# Accelerated Acquisition of Quantitative MRI Using Parametric Redundancy

By

Karthik Rao Aroor

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

(Electrical Engineering)

at the

UNIVERSITY OF WISCONSIN-MADISON

2017

Date of final oral examination: 04/04/2017

The dissertation is approved by the following members of the Final Oral Committee:

John A. Gubner, Professor, Electrical and Computer Engineering,
Andrew Alexander, Professor, Medical Physics,
Steven Kecskemeti, Scientist, Waisman Center,
Mary Elizabeth Meyerand, Professor, Biomedical Engineering,
William Sethares, Professor, Electrical and Computer Engineering,
Nagesh Adluru, Scientist, Waisman Center

© Copyright by Karthik Rao Aroor 2017

All Rights Reserved

# Abstract

## Accelerated Acquisition of Quantitative MRI Using Parametric Redundancy

Karthik Rao Aroor

*Under the Supervision of Dr. John Gubner and Dr. Andrew Alexander*

*Additional guidance provided by Dr. Steven Kecskemeti*

*At the University of Wisconsin-Madison*

Magnetic Resonance Imaging (MRI) is a non-invasive technique that can be utilized to obtain Quantitative $T_1$ images of the brain. Unfortunately, the acquisition of Quantitative MRI (Q-MRI) is an extremely slow process, and this has prevented applications of Q-MRI in many clinical situations where low scan times are critical. Current approaches to speed up the acquisition of Q-MRI are inadequate and result in images that have artifacts. This thesis develops and validates novel compressed sensing algorithms that exploit information in spatial and parametric dimensions for improved MRI $T_1$ mapping.

In the first project, two different sampling patterns are used to acquire MRI data at multiple Flip Angles (FAs). The two sampling patterns are the traditional stack-of-stars (SOS) and the recently introduced vastly undersampled isotropic projection (VIPR). It is shown that when used in parametric dimension regularized compressed sensing algorithms, at high accelerations, VIPR outperforms SOS when estimating $T_1$ maps.

The second project considers the Inversion Recovery acquisition, in which data is acquired at multiple inversion time points. Here, a novel parametric dimension compressed sensing

regularizer, Total Generalized Variation (TGV) is developed, and for this data, it is shown that using TGV as a regularizer yields better $T_1$ estimates than the current approach, which uses no regularization. TGV also outperforms the classical Total Variation (TV) regularizer by eliminating its well known staircase artifacts. In addition to continuous signals, it is also shown that TGV works well on discontinuous signals, since it does not oversmooth the discontinuities.

In the third project, we develop and validate a novel technique to estimate the $T_1$ values of all voxels in the brain simultaneously using a Total Variation (TV) regularizer, and show that this outperforms the current approach of estimating $T_1$ values of all voxels independently. Three different approaches to perform TV based $T_1$ estimation are proposed, and simulations are performed to determine the best one.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Quantitative Magnetic Resonance Imaging (Q-MRI) methods are promising for characterizing brain development, aging, and pathology [1] [2] [3] [4]. However, they are limited by the accuracy and variation of the measurements. Reductions in measurement variation would greatly improve the confidence, sensitivity, and specificity of Q-MRI for clinical and research applications.

Acquisition of Q-MRI data is a time consuming process because it involves the acquisition of multiple MRI images and thereby limits its application because subject compliance is difficult at large scan times [5] [6]. In this thesis, novel approaches to accelerate the acquisition of Q-MRI of the brain are introduced and validated.

In MRI the spatial Fourier transform of the signal of interest is measured, and we call this Fourier domain *k*-space, and the signal domain image-space [5] [6]. The length of a scan is a function of the number of *k*-space measurements, and so to lower the scan time, the number

of $k$-space samples acquired should be minimized. Our goal is to reconstruct the image-space signal accurately from a subsampled set of $k$-space measurements. To do this, the redundancies of the signal in image-space are exploited. Traditional methods to accelerate the acquisition of MRI, SENSE [7], Margosian [8], Homodyne [9], POCS [10], provide low acceleration. This is because these methods either do not completely exploit the inherent structure of the data [7] or the models used in the algorithms are not best suited for the actual data. The objective of this work is to achieve significantly higher acceleration using more sophisticated methods, namely compressed sensing algorithms. These algorithms exploit the underlying structure of the data and have a sound theoretical foundation [11]. The image-space data is composed of voxels (3D pixels), and adjacent voxels in image-space have a similar value, and this spatial redundancy is exploited in our algorithms. In addition to the three spatial dimensions, in Q-MRI there is an additional parametric dimension, which could be flip angle (FA) or time, and the redundancy in this parametric dimension is exploited as well in our algorithms. In this study we investigated two Q-MRI methods for $T_1$ estimation: Variable flip angle and Inversion recovery, whose parametric dimensions are flip angle and time respectively.

In all our projects, fully-sampled (FS) data in which all points of $k$-space are acquired are used as a reference against the results from compressed sensing algorithms.

## $\mathbf{T_1}$

All the projects in this thesis involve estimation of quantitative $T_1$ values at all brain voxels. This $T_1$ is the relaxation time constant when the magnetic moment of a voxel returns to its baseline value following a perturbation. An MR signal is produced by placing a subject in a strong magnetic field and then applying an additional radio frequency (RF) pulse that is

orthogonal to this main magnetic field. Each voxel behaves as a magnetic moment and is affected by this pulse. This magnetic moment has a transverse component in the $XY$ plane and a longitudinal component in the $Z$ plane. Once the RF pulse is applied, the longitudinal component grows exponentially as a function of time with a time constant $T_1$. An accurate $T_1$ estimate is highly desirable since this is a useful biomarker of the brain.

## 1.1 Variable Flip Angle (VFA) $T_1$ Imaging

Spoiled gradient echo (SPGR) variable flip angle (VFA) $T_1$ imaging refers to an acquisition method in which $k$-space data are acquired at multiple flip angles (FA). Each image-space voxel in the brain has a $T_1$ value, and our goal is to estimate the $T_1$ of all voxels using $k$-space data acquired at multiple FAs. The signal at a particular voxel in image-space is a function of both the $T_1$ of that voxel and the FAs, and this signal changes slowly across FAs, and so the norm of its derivative across FAs is small. To reconstruct image-space from subsampled $k$-space, a compressed sensing algorithm in which the penalty term is the $l_1$ norm of the derivative across FAs is used [12]. Once the signal in image-space is reconstructed from subsampled $k$-space, $T_1$ at all voxels can be estimated and validated that this is indeed close to the fully-sampled $T_1$. Since the performance of compressed sensing algorithms depends on the $k$-space sampling pattern, we compare the performance of different $k$-space sampling patterns. We show that a Vastly Undersampled Isotropic Projection (VIPR) $k$-space sampling pattern outperforms a stack-of-stars (SOS) $k$-space sampling pattern at large accelerations, when used in spatio-parametric compressed sensing algorithms with parametric dimension regularizers.

## 1.2 Inversion Recovery $T_1$ Imaging

Inversion Recovery $T_1$ imaging refers to an acquisition method in which $k$-space data are acquired at different inversion time points $t$. Each image-space voxel in the brain has a $T_1$ value associated with it, and our goal is to estimate $T_1$ of all voxels using the $k$-space data acquired at multiple $t$. The signal at a voxel in image-space is a function of both $t$ and the $T_1$ of that voxel, and this signal varies slowly across $t$. To reconstruct image-space from subsampled $k$-space, a compressed sensing algorithm is used in which the penalty term exploits this temporal redundancy. Typically in these kinds of problems, a Total Variation (TV) penalty term is used. However, this produces staircase/blocky artifacts, especially at high accelerations. It was recently shown that a new regularizer called Total Generalized Variation (TGV) eliminates these staircase artifacts and thus outperforms the TV when used as a spatial regularizer [13]. We develop TGV as a temporal/parametric regularizer, and test the hypothesis that TGV can outperform TV when used as a temporal regularizer. We show that TGV eliminates the staircase artifacts of TV and thus leads to better $T_1$ estimates, which in turn enables us to acquire data much faster than is currently possible.

We also show that the application of TGV is not just limited to the Inversion Recovery model, but it can be used on generic spatio-temporal datasets. We test and verify the hypothesis that TGV can be used to speed up the acquisition of a spatio-temporal dataset with temporal discontinuities, by showing that it outperforms both the current approach used for this dataset, and also the TV approach.

## 1.3 Total Variation $T_1$ Estimation

$T_1$ estimation in Variable Flip Angle (VFA) is currently done on a voxelwise basis, wherein $T_1$ of each voxel is estimated independently of the $T_1$ of other voxels. This approach however does not take into account the spatial similarity of the $T_1$ of adjacent voxels. A novel $T_1$ estimation procedure is developed, that exploits this spatial redundancy using Total Variation (TV) regularization. TV enforces the constraint that adjacent voxels in the estimated $T_1$ map will have similar $T_1$ values. We show that estimating $T_1$ values of all voxels simultaneously using a TV regularizer outperforms the current method of estimating $T_1$ values of each voxel separately, thus enabling higher quality $T_1$ maps and faster MRI acquisitions. This TV approach is shown to work when different noise models are used, thus enabling its application in different kinds of $k$-space sampling techniques. TV regularized $T_1$ estimation in VFA can be done in different ways. We discuss the three possible approaches, *pre*, *mid*, and *post*, and show that *mid* is the best.

The TV regularized $T_1$ estimation that is developed here is not just limited to VFA $T_1$ estimation, but can be applied to other quantitative imaging techniques, thus enabling us to obtain higher quality quantitative MRI maps.

# Chapter 2

# Notation

In this section, the terminology used in this thesis is discussed.

## 2.1 Voxel

In MRI, we are interested in obtaining a 3D image of the brain. The pixels in this 3D image are called voxels. While it is desirable to have voxels that are isotropic, i.e., all sides of the "cube" have the same length, it is possible that one side of this "cube" can be longer or shorter than the other two. While shorter sides are less common, longer sides are sometimes used, both because the SNR of each voxel improves and also because the scan time to acquire the MRI image is lower. Typical voxel sizes used in this thesis and in the MRI community are, $1mm^3$ and $2mm^3$. Voxel sizes larger than $4mm^3$ are not acquired because the gray matter tissue will not be discernible.

## 2.2 $k$-space

In MRI, we are interested in obtaining a 3D image of the brain. This image is represented as a scalar function defined on 3D space $x(\mathbf{r}), \mathbf{r} \in \mathbb{Z}^3$. Due to the physics underlying signal generation in MRI, $x$ is not measured directly but instead its 3D spatial Fourier transform $y$ is measured. The desired signal $x$ and the measured signal $y$ are related by

$$y(\mathbf{k}) = \sum_{\mathbf{r} \in \mathbb{Z}^3} x(\mathbf{r}) e^{-j\mathbf{k}'\mathbf{r}}, \quad \mathbf{k} \in \Omega \subset \mathbb{R}^3, \tag{2.1}$$

where $'$ is the transpose operator, and $j = \sqrt{-1}$.

The space in which $x$ lies is called image-space and the space in which $y$ lies is called $k$-space.

## 2.3 Multiple Coils

The discussion in Section 2.2 is applicable in a single coil acquisition scheme. Nowadays multiple coils are used to acquire the data. Originally introduced to increase the SNR, in recent years these have shown the ability to speed up acquisition by requiring fewer $k$-space samples to reconstruct image-space accurately. The signal $x$ is multiplied by a known coil sensitivity on a voxel-wise basis prior to taking the Fourier transform. In our work, 32 coils are used and the $k$-space sample of the $i$th coil is,

$$y(\mathbf{k}, i) = \sum_{\mathbf{r} \in \mathbb{Z}^3} x(\mathbf{r}) c(\mathbf{r}, i) e^{-j\mathbf{k}'\mathbf{r}}, \quad \mathbf{k} \in \Omega \subset \mathbb{R}^3, i = 1, \ldots, 32. \tag{2.2}$$

The coils are placed all over the brain and each coil is sensitive to a localized region of the brain around the coil, i.e., its coil sensitivity has a large amplitude near the coil and a low amplitude elsewhere. In Figures 2.1, 2.2, 2.3, a 2D slice of $x$ is shown along with the coil sensitivity $c$ of 2 different coils of the same slice.

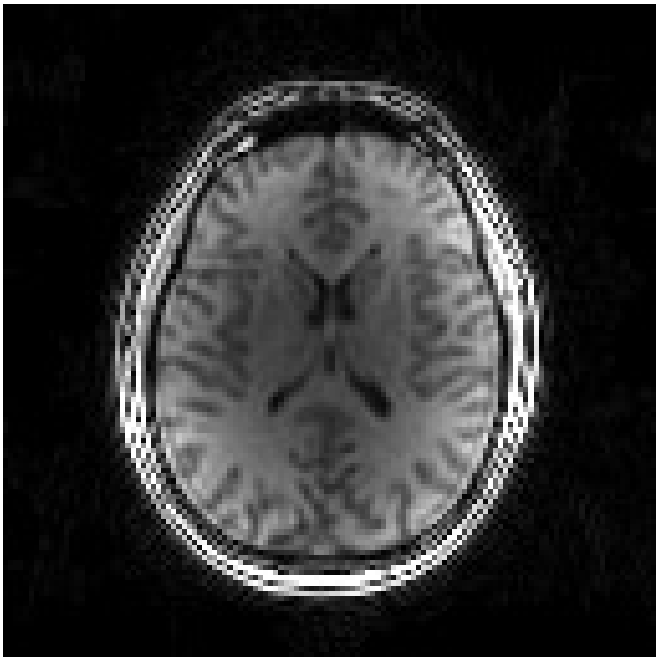

Figure 2.1: A 2D slice of image-space signal $x$. Coil sensitivities overlay perfectly on $x$ and are multiplied with $x$ prior to taking the Fourier transform to get the $k$-space signal $y$.

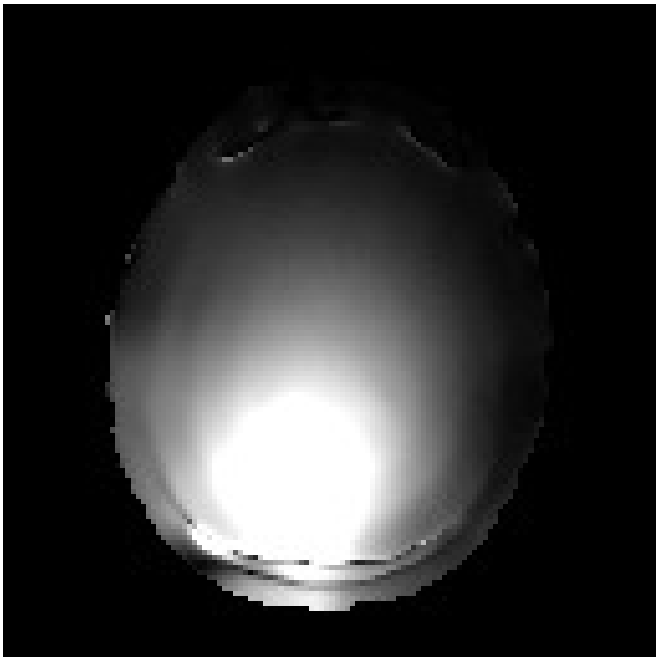

Figure 2.2: Coil Sensitivity of a coil placed near the bottom-left of the brain.

## 2.4 Parametric Dimension

We are interested in obtaining images of the brain at multiple user defined parameters $\theta$. The $\theta$ dimension is called the parametric dimension. At each $\theta$, $k$-space samples for all coils are acquired. Adding an additional index $\theta$ to (2.2), and using the fact that the coil sensitivity is the same for all $\theta$, we get

$$y(\mathbf{k},i,\theta) = \sum_{\mathbf{r}\in\mathbb{Z}^3} x(\mathbf{r},\theta)c(\mathbf{r},i)e^{-j\mathbf{k}'\mathbf{r}}, \;\; \mathbf{k}\in\Omega(\theta)\subset\mathbb{R}^3, i=1,\ldots,32. \tag{2.3}$$

Collecting all the samples from the multiple coils and from the parametric dimension, we form $y = \{y(\mathbf{k},i,\theta)\}_{\mathbf{k},i,\theta}$, $x = \{x(\mathbf{r},\theta)\}_{\mathbf{r},\theta}$, and represent the relation between $k$-space and image-

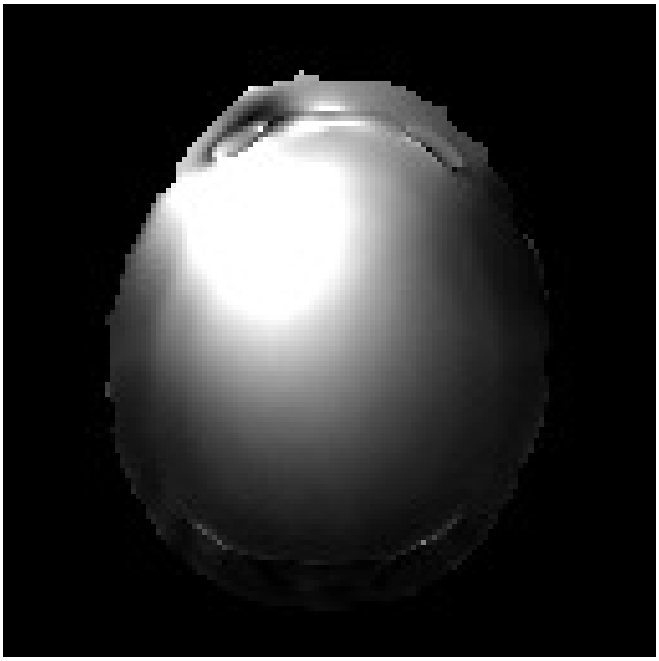

Figure 2.3: Coil Sensitivity of a coil placed near the top-left of the brain.

space more compactly by the notation $y = E(x)$. Since Fourier transform and multiplication by the coil sensitivity are linear operations, $E$ is a linear operator acting on $x$.

## 2.5 Reconstruction

If all samples in $k$-space were acquired, an inverse Fourier transform could be used to obtain $x$ from $y$. However, as described in the Introduction, since the goal is to reconstruct the signal in image-space from a subset of $k$-space samples, just taking an inverse Fourier transform on a subset of $k$-space will result in the output having artifacts. Compressed sensing methods that exploit redundancy in the signal help suppress these artifacts. These methods solve an

optimization problem of the form,

$$\min_x ||E(x) - y||^2 + P(x), \tag{2.4}$$

where $P(x)$ is the penalty term that exploits the redundancy of $x$ in both the parametric and spatial dimensions, and is different in different applications. The exact nature of $P(x)$ will be discussed per the application.

$||\cdot||$ is used to indicate the $l_2$ norm, and $||\cdot||_1$ the $l_1$ norm of the vectorized form of the argument.

