter).

While a convolutional neural network exists that is closely related to a 3D version of SRDenseNet [81], we decided to first focus on super-resolution for 2D images, as the training phase requires less computation power and less memory. However, in the future, we will make extending this approach to 3D a priority. This is especially interesting for low-field MRI as 3D acquisitions have the advantage of a higher SNR.

We trained the network on MR brain images which were acquired using high-field MRI scanners. While using low-field data might have been preferable to prevent potential modeling errors, we followed this approach because a large amount of data is required to train the network and acquiring enough low-field MR images to train the network from scratch is not feasible. By adding noise to the (reduced) k-space data corresponding to each of the images in the database, we attempted to simulate low-field MRI data. When a sufficient number of low-field MR images does become available, we could use transfer learning to tailor the network even more towards low-field MR images. In transfer learning, knowledge about one task is transferred to a related task [186]. In deep learning, it can be used when the amount of data available is insufficient to train a network to carry out the task at hand, but a large dataset or a pre-trained network is

available for a similar task, see for example Dar et al. [187] In our case, we could acquire a small dataset of low-field MR images (i.e., a few dozen) and apply transfer learning to our trained network, using this small dataset.

The potential of deep learning-based techniques to address imaging bottlenecks in the field of low-field MRI was demonstrated by Koonjoo et al. [188] They use an end-to-end deep convolutional neural network approach to boost SNR in low-field MR images acquired from highly noise-corrupted data. We note that in their paper, they describe using high-field MR data and images to train their neural network as well.

In conclusion, the main contribution of this paper is that our results show that deep learning-based methods have the potential to tackle another problem in the field of low-field MR imaging, namely how to increase the resolution of noisy LR images. We believe that, using these techniques, it is possible to bring low-field MRI technology a step closer to being used in clinical practice.

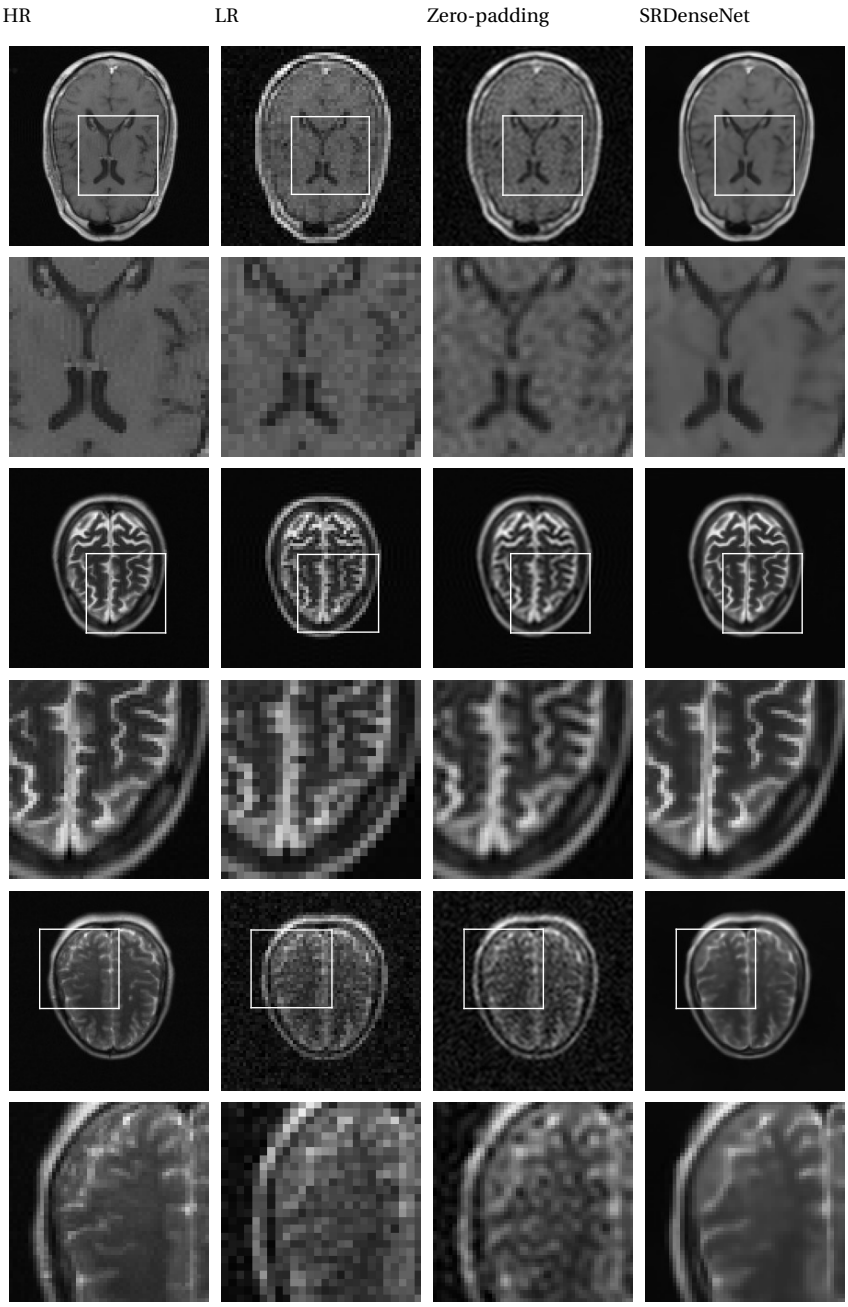


Figure 7.4: In the first column, we have three different HR images (which are high-field MR images from the database) from our test set. These HR images are our reference images, i.e., in an ideal scenario, the network would generate these exact images. The first, third and fifth rows show the full images, in the second, fourth and sixth rows we find zoomed-in versions of patches of the images in the first, third and fifth rows. The second column shows the noisy LR images corresponding to the HR images in the first column. These images are fed into the trained convolutional neural network. The third column shows the SR images obtained by zero-padding k-space, and in the fourth column, we see SR images obtained by applying SRDenseNet.

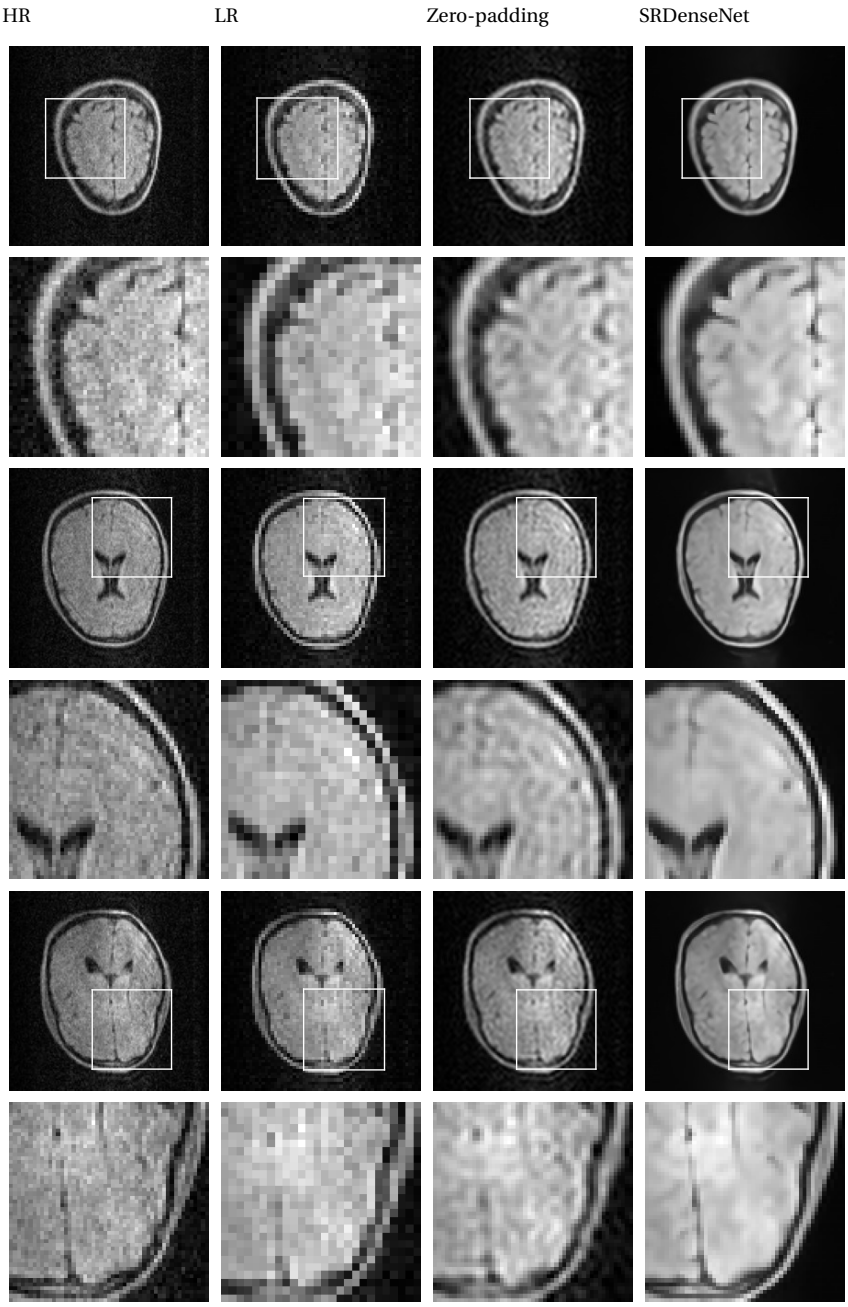


Figure 7.5: In the first column, we have three different HR slices of a brain image acquired using a low-field MRI scanner. These HR images are our reference images. The first, third and fifth rows show the full images, in the second, fourth and sixth rows we find zoomed-in versions of patches of the images in the first, third and fifth rows. The second column shows the LR images corresponding to the HR images in the first column. These images are fed into the trained convolutional neural network. The third column shows the SR images obtained by zero-padding k-space, and in the fourth column, we see SR images obtained by applying SRDenseNet.

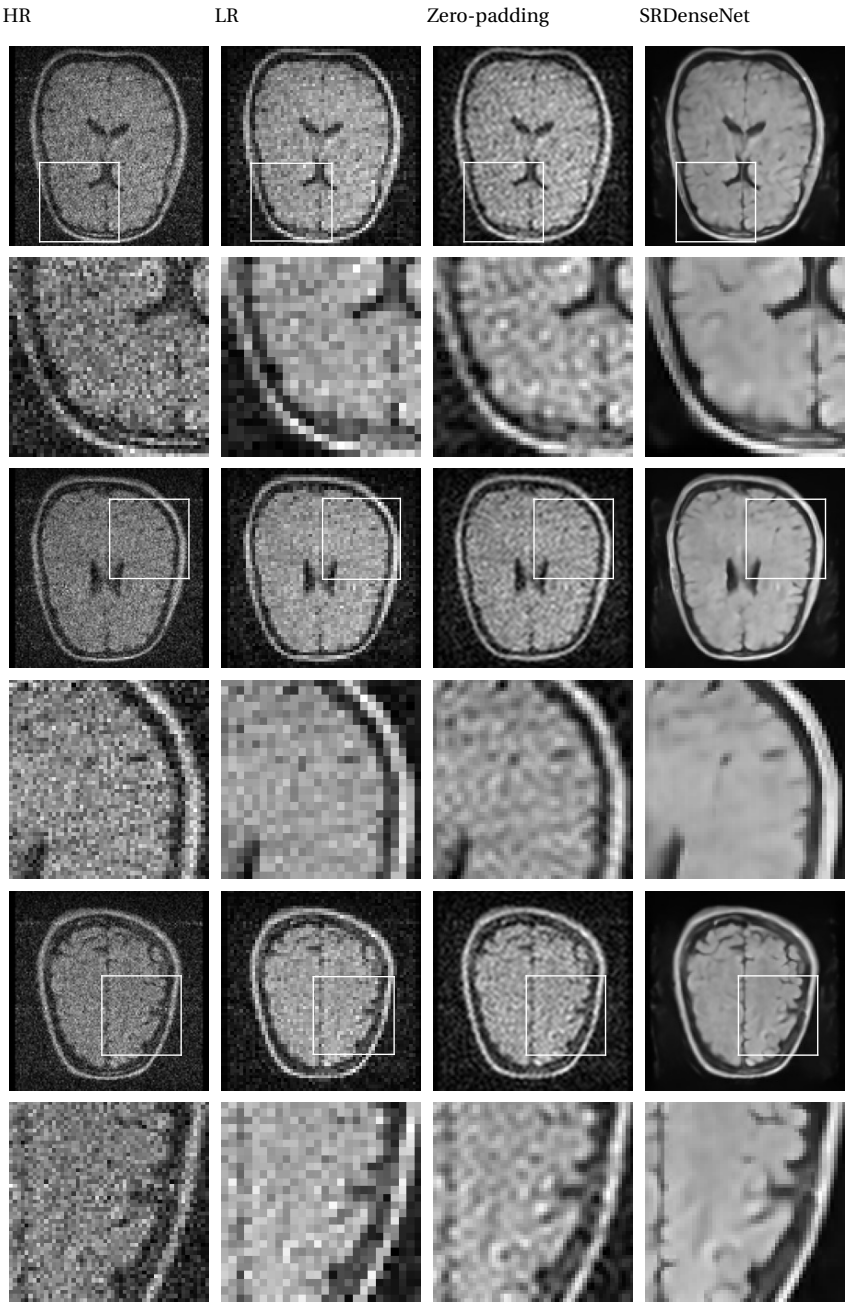


Figure 7.6: In the first column, we have three different HR slices of a different brain image acquired using a low-field MRI scanner. These HR images are our reference images. The first, third and fifth rows show the full images, in the second, fourth and sixth rows we find zoomed-in versions of patches of the images in the first, third and fifth rows. The second column shows the LR images corresponding to the HR images in the first column. These images are fed into the trained convolutional neural network. The third column shows the SR images obtained by zero-padding k-space, and in the fourth column, we see SR images obtained by applying SRDenseNet.

8

CONCLUSION

Hydrocephalus is a debilitating condition that is highly prevalent in infants in sub-Saharan Africa. The diagnostic tool of choice is an MRI scanner. However, access to MRI technology is limited in low- and middle-income countries. The research described in this dissertation was carried out as part of an interdisciplinary project that aims to develop an inexpensive, portable MRI scanner to be introduced in Uganda. This dissertation is on image reconstruction in low-field MRI, which means that we focused on the problem of creating an MR image, given a signal acquired using a low-field MRI scanner.

Within the scope of this project, several low-field MRI scanners were developed. Initially, we considered a scanner with an inhomogeneous static magnetic background field. If a sufficiently accurate map of this magnetic field is available, the inhomogeneities can be used for spatial encoding. We used a signal model that incorporates such a map to describe the relationship between signal and image. However, due to the nonbijective nature of the magnetic field, a single measurement does not yield a sufficient amount of information for a reconstruction. This problem can be remedied by acquiring several signals, after each of which the object within the scanner is either translated or rotated, and subsequently combining all information in one system of equations. In Chapter 3, we presented some preliminary results that were acquired using such a scanner.

The most promising low-field scanner developed in this project is the prototype developed at the LUMC. Its main component is a Halbach cylinder with thousands of neodymium magnets that, combined, generate a relatively homogeneous background field. Furthermore, it is equipped with three gradient coils that generate linear (or almost linear) magnetic fields that allow for spatial encoding in three directions. Therefore, the relationship between signal and image can approximately be described by a Fourier transform, as in high-field MRI.

One of the main problems in low-field MRI is the low signal-to-noise ratio. One of the ways to counteract the effects this has on the image is by posing the image reconstruction problem as a minimization problem with a data fidelity term and a regularization term, the latter of which allows us to include prior knowledge we have about the solution. In case the original problem is ill-posed, as is true for the scenario where we use

the inhomogeneities in the field for spatial encoding, adding a regularization term can make the minimization problem more well-posed as well. In Chapter 2, we introduced a variant of the conjugate gradient algorithm that can be used to solve this minimization problem efficiently for a relevant class of regularization terms. Usually, regularization is carried out in an additive manner, where an artificial regularization parameter is introduced to balance the data fidelity term and the regularization term. However, finding the right value for this parameter is usually done through extensive numerical experimentation, which can be undesirable in practice. Chapter 4 described multiplicative regularization for low-field MR imaging. By multiplying the data fidelity term by a regularizing functional, we eliminate the need for an artificial regularization parameter. We showed that this approach yields images of a quality comparable to images obtained by solving an additively regularized minimization problem, but without the need to tune an artificial regularization parameter.

Low-field MR images can end up seriously distorted due to inhomogeneities in the static magnetic background field and nonlinearities in the magnetic fields generated by the gradient coils. In Chapter 6, we trained a convolutional neural network on distorted images and their undistorted counterparts to be able to correct for distortions in low-field MR images. We applied the network to different images acquired using the LUMC low-field scanner and observed that in two out of three cases, the network yielded an image in which the distortions had been eliminated to a large extent.

In MRI, scan times are long. In contemporary MRI research, a lot of emphasis is placed on speeding up the acquisition process. For us, this is even more relevant, as the low-field scanner is tailored towards infants, for whom lying still is arguably more challenging than adults. One way to reduce the scan time, is to only acquire enough data to partially fill k-space. Chapter 5 focused on this kind of scenario. We showed that, by incorporating support information into the problem in a straightforward manner and solving the resulting system of equations using a conjugate gradient algorithm, the image quality improves. We compared our results to images acquired using a compressed sensing approach and observed that they were of a similar quality. However, our method is significantly more straightforward and does not require the tuning of parameters. Another way to reduce the scan time is by lowering the resolution of the target image. Then, one can attempt to increase the resolution as a post-processing step. In Chapter 7, we did that by training a convolutional neural network on noisy low-resolution images and their high-resolution counterparts. The resulting network was able to generate super-resolution images of a promising nature when presented with low-resolution images acquired using the LUMC low-field scanner.

8.1. OUTLOOK

Our goal is to introduce a low-field MRI scanner in Uganda, that can be used to aid in the diagnosis and treatment of hydrocephalus. Over the last four years, significant progress has been made towards attaining this goal.

Our contribution to this scanner was in the form of algorithms to be used in different scenarios. Many of them can easily be used in practice. We tested all of them thoroughly, but more research needs to be carried out when it comes to the learning-based approaches. Deep learning is very powerful and the results look promising, especially

when it is used to carry out super-resolution, but due to the lack of available data as of yet, making any definitive claims about these methods for this application is difficult. Furthermore, we trained the networks on either simulated or high-field MRI data. Acquiring appropriate low-field datasets could help tailor the networks to this particular application. We do stress that care must be taken when using images acquired using deep learning-based methods in clinical practice. While the benefits of deep learning are undeniable, we cannot exclude the possibility of a misleading output image every now and then. Therefore, we would advise to always present the original image alongside the network output, and solely use the latter as an addition to the former.

Within this project, several prototypes have been produced, with the LUMC scanner the most likely candidate to be used and/or replicated successfully in Uganda. In the near future, this scanner will be transported to Uganda to test it in the setting it was designed for. Later, clinical trials will be carried out to assess its performance versus commonly used CT scanners. With this research, we hope to have contributed towards making low-field MRI technology more accessible to all people around the world.

REFERENCES

[1] G. Dechambenoit, *Access to health care in sub-Saharan Africa,* Surgical neurology international **7** (2016).

[2] S. Geethanath and J. T. Vaughan Jr, *Accessible magnetic resonance imaging: A review,* Journal of Magnetic Resonance Imaging **49**, e65 (2019).

[3] L. Glover, *Why does an MRI cost so darn much,* Money (2014).

[4] T. O'Reilly and A. Webb, *Deconstructing and reconstructing MRI hardware,* Journal of Magnetic Resonance **306**, 134 (2019).

[5] L. Winter, H. Haopeng, A. Barghoorn, W. Hoffmann, S. Hetzer, S. Winkler, *et al.*, *Open source imaging initiative,* in *Proceedings of the International Society for Magnetic Resonance in Medicine*, Vol. 3638 (2016).

[6] R. T. Muir, S. Wang, and B. C. Warf, *Global surgery for pediatric hydrocephalus in the developing world: a review of the history, challenges, and future directions,* Neurosurgical focus **41**, E11 (2016).

[7] K. T. Kahle, A. V. Kulkarni, D. D. Limbrick Jr, and B. C. Warf, *Hydrocephalus in children,* The Lancet **387**, 788 (2016).

[8] M. C. Dewan, A. Rattani, R. Mekary, L. J. Glancz, I. Yunusa, R. E. Baticulon, G. Fieggen, J. C. Wellons, K. B. Park, and B. C. Warf, *Global hydrocephalus epidemiology and incidence: systematic review and meta-analysis,* Journal of neurosurgery **130**, 1065 (2018).

[9] J. N. Paulson, B. L. Williams, C. Hehnly, N. Mishra, S. A. Sinnar, L. Zhang, P. Ssentongo, E. Mbabazi-Kabachelor, D. S. Wijetunge, B. von Bredow, *et al.*, *Paenibacillus infection with frequent viral coinfection contributes to postinfectious hydrocephalus in Ugandan infants,* Science Translational Medicine **12** (2020).

[10] B. C. Warf, *Hydrocephalus in Uganda: the predominance of infectious origin and primary management with endoscopic third ventriculostomy,* Journal of Neurosurgery: Pediatrics **102**, 1 (2005).

[11] B. C. Warf, *Educate one to save a few. Educate a few to save many,* World neurosurgery **79**, S15 (2013).

[12] D. J. Brenner and E. J. Hall, *Computed tomography—an increasing source of radiation exposure,* New England Journal of Medicine **357**, 2277 (2007).

[13] J. Obungoloch, J. R. Harper, S. Consevage, I. M. Savukov, T. Neuberger, S. Tadigadapa, and S. J. Schiff, *Design of a sustainable prepolarizing magnetic resonance imaging system for infant hydrocephalus,* Magnetic Resonance Materials in Physics, Biology and Medicine **31**, 665 (2018).

[14] W. D. Rooney, G. Johnson, X. Li, E. R. Cohen, S.-G. Kim, K. Ugurbil, and C. S. Springer Jr, *Magnetic field and tissue dependencies of human brain longitudinal 1H2O relaxation in vivo,* Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **57**, 308 (2007).

[15] S. Lother, S. J. Schiff, T. Neuberger, P. M. Jakob, and F. Fidler, *Design of a mobile, homogeneous, and efficient electromagnet with a large field of view for neonatal low-field MRI,* Magnetic Resonance Materials in Physics, Biology and Medicine **29**, 691 (2016).

[16] R. W. Brown, Y.-C. N. Cheng, E. M. Haacke, M. R. Thompson, and R. Venkatesan, *Magnetic resonance imaging: physical principles and sequence design* (John Wiley & Sons, 2014).

[17] J. E. Williams, *Superconducting Magnets for MRI,* IEEE Transactions on Nuclear Science **31**, 994 (1984).

[18] L. L. Wald, P. C. McDaniel, T. Witzel, J. P. Stockmann, and C. Z. Cooley, *Low-cost and portable MRI,* Journal of Magnetic Resonance Imaging **52**, 686 (2020).

[19] Z.-P. Liang and P. C. Lauterbur, *Principles of magnetic resonance imaging: a signal processing perspective* (SPIE Optical Engineering Press, 2000).

[20] T. O'Reilly, W. Teeuwisse, and A. Webb, *Three-dimensional MRI in a homogenous 27cm diameter bore Halbach array magnet,* Journal of Magnetic Resonance **307**, 106578 (2019).

[21] M. Sarracanie and N. Salameh, *Low-field MRI: how low can we go? A fresh view on an old debate,* Front Phys **8**, 172 (2020).

[22] R. E. Sepponen, J. T. Sipponen, and A. Sivula, *Low field (0.02 T) nuclear magnetic resonance imaging of the brain.* Journal of computer assisted tomography **9**, 237 (1985).

[23] G. do Nascimento, M. Engelsberg, and R. De Souza, *Digital NMR imaging system for ultralow magnetic fields,* Measurement Science and Technology **3**, 370 (1992).

[24] A. Macovski and S. Conolly, *Novel approaches to low-cost MRI,* Magnetic resonance in medicine **30**, 221 (1993).

[25] R. McDermott, S. Lee, B. Ten Haken, A. H. Trabesinger, A. Pines, and J. Clarke, *Microtesla MRI with a superconducting quantum interference device,* Proceedings of the National Academy of Sciences **101**, 7857 (2004).