We are interested in estimating $T_1$ which is a function of $x$. $T_1$ is estimated from $x$ using a non-linear least squares approach.

## 2.6 $k$-space Sampling Pattern

As was mentioned earlier, our objective is to estimate image-space from a subset of $k$-space samples. Choosing this subset, called the $k$-space sampling pattern, is critical in determining the performance of the reconstruction algorithm (2.4). Since the signal in image-space varies slowly across space, i.e., it is spatially smooth, its Fourier transform has most of its energy concentrated in the low frequency region. This low frequency region is called the center of $k$-space. Since the center of $k$-space contains most of the information of the image-space signal, the center of $k$-space needs to be heavily sampled to accurately reconstruct the signal. The high frequency region in $k$-space also contains useful information about edges and fine spatial details in image-space. This high frequency region is called the edge of $k$-space. Even though the edges of $k$-space contain a much smaller portion of the overall energy compared to the

center, if the edges of $k$-space are not sampled, the reconstructed images will lose information about fine details and will appear blurry. So for a given number of $k$-space samples, a good sampling pattern will be one that places most of the samples at the center, while a few samples will be placed at the edges of $k$-space. This ensures that most of the energy, and the fine details like edges of the signal are acquired.

The $k$-space sampling pattern we choose will be either points along a Cartesian grid or points along radial lines through the origin. In radial sampling we have 2D and 3D sampling patterns. We refer to 2D as stack-of-stars (SOS) and 3D as Vastly Undersampled Isotropic Projection (VIPR). In SOS, the brain is divided into 2D slices. For each slice encode, the $k$-space sampling pattern is radial lines in a 2D plane passing through the origin. To subsample $k$-space, a subset of these radial lines is taken, and the same radial sampling pattern is used for all slices. The fully-sampled (FS) SOS data has $\frac{\pi n}{2}$ radial lines in each 2D plane, where $n$ is the size of the volume in image-space. For example, if the object of interest is a cube of size $256^3$, then $n = 256$. To obtain an acceleration of $a$, we take $\frac{\pi n}{2a}$ radial lines. In VIPR, the $k$-space sampling pattern is radial lines in 3D passing through the origin. To subsample $k$-space, we take a subset of these radial lines. The FS VIPR data has $\frac{\pi n^2}{2}$ radial lines. To obtain an acceleration of $a$, we take $\frac{\pi n^2}{2a}$ radial lines.

SOS and VIPR $k$-space sampling patterns are illustrated in Figure 2.4.

The output obtained using all samples of $k$-space, setting $P(x) = 0$, and solving (2.4) is called the fully-sampled (FS) output, the output obtained using a subset of $k$-space and solving (2.4) with appropriate (non-zero) choice of $P(x)$ is called the compressed sensing (CS) output. The output obtained using a subset of $k$-space and solving (2.4) by setting $P(x) = 0$ is called Sensitivity Encoding (SENSE). Zero-Filled (ZF) output is obtained by using a subset of $k$-space,

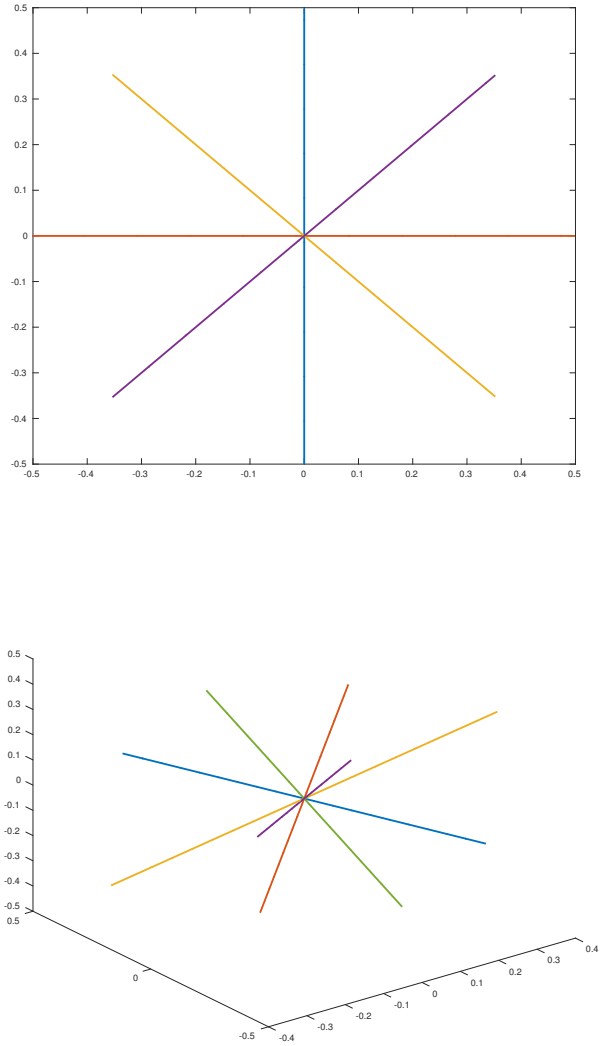

Figure 2.4: Radial *k*-space sampling patterns, (a) Stack-of-Stars (SOS), (b) Vastly Undersampled Isotropic Projection (VIPR). Each of these radial lines is called a projection.

setting all the samples to zero and taking the inverse Fourier transform. Conceptually ZF and SENSE solve the same problem and therefore should give the same solution. The FS output is used as a benchmark to compare the CS, SE, ZF results to.

# Chapter 3

# Variable Flip Angle (VFA) $T_1$ Imaging

Variable Flip Angle (VFA) $T_1$ imaging refers to a spoiled gradient echo (SPGR) acquisition in which $k$-space data are acquired at multiple flip angles (FA) [5]. Our goal in this project is to estimate the $T_1$ of each voxel in the brain using subsampled $k$-space data acquired at multiple FAs. The signal at a particular voxel in image-space is a function of both the $T_1$ of that voxel and the FA. This signal changes slowly across FA, and so its derivative across FA is bounded [12]. To reconstruct image-space from subsampled $k$-space, a compressed sensing algorithm is used in which the penalty term is the $l_1$ norm of the derivative across FAs. Once the signal in image-space is reconstructed from subsampled $k$-space, $T_1$ at all voxels can be estimated and validated that this is indeed close to the true $T_1$, which we take to be fully-sampled $T_1$.

There are other approaches to do regularization in compressed sensing MRI reconstructions, but all of these suffer from artifacts because of the mismatch between the assumptions of the regularization term and what is present in the actual data. Spatial regularization using

Total Variation produce staircase or blocky artifacts, while quadratic Tikhonov regularization produce blurry artifacts, and some parametric dimension regularization techniques make strict assumptions and only work well on certain kinds of datasets. Even though compressed sensing with spatial regularization is not necessarily the best approach to speed up the acquisition of MRI, the theory for this approach is well developed and it provides a number of guarantees like the number of samples needed to achieve a certain accuracy, the kinds of regularizers that can be used etc. [11]. Our approach of using a regularizer in the parametric dimension (i.e., across FA) has no theoretical guarantees whatsoever. It is unclear if this approach will even converge to a reasonable solution. Nevertheless, in our case, this approach worked well in practice.

Despite the well developed compressed sensing theory, a major shortcoming that hinders its wide spread practical applicability is that there is no optimal approach to choose the regularization parameter $\lambda$ given in (3.2). Only heuristic approaches to choosing $\lambda$ exist and we use the approach described in [14].

The setup described above of using the $l_1$ norm of the derivative across FAs as a regularizer was shown to achieve accelerated acquisition when a stack-of-stars (SOS) radial $k$-space sampling pattern was used [12]. It is known that in CS algorithms, the performance depends upon the $k$-space sampling pattern [11], although no analytical expressions exist that describe the exact nature of this dependence. We test the hypothesis that for a given scan duration, the Vastly Undersampled Isotropic Projection (VIPR) $k$-space sampling pattern achieves lower $T_1$ estimation error than SOS. VIPR has the advantage of larger incoherence because it spreads the undersampling artifacts in three dimensions compared to two dimensions for SOS, and larger incoherence is known to typically give better CS performance [11]. However, the anisotropic field of view (FOV) of the brain (the shape of the brain is not a perfect sphere) means that SOS has an initial advantage since unlike VIPR, it can be easily tailored to acquire data more

consistent with the shape of the brain, and therefore needs to acquire less data than VIPR, i.e., in VIPR, the same number of voxels need to be acquired in the $x, y, z$ dimensions whereas in SOS, fewer voxels can be acquired in $z$ dimension thus requiring a lower scan time. We investigate this dynamic of these two competing forces, namely incoherent artifacts versus anisotropic FOV. VIPR also has an advantage of increased robustness to motion compared to SOS. But since this benefit is hard to quantify, we do not compare these two techniques based on their robustness to motion.

## 3.1 Methods

### 3.1.1 Signal Model

In VFA, spoiled gradient echo (SPGR) data are acquired at several FAs. The relation between the signal in image-space and the various quantities of interest is now described. The image-space signal at a voxel $\mathbf{r}$ is given by,

$$x(\mathbf{r}, \alpha) = \frac{M_0(\mathbf{r})(1 - E_1(\mathbf{r}))\sin(\alpha)}{1 - E_1(\mathbf{r})\cos(\alpha)}, \tag{3.1}$$

where $\alpha$ is the flip angle (FA), $E_1(\mathbf{r}) = \exp(-\frac{T_R}{T_1(\mathbf{r})})$, $T_R$ is the repetition time which is known to the user and is constant for all FAs, $M_0(\mathbf{r})$ is the proton density. In Figure 3.1 the plot of $x(\mathbf{r}, \alpha)$ vs. $\alpha$ for different $T_1(\mathbf{r})$ is shown. For the range of possible $T_1(\mathbf{r})$ in the brain, $x(\mathbf{r}, \alpha)$ varies slowly with respect to $\alpha$, and this redundancy in the parametric dimension will be used later to construct regularizers in the parametric dimension.

$k$-space data is acquired as described in Section 2.4 with the parametric dimension $\theta$ being the flip angle (FA) $\alpha$. Here the acquisition time can be lowered by subsampling $k$-space at each

FA, or by acquiring fewer FAs or both. FA is a user defined parameter, and both the number and the values of the FAs are chosen to minimize scan time, while estimating $T_1$ accurately.

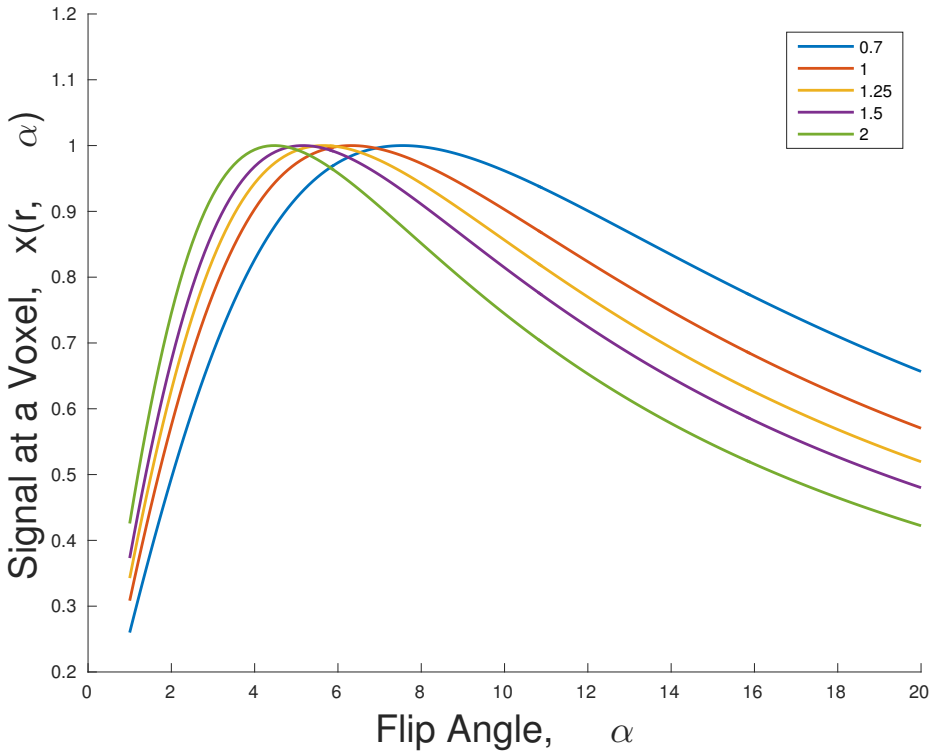


Figure 3.1: For a particular $\mathbf{r}$, a plot of $x(\mathbf{r},\alpha)$ (3.1) vs. $\alpha$ for the range of possible $T_1(\mathbf{r})$ in the brain. In all the curves, $x(\mathbf{r},\alpha)$ varies slowly w.r.t. $\alpha$ and hence its derivative w.r.t. $\alpha$ is small. For visualization, the different curves are normalized so that the peaks of $x$ for all $T_1$ are the same.

### 3.1.2 Dataset

In-vivo data of a healthy male adult subject is acquired in a GE scanner whose magnetic field strength is 3T. This is a multiple coil acquisition in which 32 coils uniformly distributed across the brain were used. The voxel size in image-space is $2\text{mm}^3$. The repetition time used is $T_R = 6.1ms$.

### 3.1.3 Estimating $x$

A goal of this project is to estimate image-space from a subset of $k$-space. For the VIPR (SOS) case, the $k$-space at each FA is a subsampled VIPR (SOS) sampling pattern as described in Section 2.6. To estimate image-space $x$ from the subsampled $k$-space $y$, we solve the following optimization problem,

$$\min_x ||E(x)-y||^2 + \lambda ||T(x)||_1, \tag{3.2}$$

where $E, x, y$ are described in Section 2.4, $T(x) = \{(T(x))(\mathbf{r},\alpha)\}_{\mathbf{r},\alpha}$ is a function defined on 4D grid (3 spatial dimensions and 1 FA dimension), and for a particular voxel $\mathbf{r}$, $(T(x))(\mathbf{r},\alpha)$ is the discrete derivative of $x(\mathbf{r},\alpha)$ with respect to $\alpha$, and $\lambda$ is a regularization parameter. As discussed earlier, the slowly varying nature of $x$ with respect to $\alpha$ means that $T(x)$ has a small norm and so compressed sensing [15] suggests that a near perfect reconstruction of $x$ can be obtained as the solution to the optimization problem (3.2).

### 3.1.4 Estimating $T_1, M_0$

Once $x$ is estimated, $M_0(\mathbf{r})$ and $T_1(\mathbf{r})$ are estimated using $x(\mathbf{r},\alpha)$ as follows. Rearranging the terms in (3.1) we get,

$$\frac{x(\mathbf{r},\alpha)}{\sin(\alpha)} = \frac{x(\mathbf{r},\alpha)}{\tan(\alpha)} E_1(\mathbf{r}) + M_0(\mathbf{r})(1-E_1(\mathbf{r})). \tag{3.3}$$

This can be solved for $E_1(\mathbf{r})$ and $M_0(\mathbf{r})$ by acquiring data at multiple $\alpha$, and using least squares. $T_1(\mathbf{r})$ can then be estimated from $E_1(\mathbf{r})$ using the relation $E_1(\mathbf{r}) = \exp(-\frac{T_R}{T_1(\mathbf{r})})$.

In the noise free FS case, just two FAs are needed to accurately estimate $T_1$ and $M_0$. In the pres-

ence of noise, more FAs result in better estimates of $T_1$ and $M_0$. Even though $M_0$ is estimated using the above approach, only the accuracy of $T_1$ is validated because this is the quantity of interest to most researchers.

In the Appendix, we show that in the presence of noise, the expected value of the $T_1$ estimate in the presence of noise is larger than than the true $T_1$ for all choices of acquisition parameters like flip angles ($\alpha$), repetition time ($T_R$), noise standard deviation ($\sigma$). This means that the estimate of $T_1$ should be divided by a correction factor so that an unbiased estimate is obtained. This has the additional advantage of reducing the variance of the estimate, and therefore the overall mean square error of the estimate is reduced when this correction factor is used. This correction factor can be estimated using Monte-Carlo simulations if the standard deviation of the noise $\sigma$ is known, and this $\sigma$ can be estimated using data outside the brain where only noise is present.

### 3.1.5 Choosing the Flip Angles

The flip angles (FA) are a user defined parameter, and we have control on both the number and the values of the FAs. Our goal is to choose FAs so as to accurately estimate $T_1$ while minimizing the scan time. Acquiring a large number of FAs leads to very accurate $T_1$ estimates while at the same time increasing the scan time. Even though in theory any positive FA can be used, in practice the maximum FA is limited to 20 degrees since the SNR is low above this value. For a given number of $k$-space samples, the acquisition time for any FA in the range $[1, 2, \ldots, 20]$ is the same. For the CS algorithm to work well, the derivative needs to be sparse, and so the FAs will be chosen relatively closely spaced. In Figure 3.1, notice that the signal for different $T_1$ varies rapidly for FA in the range 3 to 10. So to get a good approximation of

the derivative, more FAs need to be chosen in this range while FAs outside this range can be coarsely sampled since the signal is approximately linear. To avoid acquiring a large number of FA in the range of 3 to 10, we avoid taking any FA in this range. This is justified because for all $T_1$, the signal value at FA = 3 is close to that at FA = 10 for all $T_1$, and hence the signal change between these two FA is small for all $T_1$. In our experiments, FA $= [2,3,10,12,13,15]^0$ are acquired. For simplicity, the same number of $k$-space samples are chosen at each FA, even though we have the flexibility of choosing them to be different.

### 3.1.6 Validation

The accuracy of the $T_1$ estimate is validated by comparing $T_1$ obtained using a reconstruction technique (R) ($T_1^{(R)}$) with that obtained using fully sampled (FS) ($T_1^{(FS)}$) data. R could be either the VIPR or SOS case. The error is measured by,

$$e = 100 \times \sum_{\mathbf{r}} \left| \frac{T_1^{(R)}(\mathbf{r}) - T_1^{(FS)}(\mathbf{r})}{T_1^{(FS)}(\mathbf{r})} \right|, \tag{3.4}$$

where the summation is over all voxels in the white and gray matter in the brain. Voxels in the CSF tissue are not included in this computation since these voxels are not of interest in most clinical applications. CSF voxels are identified as those for which $T_1^{(FS)}$ is greater than 3000ms. Acquisition parameters like FA, number of $k$-space samples, etc. are chosen such that this error is less than 10%.

## 3.2 Results

The image-space is reconstructed from subsampled $k$-space and then $T_1$ maps are estimated from this reconstructed image-space signal. This procedure is done for both VIPR and SOS $k$-space sampling patterns for different acceleration factors. Here acceleration refers to the amount of $k$-space undersampling w.r.t. the fully sampled radial (VIPR or SOS) dataset. For example in VIPR if acceleration = 4, then the number of $k$-space samples used is $\frac{1}{4}$th of the fully sampled VIPR dataset. In Figure 3.2 the error is plotted as a function of the acceleration for the VIPR and SOS cases. For moderate to large accelerations, the error in VIPR is lower than SOS while for low accelerations the reverse is true. The error measure used is the relative percentage error given by (3.4). In Figure 3.3 the error is plotted as a function of the scan time (in seconds) in VIPR and SOS cases. For most scan times, the error in SOS is lower than VIPR, while for very short scan times, the reverse is true. In Figure 3.4 - 3.6, the $T_1$ maps of VIPR and SOS $k$-space sampling patterns at a number of different scan times, and the FS $T_1$ map are shown. The difference between the $T_1$ maps of different $k$-space sampling patterns at a number of different scan times, and the FS $T_1$ map are shown in Figures 3.7 - 3.9. From Figure 3.3, for most scan times, SOS has a lower error than VIPR. This is also evident in the $T_1$ maps of Figure 3.4 - 3.6 and 3.7 - 3.9.

## 3.3 Discussion

Here a compressed sensing (CS) approach to speed up the acquisition of VFA data is discussed and the $T_1$ estimated in this accelerated acquisition is verified to be close to the fully-sampled (FS) $T_1$. In [12], the CS reconstruction approach discussed above was tested only for the SOS

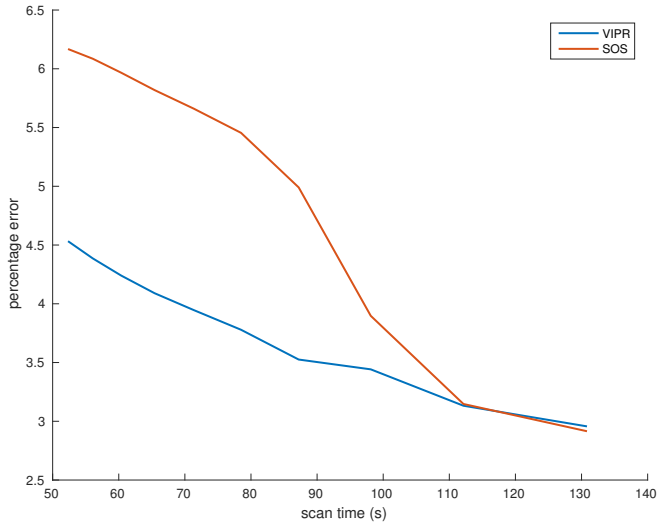


Figure 3.2: Error in $T_1$ estimates using VIPR and SOS radial $k$-space sampling strategies vs. acceleration i.e. $k$-space undersampling factor. Error measure used is the relative percentage error w.r.t. the ground truth and is computed using (3.4).

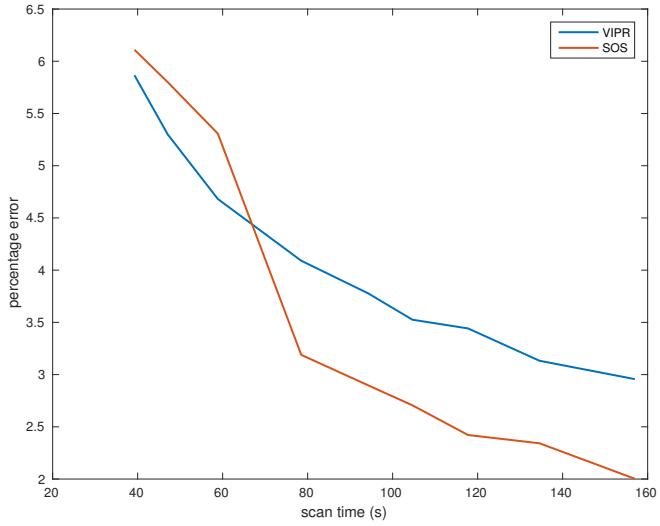


Figure 3.3: Error in $T_1$ estimates using VIPR and SOS radial $k$-space sampling strategies vs. scan time (in seconds). Error measure used is the relative percentage error w.r.t. the ground truth and is computed using (3.4).

$k$-space sampling case. Here we investigated if results can be improved by introducing undersampling in the third dimension using VIPR. In CS algorithms, $k$-space sampling patterns that

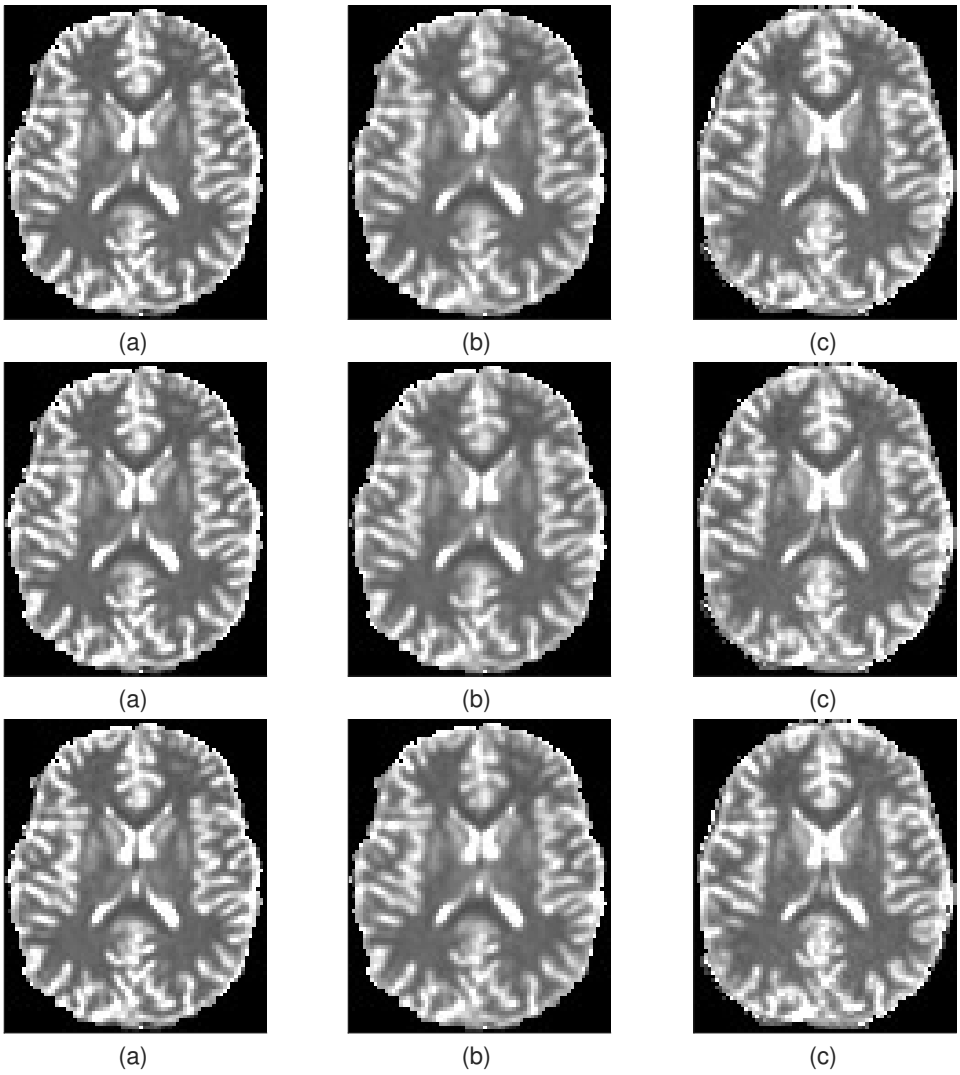


Figure 3.4: $T_1$ maps of different $k$-space sampling patterns (a) FS (b) VIPR (c) SOS. The scan time for both the subsampled VIPR and SOS datasets was identical. The three rows correspond to scan times of 156, 134, 117 seconds respectively.

are more incoherent give better performance [11]. As discussed earlier, we expected VIPR to be more incoherent than SOS. This incoherence explains why in Figure 3.3, VIPR outperforms SOS at low scan times. Despite the lower incoherence of SOS, in Figure 3.3 SOS outperforms VIPR at high scan times because the anisotropic field of view (FOV) of the brain means that fewer axial slices can be acquired in SOS while still achieving whole brain coverage, whereas

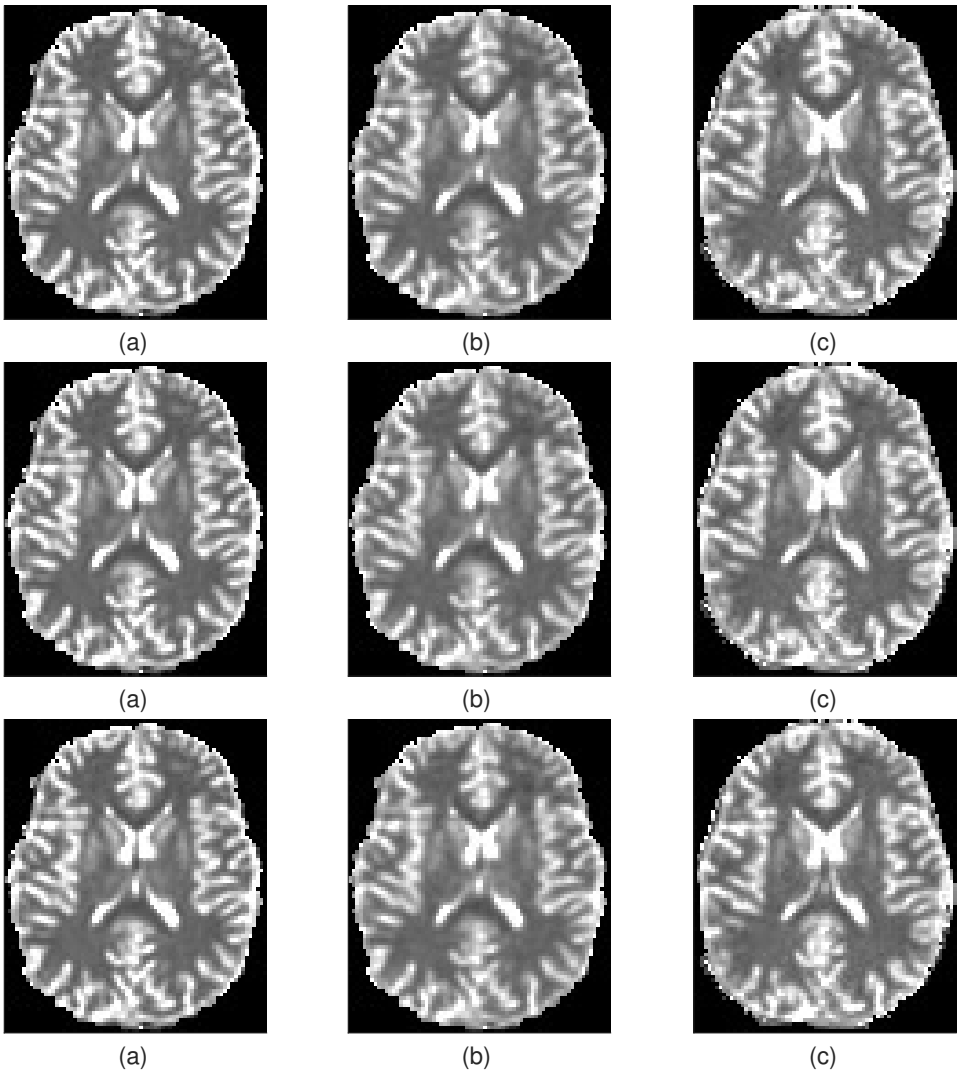


Figure 3.5: $T_1$ maps of different *k*-space sampling patterns (a) FS (b) VIPR (c) SOS. The scan time for both the subsampled VIPR and SOS datasets was identical. The three rows correspond to scan times of 104, 94, 78 seconds respectively.

in VIPR it is difficult to achieve asymmetrical FOV acquisition. However, at low scan times wherein very few *k*-space samples are acquired, the extra incoherence of the 3D sampling pattern of VIPR is more advantageous. If the object of interest is more spherical, then it is likely that VIPR will have lower error than SOS for all scan times since performance difference between the two cases will be based just on the incoherence of the sampling pattern and not on

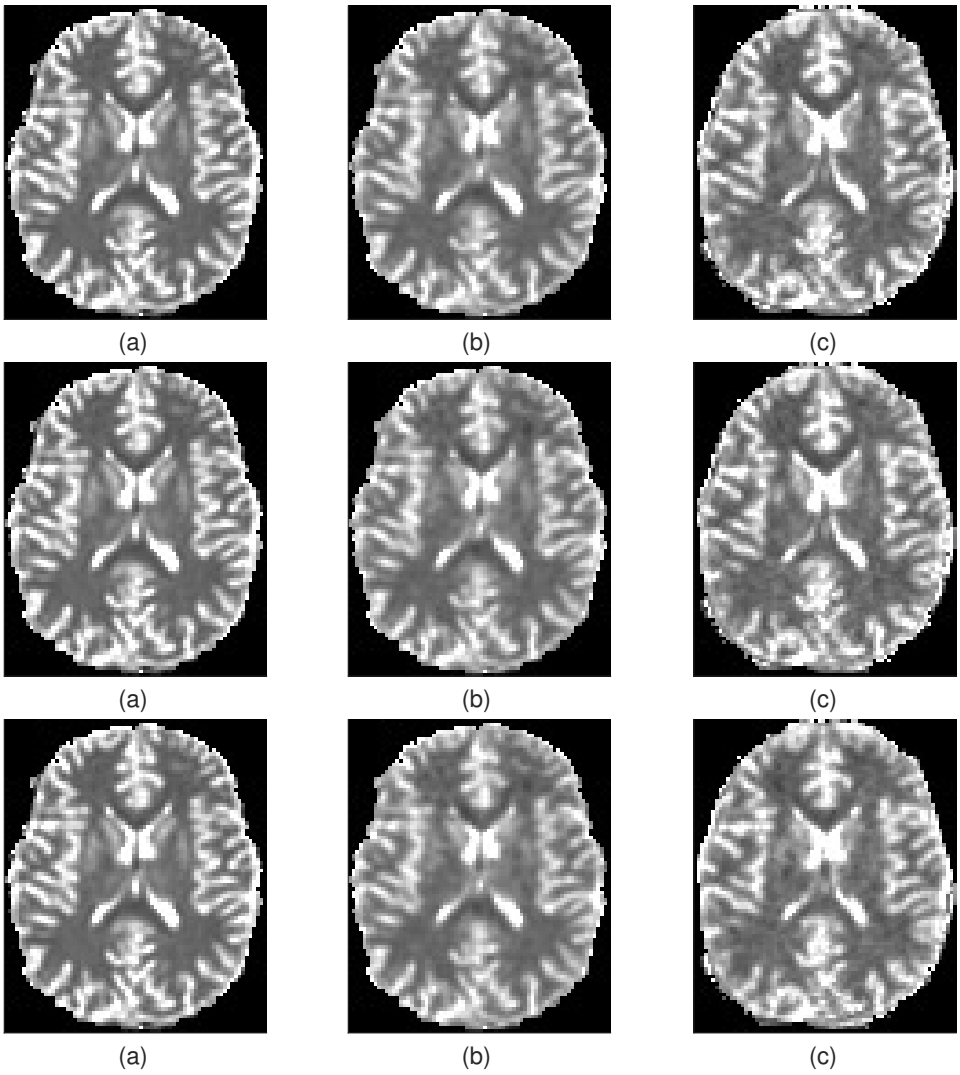


Figure 3.6: $T_1$ maps of different $k$-space sampling patterns (a) FS (b) VIPR (c) SOS. The scan time for both the subsampled VIPR and SOS datasets was identical. The three rows correspond to scan times of 58, 47, 39 seconds respectively.

FOV. Even though SOS outperforms VIPR at high scan times, its SNR is lower than VIPR since fewer $k$-space samples are acquired, and so the "quality" of the fully-sampled $T_1$ in SOS is worse than that of the fully-sampled $T_1$ of VIPR, and the same holds for the subsampled cases. So the earlier result that SOS sometimes outperforms VIPR needs to be interpreted judiciously. Sub-sampling in SOS was achieved by choosing fewer radial projections in each

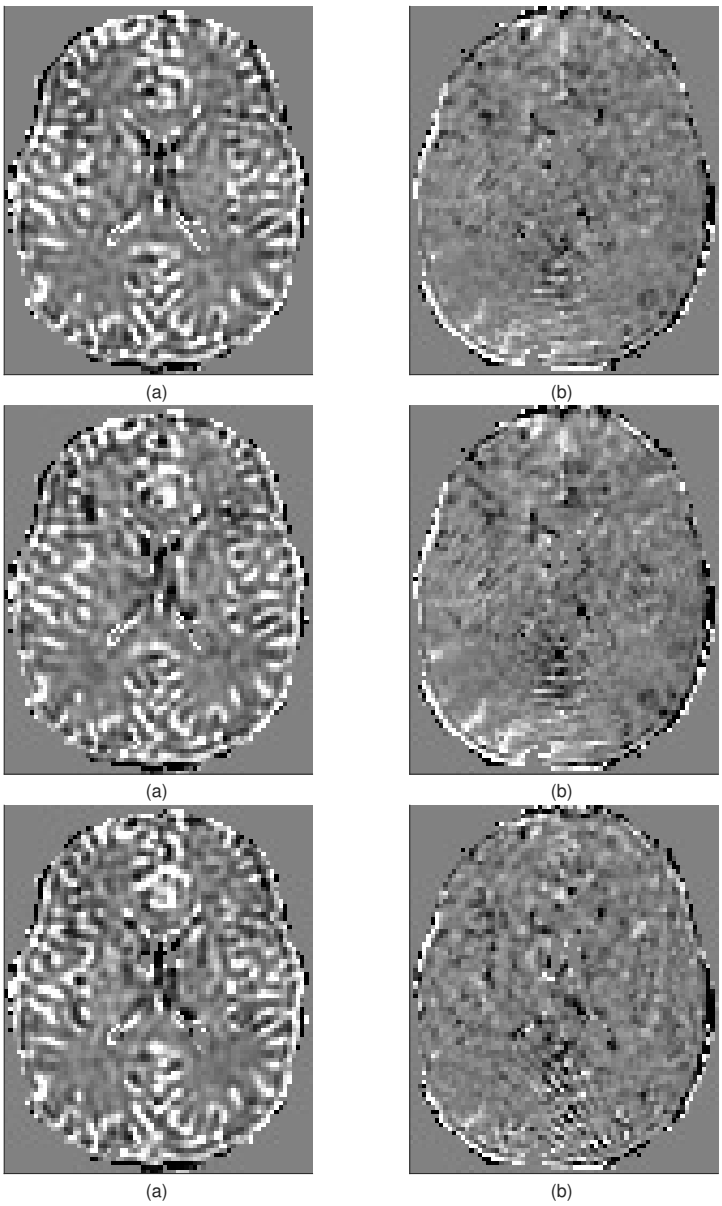


Figure 3.7: Difference between the $T_1$ maps of subsampled and fully-sampled cases for (a) VIPR (b) SOS. The scan time for both the subsampled VIPR and SOS datasets was identical. The three rows correspond to scan times of 156, 134, 117 seconds respectively.

plane while fully sampling the slice direction $z$. Also for each slice, the projections were exactly the same. Future work for improving SOS will involve subsampling the slice direction and also dithering the $k$-space projections for different slices to achieve larger incoherence.