[26] I. Savukov and T. Karaulanov, *Magnetic-resonance imaging of the human brain with an atomic magnetometer,* Applied physics letters **103**, 043703 (2013).

[27] V. S. Zotev, A. N. Matlachov, P. L. Volegov, H. J. Sandin, M. A. Espy, J. C. Mosher, A. V. Urbaitis, S. G. Newman, and R. H. Kraus, *Multi-channel SQUID system for MEG and ultra-low-field MRI,* IEEE transactions on applied superconductivity **17**, 839 (2007).

[28] M. Sarracanie, C. D. LaPierre, N. Salameh, D. E. Waddington, T. Witzel, and M. S. Rosen, *Low-cost high-performance MRI,* Scientific reports **5**, 1 (2015).

[29] K. Halbach, *Design of permanent multipole magnets with oriented rare earth cobalt material,* Nuclear instruments and methods **169**, 1 (1980).

[30] C. Z. Cooley, J. P. Stockmann, B. D. Armstrong, M. Sarracanie, M. H. Lev, M. S. Rosen, and L. L. Wald, *Two-dimensional imaging in a lightweight portable MRI scanner without gradient coils,* Magnetic resonance in medicine **73**, 872 (2015).

[31] C. Z. Cooley, M. W. Haskell, S. F. Cauley, C. Sappo, C. D. Lapierre, C. G. Ha, J. P. Stockmann, and L. L. Wald, *Design of sparse Halbach magnet arrays for portable MRI using a genetic algorithm,* IEEE transactions on magnetics **54**, 1 (2017).

[32] C. Z. Cooley, P. C. McDaniel, J. P. Stockmann, S. A. Srinivas, S. F. Cauley, M. Śliwiak, C. R. Sappo, C. F. Vaughn, B. Guerin, M. S. Rosen, *et al.*, *A portable scanner for magnetic resonance imaging of the brain,* Nature Biomedical Engineering , 1 (2020).

[33] Z. H. Ren, W. C. Mu, and S. Y. Huang, *Design and optimization of a ring-pair permanent magnet array for head imaging in a low-field portable MRI system,* IEEE Transactions on Magnetics **55**, 1 (2018).

[34] Z. H. Ren, J. Gong, and S. Y. Huang, *An irregular-shaped inward-outward ring-pair magnet array with a monotonic field gradient for 2D head imaging in low-field portable MRI,* IEEE Access **7**, 48715 (2019).

[35] M. L. de Leeuw den Bouter, D. B. Gecmen, A. E. Meijer, D. H. de Gans, L. N. M. Middelplaats, R. F. Remis, and M. B. van Gijzen, *Description of a Low-field MRI Scanner Based on Permanent Magnets,* CVCS 2020 Colour and Visual Computing Symposium 2020 Proceedings of the 10th Colour and Visual Computing Symposium 2020 (CVCS 2020) Gjøvik, Norway, and Virtual, September 16-17, 2020 **2688** (2020).

[36] B. De Vos, P. Fuchs, T. O'Reilly, A. Webb, and R. Remis, *Gradient Coil Design and Realization for a Halbach-Based MRI System,* IEEE Transactions on Magnetics **56**, 1 (2020).

[37] T. O'Reilly, W. M. Teeuwisse, D. de Gans, K. Koolstra, and A. G. Webb, *In vivo 3D brain and extremity MRI at 50 mT using a permanent magnet Halbach array,* Magnetic resonance in medicine **85**, 495 (2021).

[38] G. Eidmann, R. Savelsberg, P. Blümler, and B. Blümich, *The NMR MOUSE, a mobile universal surface explorer,* Journal of Magnetic Resonance, Series A **122**, 104 (1996).

[39] B. Blümich, V. Anferov, S. Anferova, M. Klein, R. Fechete, M. Adams, and F. Casanova, *Simple NMR-mouse with a bar magnet,* Concepts in Magnetic Resonance: An Educational Journal **15**, 255 (2002).

[40] Z. He, W. He, J. Wu, and Z. Xu, *The novel design of a single-sided MRI probe for assessing burn depth,* Sensors **17**, 526 (2017).

[41] M. Greer, C. Chen, and S. Mandal, *An easily reproducible, hand-held, single-sided, MRI sensor,* Journal of Magnetic Resonance **308**, 106591 (2019).

[42] P. McDonald, A. Akhmerov, L. Backhouse, and S. Pitts, *Magnetic resonance profiling of human skin in vivo using GARField magnets,* Journal of pharmaceutical sciences **94**, 1850 (2005).

[43] B. Manz, A. Coy, R. Dykstra, C. Eccles, M. Hunter, B. Parkinson, and P. Callaghan, *A mobile one-sided NMR sensor with a homogeneous magnetic field: The NMR-MOLE,* Journal of Magnetic Resonance **183**, 25 (2006).

[44] J. A. Fessler, *Model-based image reconstruction for MRI,* IEEE signal processing magazine **27**, 81 (2010).

[45] P. C. Hansen, *Discrete inverse problems: insight and algorithms* (SIAM, 2010).

[46] A. S. Lundervold and A. Lundervold, *An overview of deep learning in medical imaging focusing on MRI,* Zeitschrift für Medizinische Physik **29**, 102 (2019).

[47] D. Liang, J. Cheng, Z. Ke, and L. Ying, *Deep magnetic resonance image reconstruction: Inverse problems meet neural networks,* IEEE Signal Processing Magazine **37**, 141 (2020).

[48] J. Zhang and B. Ghanem, *ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing,* in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2018) pp. 1828–1837.

[49] Y. Yang, J. Sun, H. Li, and Z. Xu, *ADMM-CSNet: A deep learning approach for image compressive sensing,* IEEE transactions on pattern analysis and machine intelligence **42**, 521 (2018).

[50] J. Adler and O. Öktem, *Learned primal-dual reconstruction,* IEEE transactions on medical imaging **37**, 1322 (2018).

[51] K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock, and F. Knoll, *Learning a variational network for reconstruction of accelerated MRI data,* Magnetic resonance in medicine **79**, 3055 (2018).

[52] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, *Image reconstruction by domain-transform manifold learning,* Nature **555**, 487 (2018).

[53] S. C. Park, M. K. Park, and M. G. Kang, *Super-resolution image reconstruction: a technical overview,* IEEE signal processing magazine **20**, 21 (2003).

[54] E. Van Reeth, I. W. Tham, C. H. Tan, and C. L. Poh, *Super-resolution in magnetic resonance imaging: a review,* Concepts in Magnetic Resonance Part A **40**, 306 (2012).

[55] Å. Björck, *Numerical methods for least squares problems* (SIAM, 1996).

[56] D. Orban and M. Arioli, *Iterative Solution of Symmetric Quasi-Definite Linear Systems*, Vol. 3 (SIAM, 2017).

[57] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, Vol. 49 (NBS, 1952).

[58] Y. Saad, *Iterative methods for sparse linear systems* (SIAM, 2003).

[59] M. Hanke, *Conjugate gradient type methods for ill-posed problems*, Vol. 327 (CRC Press, 1995).

[60] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, Vol. 375 (Springer Science & Business Media, 1996).

[61] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst, *Templates for the solution of linear systems: building blocks for iterative methods* (SIAM, 1994).

[62] M. A. Saunders, *Solution of sparse rectangular systems using LSQR and CRAIG,* BIT Numerical Mathematics **35**, 588 (1995).

[63] E. J. Craig, *The N-step iteration procedures,* Studies in Applied Mathematics **34**, 64 (1955).

[64] J. King, *A minimal error conjugate gradient method for ill-posed problems,* Journal of optimization theory and applications **60**, 297 (1989).

[65] A. S. Nemirovskii, *The regularizing properties of the adjoint gradient method in ill-posed problems,* USSR Computational Mathematics and Mathematical Physics **26**, 7 (1986).

[66] M. Hanke, *The minimal error conjugate gradient method is a regularization method,* Proceedings of the American Mathematical Society **123**, 3487 (1995).

[67] J. A. Fessler, S. Lee, V. T. Olafsson, H. R. Shi, and D. C. Noll, *Toeplitz-based iterative image reconstruction for MRI with correction for magnetic field inhomogeneity,* IEEE Transactions on Signal Processing **53**, 3393 (2005).

[68] K. P. Pruessmann, M. Weiger, P. Börnert, and P. Boesiger, *Advances in sensitivity encoding with arbitrary k-space trajectories,* Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **46**, 638 (2001).

[69] B. P. Sutton, D. C. Noll, and J. A. Fessler, *Fast, iterative image reconstruction for MRI in the presence of field inhomogeneities,* IEEE transactions on medical imaging **22**, 178 (2003).

[70] K. P. Pruessmann, *Encoding and reconstruction in parallel MRI,* NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In vivo **19**, 288 (2006).

[71] S. Ramani and J. A. Fessler, *Parallel MR image reconstruction using augmented lagrangian methods,* IEEE Transactions on Medical Imaging **30**, 694 (2010).

[72] J. C. Ye, S. Tak, Y. Han, and H. W. Park, *Projection reconstruction MR imaging using FOCUSS,* Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **57**, 764 (2007).

[73] A. N. Tikhonov, *On the solution of ill-posed problems and the method of regularization,* in *Doklady Akademii Nauk*, Vol. 151 (Russian Academy of Sciences, 1963) pp. 501–504.

[74] H. Weyl, *Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung),* Mathematische Annalen **71**, 441 (1912).

[75] A. Buccini and L. Reichel, *An $\ell^2 - \ell^q$ Regularization Method for Large Discrete Ill-Posed Problems,* Journal of Scientific Computing **78**, 1526 (2019).

[76] R. H. Chan and H.-X. Liang, *Half-Quadratic Algorithm for $\ell_p - \ell_q$ Problems with Applications to TV-ℓ_1 Image Restoration and Compressive Sensing,* in *Efficient algorithms for global optimization methods in computer vision* (Springer, 2014) pp. 78–103.

[77] G. Huang, A. Lanza, S. Morigi, L. Reichel, and F. Sgallari, *Majorization–minimization generalized Krylov subspace methods for $\ell_p - \ell_q$ optimization applied to image restoration,* BIT Numerical Mathematics **57**, 351 (2017).

[78] A. Lanza, S. Morigi, L. Reichel, and F. Sgallari, *A Generalized Krylov Subspace Method for $\ell_p - \ell_q$ Minimization,* SIAM Journal on Scientific Computing **37**, S30 (2015).

[79] T. Goldstein and S. Osher, *The split Bregman method for L1-regularized problems,* SIAM journal on imaging sciences **2**, 323 (2009).

[80] M. Elad, *Sparse and redundant representations: from theory to applications in signal and image processing* (Springer Science & Business Media, 2010).

[81] C.-N. Chen and D. I. Hoult, *Biomedical magnetic resonance technology* (Hilger, 1989).

[82] R. Tibshirani, *Regression shrinkage and selection via the LASSO,* Journal of the Royal Statistical Society. Series B (Methodological) , 267 (1996).

[83] J. Hennig, A. M. Welz, G. Schultz, J. Korvink, Z. Liu, O. Speck, and M. Zaitsev, *Parallel imaging in non-bijective, curvilinear magnetic field gradients: a concept study,* Magnetic Resonance Materials in Physics, Biology and Medicine **21**, 5 (2008).

[84] M. Haas, P. Ullmann, J. Schneider, H. Post, W. Ruhm, J. Hennig, and M. Zaitsev, *PexLoc—Parallel excitation using local encoding magnetic fields with nonlinear and nonbijective spatial profiles,* Magnetic resonance in medicine **70**, 1220 (2013).

[85] G. Schultz, *Magnetic resonance imaging with nonlinear gradient fields: signal encoding and image reconstruction* (Springer Science & Business Media, 2013).

[86] G. Schultz, D. Gallichan, H. Weber, W. R. Witschey, M. Honal, J. Hennig, and M. Zaitsev, *Image reconstruction in k-space from MR data encoded with ambiguous gradient fields,* Magnetic resonance in medicine **73**, 857 (2015).

[87] G. Schultz, P. Ullmann, H. Lehr, A. M. Welz, J. Hennig, and M. Zaitsev, *Reconstruction of MRI data encoded with arbitrarily shaped, curvilinear, nonbijective magnetic fields,* Magnetic resonance in medicine **64**, 1390 (2010).

[88] F. Casanova, J. Perlo, and B. Blümich, *Single-sided NMR,* in *Single-Sided NMR* (Springer, 2011) pp. 1–10.

[89] B. Blümich, J. Perlo, and F. Casanova, *Mobile single-sided NMR,* Progress in Nuclear Magnetic Resonance Spectroscopy **52**, 197 (2008).

[90] M. Prammer, *NMR in well logging and hydrocarbon exploration,* Applied Magnetic Resonance **25**, 637 (2004).

[91] O. V. Petrov, J. Hay, I. V. Mastikhin, and B. J. Balcom, *Fat and moisture content determination with unilateral NMR,* Food research international **41**, 758 (2008).

[92] E. Veliyulin, I. V. Mastikhin, A. E. Marble, and B. J. Balcom, *Rapid determination of the fat content in packaged dairy products by unilateral NMR,* Journal of the Science of Food and Agriculture **88**, 2563 (2008).

[93] F. Casanova and B. Blümich, *Two-dimensional imaging with a single-sided NMR probe,* Journal of Magnetic Resonance **163**, 38 (2003).

[94] J. Perlo, F. Casanova, and B. Blümich, *3D imaging with a single-sided sensor: an open tomograph,* Journal of Magnetic Resonance **166**, 228 (2004).

[95] P. McDonald, P. Aptaker, J. Mitchell, and M. Mulheron, *A unilateral NMR magnet for sub-structure analysis in the built environment: The Surface GARField,* Journal of Magnetic Resonance **185**, 1 (2007).

[96] A. E. Marble, I. V. Mastikhin, B. G. Colpitts, and B. J. Balcom, *A compact permanent magnet array with a remote homogeneous field,* Journal of Magnetic Resonance **186**, 100 (2007).

[97] H. Chang, D.-Y. Yeung, and Y. Xiong, *Super-resolution through neighbor embedding,* in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, Vol. 1 (IEEE, 2004) pp. I–I.

[98] Z. Xu, K. Meng, and J. Cheng, *Highly uniform single-sided portable NMR sensor and its application in assessing the aging level of silicone rubber insulators,* International Journal of Applied Electromagnetics and Mechanics **47**, 777 (2015).

[99] S. Utsuzawa and E. Fukushima, *Unilateral NMR with a barrel magnet,* Journal of Magnetic Resonance **282**, 104 (2017).

[100] A. Bashyam, C. J. Frangieh, S. Raigani, J. Sogo, R. T. Bronson, K. Uygun, H. Yeh, D. A. Ausiello, and M. J. Cima, *A portable single-sided magnetic-resonance sensor for the grading of liver steatosis and fibrosis,* Nature Biomedical Engineering **5**, 240 (2021).

[101] J. L. Paulsen, L. S. Bouchard, D. Graziani, B. Blümich, and A. Pines, *Volume-selective magnetic resonance imaging using an adjustable, single-sided, portable sensor,* Proceedings of the National Academy of Sciences **105**, 20601 (2008).

[102] S. Tewari, S. Yousefi, and A. Webb, *Deep neural-network based optimization for the design of a multi-element surface magnet for mri applications (*under review*,* Inverse Problems (2021).

[103] J. A. Fessler, *Model-based image reconstruction for MRI,* IEEE Signal Processing Magazine **27**, 81 (2010).

[104] M. L. de Leeuw den Bouter, M. B. van Gijzen, and R. F. Remis, *Conjugate gradient variants for ℓ_p-regularized image reconstruction in low-field MRI,* SN Applied Sciences **1**, 1736 (2019).

[105] C. A. Bouman, *Model based image processing,* Purdue University (2013).

[106] J. A. Fessler, *Optimization methods for magnetic resonance image reconstruction: Key models and optimization algorithms,* IEEE Signal Processing Magazine **37**, 33 (2020).

[107] A. Janke, H. Zhao, G. J. Cowin, G. J. Galloway, and D. M. Doddrell, *Use of spherical harmonic deconvolution methods to compensate for nonlinear gradient effects on MRI images,* Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **52**, 115 (2004).

[108] K. T. Block, M. Uecker, and J. Frahm, *Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint,* Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **57**, 1086 (2007).

[109] K. Koolstra, J. Van Gemert, P. Börnert, A. Webb, and R. Remis, *Accelerating compressed sensing in parallel imaging reconstructions using an efficient circulant preconditioner for cartesian trajectories,* Magnetic Resonance in Medicine **81**, 670 (2019).

[110] J. Liu, Y. Sun, X. Xu, and U. S. Kamilov, *Image restoration using total variation regularized deep image prior,* in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, 2019) pp. 7715–7719.

[111] S. Ma, W. Yin, Y. Zhang, and A. Chakraborty, *An efficient algorithm for compressed MR imaging using total variation and wavelets,* in *2008 IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2008) pp. 1–8.

[112] P. C. Hansen, J. G. Nagy, and D. P. O'Leary, *Deblurring images: matrices, spectra, and filtering* (Society for Industrial and Applied Mathematics, Philadelphia, 2006).

[113] P. M. van den Berg, A. Abubakar, and J. T. Fokkema, *Multiplicative regularization for contrast profile inversion,* Radio Sci. **38**, 23.1 (2003).

[114] P. M. van den Berg and A. Abubakar, *Contrast Source Inversion Method: State of Art,* Pr. Electromagn. Res. **34**, 189 (2001).

[115] P. M. van den Berg and A. Abubakar, *Optical Microscopy Imaging Using the Contrast Source Inversion Method,* J. Mod. Optic. **57**, 756 (2010).

[116] E. Balidemaj, C. A. van den Berg, J. Trinks, A. L. van Lier, A. J. Nederveen, L. J. Stalpers, H. Crezee, and R. F. Remis, *CSI-EPT: a contrast source inversion approach for improved MRI-based electric properties tomography,* IEEE T. Med. Imaging **34**, 1788 (2015).

[117] J. Nocedal and S. J. Wright, *Numerical optimization* (Springer, 2006).

[118] A. Abubakar, P. M. Van Den Berg, T. M. Habashy, and H. Braunisch, *A multiplicative regularization approach for deblurring problems,* IEEE Transactions on Image Processing **13**, 1524 (2004).

[119] P. M. van den Berg, A. L. van Broekhoven, and A. Abubakar, *Extended Contrast Source Inversion,* Inverse Probl. **15**, 1325 (1999).

[120] E. Gorgin, *An Analysis of Multiplicative Regularization, Dissertation* (Michigan Technological University, 2015).

[121] P. R. Haffinger, *Seismic broadband full waveform inversion by shot/receiver refocusing,* Ph.D. thesis, Delft University of Technology (2013).

[122] J. B. Rhebergen, P. M. van den Berg, and T. M. Habashy, *Iterative reconstruction of images from incomplete spectral data,* Inverse problems **13**, 829 (1997).

[123] A. Papoulis, *A new algorithm in spectral analysis and band-limited extrapolation,* IEEE Transactions on Circuits and systems **22**, 735 (1975).

[124] F. J. Simons and D. V. Wang, *Spatiospectral concentration in the Cartesian plane,* GEM-International Journal on Geomathematics **2**, 1 (2011).

[125] E. J. Candès and M. B. Wakin, *An introduction to compressive sampling,* IEEE signal processing magazine **25**, 21 (2008).

[126] J. P. Marques, F. F. Simonis, and A. G. Webb, *Low-field MRI: An MR physics perspective,* Journal of Magnetic Resonance Imaging **49**, 1528 (2019).

[127] N. Vaswani and W. Lu, *Modified-CS: Modifying compressive sensing for problems with partially known support,* IEEE Transactions on Signal Processing **58**, 4595 (2010).

[128] R. E. Carrillo, L. F. Polania, and K. E. Barner, *Iterative algorithms for compressed sensing with partially known support,* in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing* (IEEE, 2010) pp. 3654–3657.

[129] L. Jacques, *A short note on compressed sensing with partially known signal support,* Signal Processing **90**, 3308 (2010).

[130] D. Slepian and H. O. Pollak, *Prolate spheroidal wave functions, Fourier analysis and uncertainty—I,* Bell System Technical Journal **40**, 43 (1961).

[131] L. A. Shepp and B. F. Logan, *The Fourier reconstruction of a head section,* IEEE Transactions on nuclear science **21**, 21 (1974).

[132] B. A. Landman, A. J. Huang, A. Gifford, D. S. Vikram, I. A. L. Lim, J. A. Farrell, J. A. Bogovic, J. Hua, M. Chen, S. Jarso, *et al.*, *Multi-parametric neuroimaging reproducibility: a 3-T resource study,* Neuroimage **54**, 2854 (2011).

[133] W. Brouw, *Aperture synthesis,* in *Image processing techniques in astronomy* (Springer, 1975) pp. 301–307.

[134] H. Schomberg and J. Timmer, *The gridding method for image reconstruction by Fourier transformation,* IEEE transactions on medical imaging **14**, 596 (1995).

[135] I. Daubechies, M. Defrise, and C. De Mol, *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint,* Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences **57**, 1413 (2004).

[136] S. Bhave, S. G. Lingala, C. P. Johnson, V. A. Magnotta, and M. Jacob, *Accelerated whole-brain multi-parameter mapping using blind compressed sensing,* Magnetic resonance in medicine **75**, 1175 (2016).

[137] S. U. H. Dar, M. Yurt, M. Shahdloo, M. E. Ildız, B. Tınaz, and T. Çukur, *Prior-guided image reconstruction for accelerated multi-contrast MRI via generative adversarial networks,* IEEE Journal of Selected Topics in Signal Processing **14**, 1072 (2020).

[138] S. U. H. Dar, M. Özbey, A. B. Çatlı, and T. Çukur, *A transfer-learning approach for accelerated MRI using deep neural networks,* Magnetic resonance in medicine **84**, 663 (2020).

[139] B. Gözcü, R. K. Mahabadi, Y.-H. Li, E. Ilıcak, T. Cukur, J. Scarlett, and V. Cevher, *Learning-based compressive MRI,* IEEE transactions on medical imaging **37**, 1394 (2018).

[140] D. Liang, B. Liu, J. Wang, and L. Ying, *Accelerating SENSE using compressed sensing,* Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **62**, 1574 (2009).

[141] M. Lustig, D. Donoho, and J. M. Pauly, *Sparse MRI: The application of compressed sensing for rapid MR imaging,* Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **58**, 1182 (2007).

[142] L. B. Montefusco, D. Lazzaro, S. Papi, and C. Guerrini, *A fast compressed sensing approach to 3D MR image reconstruction,* IEEE transactions on medical imaging **30**, 1064 (2010).

[143] M. Shahdloo, E. Ilicak, M. Tofighi, E. U. Saritas, A. E. Çetin, and T. Çukur, *Projection onto epigraph sets for rapid self-tuning compressed sensing MRI,* IEEE transactions on medical imaging **38**, 1677 (2018).

[144] L. Sun, Y. Wu, Z. Fan, X. Ding, Y. Huang, and J. Paisley, *A deep error correction network for compressed sensing MRI,* BMC Biomedical Engineering **2**, 1 (2020).

[145] A. C. Zelinski, V. K. Goyal, and E. Adalsteinsson, *Simultaneously sparse solutions to linear inverse problems with multiple system matrices and a single observation vector,* SIAM Journal on Scientific Computing **31**, 4533 (2010).

[146] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, *SENSE: sensitivity encoding for fast MRI,* Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **42**, 952 (1999).

[147] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, *Generalized autocalibrating partially parallel acquisitions (GRAPPA),* Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **47**, 1202 (2002).

[148] M. de Leeuw den Bouter, M. van Gijzen, and R. Remis, *Low-field magnetic resonance imaging using multiplicative regularization,* Magnetic Resonance Imaging **75**, 21 (2021).

[149] A. Janke, H. Zhao, G. J. Cowin, G. J. Galloway, and D. M. Doddrell, *Use of spherical harmonic deconvolution methods to compensate for nonlinear gradient effects on MRI images,* Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **52**, 115 (2004).

[150] A. Macovski, *Volumetric NMR imaging with time-varying gradients,* Magnetic Resonance in Medicine **2**, 29 (1985).