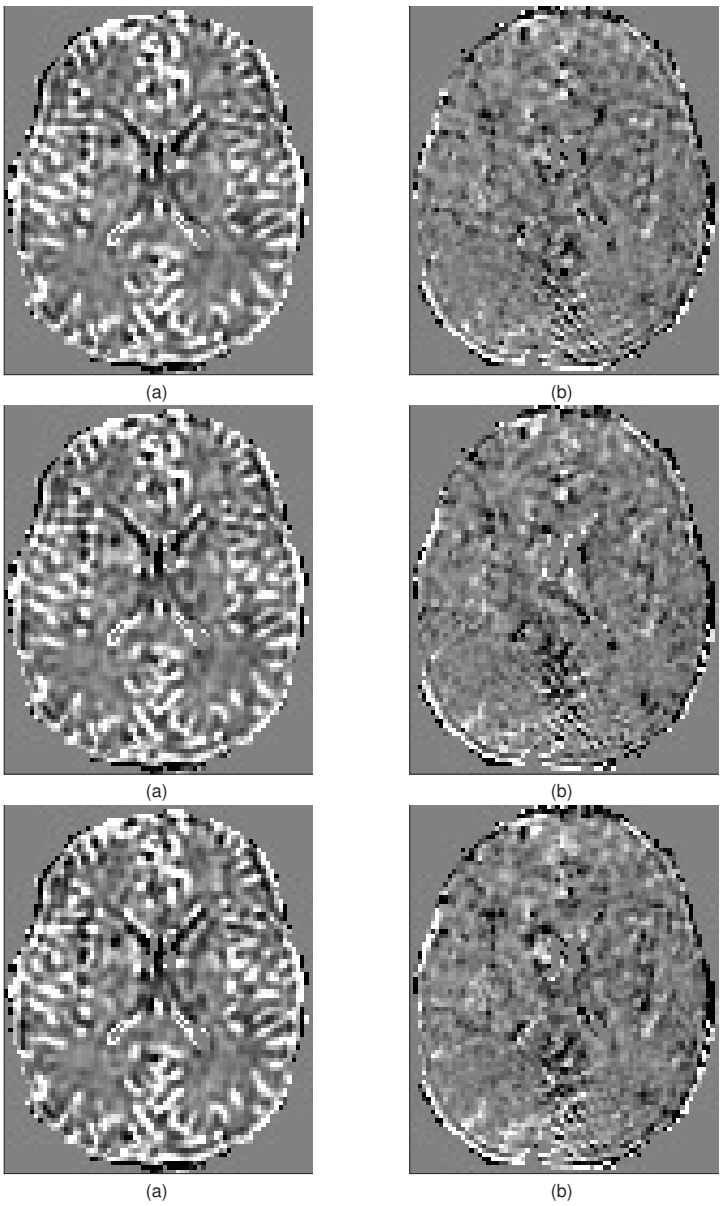


Figure 3.8: Difference between the $T_1$ maps of subsampled and fully-sampled cases for (a) VIPR (b) SOS. The scan time for both the subsampled VIPR and SOS datasets was identical. The three rows correspond to scan times of 104, 94, 78 seconds respectively.

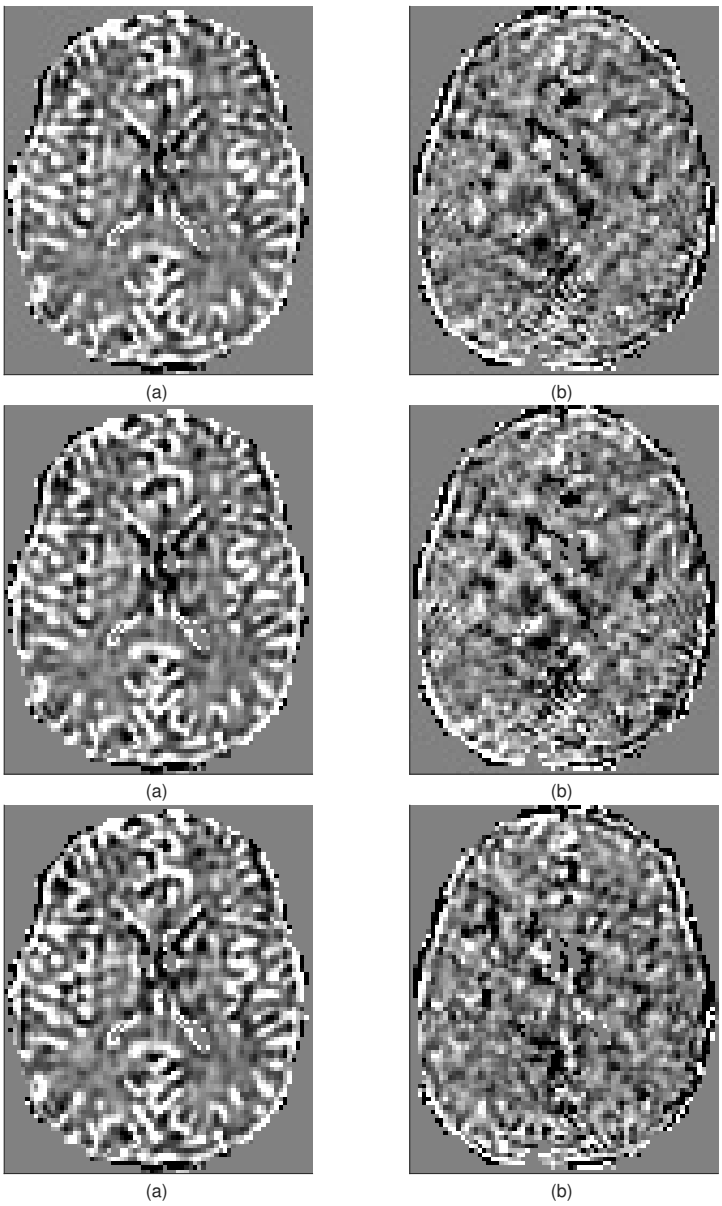


Figure 3.9: Difference between the $T_1$ maps of subsampled and fully-sampled cases for (a) VIPR (b) SOS. The scan time for both the subsampled VIPR and SOS datasets was identical. The three rows correspond to scan times of 58, 47, 39 seconds respectively.

# Chapter 4

# Inversion Recovery $T_1$ Imaging

Inversion Recovery (IR) $T_1$ imaging is an acquisition technique to perform quantitative $T_1$ estimation of all voxels in the brain. In this setup, $k$-space data are acquired at multiple inversion time points $t$, and our goal is to estimate $T_1$ of all voxels in image-space. At each of these inversion time points in image-space, the contrast between various brain tissues is different. In addition to estimating $T_1$, one other advantage of this acquisition method is that images of the brain can be obtained with the signals from specific tissues being equal to zero ("nulled"), and this can help in clinical diagnosis. Our primary goal in this project is to estimate $T_1$ of all voxels using the $k$-space data acquired at multiple inversion times $t$.

There are many approaches to acquire Inversion Recovery data, and we use the MPnRAGE model [16] that is also robust to head motion. The robustness is due to the fact that here separable "navigator" measurements are made every two seconds which can be used to measure and correct motion, and so the subject needs to remain still for just two seconds [16], whereas in other Inversion Recovery techniques (for example MP-RAGE), the subject needs to remain

still for the entire duration of the scan, which is several minutes. This makes MPnRAGE extremely attractive for applications involving a lot of motion like imaging small children.

Inversion Recovery is very time consuming due to the acquisition of multiple MRI images at different inversion time points $t$. Compressed sensing (CS) algorithms with spatial regularizers help to accelerate this acquisition but have limited benefit due to lack of spatial redundancy. However, in the MPnRAGE model, there is a large amount of redundancy in temporal dimension, offering the possibility of exploiting this redundancy using a regularizer along the temporal dimension. Other studies involving spatio-temporal MRI datasets have shown that temporal regularization outperforms spatial regularization [12] because of larger temporal redundancy.

One approach to exploit the temporal redundancy would be to use a temporal total variation (TV) regularizer. However, its utility is limited because TV produces staircase artifacts [13], i.e., it produces an output that is piecewise constant. Total generalized variation (TGV) has been used successfully as a spatial regularizer to eliminate the staircase artifacts typically observed with TV [13]. Spatial TGV has all of the advantages of TV in terms of enforcing spatial consistency of the signal, preserving edges, while eliminating drawbacks of TV like staircase artifacts [13]. This work attempts to extend many of these results in the temporal dimension. Here TGV is developed as a temporal regularizer that can be used in CS algorithms to speed up the acquisition of spatio-temporal MRI, and produce signals that do not have the staircase artifacts of a temporal TV regularizer. More specifically we hypothesize that TGV can be used as a temporal regularizer to speed up the acquisition of $T_1$ maps of the brain using the MPnRAGE protocol, and that TGV outperforms TV by eliminating the staircase artifacts.

Two kinds of datasets are used to validate the performance of the algorithm. Numerical phan-

toms and an in-vivo dataset of a healthy adult subject are used to compare the proposed TGV method to TV and Zero-Filled (ZF) reconstruction. All these different reconstruction procedures reconstruct the image-space signal from subsampled $k$-space, and once the image-space signal is reconstructed, $T_1$ at all voxels can be estimated from this and validated that this is indeed close to the true $T_1$ which we take to be the fully-sampled (FS) $T_1$.

TGV, in addition to performing well on the MPnRAGE model in which the signal is a continuous function that slowly varies across time, performs well on signals that have discontinuities. To test this hypothesis, a synthetic dataset that has discontinuities in its time series is used, and the performance of TGV is compared to the other reconstruction techniques described above.

## 4.1 Methods

To test the performance of the TGV algorithm, three kinds of simulation setups are used viz; phantom, in-vivo MPnRAGE models to test the performance on quantitative $T_1$ estimation, and a phantom discontinuous model to test the ability of TGV to preserve discontinuities.

**Phantom MPnRAGE**

A numerical phantom was used whose $T_1$ and proton density values are obtained by scanning the same subject 8 times using the MPnRAGE sequence [16] and averaging the $T_1$ and proton density values respectively. This gives $T_1$ and proton density maps that have a high SNR. Using these $T_1$ and proton density values, synthetic relaxometry data of a brain scan were generated using a 2D radial Look-Locker sequence, a.k.a. MPnRAGE model with 386 inversion time frames, and the image-space signal at each voxel is given by (4.1). At each time point, the

Fourier transform of the image-space signal is computed at the radial $k$-space locations shown in Figure 2.4 to get the $k$-space data. The $k$-space locations are different at all time points. The radial lines in $k$-space are interleaved both within a single inversion recovery acquisition and also across multiple inversion recovery acquisitions [16]. This means that both at a particular inversion time point $t$, and across neighboring inversion time points, the radial $k$-space projections are maximally separated from one another, similar to that shown in Figure 4.2.

The voxel size of this phantom is $1\text{mm}^3$. The $T_R$ used is 7ms and this determines the time duration between successive inversion times as described later. A flip angle of 4 was used. This is a single coil simulation setup in which the coil sensitivity is equal to one for all voxels. The $T_1$ maps used to create the phantom were acquired in a scanner whose magnetic field strength was 3T.

**In-Vivo MPnRAGE**

In-vivo data of a healthy male subject was acquired using a 2D radial Look-Locker sequence, a.k.a. MPnRAGE [16] that produces 331 inversion time points. At each inversion time point, 202 radial $k$-space projections were acquired, and this is our fully-sampled (FS) dataset. The radial $k$-space projections are interleaved similar to the phantom case described earlier. Subsampled data was chosen by selecting a subset of these radial projections.

The in-plane resolutions of the voxel is $1\text{mm}^3$ and the slice thickness is $3\text{mm}^3$. The $T_R$ used is 5.2ms and this determines the time duration between successive inversion times as described later. This is a multiple coil acquisition in which 8 coils are used. The scanner magnetic field strength is 3T.

**Temporally Discontinuous Phantom**

To illustrate the ability of TGV to preserve discontinuities, a synthetic phantom with an image-space time course having occasional sudden jumps was used. This was generated as follows. First data was generated using the phantom MPnRAGE model as before. Then the signal amplitudes for time points between 100 to 200 are doubled. This produces the sharp discontinuities as shown in Figure 4.16. The radial $k$-space projections are interleaved as described earlier.

### 4.1.1 MPnRAGE Signal Model

In both the phantom and in-vivo MPnRAGE cases, as described in [16], the signal at a voxel $\mathbf{r}$ in image-space is given by,

$$x(\mathbf{r},t) = A(\mathbf{r}) - B(\mathbf{r})\exp(-\frac{t}{T_1(\mathbf{r})}), \tag{4.1}$$

where $A(\mathbf{r}),\ B(\mathbf{r}),\ T_1(\mathbf{r})$ are unknown and need to be estimated for all $\mathbf{r}$ in the brain. This estimation is done using the nonlinear least-squares approach described in [16]. Inversion time $t$ is a discrete valued, uniformly spaced parameter, and the spacing between adjacent $t$ is the repetition time $T_R$ which is 7ms for the phantom dataset and 5.2ms for the in-vivo dataset. In the phantom case, 386 such time points are used and in the in-vivo case, 331 such time points are used. Typically the ratio $\frac{A(\mathbf{r})}{B(\mathbf{r})} \approx \frac{1}{2}$. In Figure 4.1 the plot of $x(\mathbf{r},t)$ vs. $t$ for the range of possible $T_1(\mathbf{r})$ in the brain is shown. For all $T_1$ of interest, $x$ varies slowly w.r.t. $t$ since the sampling period is a few ms, and this means that $x$ has a lot of temporal redundancy which is exploited in our algorithm.

$k$-space data is acquired as described in Section 2.4 with the parametric dimension $\theta$ being

time $t$, i.e. at each time point $t$, $k$-space radial projections as shown in Figure 4.2 are acquired, with the number of radial projections at each time point being identical. In the subsampled dataset, the same number of time points are acquired but at each time point, fewer $k$-space radial projections are acquired.

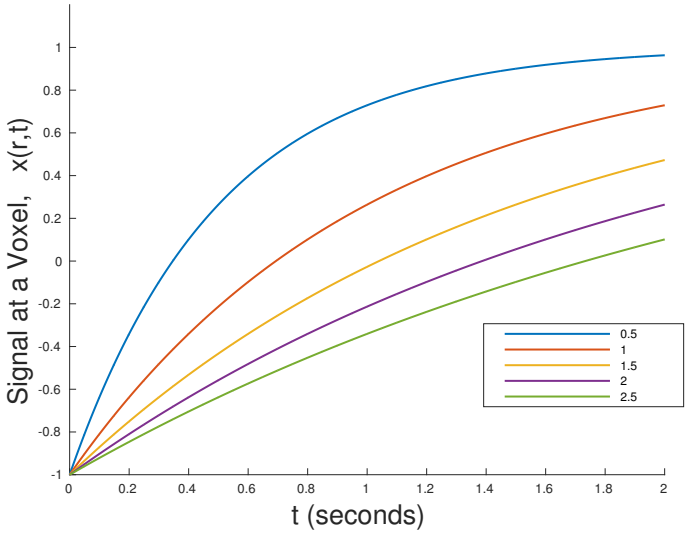


Figure 4.1: For $A(\mathbf{r}), B(\mathbf{r})$ fixed, using the MPnRAGE model equation (4.1), a plot of $x(\mathbf{r},t)$ vs. $t$ (seconds) for the range of possible $T_1(\mathbf{r})$ in the brain. Clearly $x(\mathbf{r},t)$ varies slowly w.r.t. $t$ for all $T_1(\mathbf{r})$ since the spacing between two adjacent time points is approximately 5ms.