[151] A. Matakos and J. A. Fessler, *Joint estimation of image and fieldmap in parallel MRI using single-shot acquisitions,* in *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (IEEE, 2010) pp. 984–987.

[152] Z. Akkus, A. Galimzianova, A. Hoogi, D. L. Rubin, and B. J. Erickson, *Deep learning for brain MRI segmentation: state of the art and future directions,* Journal of digital imaging **30**, 449 (2017).

[153] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo, *Deep learning for undersampled MRI reconstruction,* Physics in Medicine & Biology **63**, 135007 (2018).

[154] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, *Image reconstruction by domain-transform manifold learning,* Nature **555**, 487 (2018).

[155] C. D. Eccles, S. Crozier, M. Westphal, and D. M. Doddrell, *Temporal spherical-harmonic expansion and compensation of eddy-current fields produced by gradient pulses,* Journal of Magnetic Resonance, Series A **103**, 135 (1993).

[156] D. Hoult, *Shimming on spatially localized signals,* Journal of magnetic resonance **73**, 174 (1987).

[157] S. Langlois, M. Desvignes, J.-M. Constans, and M. Revenu, *MRI geometric distortion: a simple approach to correcting the effects of non-linear gradient fields,* Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine **9**, 821 (1999).

[158] J. P. Stockmann and L. L. Wald, *In vivo B0 field shimming methods for MRI at 7 T,* NeuroImage **168**, 71 (2018).

[159] O. Ronneberger, P. Fischer, and T. Brox, *U-net: Convolutional networks for biomedical image segmentation,* in *International Conference on Medical image computing and computer-assisted intervention* (Springer, 2015) pp. 234–241.

[160] C. Zhang and Y. Xing, *CT artifact reduction via U-net CNN,* in *Medical Imaging 2018: Image Processing*, Vol. 10574 (International Society for Optics and Photonics, 2018) p. 105741R.

[161] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization,* arXiv preprint arXiv:1412.6980 (2014).

[162] M. Jaderberg, K. Simonyan, A. Zisserman, *et al.*, *Spatial transformer networks,* in *Advances in neural information processing systems* (2015) pp. 2017–2025.

[163] N. S. Detlefsen, O. Freifeld, and S. Hauberg, *Deep diffeomorphic transformer networks,* in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2018) pp. 4403–4412.

[164] R. Keys, *Cubic convolution interpolation for digital image processing,* IEEE transactions on acoustics, speech, and signal processing **29**, 1153 (1981).

[165] C. E. Duchon, *Lanczos filtering in one and two dimensions,* Journal of Applied Meteorology and Climatology **18**, 1016 (1979).

[166] S. Baker and T. Kanade, *Limits on super-resolution and how to break them,* IEEE Transactions on Pattern Analysis and Machine Intelligence **24**, 1167 (2002).

[167] S. Dai, M. Han, W. Xu, Y. Wu, Y. Gong, and A. K. Katsaggelos, *Softcuts: a soft edge smoothness prior for color image super-resolution,* IEEE Transactions on Image Processing **18**, 969 (2009).

[168] A. Marquina and S. J. Osher, *Image super-resolution by TV-regularization and Bregman iteration,* Journal of Scientific Computing **37**, 367 (2008).

[169] F. Shi, J. Cheng, L. Wang, P.-T. Yap, and D. Shen, *LRTV: MR image super-resolution with low-rank and total variation regularizations,* IEEE transactions on medical imaging **34**, 2459 (2015).

[170] J. Sun, Z. Xu, and H.-Y. Shum, *Image super-resolution using gradient profile prior,* in *2008 IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2008) pp. 1–8.

[171] W. T. Freeman, T. R. Jones, and E. C. Pasztor, *Example-based super-resolution,* IEEE Computer graphics and Applications **22**, 56 (2002).

[172] J. Yang, J. Wright, T. S. Huang, and Y. Ma, *Image super-resolution via sparse representation,* IEEE transactions on image processing **19**, 2861 (2010).

[173] R. Zeyde, M. Elad, and M. Protter, *On single image scale-up using sparse-representations,* in *International conference on curves and surfaces* (Springer, 2010) pp. 711–730.

[174] C. Dong, C. C. Loy, K. He, and X. Tang, *Learning a deep convolutional network for image super-resolution,* in *European conference on computer vision* (Springer, 2014) pp. 184–199.

[175] T. Tong, G. Li, X. Liu, and Q. Gao, *Image super-resolution using dense skip connections,* in *Proceedings of the IEEE international conference on computer vision* (2017) pp. 4799–4807.

[176] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, *et al.*, *Photo-realistic single image super-resolution using a generative adversarial network,* in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2017) pp. 4681–4690.

[177] Z. Wang, J. Chen, and S. C. Hoi, *Deep learning for image super-resolution: A survey,* IEEE transactions on pattern analysis and machine intelligence (2020).

[178] C.-H. Pham, A. Ducournau, R. Fablet, and F. Rousseau, *Brain MRI super-resolution using deep 3D convolutional networks,* in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)* (IEEE, 2017) pp. 197–200.

[179] E. M. Masutani, N. Bahrami, and A. Hsiao, *Deep learning single-frame and multi-frame super-resolution for cardiac MRI,* Radiology **295**, 552 (2020).

[180] Y. Chen, F. Shi, A. G. Christodoulou, Y. Xie, Z. Zhou, and D. Li, *Efficient and accurate MRI super-resolution using a generative adversarial network and 3D multi-level densely connected network,* in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, 2018) pp. 91–99.

[181] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, *Densely connected convolutional networks,* in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2017) pp. 4700–4708.

[182] F. Knoll, J. Zbontar, A. Sriram, M. J. Muckley, M. Bruno, A. Defazio, M. Parente, K. J. Geras, J. Katsnelson, H. Chandarana, *et al.*, *fastMRI: A publicly available raw k-space and DICOM dataset of knee images for accelerated MR image reconstruction using machine learning,* Radiology: Artificial Intelligence **2**, e190007 (2020).

[183] J. Zbontar, F. Knoll, A. Sriram, T. Murrell, Z. Huang, M. J. Muckley, A. Defazio, R. Stern, P. Johnson, M. Bruno, *et al.*, *fastMRI: An open dataset and benchmarks for accelerated MRI,* arXiv preprint arXiv:1811.08839 (2018).

[184] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, *et al.*, *Tensorflow: A system for large-scale machine learning,* in *12th USENIX symposium on operating systems design and implementation (OSDI 16)* (2016) pp. 265–283.

[185] S. Ravishankar and Y. Bresler, *MR image reconstruction from highly undersampled k-space data by dictionary learning,* IEEE transactions on medical imaging **30**, 1028 (2010).

[186] L. Torrey and J. Shavlik, *Transfer learning,* in *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques* (IGI global, 2010) pp. 242–264.

[187] S. U. H. Dar, M. Özbey, A. B. Çatlı, and T. Çukur, *A transfer-learning approach for accelerated MRI using deep neural networks,* Magnetic resonance in medicine **84**, 663 (2020).

[188] N. Koonjoo, B. Zhu, G. C. Bagnall, D. Bhutto, and M. Rosen, *Boosting the signal-to-noise of low-field MRI with deep learning image reconstruction,* Scientific reports **11**, 1 (2021).

[189] P. Volegov, J. Mosher, M. Espy, and R. Kraus Jr, *On concomitant gradients in low-field MRI,* Journal of Magnetic Resonance **175**, 103 (2005).

A

OPTIMALITY PROPERTY OF GCGLS AND GCGME

Suppose we have a linear system of equations $\mathbf{Lu} = \mathbf{f}$ with solution $\mathbf{u}^*$. $\mathbf{L}$ is a hermitian positive definite (HPD) matrix. Then, at iteration k, the conjugate gradient method finds $\mathbf{u}_k$ such that $||\mathbf{u}_k - \mathbf{u}^*||_{\mathbf{L}}$, the error induced by the system matrix $\mathbf{L}$, is minimized over the Krylov subspace $\mathbf{u}_0 + \mathscr{K}_k(\mathbf{L}, \mathbf{f}) := \mathbf{u}_0 + \text{span}\{\mathbf{f}, \mathbf{Lf}, \mathbf{L}^2\mathbf{f}, ..., \mathbf{L}^{k-1}\mathbf{f}\}$. This means that, in every iteration, GCGLS minimizes

$$\begin{aligned} &\min_{\mathbf{x}_k} (\mathbf{x}_k - \mathbf{x})^H (\mathbf{A}^H \mathbf{C}^{-1} \mathbf{A} + \tau \mathbf{R})(\mathbf{x}_k - \mathbf{x}) \Leftrightarrow \\ &\min_{\mathbf{x}_k} (\mathbf{r} - \mathbf{r}_k)^H \mathbf{C} (\mathbf{r} - \mathbf{r}_k) + \tau (\mathbf{x}_k - \mathbf{x})^H \mathbf{R} (\mathbf{x}_k - \mathbf{x}) \,, \quad \mathbf{r}_k = \mathbf{C}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}_k) \Leftrightarrow \\ &\min_{\mathbf{x}_k} \|\mathbf{r} - \mathbf{r}_k\|_{\mathbf{C}}^2 + \tau \|\mathbf{x} - \mathbf{x}_k\|_{\mathbf{R}}^2 \,, \quad \mathbf{r}_k = \mathbf{C}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}_k) \end{aligned} \tag{A.1}$$

for $\mathbf{x}_k - \mathbf{x}_0 \in \mathscr{K}_k(\mathbf{A}^H \mathbf{C}^{-1} \mathbf{A} + \tau \mathbf{R}, \mathbf{A}^H \mathbf{C}^{-1} \mathbf{r}_0 + \tau \mathbf{R} \mathbf{x}_0)$. For every iteration of GCGME, the following holds:

$$\begin{aligned} &\min_{\mathbf{r}_k} (\mathbf{r} - \mathbf{r}_k)^H (\frac{1}{\tau} \mathbf{A} \mathbf{R}^{-1} \mathbf{A}^H + \mathbf{C})(\mathbf{r} - \mathbf{r}_k) \Leftrightarrow \\ &\min_{\mathbf{r}_k} \tau (\mathbf{x} - \mathbf{x}_k)^H \mathbf{R} (\mathbf{x} - \mathbf{x}_k) + (\mathbf{r} - \mathbf{r}_k)^H \mathbf{C} (\mathbf{r} - \mathbf{r}_k), \quad \mathbf{x}_k = \frac{1}{\tau} \mathbf{R}^{-1} \mathbf{A}^H \mathbf{r}_k \\ &\min_{\mathbf{r}_k} \|\mathbf{r} - \mathbf{r}_k\|_{\mathbf{C}}^2 + \tau \|\mathbf{x} - \mathbf{x}_k\|_{\mathbf{R}}^2 \,, \quad \mathbf{x}_k = \frac{1}{\tau} \mathbf{R}^{-1} \mathbf{A}^H \mathbf{r}_k \end{aligned} \tag{A.2}$$

with $\mathbf{r}_k - \mathbf{r}_0 \in \mathscr{K}_k(\mathbf{r}_0, \frac{1}{\tau} \mathbf{A} \mathbf{R}^{-1} \mathbf{A}^H + \mathbf{C})$. Note that GCGLS and GCGME minimize the same weighted combination of the errors in the residual and in the solution, but over different subspaces and under different constraints.