### 4.1.2 Estimating $x$

The TGV algorithm to estimate the image-space signal $x$ from subsampled $k$-space data $y$ enforces smoothness of $x$ across time $t$ and is given by the following optimization problem,

$$\min_x ||E(x)-y||^2 + \min_z \ \lambda_1 \sum_{\mathbf{r}} ||\frac{\partial x(\mathbf{r},t)}{\partial t} - z(\mathbf{r},t)||_1 + \lambda_0 \sum_{\mathbf{r}} ||\frac{\partial z(\mathbf{r},t)}{\partial t}||_1, \tag{4.2}$$

where $E, x, y$ are described in the Section 2.4, $z$ is a dummy variable in image-space and has the same dimensions as $x$, and $\lambda_1, \lambda_0$ are regularization parameters. The penalty term in (4.2)

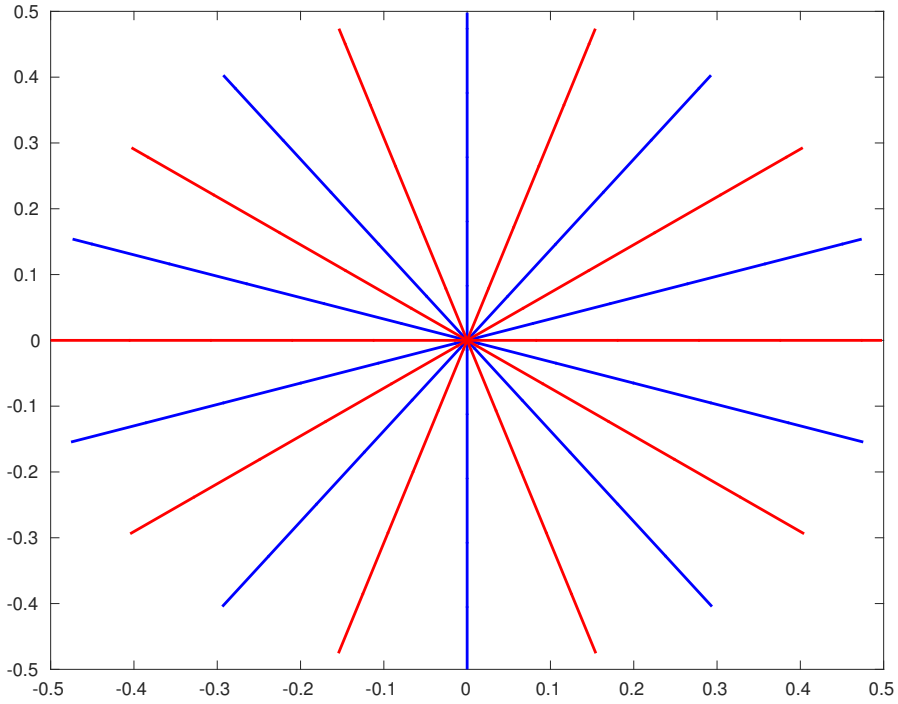


Figure 4.2: Interleaved $k$-space sampling pattern used in MPnRAGE $T_1$ acquisition for two adjacent time points. The $k$-space sampling pattern for the first time point is in blue and for the second time point is in red. The $k$-space sampling pattern for each time point has five radial projections. The interleaved sampling pattern results in the projections being maximally separated both within a time point and also across time points.

is $\mathrm{TGV}(x)$, and as described in [13], TGV enforces a smoothness of $x(\mathbf{r},t)$ across time $t$ for all voxels $\mathbf{r}$, and does not have staircase artifacts because of the following. For a particular $\mathbf{r}$, in smooth regions of $x(\mathbf{r},t)$, the second derivative $\frac{\partial^2 x(\mathbf{r},t)}{\partial t^2}$ is locally small. It is therefore likely that the minimizer of $\mathrm{TGV}(x)$ will be $z(\mathbf{r},t) = \frac{\partial x(\mathbf{r},t)}{\partial t}$ in these regions, resulting in $\mathrm{TGV}(x) = ||\frac{\partial^2 x(\mathbf{r},t)}{\partial t^2}||_1$ in smooth regions and hence does not produce staircase artifacts. On the other hand, in the neighborhood of edges, $\frac{\partial^2 x(\mathbf{r},t)}{\partial t^2}$ will be much larger than $\frac{\partial x(\mathbf{r},t)}{\partial t}$, so it is beneficial for the minimization of $\mathrm{TGV}(x)$ to choose $z = 0$, resulting in $\mathrm{TGV}(x) = ||\frac{\partial x(\mathbf{r},t)}{\partial t}||_1 = \mathrm{TV}(x)$ near edges. Since TV is known to preserve edges [17], it is expected that TGV preserves edges as well.

Like TGV, the TV approach enforces the constraint that the image-space signal at any two adjacent time points are similar. Here, for each voxel $\mathbf{r}$, we compute the derivative of $x(\mathbf{r},t)$ w.r.t. $t$, then compute the $l_1$ norm of this derivative, and use this as a regularizer in an optimization problem. This enforces the constraint that the derivative of the image-space signal at each voxel is sparse, thereby enforcing the constraint that the image-space signal at any two adjacent time points are similar. To estimate the TV solution, we solve the following optimization problem,

$$\min_x ||E(x)-y||^2 + \lambda \sum_{\mathbf{r}} ||\frac{\partial x(\mathbf{r},t)}{\partial t}||_1, \tag{4.3}$$

To reconstruct image-space $x$ from subsampled $k$-space $y$, (4.2) and (4.3) are solved using the primal-dual algorithm described in [13] to obtain the TGV and TV solutions respectively.

### 4.1.3 Validation

For the different reconstruction approaches, even though image-space signal $x$ is estimated, the performance of the algorithm is not measured in terms of the accuracy of the image-space signal $x$ but in terms of accuracy of $T_1$ which is a function of $x$. Once $x$ is reconstructed, using the model equation, parameters $A(\mathbf{r})$, $B(\mathbf{r})$, and $T_1(\mathbf{r})$ are estimated for all $\mathbf{r}$ voxel-wise using the non-linear least squares approach described in [16]. The FS dataset was used as a ground truth to compare the TGV, TV and ZF results to. The error is measured as (3.4), where the summation is over all voxels in the brain, and $T_1^{(FS)}$ and $T_1^{(R)}$ are the $T_1$ values of FS and of different reconstruction techniques (i.e., TGV, TV, and ZF) respectively. The largest acceleration that gives an error that is less than 10% is an optimal acceleration factor. This validation approach is used for both the phantom and in-vivo MPnRAGE datasets.

For the discontinuous phantom case, the procedure discussed earlier for the reconstruction of image-space from $k$-space is carried out and validation is done visually, by looking at the time series of different voxels. A good reconstruction procedure would perform well in both smooth regions of the signal and at discontinuities as well.

## 4.2 Results

The performance of different reconstruction techniques is compared by numerically evaluating the accuracy of the $T_1$ maps, visually inspecting the accuracy of the reconstructed source images $x$ and the $T_1$ maps, and visually inspecting the voxel time series. Three different simulation setups are discussed, namely, a phantom MPnRAGE model, an in-vivo MPnRAGE model, and a phantom discontinuous model.

**Phantom MPnRAGE**

In the phantom MPnRAGE model, image-space is reconstructed from subsampled $k$-space and then $T_1$ maps are estimated from this image-space signal. The $T_1$ maps estimated are compared with the ground truth $T_1$ maps and the error is computed. In Figure 4.3 the error of the different reconstruction techniques is plotted as a function of the scan time. TGV has the lowest error at all accelerations followed by ZF and then TV. The separation between the curves reduces as the scan time increases because we move closer to a FS acquisition and so all cases have a near perfect reconstruction. The time series at an arbitrarily chosen white matter voxel for the different reconstruction techniques namely FS, TGV, TV, ZF is shown in Figure 4.4. The $y$-axis in this figure is in arbitrary units (AU) and the $x$-axis is the time point / frame number. For the subsampled cases, an acceleration, i.e., the $k$-space undersampling factor of 133 is used to

generate the curves in Figure 4.4. This acceleration corresponds to a scan time of 8 seconds for each 2D slice. Here ZF looks like the FS with Gaussian noise added, TV has staircase artifacts and TGV eliminates these staircase artifacts. The same voxel time series curves with the FS time series curve overlaid are plotted in Figure 4.5. This helps to better examine the similarity between the true time series and the time series of different subsampled reconstructions. In this figure TGV is closest to the FS followed by TV and ZF. Despite the fact that the time series of TV is closer to FS than ZF, Figures 4.3 and 4.7 indicate that for the $T_1$ maps estimated from these time series, ZF is closer to FS than TV. This is due to the fact that $T_1$ fitting is not robust to non-Gaussian noise artifacts, and since TV has staircase artifacts, the $T_1$ estimated will be extremely erroneous. The $T_1$ maps of different reconstruction techniques are shown in Figure 4.6. The same acceleration factor as earlier is used to generate this figure. The $T_1$ map of TGV is closest to FS and also preserves information about fine structures in the $T_1$ map. Even though the ZF source images have more streak artifacts than TV, the $T_1$ map of ZF is still closer to the ground truth (i.e., FS) $T_1$ than TV. This is also evident in the difference map shown in Figure 4.7. The difference of the $T_1$ maps of the different reconstruction techniques and the ground truth, i.e., FS is shown in Figure 4.7. The same acceleration factor as earlier is used to generate this figure. In this figure, TGV is closest to the FS followed by TV and ZF. This follows from the error plot shown in Figure 4.3. The estimated source images $x$ are shown in Figure 4.8. These are the solutions to (4.2), (4.3) for TGV and TV respectively, and the ZF and FS are computed as described earlier. This $x$ is a time series of volumes and in Figure 4.8 the $180^{\text{th}}$ time point is chosen because it produces the best contrast between different tissues, and potential artifacts like blurring of edges between different tissue boundaries can be easily observed. In this figure (and in unpublished figures of other time points), the TGV and TV images look very similar to the FS while the ZF is extremely noisy and looks nothing like the original. Despite the extremely poor quality of ZF relative to TV, the $T_1$ map of ZF is better

than TV as discussed previously.

Thus, Figures 4.4-4.8 indicate that the temporal TGV regularizer performs the best among all reconstruction techniques since it outperforms other reconstructions on all fronts, $T_1$ map, voxel time series, and source images $x$.

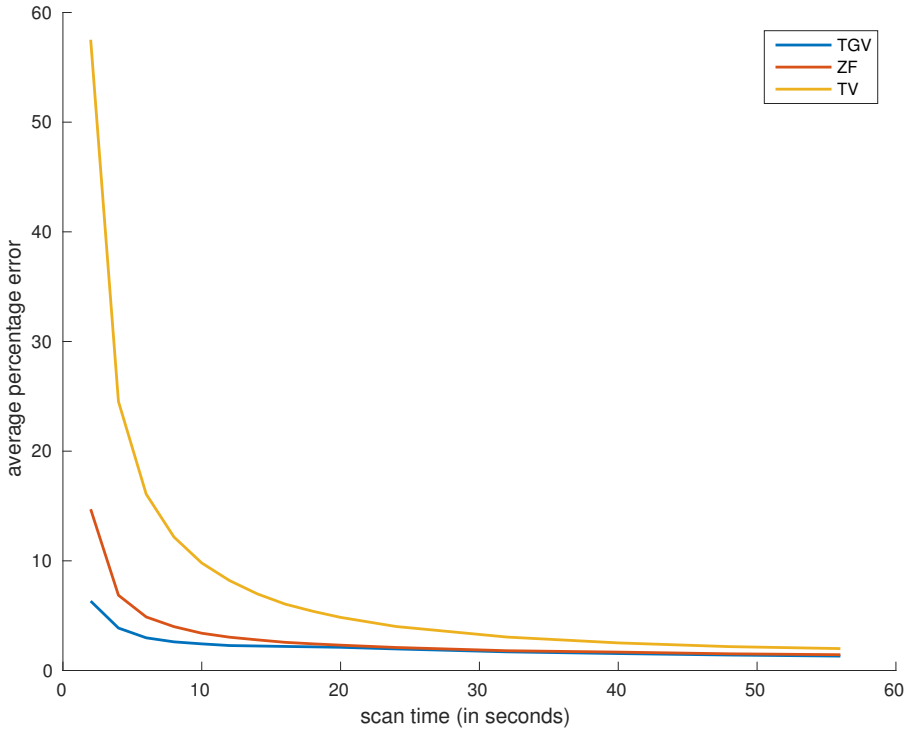


Figure 4.3: For the phantom MPnRAGE dataset, for a single 2D slice, $T_1$ error as a function of scan time (in seconds) for TGV, TV and ZF reconstruction techniques. TGV is obtained by solving (4.2) with $\lambda_1, \lambda_0 \neq 0$ and ZF is obtained using the approach described in Section 2.6 The error measure used is the relative error and is computed using (3.4).

**In-vivo**

In the in-vivo MPnRAGE model, image-space is reconstructed from subsampled $k$-space and then $T_1$ maps are estimated from this image-space signal. The $T_1$ maps estimated are compared with the ground truth $T_1$ maps and the error is computed. In Figure 4.9 the error of the different

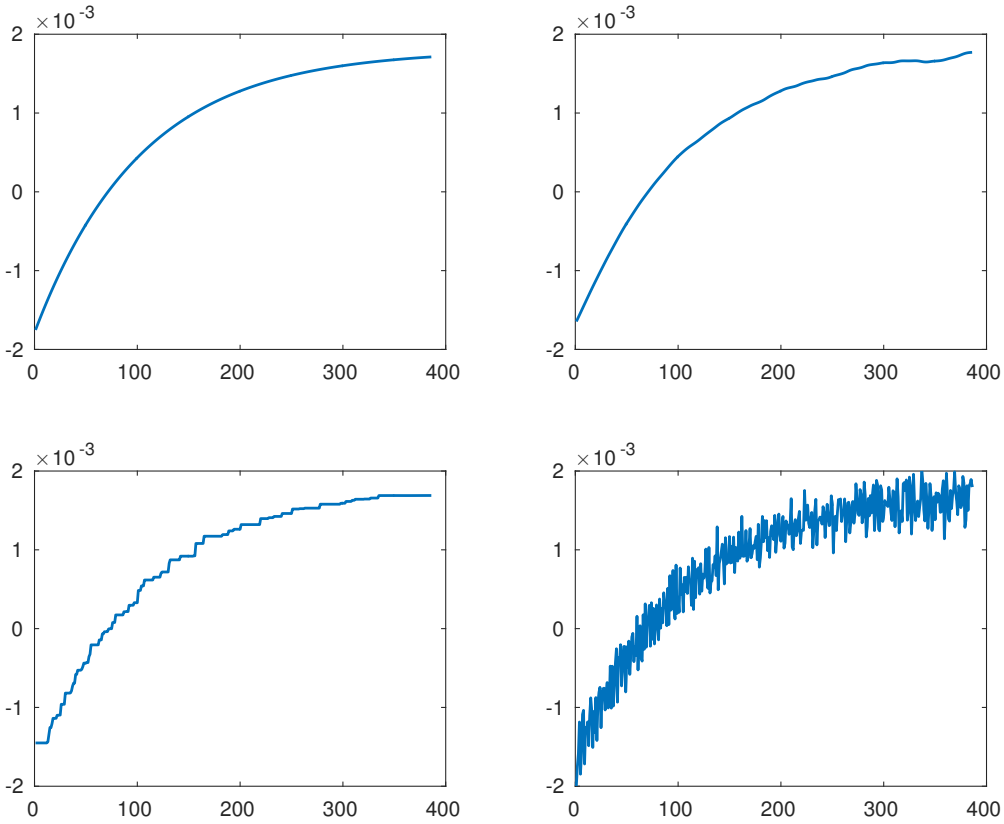


Figure 4.4: For the phantom MPnRAGE dataset, time series at a white matter voxel for different reconstruction techniques, (a) FS, (b) TGV, (c) TV, (d) ZF. An acceleration factor of 133 is used for the subsampled cases.

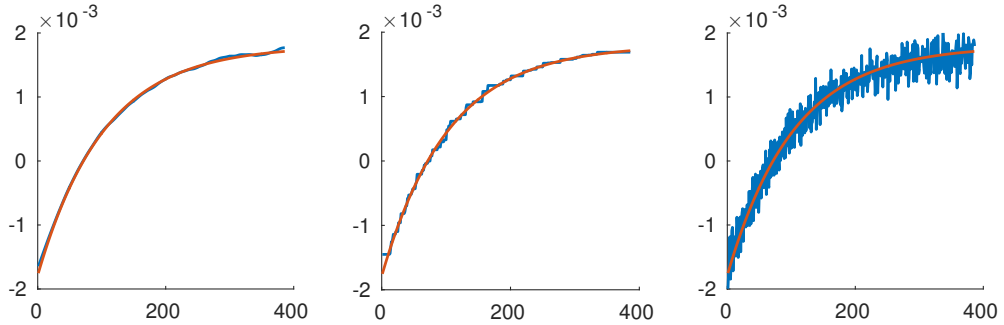


Figure 4.5: For the phantom MPnRAGE dataset, time series at a white matter voxel for different reconstruction techniques, (a) TGV, (b) TV, (c) ZF, overlaid with FS (in red). An acceleration factor of 133 is used for the subsampled cases.

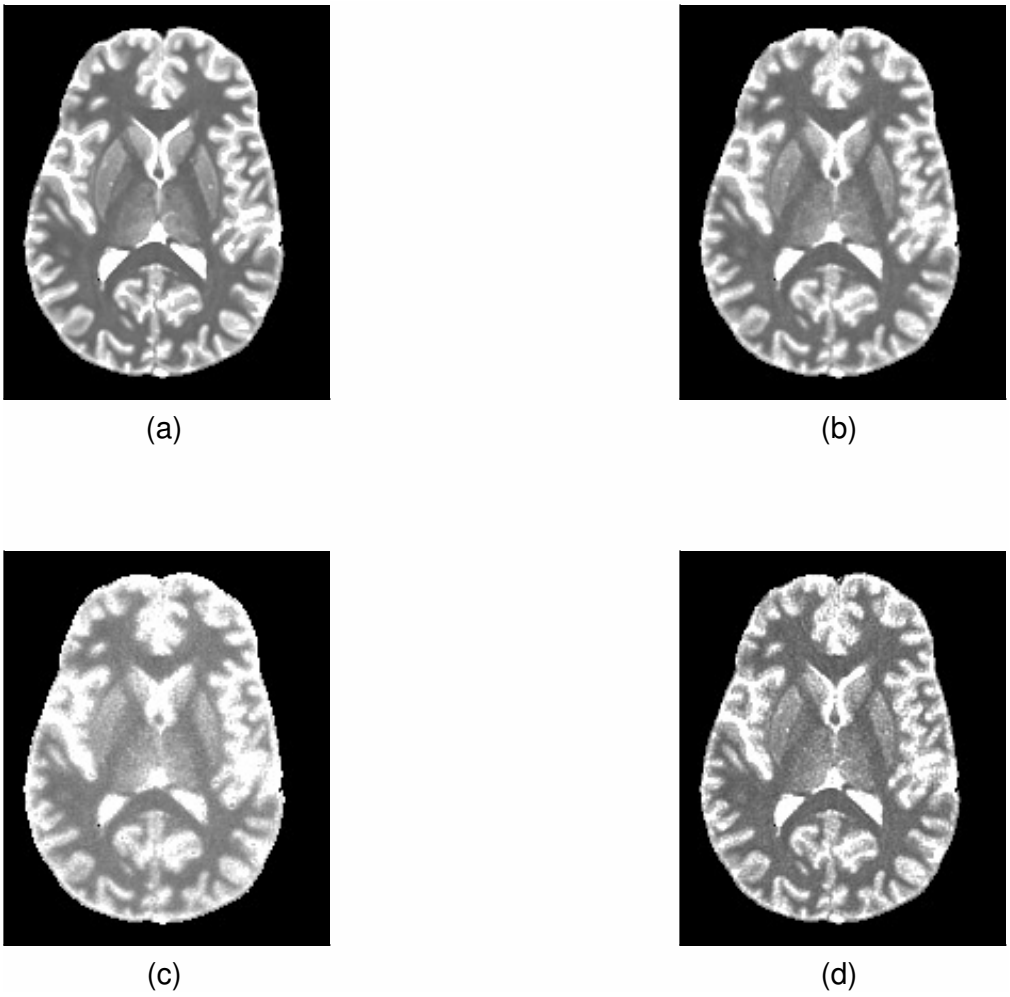


Figure 4.6: For the phantom MPnRAGE dataset, $T_1$ maps (in ms) for different reconstruction techniques, (a) FS, (b) TGV, (c) TV, (d) ZF. An acceleration factor of 133 is used for the subsampled cases.

reconstruction techniques is plotted as a function of the scan time. TGV has a lower error at all accelerations compared to ZF. The time series at an arbitrarily chosen white matter voxel for the different reconstruction techniques namely FS, TGV, ZF is shown in Figure 4.10. The $y$-axis in this figure is in arbitrary units (AU) and the $x$-axis is the time point / frame number. For the subsampled cases, an acceleration, i.e., the $k$-space undersampling factor of 33 is used to generate the curves in Figure 4.10. Here ZF looks like the FS with Gaussian noise added

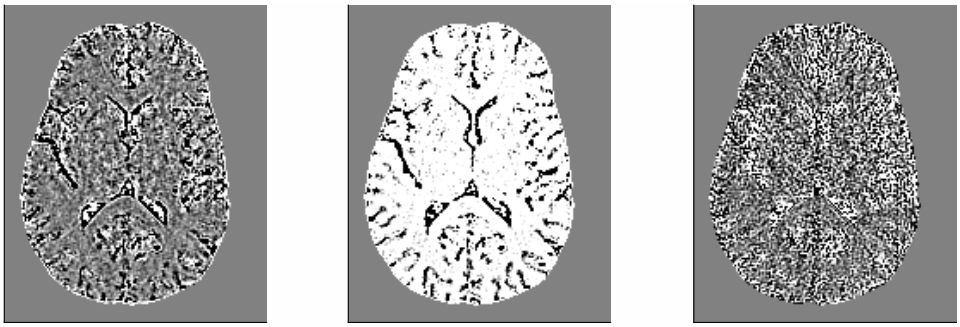


Figure 4.7: For the phantom MPnRAGE dataset, difference (in ms) between the $T_1$ of different reconstruction techniques, and the ground truth $T_1$. (a) TGV – FS, (b) TV – FS, (c) ZF – FS. An acceleration factor of 133 is used for the subsampled cases.

because of the interleaved radial $k$-space sampling, and TGV has no staircase artifacts. The same voxel time series curves with the FS time series curve overlaid are plotted in Figure 4.11. The $T_1$ maps of different reconstruction techniques is shown in Figure 4.12. The difference of the $T_1$ maps of the different reconstruction techniques and the ground truth, i.e., FS is shown in Figure 4.13. The same acceleration factor as earlier is used to generate these figures. Despite the fact that the error plot in Figure 4.9 indicates that TGV outperforms ZF, visual inspection of the $T_1$ maps, does not clearly indicate if TGV outperforms ZF. This is possibly due to the fact that low SNR of the FS image results in small differences being buried in noise. The denoised source images are shown in Figure 4.14 and are obtained as follows. First (4.2) is solved for $x$. Using this solution $x$, (4.1) is solved for $A, B, T_1$. Using these estimates, the source images $x$ are regenerated using (4.1). Because the source images are generated using this denoising procedure, these are called denoised source images. This denoised $x$ is a time series of volumes and in Figure 4.14 the 150th time point is chosen because it produces the best contrast between different tissues and potential artifacts like blurring of edges between different tissue boundaries can be easily observed. In this figure (and in unpublished figures of other time points), the TGV looks very similar to the FS while the ZF is extremely noisy

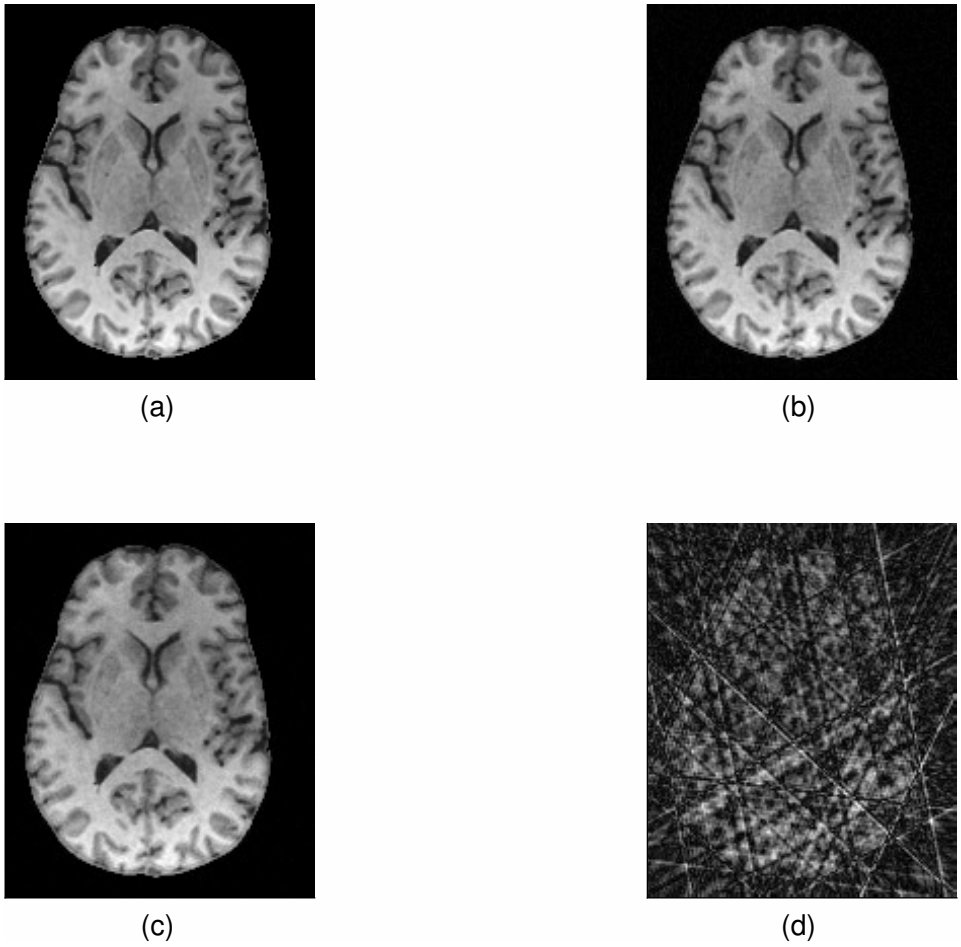


Figure 4.8: For the phantom MPnRAGE dataset, source images $x$ (in AU) of different reconstruction techniques, (a) FS, (b) TGV, (c) TV, (d) ZF. An acceleration factor of 133 is used for the subsampled cases.

and looks nothing like the original. Despite the extremely poor quality of the source images of ZF relative to TGV, the $T_1$ map of ZF is much closer to TGV than the source images suggest. The difference between the source images and the ground truth source images in Figure 4.15. Here clearly TGV outperforms ZF since the difference image in case of TGV is approximately Gaussian noise while in the ZF case, the difference image has a lot of brain structure. So in applications in which just the source images are desired, TGV is still preferable to ZF.

Thus, Figures 4.9-4.15 indicate that for in-vivo data, temporal TGV regularizer outperforms ZF on all fronts, $T_1$ map, voxel time series, and source images $x$.

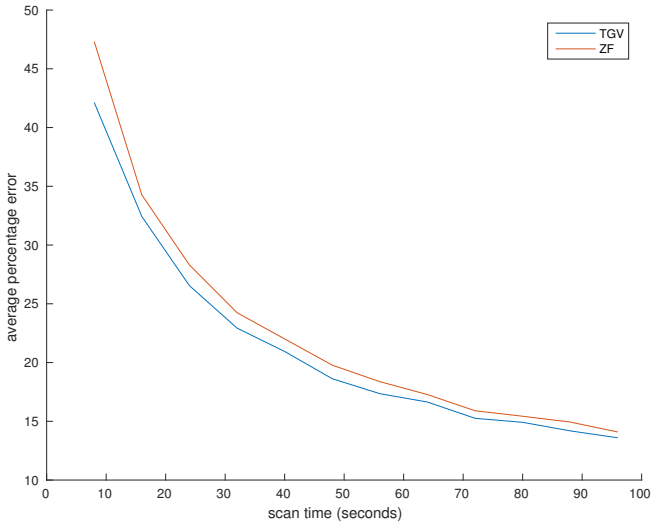


Figure 4.9: For the in-vivo MPnRAGE dataset, plot of the percentage error in the $T_1$ values using different reconstruction techniques, TGV and ZF, versus scan time. The error measure used is the absolute relative error averaged across the brain and is given by (3.4).

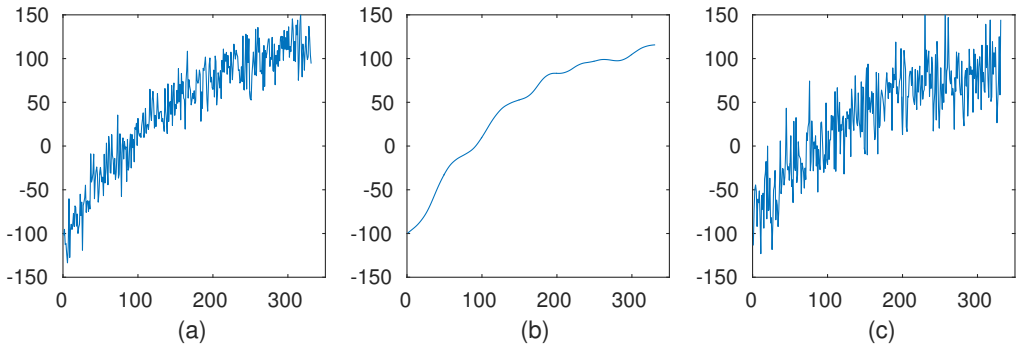


Figure 4.10: For the in-vivo MPnRAGE dataset, the time series at a white matter voxel for different reconstruction methods in the MPnRAGE case. (a) FS (b) TGV (c) ZF. The $y$-axis for all (a) - (c) is signal intensity (AU) and the $x$-axis is frame number (time). An acceleration factor of 33 is used for the subsampled cases.

### Temporally Discontinuous Phantom

For the phantom with discontinuities, voxel time series of FS, TGV, TV, ZF are shown in Fig-

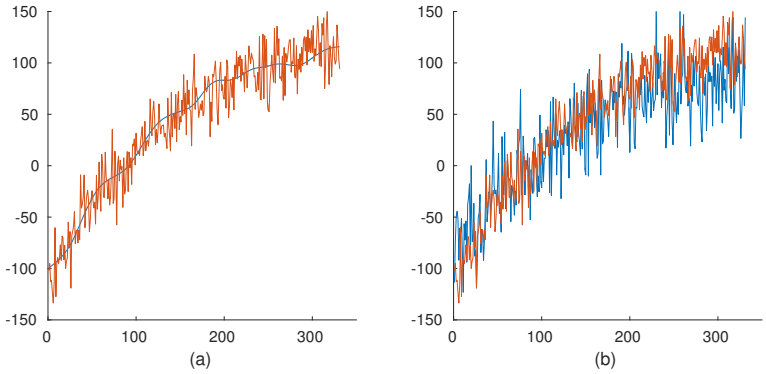


Figure 4.11: For the in-vivo MPnRAGE dataset, the time series at a white matter voxel for different reconstruction methods in the MPnRAGE case. (a) TGV (b) ZF, overlaid with FS (in red). The $y$-axis for all (a) - (c) is signal intensity (AU) and the $x$-axis is frame number (time). An acceleration factor of 33 is used for the subsampled cases.


Figure 4.12: For the in-vivo MPnRAGE dataset, $T_1$ maps (ms) of different reconstruction techniques, (a) FS (b) TGV (c) ZF. An acceleration factor of 33 is used for the subsampled cases.

ure 4.16. The true signal (FS) has 3 smooth regions where the signal values change slowly with time, and 2 sudden jumps (edges) where the signal value at adjacent time points are significantly different. Typically, since temporal regularizers enforce the constraint that signal values at adjacent time points are close to one another, sharp edges tend to get smoothed out. TV regularizers however do not smooth out the edges [18], and since TGV behaves like TV near the edges [13], TGV also does not smooth out the edges. However, in the smooth regions,

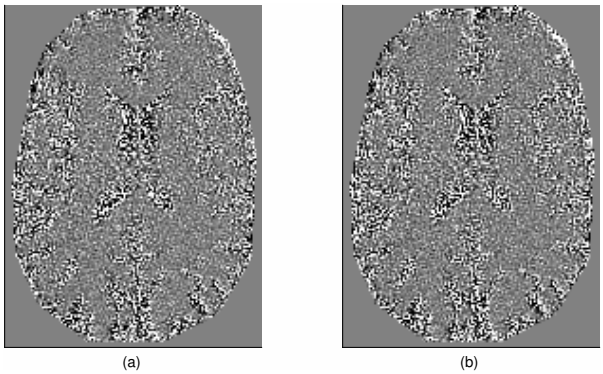


Figure 4.13: For the in-vivo MPnRAGE dataset, difference (in ms) between the $T_1$ maps of different reconstruction techniques, and the ground truth (i.e., FS) $T_1$ maps (a) TGV - FS (b) ZF - FS. An acceleration factor of 33 is used for the subsampled cases.


Figure 4.14: For the in-vivo MPnRAGE dataset, "denoised" source images $x$ (in AU) of different reconstruction techniques, (a) FS (b) TGV (c) ZF. An acceleration factor of 33 is used for the subsampled cases.

TGV and TV behave differently [13]. In the smooth regions, the TV result has staircase artifacts while the TGV result looks very similar to the FS. The estimated source images $x$ at the 180th inversion time point are shown in Figure 4.18. These are the solutions to (4.2), (4.3) for TGV and TV and the ZF and FS computed as described earlier. This $x$ is a time series of volumes and in Figure 4.18 the 180th time point is chosen because it produces the best contrast between different tissues and potential artifacts like blurring of edges between different tissue boundaries can be easily observed. In this figure (and in unpublished figures of other time

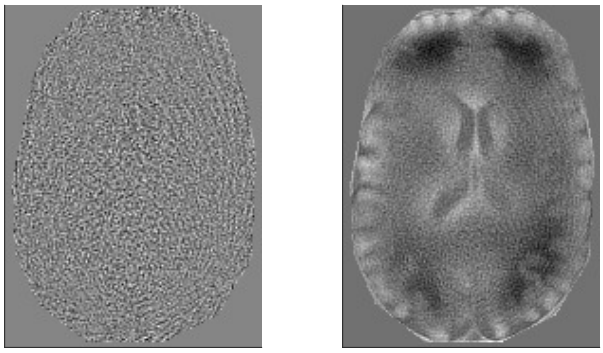


Figure 4.15: For the in-vivo MPnRAGE dataset, difference (in AU) between the source images $x$ of different reconstruction techniques and the ground truth, (a) TGV – FS, (b) ZF – FS. An acceleration factor of 33 is used for the subsampled cases.

points), the TGV and TV images look very similar to the FS while the ZF is extremely noisy and looks nothing like the original. Unlike the previous cases estimation, here a $T_1$ estimation is not performed and so the ZF cannot be improved by computing a function that is robust to Gaussian noise, and so in this case both TGV and TV outperform ZF, while TGV is still better than TV.

## 4.3 Discussion

MPnRAGE is a recently introduced technique to do quantitative $T_1$ estimation [16]. There, a ZF reconstruction technique was used, which did not exploit the redundancy in the temporal dimension. Here an even higher acceleration is achieved using a TGV regularizer in the compressed sensing framework. The fine temporal sampling of MPnRAGE means that the signal at each voxel is slowly varying in time, and this temporal redundancy is exploited by TGV to give a smooth output shown in Figure 4.4 and 4.10. Even though the ZF time series looks noisy, as shown in Figure 4.4 and 4.10, ZF gives a relatively small $T_1$ error even at high

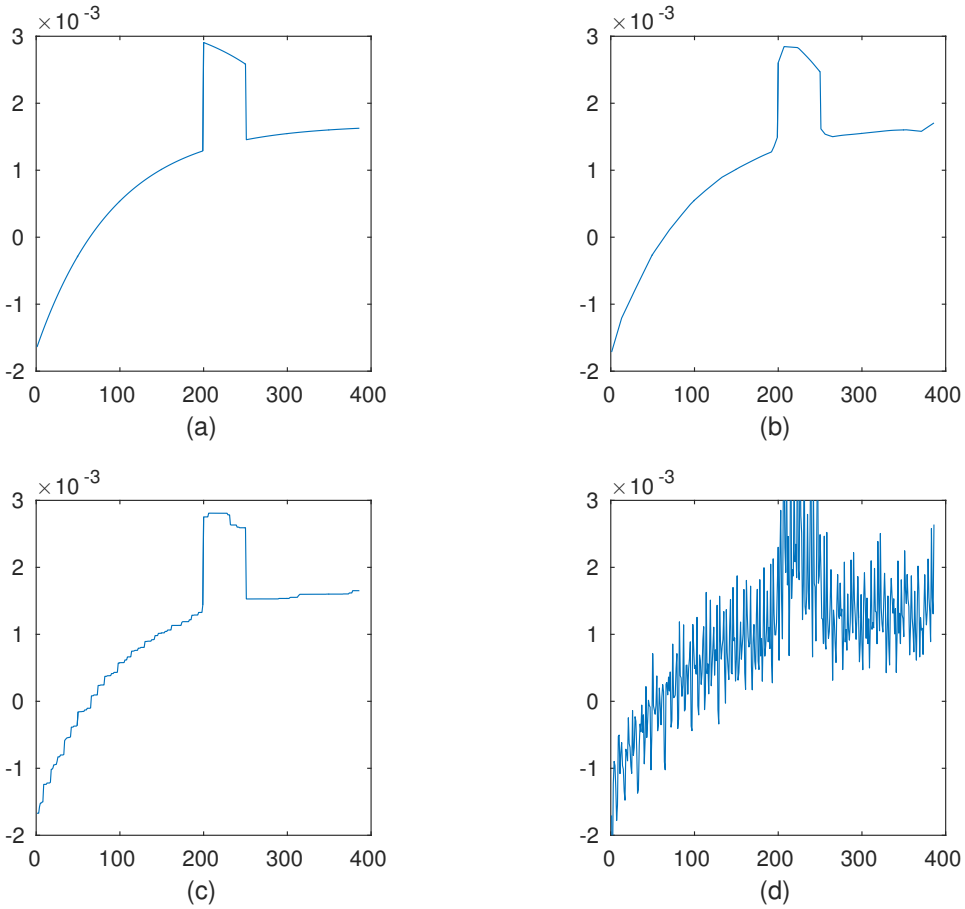


Figure 4.16: For the phantom discontinuous case, time series at a white matter voxel using different reconstruction techniques, (a) FS (b) TGV (c) TV (d) ZF. The $y$-axis for all (a) - (d) is signal intensity (AU) and the $x$-axis is frame number (time). An acceleration factor of 133 is used for the subsampled cases.

accelerations because the interleaved radial undersampling artifacts approximate zero-mean additive noise that is generally tolerated well in the $T_1$ fitting procedure, whereas if a different $k$-space sampling pattern is used in which the undersampling artifacts are not zero-mean, then $T_1$ fitting of ZF gives much worse results. The results of these suboptimal sampling patterns are not reported here. TGV reduces the variance of this noise producing better $T_1$ estimates than ZF. While TV also exploits the temporal redundancy and reduces the variance of the noise, its output is piecewise constant, which is not the underlying true signal, and hence gives worse $T_1$

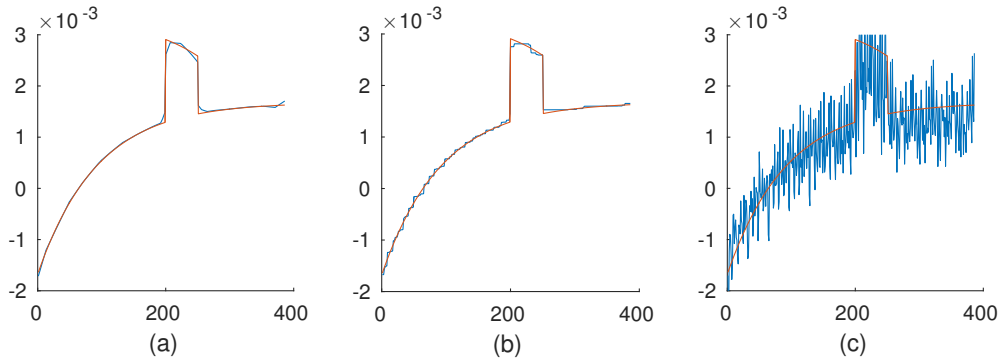


Figure 4.17: For the phantom discontinuous case, time series at a white matter voxel using different reconstruction techniques, (a) TGV (b) TV (c) ZF. The $y$-axis for all (a) - (c) is signal intensity (AU) and the $x$-axis is frame number (time). An acceleration factor of 133 is used for the subsampled cases.

estimates than TGV.