B

COMPARISON OF THE CONDITION NUMBERS OF GCGLS AND GCGME: A SIMPLE CASE

In this section we consider a very simple but illustrative case that allows us to analyze the condition numbers, and hence the convergence speed, of GCGME and GCGLS. We demonstrate, depending on the regularization parameter, which method is to be preferred. We set $\mathbf{A} = \mathscr{F}$ and the noise is assumed to be white noise, so $\mathbf{C} = \mathbf{I}$. We define the regularization matrix to be the discretized 2D Laplacian $\mathbf{L}$ complemented with Dirichlet boundary conditions. Choosing the regularization matrix in this way means that large jumps in the reconstructed image $\mathbf{x}$ are discouraged. In that case, GCGLS solves

$$(\mathbf{I} + \tau\mathbf{L})\,\mathbf{x} = \mathscr{F}^*\mathbf{v}, \tag{B.1}$$

where $\mathscr{F}^* = \mathscr{F}^{-1}$ is the inverse 2D Fourier Transform. GCGME solves

$$\left(\frac{1}{\tau}\mathscr{F}\mathbf{L}^{-1}\mathscr{F}^* + \mathbf{I}\right)\mathbf{r} = \mathbf{b} \quad , \tag{B.2}$$

$$\mathbf{x} = \frac{1}{\tau}\mathbf{L}^{-1}\mathscr{F}^*\mathbf{r}.$$

The convergence speed of GCGLS and GCGME depends on the condition number of the matrices $\mathbf{I} + \tau\mathbf{L}$ and $\frac{1}{\tau}\mathscr{F}\mathbf{L}^{-1}\mathscr{F}^* + \mathbf{I}$, respectively. The eigenvalues of the Laplacian $\mathbf{L}$ are

This appendix is based on the article:

M.L. de Leeuw den Bouter, M.B. van Gijzen, and R.F. Remis, *CG Variants for General-Form Regularization with an Application to Low-Field MRI*, Numerical Mathematics and Advanced Applications ENUMATH 2019: European Conference, Egmond aan Zee, The Netherlands, September 30-October 4, 673-681 (2021).

well-known and hence we can find explicit expressions for the condition numbers. For GCGLS, we have

$$\kappa_2\left(\mathbf{I}+\tau\mathbf{L}\right)=\frac{1+8\tau\cos^2\left(\frac{\pi}{2}\frac{1}{N+1}\right)}{1+8\tau\sin^2\left(\frac{\pi}{2}\frac{1}{N+1}\right)}. \tag{B.3}$$

Here, we assume that our image consists of $N\times N$ pixels. For GCGME, we make use of the fact that $\mathscr{F}\mathbf{L}^{-1}\mathscr{F}^*$ is a similarity transformation and therefore has the same eigenvalues as $\mathbf{L}^{-1}$, yielding

$$\kappa_2\left(\frac{1}{\tau}\mathscr{F}\mathbf{L}^{-1}\mathscr{F}^*+\mathbf{I}\right)=\frac{1+\frac{1}{8\tau\sin^2\left(\frac{\pi}{2}\frac{1}{N+1}\right)}}{1+\frac{1}{8\tau\cos^2\left(\frac{\pi}{2}\frac{1}{N+1}\right)}}. \tag{B.4}$$

These condition numbers can be shown to be equal when

$$\tau^*=\frac{1}{8\cos\left(\frac{\pi}{2(N+1)}\right)\sin\left(\frac{\pi}{2(N+1)}\right)}. \tag{B.5}$$

Figure B.1 shows a plot of the condition numbers as a function of the value of the regularization parameter τ, in case $N=128$. We observe that when $\tau<\tau^*$, GCGLS is handling a matrix with a smaller condition number, whereas the GCGME system has a smaller condition number when $\tau>\tau^*$ is large. Therefore, we expect GCGME to attain faster convergence for large τ.

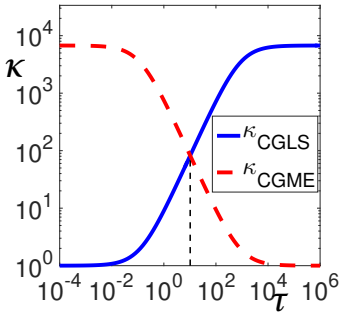


Figure B.1: Condition numbers of the GCGLS matrix $\mathbf{I}+\tau\mathbf{L}$ and the GCGME matrix $\frac{1}{\tau}\mathscr{F}\mathbf{L}^{-1}\mathscr{F}^*+\mathbf{I}$ as a function of the value of the regularization parameter τ.

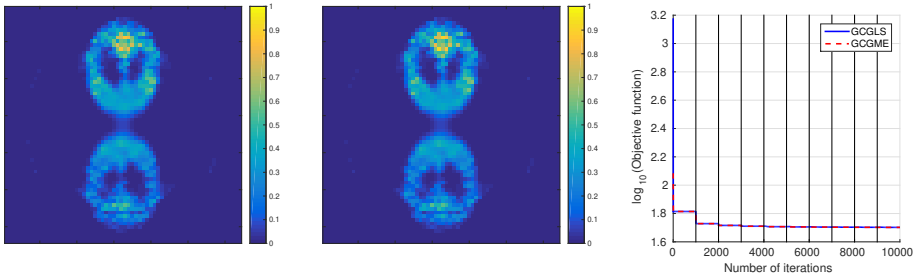
C

INCREASING THE NUMBER OF CG ITERATIONS PER IRLS ITERATION

In Section 2.3, minimization problem

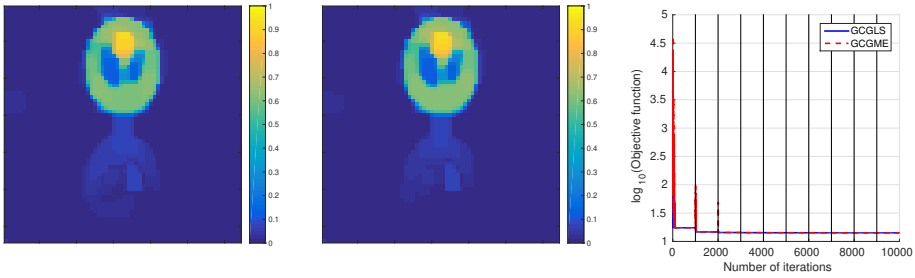
$$\min_{\mathbf{x}} \frac{1}{2}||\mathbf{Ax}-\mathbf{b}||_2^2 + \tau||\mathbf{Fx}||_1$$

is solved for two different regularization operators $\mathbf{F}$. The number of (inner) CG iterations per (outer) IRLS iteration was set to 10. The results seem to indicate that GCGME converges to a lower value of the objective function for all cases. However, here we increase the number of CG iterations to 1000 and we show that the convergence of GCGLS is simply very slow compared to the convergence of GCGME, and 10 iterations is not enough for GCGLS to make any visible progress towards the minimum. In Figures C.1 and C.2, we see that both methods converge to the same objective function value.



(a) GCGLS result (b) GCGME result (c) Objective function value

Figure C.1: Reconstruction results with $\mathbf{F} = \mathbf{I}$, $\tau = 1.5 \cdot 10^{-1}$ and 1000 CG iterations per IRLS iteration. In the rightmost figure, the value of objective function (2.47) is plotted as a function of the iteration number. The vertical black lines indicate the start of a new IRLS iteration



(a) GCGLS result (b) GCGME result (c) Objective function value

Figure C.2: Reconstruction results with $\mathbf{F} = \mathbf{T}$, $\tau = 1 \cdot 10^{-2}$ and 1000 CG iterations per IRLS iteration. In the rightmost figure, the value of objective function (2.50) is plotted as a function of the iteration number. The vertical black lines indicate the start of a new IRLS iteration

D

SIGNAL MODEL FOR GENERAL FIELDS

In this analysis, we ignore relaxation effects. We follow the derivation in [189] and adapt it to our case. We are interested in the magnetization $\mathbf{m}(\mathbf{r}, t)$. Denoting the magnetic field by $\mathbf{B}$, the Bloch equation looks as follows

$$\frac{d\mathbf{m}(\mathbf{r}, t)}{dt} = \gamma[\mathbf{m}(\mathbf{r}, t) \times \mathbf{B}(\mathbf{r}, t)]. \tag{D.1}$$

We suppose that the orientation of $\mathbf{B}$ does not change with time, so we can write $\mathbf{B}(\mathbf{r}, t) = \mathbf{B}(\mathbf{r}) = B(\mathbf{r})\hat{\mathbf{e}}(\mathbf{r})$, with $B(\mathbf{r})$ the magnitude of the magnetic field and $\hat{\mathbf{e}}(\mathbf{r})$ its orientation. Then, the Bloch equation can be rewritten:

$$\frac{d\mathbf{m}(\mathbf{r}, t)}{dt} = -\gamma\mathbf{B}(\mathbf{r})[\hat{\mathbf{e}}(\mathbf{r}) \times \mathbf{m}(\mathbf{r})] \tag{D.2}$$

$$= -\omega(\mathbf{r})[\mathbf{A}(\mathbf{r})\mathbf{m}(\mathbf{r}, t)]. \tag{D.3}$$

Here, $\omega(\mathbf{r}) = \gamma B(\mathbf{r})$ is the Larmor frequency and $\mathbf{A}(\mathbf{r})$ represents the cross product in matrix form:

$$\mathbf{A}(\mathbf{r}) = \begin{pmatrix} 0 & -e_z(\mathbf{r}) & e_y(\mathbf{r}) \\ e_z(\mathbf{r}) & 0 & -e_x(\mathbf{r}) \\ -e_y(\mathbf{r}) & e_x(\mathbf{r}) & 0 \end{pmatrix}. \tag{D.4}$$

The orientation of $\hat{\mathbf{e}}(\mathbf{r})$ is determined by the total applied magnetic field. The eigenvalues of $\mathbf{A}$ are $\lambda_{-1} = -i$, $\lambda_0 = 0$ and $\lambda_1 = i$, where i is the imaginary unit. The eigenvectors are

$$\mathbf{v}_{-1} = \begin{pmatrix} -e_x(\mathbf{r})e_y(\mathbf{r}) - ie_z(\mathbf{r}) \\ 1 - e_y^2(\mathbf{r}) \\ -e_y(\mathbf{r})e_z(\mathbf{r}) + ie_x(\mathbf{r}) \end{pmatrix}, \ \mathbf{v}_0 = \hat{\mathbf{e}}(\mathbf{r}), \ \mathbf{v}_1 = \begin{pmatrix} -e_x(\mathbf{r})e_y(\mathbf{r}) + ie_z(\mathbf{r}) \\ 1 - e_y^2(\mathbf{r}) \\ -e_y(\mathbf{r})e_z(\mathbf{r}) - ie_x(\mathbf{r}) \end{pmatrix}. \tag{D.5}$$

Denoting the initial magnetization $\mathbf{m}(\mathbf{r},0)$ by $\mathbf{m}_0(\mathbf{r})$, $\mathbf{m}(\mathbf{r},t)$ is described by

$$\mathbf{m}(\mathbf{r},t) = \sum_{l=-1}^{1} e^{-il\omega(\mathbf{r})t}\mathbf{v}_l(\mathbf{r})(\mathbf{v}_l(\mathbf{r})^*\mathbf{m}_0(\mathbf{r})). \tag{D.6}$$

The precession of the magnetization produces a magnetic flux in the receiver coil, described by

$$\Phi(t) = \int_{\mathbf{r}\in\mathbb{D}} \mathbf{W}(\mathbf{r})\cdot\mathbf{m}(\mathbf{r},t)\, d\mathbf{r}, \tag{D.7}$$

where $\mathbb{D}$ the imaging domain and $\mathbf{W}$ is the receive field. Then, Faraday's law of induction states that the voltage induced in the receive coil is given by [19]:

$$V(t) = -\frac{\partial\Phi(t)}{\partial t} = -\frac{\partial}{\partial t}\int_{\mathbf{r}\in\mathbb{D}} \mathbf{W}(\mathbf{r})\cdot\mathbf{m}(\mathbf{r},t)\, d\mathbf{r}. \tag{D.8}$$

Substituting equation (D.6) into equation (D.8) yields

$$V(t) = -\frac{\partial}{\partial t}\int_{\mathbf{r}\in\mathbb{D}} \sum_{l=-1}^{1} e^{-il\omega(\mathbf{r})t}(\mathbf{W}(\mathbf{r})^*\mathbf{v}_l(\mathbf{r}))(\mathbf{v}_l^*(\mathbf{r})\mathbf{m}(\mathbf{r},t))\, d\mathbf{r}. \tag{D.9}$$

Our quantity of interest is the spin density $\rho(\mathbf{r})$. We know that for the magnitude of the initial magnetization, the following holds:

$$m_0(\mathbf{r}) = \rho(\mathbf{r})\frac{\gamma\hbar^2}{4kT}\omega(\mathbf{r}). \tag{D.10}$$