As shown in Figure 4.3, the phantom simulations gave low errors for reasonably small scan times. This is despite the fact that the phantom simulations were performed with a single coil. In practice, 32 coils are used and so we can potentially get an 8 fold reduction in scan time [7]. In addition we use a stack-of-stars $k$-space sampling pattern. Further reductions in scan time can be obtained using a 3D radial sampling pattern based on the discussion in Section 3.

The individual source images $x$ shown in Figures 4.8, 4.14, 4.15 indicate that, TGV is better than ZF since the difference between TGV and FS is Gaussian noise while the difference between ZF and FS shows a lot of the brain structure. Even though we are primarily interested in estimating $T_1$, good reconstruction of source images $x$ is also important. This is because if for example the datasets of different subjects need to be registered together, then using $T_1$ images for this purpose might not necessarily be optimal, and it is possible that the source images $x$ give better results. Another potential application is that sometimes it is desirable to look at the image of the brain with certain tissues nulled out i.e. the signal intensity at a particular tissue is zero and the other tissues are non-zero, and the source images $x$ enable us to

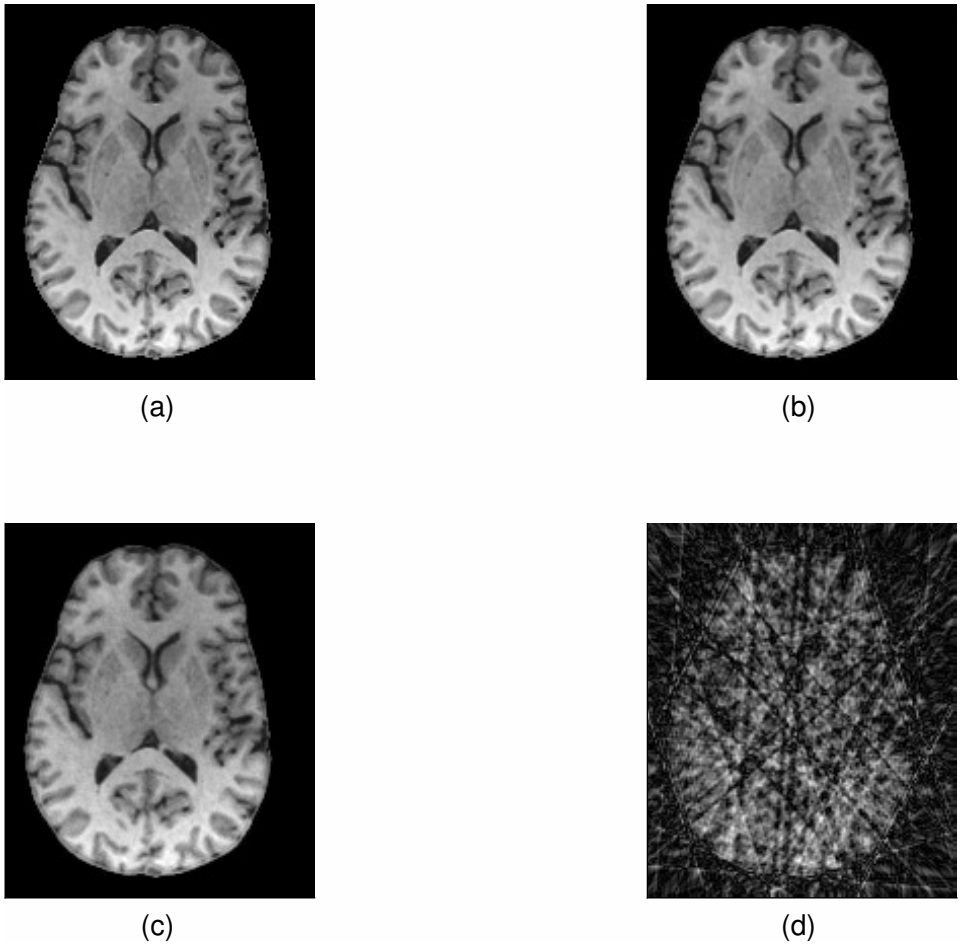


Figure 4.18: For the phantom discontinuous case, source images $x$ (in AU) of different reconstruction techniques, (a) FS, (b) TGV, (c) TV, (d) ZF. An acceleration factor of 133 is used for the subsampled cases.

obtain these kinds of images. There might be numerous other applications in which the source images $x$ might be more useful than $T_1$ and therefore accurate reconstruction of these images is a desirable feature.

For the in-vivo MPnRAGE case, even though the individual source images $x$ shown in Figures 4.14, 4.15 and the $T_1$ error plots shown in Figure 4.3 show that TGV is better than ZF, the $T_1$ maps shown in Figures 4.6, 4.7 do not clearly show that TGV is better than ZF. This could be

because the $T_1$ maps are so noisy that the difference is buried in noise. Here the $T_1$ maps are noisy even for the FS case because the FS dataset was acquired in 10 minutes. A longer FS scan would potentially show that even visually, $T_1$ of TGV is better than ZF. Nevertheless, the parameters $A$, $B$ estimated as part of the $T_1$ estimation process indicate that TGV is closer to the FS than ZF. These $A$, $B$ figures were not presented for brevity.

For a given acceleration, in-vivo gave higher error than the phantom as shown in Figures 4.3 and 4.9. This is because the SNR for in-vivo was much lower than that of the phantom. This is evident in the FS voxel time series shown in Figure 4.10 as opposed to the exponential signal given by (4.1) and shown in Figure 4.4. We performed phantom simulations (unpublished results) in which the SNR was comparable to that of in-vivo, and in this case phantom and in-vivo gave similar errors. To increase the SNR of in-vivo, a longer scan is needed. In this case, while the subsampled data can be acquired in a realistic scan time, the fully-sampled scan would be about an hour long which is impractical, and so we cannot compare the subsampled $T_1$ maps to any ground truth.

For the discontinuous signal of Figure 4.16, TGV outperforms TV in the smooth regions of the time series, and is comparable to TV at the discontinuities because of reasons given in the description of TGV in the Methods section.

# Chapter 5

# Total Variation $T_1$ Estimation

In the Variable Flip Angle (VFA) setup described in Section 3, the goal is to estimate the $T_1$ at all voxels using the images collected at multiple flip angles. $T_1$ estimation is currently done on a voxelwise basis, wherein $T_1$ of each voxel is estimated independently of the $T_1$ of other voxels. In the noise-free setting, this approach is adequate as the $T_1$ estimation is perfect. However, MRI data is acquired in a noisy setting and this is due to the following: In MRI, the signal is generated by polarizing magnetic moments of hydrogen protons in the water molecules of a voxel using a strong magnetic field. Increased magnetic moments in a voxel aligned with the field will increase the SNR. Two factors that influence this signal strength are the voxel size and scanner field strength. Smaller voxel sizes and lower magnetic field strengths lead to lower SNR, and in these cases the current approaches to $T_1$ estimation may be inadequate. One shortcoming of the current $T_1$ estimation approach is that it does not take into account the spatial similarity of the $T_1$ of adjacent voxels. Here, a novel $T_1$ estimation procedure is developed that exploits this spatial redundancy using Total Variation (TV) regularization. TV enforces the constraint that adjacent voxels in the estimated $T_1$ map will have similar $T_1$ values.

While there are many approaches to enforce the spatial similarity constraint, approaches like Tikhonov regularization suffer from the drawback of oversmoothing and thus blurring edges, which means that contrast between different tissues is reduced. However, TV regularization has the ability to denoise images while retaining edges and thus maintain the same resolution as the original [17] [18]. In addition to TV, there are other spatial regularizers like Total Generalized Variation (TGV) and Wavelet that also preserve edges, but here just the TV based regularization is discussed.

Traditionally TV denoising is done on a single image. Pioneering work on denoising a 2D image while retaining sharp edges was done using TV in [17]. This was followed by a number of different studies that minimized the same objective function as [17] using different optimization procedures. A major improvement in TV denoising came when the Split-Bregman algorithm [19] was introduced. Here the objective function was a slightly relaxed version of the original objective function used in [17], and a different optimization procedure was used. These changes resulted in more accurate solutions, and in much faster convergence. All these studies assumed that the observed signal is a noisy version of a single 2D image, and small variants of this basic model. However, in VFA, we have multiple images at different flip angles, and the quantity of interest, $T_1$ is a complicated non-linear function of these images. Despite the complicated nature of this dependence, a novel approach is developed here to combine these images and apply a regularization on the $T_1$.

The spatial redundancy of MRI is not very high [11], and therefore using a spatial regularizer to do $T_1$ denoising is not necessarily the best approach, and using a regularizer along the FA dimension as described in Section 3 and [12] might give better results. However, applying a regularization along FA needs a non-trivial number of FAs and so if say just two or three FAs are acquired, this regularization along the FA dimension will fail and here spatial regularization

will be a better approach. Acquiring just two or three FAs will typically be done when a high resolution dataset is acquired since here acquiring more FAs will make the scan time probihitively large.

Phantom simulations are performed in which noise free $T_1$ and $M_0$ maps are taken, and using these, VFA source images at different flip angles are generated using the VFA signal model. Noise is then added to these source images and $T_1$ estimation is performed on these noisy source images using either the new TV regularization procedure or the usual non-linear least squares (NLLS). We show that TV can reduce artifacts in two different noise generation models, one in which the noise is independent additive Gaussian noise at all voxels, and the other in which noise is due to $k$-space undersampling artifacts.

For $T_1$ estimation in VFA, TV regularization can be performed in three different ways, *pre*, *mid*, and *post*. In *pre*, TV denoising is done on the VFA source images followed by the usual NLLS voxelwise $T_1$ fitting; in *mid*, TV denoising is performed both on $T_1$ and $M_0$ while simultaneously estimating $T_1, M_0$; and in *post*, $T_1$ fitting is done on the raw VFA source images and then TV denoising is done on this $T_1$ map. We hypothesize and later show that *mid* is the best modelling procedure, since it performs well both in smooth regions and is also able to recover small lesions, while *pre* does not perform well in smooth regions and *post* is unable to recover lesions.

For most simulations, an actual brain phantom is used. In a few simulations in which the different TV regularization methods, *pre*, *mid*, *post*, are compared, a Shepp-Logan phantom is used. This Shepp-Logan phantom enables us to see the contrast between different methods better than the brain phantom.

## 5.1 Methods

For most of the simulations, a high SNR brain phantom is used. This was obtained by taking 8 different 1mm$^3$ in-vivo fully-sampled VFA scans of the same subject, estimating $T_1$ of each, registering the $T_1$ images together, and then averaging the registered $T_1$ maps. This gives a $T_1$ map with a very high SNR and hence is a good model for the ground truth signal. The same process is repeated for $M_0$. In a few simulations, a Shepp-Logan phantom is used since subtle artifacts can be better observed than the earlier brain phantom. The Shepp-Logan phantom obtained in MATLAB is rescaled such that the values in the phantom are within the domain bounds (700 to 3000 ms) of actual brain $T_1$ values. The same process is repeated for $M_0$. In both phantoms, once the $T_1$ and $M_0$ maps are obtained, these are used to generate the source images $x$ using the VFA equation,

$$x(\mathbf{r},\alpha) = \frac{M_0(\mathbf{r})(1-E_1(\mathbf{r}))\sin(\alpha)}{1-E_1(\mathbf{r})\cos(\alpha)},$$

where $\mathbf{r}$ is a voxel location, $\alpha$ is the flip angle (FA), $E_1(\mathbf{r}) = \exp(-\frac{T_R}{T_1(\mathbf{r})})$, and $T_R$ is a user defined parameter which is constant for all FAs.

Noise is then added to these source images and TV based $T_1$ estimation is performed on these noisy source images. Two noise generation models are considered. In the first model, i.i.d. Gaussian noise is added to the source images $x$ at each voxel at all flip angles. In the second model, the source images are transformed into $k$-space, 2D radial undersampling is performed in $k$-space, and the source images are reconstructed using zero-filling. These reconstructed source images are noisy versions of the original noise-free source images. In both cases, the same optimization procedures discussed below are used to estimate $T_1$ from the noisy source images.

TV regularization is used to eliminate noise and can be performed in three different ways, *pre*, *mid*, and *post*.

In *pre*, first the TV denoising is performed on the individual source images to obtain a set of denoised images. This is achieved by solving the following optimization problem. For each FA $\alpha$, if $x(\cdot,\alpha)$ represents the noisy source image ($x(\cdot,\alpha)$ is a signal defined on a 3d Cartesian grid), the optimization problem

$$\min_{u(\cdot,\alpha)} ||u(\cdot,\alpha)-x(\cdot,\alpha)||^2+\lambda\ \mathrm{TV}(u(\cdot,\alpha)),$$

is solved to get the denoised image $u(\cdot,\alpha)$. Once this is solved for different FA $\alpha$, $T_1$ is estimated from the solution $u$ using the standard approach, i.e., on a voxelwise basis using NLLS.

In *mid*, $T_1$ fitting is performed using the VFA signal model equation as data fidelity while simultaneously performing a TV regularization on $T_1$ and $M_0$ as follows.

$$\min_{T_1,M_0}\sum_{\mathbf{r},\alpha}\left|x(\mathbf{r},\alpha)-\frac{M_0(\mathbf{r})(1-E_1(\mathbf{r}))\sin(\alpha)}{1-E_1(\mathbf{r})\cos(\alpha)}\right|^2+\lambda\ \mathrm{TV}(T_1)+\lambda\ \mathrm{TV}(M_0),$$

This optimization problem can be solved using gradient descent. In the above optimization problem, different $\lambda$ could have been used for regularizing $T_1$, $M_0$, but we found that using the same $\lambda$ gave good results. This is possibly because in our simulation setup, both $T_1$, $M_0$ have the same domain bounds. In general, the $T_1$, $M_0$ will not have the same domain bounds. In this case, first $T_1$, $M_0$ can be estimated voxelwise using NLLS. From these estimates for all voxels, regularization parameter $\lambda_2$ can be estimated such that $\lambda_1\ \mathrm{mean}(T_1)=\lambda_2\ \mathrm{mean}(M_0)$, and $\lambda_1$ is estimated using the approach described below to estimate $\lambda$. These are used in the regularization term in the above optimization problem as $\lambda_1\ \mathrm{TV}(T_1)+\lambda_2\ \mathrm{TV}(M_0)$.

In *post*, the $T_1$ fitting is performed on the noisy source images $x$ on a voxel-wise basis using

NLLS to get the estimate $T_1^{NLLS}$. Then TV denoising is performed on these noisy $T_1^{NLLS}$ images by solving the following optimization problem,

$$\min_{T_1} ||T_1 - T_1^{NLLS}||^2 + \lambda \text{ TV}(T_1).$$

For most simulations, for simplicity, the same flip angle is used for all voxels within a source image. However, in practice, the flip angle experienced by each voxel will be slightly different since it is modulated by the $B_1$ map. To compare *pre* and *mid*, we do not use the same flip angle for all voxels but each voxel experiences its own flip angle based on the $B_1$ map shown in Figure 5.7. $B_1$ maps of this nature resemble $B_1$ maps observed in practice. Typically, near the center of the brain, $B_1$ map has its highest value of about 1.2, and it decays towards the edges of the brain where it has its lowest value of about 0.7.

To show that TV regularized $T_1$ estimation outperforms NLLS, the brain phantom is used, and *mid* TV approach is used since it is the best. To compare *pre*, *mid*, *post*, a Shepp-Logan phantom was used. To solve all the above optimization problems, the regularization parameter lambda is chosen using the approach described in [14].

## 5.2 Results

Despite the ROIs being really small, the standard deviation is not zero because in actual brains, $T_1$ is not perfectly homogenous.

Noise is added to the phantom VFA images, and $T_1$ is estimated using the TV regularization approach. These TV regularized $T_1$ estimates are compared to the noise free $T_1$ and also to the

state-of-the-art $T_1$ estimation procedure, NLLS. For the brain phantom, the mean and standard deviation of the $T_1$ values in 8 ROI are computed. The ROI shown in Figure 5.8 were used in [16], and the performance in these ROI are representative of the performance in the entire brain. For all simulations, the mean of the $T_1$ estimates in these ROI were almost identical for the 3 cases, noise free, TV, NLLS, and therefore it is not plotted. However, the standard deviation of the the different $T_1$ estimates vary and are given in Tables 1-5 for the 8 ROI in the Gaussian noise case and the radial undersampled $k$-space case. The standard deviation of $T_1$ using TV is closer than NLLS to the noise free case. In addition, the standard deviation of TV is typically not lower than the noise free case which indicates that there is no loss of resolution. $T_1$ maps for the Gaussian noise case for different SNR levels are shown in Figure 5.1 - 5.2. These figures show that the NLLS is noisy, while TV reduces this noise without blurring the image and is closer to the noise free case than NLLS. $T_1$ maps for the radial undersampling case for different accelerations are shown in Figure 5.3 - 5.4. These figures show that the NLLS is noisy and affected by streak artifacts, while TV reduces this noise without blurring the image and is closer to the noise free case than NLLS.