Here, $\hbar = \frac{h}{2\pi}$ where h is Planck's constant, k is Boltzmann's constant and T is the temperature. We will assume the temperature to be constant. Using equation (D.10), but ignoring the constants, we can rewrite equation (D.9):

$$V(t) = -\frac{\partial}{\partial t}\int_{\mathbf{r}\in\mathbb{D}} \sum_{l=-1}^{1} \rho(\mathbf{r})\omega(\mathbf{r})e^{-il\omega(\mathbf{r})t}(\mathbf{W}(\mathbf{r})^*\mathbf{v}_l(\mathbf{r}))(\mathbf{v}_l(\mathbf{r})^*\hat{\mathbf{m}}(\mathbf{r},t))\, d\mathbf{r}, \tag{D.11}$$

where $\hat{\mathbf{m}}(\mathbf{r},t)$ is the orientation of the initial magnetization. Interchanging the order of the operations, we get

$$V(t) = -\int_{\mathbf{r}\in\mathbb{D}} \sum_{l=-1}^{1} \frac{\partial}{\partial t}\rho(\mathbf{r})\omega(\mathbf{r})e^{-il\omega(\mathbf{r})t}(\mathbf{W}(\mathbf{r})^*\mathbf{v}_l)(\mathbf{v}_l^*\hat{\mathbf{m}}(\mathbf{r},t))\, d\mathbf{r} \tag{D.12}$$

$$= \int_{\mathbf{r}\in\mathbb{D}} \sum_{l=-1}^{1} il\rho(\mathbf{r})\omega^2(\mathbf{r})e^{-il\omega(\mathbf{r})t}(\mathbf{W}(\mathbf{r})^*\mathbf{v}_l(\mathbf{r}))(\mathbf{v}_l(\mathbf{r})^*\hat{\mathbf{m}}(\mathbf{r},t))\, d\mathbf{r} \tag{D.13}$$

$$= \int_{\mathbf{r}\in\mathbb{D}} -i\rho(\mathbf{r})\omega^2(\mathbf{r})e^{i\omega(\mathbf{r})t}(\mathbf{W}(\mathbf{r})^*\mathbf{v}_{-1}(\mathbf{r}))(\mathbf{v}_{-1}(\mathbf{r})^*\hat{\mathbf{m}}(\mathbf{r},t))$$

$$+ i\rho(\mathbf{r})\omega^2(\mathbf{r})e^{-i\omega(\mathbf{r})t}(\mathbf{W}(\mathbf{r})^*\mathbf{v}_1(\mathbf{r}))(\mathbf{v}_1(\mathbf{r})^*\hat{\mathbf{m}}(\mathbf{r},t))d\mathbf{r} \tag{D.14}$$

When quadrature phase sensitive detection is used, the signal is multiplied by a factor of $e^{i\omega_0 t}$, where ω_0 is the demodulation frequency which typically corresponds to the frequency in the center of the field, and subsequently passes through a low-pass filter. Only the term including $\mathbf{v}_1$ will make it through the low-pass filter. Now, including the coil sensitivity $c(\mathbf{r})$, the final signal is given by

$$S(t) = \int_{\mathbf{r}\in\mathbb{D}} i\rho(\mathbf{r})c(\mathbf{r})\omega^2(\mathbf{r})e^{-i(\omega(\mathbf{r})-\omega_0)t}(\mathbf{W}(\mathbf{r})^*\mathbf{v}_1(\mathbf{r}))(\mathbf{v}_1(\mathbf{r})^*\hat{\mathbf{m}}(\mathbf{r},t))d\mathbf{r}. \tag{D.15}$$

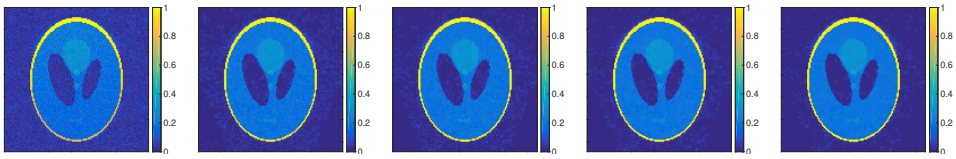
We note that $(\mathbf{W}(\mathbf{r})^*\mathbf{v}_1(\mathbf{r}))(\mathbf{v}_1(\mathbf{r})^*\hat{\mathbf{m}}(\mathbf{r},t))$ is equivalent to a voxel-wise weighting.

E

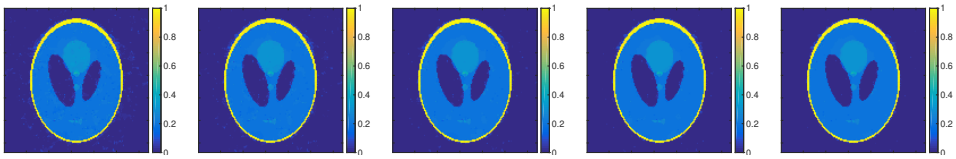
INCREASING THE NUMBER OF ITERATIONS FOR MULTIPLICATIVE REGULARIZATION

In Figs. E.1 and E.2, the Shepp-Logan reconstructions (for an SNR of 20 and an SNR of 5) are shown for a larger number of iterations. (To obtain the results shown in Figures 4.2d and 4.4d, we stopped after 50 iterations.) We see that in the low SNR case, the result becomes very blurry and we lose all the small structures in the image if we iterate too long. In the case of a higher SNR of 20, some oversmoothing is taking place when the iteration number exceeds 50 but much less severely than in the low SNR case. We see something similar happening in Figs. E.3 and E.4, where the apple and melon reconstructions are shown for a larger number of iterations. We note that, for the apple experiment, increasing the number of iterations to 45 yields an image that is slightly blurrier than the image obtained after 30 iterations, which is the result that is shown in Fig. 4.6b. However, for the melon, we see that when the number of iterations exceeds approximately 20 (which is the number of iterations used for the reconstruction in Fig. 4.8b, the amount of blur in the image becomes excessive.

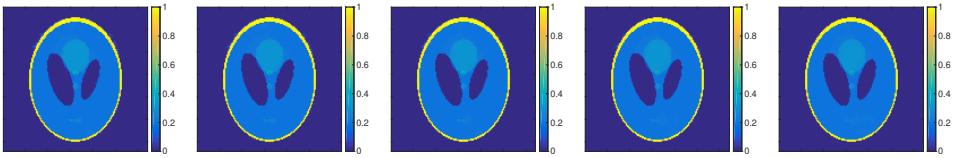
These observations suggest that stopping the algorithm before full convergence is attained yields a solution of better quality. This is especially important in case of a low SNR.



(a) Initial guess (b) Iteration 5 (c) Iteration 10 (d) Iteration 15 (e) Iteration 20

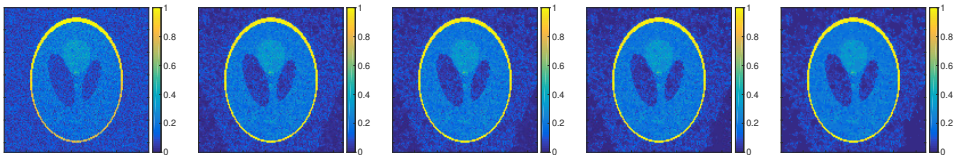


(f) Iteration 25 (g) Iteration 30 (h) Iteration 35 (i) Iteration 40 (j) Iteration 45

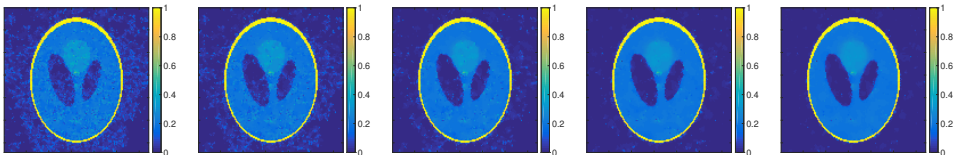


(k) Iteration 50 (l) Iteration 55 (m) Iteration 60 (n) Iteration 65 (o) Iteration 70

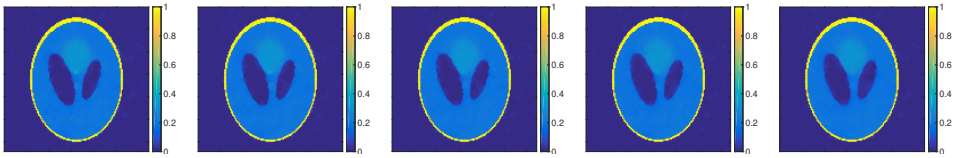
Figure E.1: Reconstruction of the Shepp-Logan phantom, with an SNR of 20, for different iteration numbers.



(a) Initial guess (b) Iteration 5 (c) Iteration 10 (d) Iteration 15 (e) Iteration 20



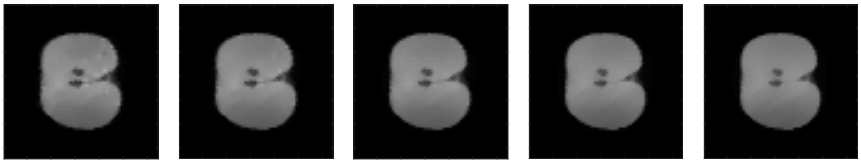
(f) Iteration 25 (g) Iteration 30 (h) Iteration 35 (i) Iteration 40 (j) Iteration 45



(k) Iteration 50 (l) Iteration 55 (m) Iteration 60 (n) Iteration 65 (o) Iteration 70

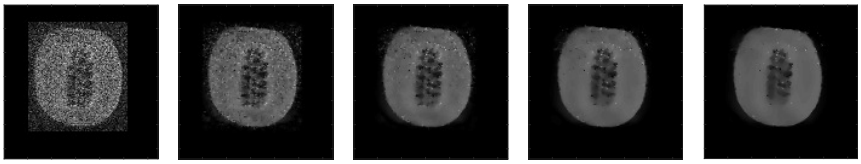
Figure E.2: Reconstruction of the Shepp-Logan phantom, with an SNR of 5, for different iteration numbers.

(a) Initial guess (b) Iteration 5 (c) Iteration 10 (d) Iteration 15 (e) Iteration 20

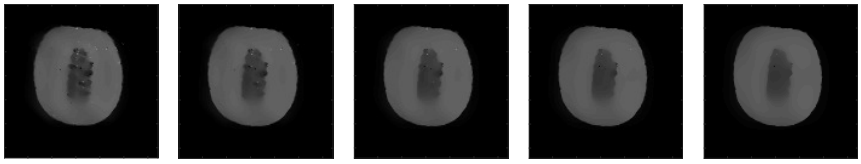


(f) Iteration 25 (g) Iteration 30 (h) Iteration 35 (i) Iteration 40 (j) Iteration 45

Figure E.3: Reconstruction of the 35th slice of the apple for different iteration numbers.

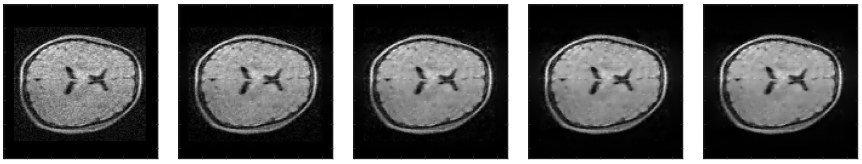


(a) Initial guess (b) Iteration 5 (c) Iteration 10 (d) Iteration 15 (e) Iteration 20

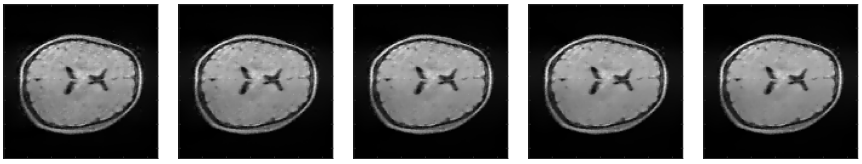


(f) Iteration 25 (g) Iteration 30 (h) Iteration 35 (i) Iteration 40 (j) Iteration 45

Figure E.4: Reconstruction of the 64th slice of the melon for different iteration numbers.



(a) Initial guess (b) Iteration 5 (c) Iteration 10 (d) Iteration 15 (e) Iteration 20



(f) Iteration 25 (g) Iteration 30 (h) Iteration 35 (i) Iteration 40 (j) Iteration 45

Figure E.5: Reconstruction of the 34th slice of the brain for different iteration numbers.

ACKNOWLEDGEMENTS

First of all, I would like to express my gratitude to NWO-WOTRO, for funding this research. A big thank you to the people of the TU Delft | Global Initiative. I am very grateful for your support (and the nice lunches you provided on a monthly basis). You always showed a lot of interest in and enthusiasm for the project, which was really inspiring.

Next on the list are my two amazing promotors. Martin and Rob, I do not think a PhD candidate can have better supervisors than the two of you. I consider myself truly blessed. Without you, the past four years would have been a lot less enjoyable.

Martin, thank you for the trust you placed in me all these years, the space you gave me and of course the Christmas lunches Marielba and you hosted on a yearly basis. Over the years, you have become like a second father to me. The well-being of your students is the most important thing to you. I know that your door is always open and that I can talk to you about anything.

Thank you, Rob, for coming up with inspiring new ideas at exactly the right time. Collaborating with you truly gave me a boost, which was in part due to your enormous enthusiasm. During our meetings we talked about everything and nothing.

Thanks to all other members of the low-field MRI team of the Leiden University Medical Center and Delft University of Technology, Pennsylvania State University and Mbarara University of Technology, for the great collaboration. Specifically, I would like to acknowledge Tom O'Reilly and Andrew Webb, for making this project a success, Johnes Obungoloch, for welcoming us to Uganda, and Sumit Tewari, for all your hard work and expertise on the hand-held scanner project. And thank you Kirsten, for all those times I bombarded you with MRI-related questions and you answered patiently and, of course, for suggesting to travel to Sri Lanka together when I really needed it.

Kees Lemmens, system administrator extraordinaire, thank you for all the patience and kindness you showed me, formerly (?) a total Linux noob, when I tried to find my way around the cluster. Also, for that time I accidentally deleted all my files from the cluster and you saved me by managing to retrieve them. I learned a lot from you.

We are very lucky to have department head Kees Vuik who makes everyone feel appreciated and welcome. Also, thanks to Deborah for providing support in such a capable manner.

I am glad I got to (co-)supervise a number of students. It truly added to the experience of doing a PhD. This list of acknowledgments would not be complete if I didn't specifically mention Dilan, who is full of wisdom and humor (and has very cool personal style!). I am glad we stayed in touch even after you received your M.Sc. degree.

To my colleagues Jochen, Mousa, Prajakta, Manuel, Reinaldo, Dennis, Behrouz, Baljaa, Owen, Jiao, Shuaiqiang, Luis, Thomas, Hugo, Kristof, Linlin, Deepesh, Alexander, Carolina, Merel, Leila, Ye, Kewang, Jun, Jinqiang, Manuel, Elisa, Giorgio, Daniel and Buu-Van, thanks for the friendliness and warmth. Over the years, our lunch topics became less nerdy and increasingly entertaining. It's been a joy working with you.

Marieke, Vandana and Anne, I am really happy we became friends during our PhDs. We had many cups of coffee together, over which we discussed everything and nothing. With you, I always found a safe space to vent. And of course, thanks for all the laughter! During the last few months of my PhD, I was lucky enough to get to spend a lot of time with Marieke and Vandana, both of whom are incredibly kind and caring people who are an inspiration to me.

To my office mates Qiyao (Alice), Xiujie (Xiaoshan), Mohamed (Mo) and Gabriela (Gaby) and adopted 20% office mate Roel, the awesome people who made our office the coolest office: I am so grateful I got to spend so much time with you. It's much better to come to work knowing you'll get to see your friends all day. Thanks for all the activities we did together, like go karting, boat trips, cheese fondue, etc, etc. Alice, thanks for the companionship. Roel, thank you for being so level-headed and for giving good advice. Gaby, thanks for getting everyone together for cool events. Thank you Mo, for your hospitality and for buying the good fondue supplies. And Xiaoshan, thank you for being extremely sweet and supportive all the time. I couldn't have wished for a better academic sister!

When the math department only had flexible office spaces without any equipment, our computer science neighbors Otto and Thomas were kind enough to lend Alice and me monitors, keyboards and mouses everyday. We really appreciated that and it was nice to get to know you guys when the four of us were basically the only people in the office during the pandemic.

Yvonne, thanks for all the therapeutic tennis sessions and for your boundless energy. Nicole, I really enjoyed our catching up sessions every Friday at lunch. Kian, I am glad you decided to do a PhD in Delft after all! It was good to have a coffee with you every now and then. Anneline and Mireille, I know we haven't seen each other very often the past two years but thanks for being there for me when I needed it. Ruben, I don't think you know exactly how much you helped me with my struggles. To my boxing buddy Anne: thank you for all those hours that made me forget about everything.

Jenny, we got to know each other when you started your M.Sc. graduation project in the same department and I am so happy you came into my life at that time. Being in your presence truly is a joy, not least because of your crazy sense of humor. Michelle, you might be the sweetest person I have ever met. Thank you for your company during the lockdown, for all the walls you gladly and selflessly helped me paint and of course for designing my thesis cover. Here's to many more years of living in Rotterdam as neighbors!

Andrea, we have not been together that long but this list definitely would not be complete without you. I know you think you did not really help me with my PhD but I should think that making me happy definitely counts for something.

Last but certainly not least I want to thank my unofficial and my official family. Jelle, Maaike, Abe and Noortje, thanks for all the fun, the escape rooms and the good times in both Switzerland and the Netherlands. Tommy, you know why. Mom, Dad, Mats, and Nyx: thanks for giving me a place of rest and love I can always return to. I feel like this paragraph is too short but words cannot express how much you mean to me.

Curriculum Vitæ

Merel Lisanne de Leeuw den Bouter

09-11-1992 Born in Capelle aan den IJssel, The Netherlands

Education

2011–2014 B.Sc. in Applied Mathematics (*cum laude*)
Delft University of Technology, Delft, The Netherlands

2011–2014 B.Sc. in Applied Physics (*cum laude*)
Delft University of Technology, Delft, The Netherlands

2015–2017 M.Sc. in Applied Mathematics (*cum laude*)
Delft University of Technology, Delft, The Netherlands (2015-2017)
Massachusetts Institute of Technology, Cambridge, Massachusetts, USA (2016)

2017–2021 Ph.D. in Numerical Analysis
Delft University of Technology, Delft, The Netherlands

PUBLICATIONS AND ACADEMIC ACTIVITIES

LIST OF PUBLICATIONS

- M.L. de Leeuw den Bouter, G. Ippolito, T.P.A. O'Reilly, R.F. Remis, M.B. van Gijzen, and A.G. Webb, *Deep learning-based single image super-resolution for low-field MR brain images*, submitted to Nature Scientific Reports, December 2021.

- C.M.W. Goedmakers, L.M. Pereboom, J.W. Schoones, M.L. de Leeuw den Bouter, R.F. Remis, M. Staring, C.L.A. Vleggeert-Lankamp, *Machine and Deep Learning for Image Analysis in the Cervical Spine: Systematic Review of Available Models and Methods*, submitted to Computer Methods and Programs in Biomedicine, December 2021.

- M.L. de Leeuw den Bouter, P.M. van den Berg and R.F. Remis, *Inversion of incomplete spectral data using support information with an application to magnetic resonance imaging*, Journal of Physics Communications **5**, 055006 (2021).

- M.L. de Leeuw den Bouter, M.B. van Gijzen and R.F. Remis, *CG Variants for General-Form Regularization with an Application to Low-Field MRI*, Numerical Mathematics and Advanced Applications ENUMATH 2019: European Conference, Egmond aan Zee, The Netherlands, September 30-October 4, Springer International Publishing, 673-681 (2021).

- M.L. de Leeuw den Bouter, M.B. van Gijzen and R.F. Remis, *Low-field magnetic resonance imaging using multiplicative regularization*, Magnetic Resonance Imaging **75**, 21-33 (2021).

- M.L. de Leeuw den Bouter, D.B. Geçmen, A.E. Meijer, D.H. de Gans, L.N.M. Middelplaats, R.F. Remis and M.B. van Gijzen, *Description of a Low-field MRI Scanner Based on Permanent Magnets*, CEUR Workshop Proceedings **2688** (2020).

- M.L. de Leeuw den Bouter, M.B. van Gijzen and R.F. Remis, *Conjugate gradient variants for ℓ_p-regularized image reconstruction in low-field MRI*, SN Applied Sciences **1**, 1736 (2019).

CONTRIBUTIONS AT CONFERENCES AND SYMPOSIA

ORAL PRESENTATIONS

- *Low-field MR imaging using multiplicative regularization.* Presented at: URSI GASS, online, August 28 - September 4, 2021.

- *Low-field MR imaging using multiplicative regularization.* Presented at: ISMRM Annual Meeting, online, May 15-20, 2021.

- *Correcting image distortions caused by gradient nonlinearities in low-field MRI using a neural network.* Presented at: ESMRMB Annual Meeting, online, September 30 - October 2, 2020.

- *CG variants for general-form regularization with an application to low-field MRI*. Presented at: ENUMATH, Egmond aan Zee, The Netherlands, October 2, 2019.
- *CG variants for general-form regularization with an application to low-field MRI* (invited). Presented at: SCS Annual Spring Meeting, Antwerp, Belgium, May 17, 2019.
- *Joint image reconstruction and field map estimation in a low-field MRI setting*. Presented at: ISMRM Benelux Annual Meeting, Leiden, January 17, 2019.
- *MRI for low-resource settings*. Presented at: TU Delft | Global Initiative Healthcare4all symposium, Delft, The Netherlands, November 26, 2018.

SCIENTIFIC POSTERS

- *Super-resolution image reconstruction in low-field MRI using deep learning*. Presented at: forty-fifth SCS Woudschoten Conference, Zeist, The Netherlands, October 7, 2021.
- *Correcting image distortions caused by gradient nonlinearities in low-field MRI using a neural network*. Presented at: ESMRMB Annual Meeting, online, September 30 - October 2, 2020.
- *Joint image reconstruction and field map estimation: Deep learning in low-field MRI*. Presented at: forty-fourth SCS Woudschoten Conference, Zeist, The Netherlands, October 10, 2019.
- *A comparison of CGLS and CGME for low-field MRI*. Presented at: forty-third SCS Woudschoten Conference, Zeist, The Netherlands, October 4, 2018.
- *p-norm regularization in low-field MRI*. Presented at: SIAM Annual Meeting, Portland, United States of America, July 10, 2018.
- *Image reconstruction in low-field MRI*. Presented at: ISMRM Annual Meeting, Paris, France, June 21, 2018.
- *Image reconstruction in low-field MRI*. Presented at: SIAM Conference on Imaging Science, Bologna, Italy, June 5-6, 2018.

AWARDS

- 3rd poster prize at the forty-third SCS Woudschoten Conference, Zeist, The Netherlands, October 4, 2018.
- 1st poster prize at the TU Delft EEMCS PhD Day, Delft, The Netherlands, October 31, 2017.

TEACHING ACTIVITIES AT TU DELFT

- WI1421LR: Calculus 1 (2020-2021 Q1-Q2)
- CTB2000: Numerical Methods Lab (2019-2020 Q4)
- TW2060: Numerical Methods 1 (2019-2020 Q3-Q4)
- CTB2000: Numerical Methods Lab (2018-2019 Q4)

- TW2060: Numerical Methods 1 (2018-2019 Q3-Q4)
- TW2060: Numerical Methods 1 (2017-2018 Q3-Q4)

SUPERVISED WORKS

- M.C. Zijta, *Automatic Segmentation of Low-Field MRI Brain Scans by Integrating Analytic and Deep Learning Techniques.* M.Sc. thesis, TU Delft, 2021.
- K. van der Werff (cum laude), *Image reconstruction for multicoil low-field MRI.* M.Sc. thesis, TU Delft, 2021.
- P.W. Oppelaar, *Model based image reconstruction for low-field hand-held MRI.* B.Sc. thesis, TU Delft, 2021.
- L. Pereboom, *Prediction of Invasive Cervical Spine Surgery Success by a Convolutional Neural Network Algorithm.* M.Sc. thesis, TU Delft, 2021.
- G. Ippolito (cum laude), *Super resolution Techniques Applied to Low-Field MRI.* M.Sc. thesis, TU Delft, 2020.
- M. Macarulla Rodriguez, *Low-field MR Imaging Using a Nonuniform Fast Fourier Transform.* M.Sc. thesis, TU Delft, 2020.
- A. Francke, *Data-driven image reconstruction for Low-Field MRI*, B.Sc. thesis, TU Delft, 2020.
- D.B. Geçmen, *Deep Learning Techniques for Low-Field MRI.* M.Sc. thesis, TU Delft, 2020.