To compare the different TV estimation procedures, viz *pre*, *mid*, and *post*, a Shepp-Logan Phantom with a small lesion is used. This lesion is a small area in the brain in which the $T_1$ value is slightly different from its neighborhood. We use a lesion whose $T_1$ is 1100ms while the $T_1$ of the neighborhood is 1000ms. $T_1$ maps of *mid* and *post* are shown in Figure 5.5. This shows that *mid* recovers the lesion while *post* does not. $T_1$ maps of *mid* and *pre* are shown in Figure 5.6. Here the FA at different voxels are modulated by the $B_1$ map shown in Figure 5.7. This shows that the *mid* is better than *pre* both in the smooth regions and also in the lesion when there is an influence of $B_1$ on the FA.

| ROI | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| | 23 | 33 | 23 | 20 | 23 | 31 | 85 | 81 |

Table 5.1: Standard Deviation of $T_1$ in the ROIs of Figure 5.8 in the noise free case.

| SNR \ ROI | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1200 | 32 | 33 | 27 | 26 | 29 | 31 | 83 | 81 |
| 1000 | 30 | 34 | 24 | 27 | 23 | 31 | 97 | 67 |
| 800 | 28 | 33 | 25 | 31 | 25 | 29 | 92 | 62 |
| 600 | 32 | 35 | 36 | 41 | 34 | 35 | 82 | 73 |

Table 5.2: Standard Deviation of $T_1$ in the ROIs of Figure 5.8 when an additive Gaussian noise model and a TV based $T_1$ estimation is used.

| SNR \ ROI | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1200 | 52 | 49 | 41 | 49 | 52 | 47 | 103 | 102 |
| 1000 | 56 | 51 | 48 | 59 | 60 | 55 | 123 | 99 |
| 800 | 65 | 66 | 49 | 77 | 69 | 71 | 143 | 116 |
| 600 | 88 | 87 | 81 | 103 | 95 | 78 | 131 | 134 |

Table 5.3: Standard Deviation of $T_1$ in the ROIs of Figure 5.8 when an additive Gaussian noise model and a NLLS based $T_1$ estimation is used.

| ACC \ ROI | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 2.5 | 23 | 30 | 21 | 26 | 36 | 26 | 81 | 76 |
| 3.0 | 29 | 34 | 21 | 30 | 31 | 28 | 96 | 81 |
| 3.5 | 26 | 41 | 41 | 41 | 27 | 28 | 99 | 92 |
| 4.0 | 38 | 37 | 21 | 34 | 21 | 33 | 86 | 102 |

Table 5.4: Standard Deviation of $T_1$ in the ROIs of Figure 5.8 when the source of noise is radially undersampled $k$-space and a TV based $T_1$ estimation is used.

| ACC \ ROI | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 2.5 | 35 | 37 | 31 | 39 | 51 | 34 | 88 | 84 |
| 3.0 | 46 | 48 | 36 | 50 | 48 | 44 | 110 | 98 |
| 3.5 | 53 | 67 | 72 | 70 | 53 | 57 | 122 | 119 |
| 4.0 | 75 | 68 | 57 | 69 | 59 | 75 | 116 | 142 |

Table 5.5: Standard Deviation of $T_1$ in the ROIs of Figure 5.8 when the source of noise is radially undersampled $k$-space and a NLLS based $T_1$ estimation is used.

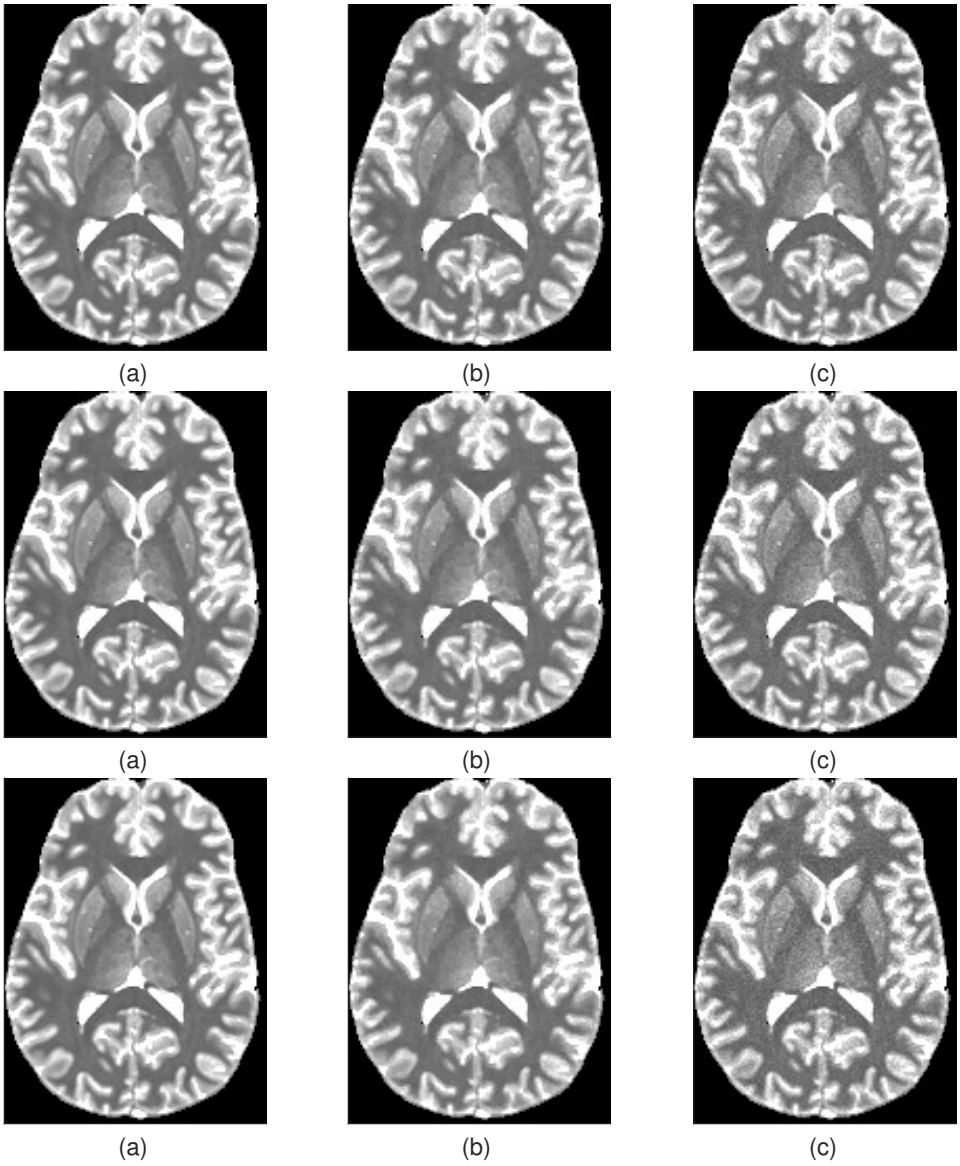


Figure 5.1: $T_1$ maps for the additive Gaussian noise case, (a) Noise free, (b) TV based $T_1$ estimation, (c) NLLS $T_1$ estimation. The 3 rows correspond to SNR of 1200, 1000, and 800 respectively.

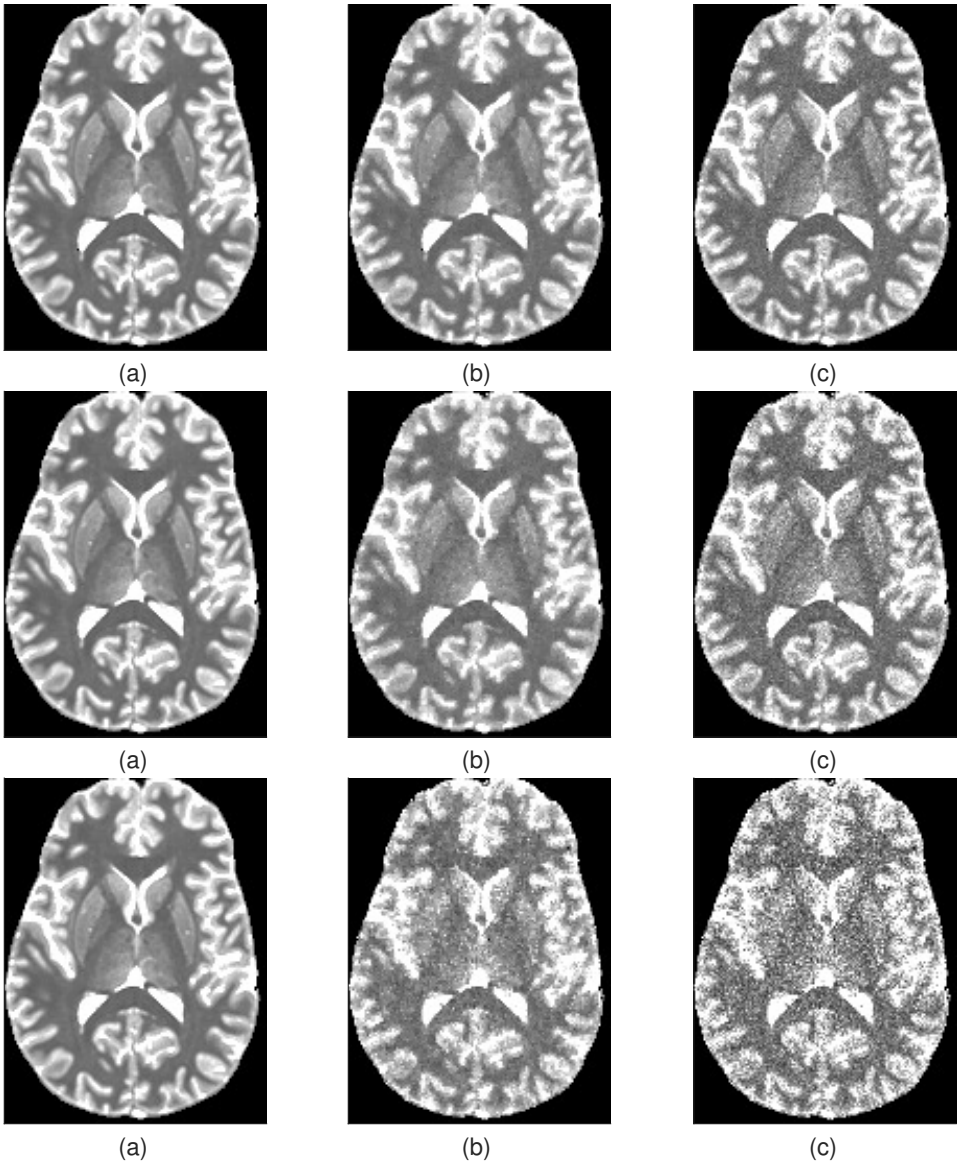


Figure 5.2: $T_1$ maps for the additive Gaussian noise case, (a) Noise free, (b) TV based $T_1$ estimation, (c) NLLS $T_1$ estimation. The 3 rows correspond to SNR of 600, 400, and 200 respectively.

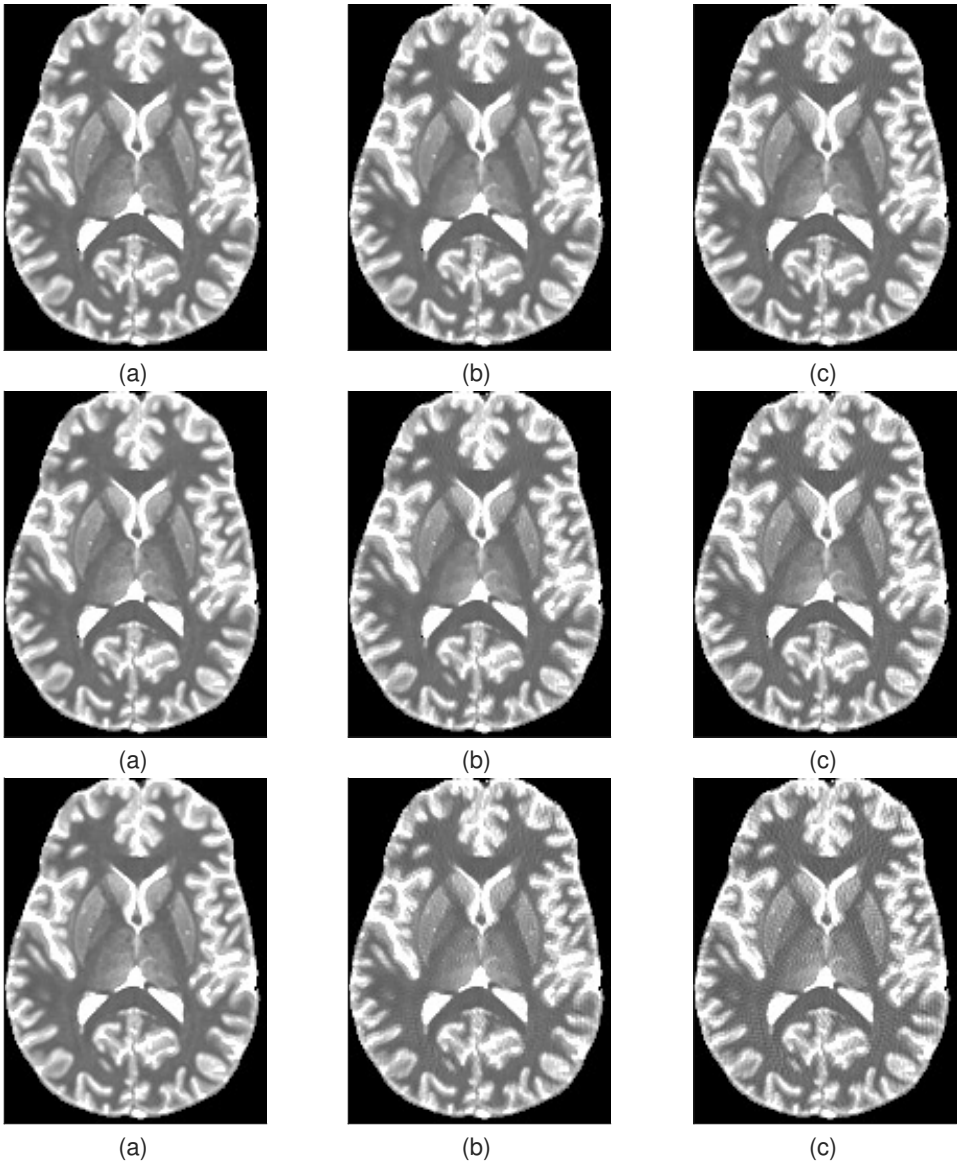


Figure 5.3: $T_1$ maps for the radial $k$-space undersampling case, (a) Noise free, (b) TV based $T_1$ estimation, (c) NLLS $T_1$ estimation. The 3 rows correspond to undersampling factors of 2.5, 3, and 3.5 respectively.

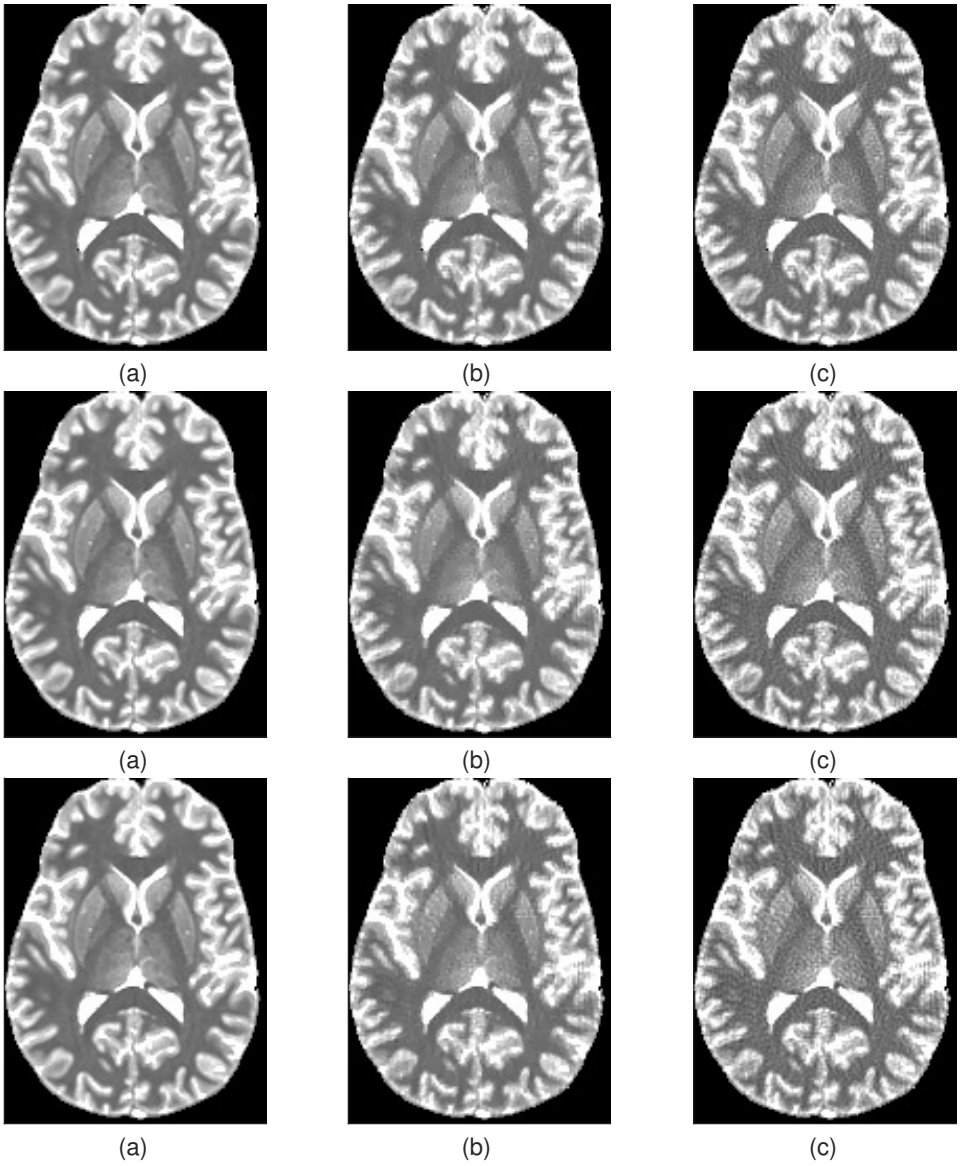


Figure 5.4: $T_1$ maps for the radial $k$-space undersampling case, (a) Noise free, (b) TV based $T_1$ estimation, (c) NLLS $T_1$ estimation. The 3 rows correspond to undersampling factors of 4, 4.5, and 5 respectively.

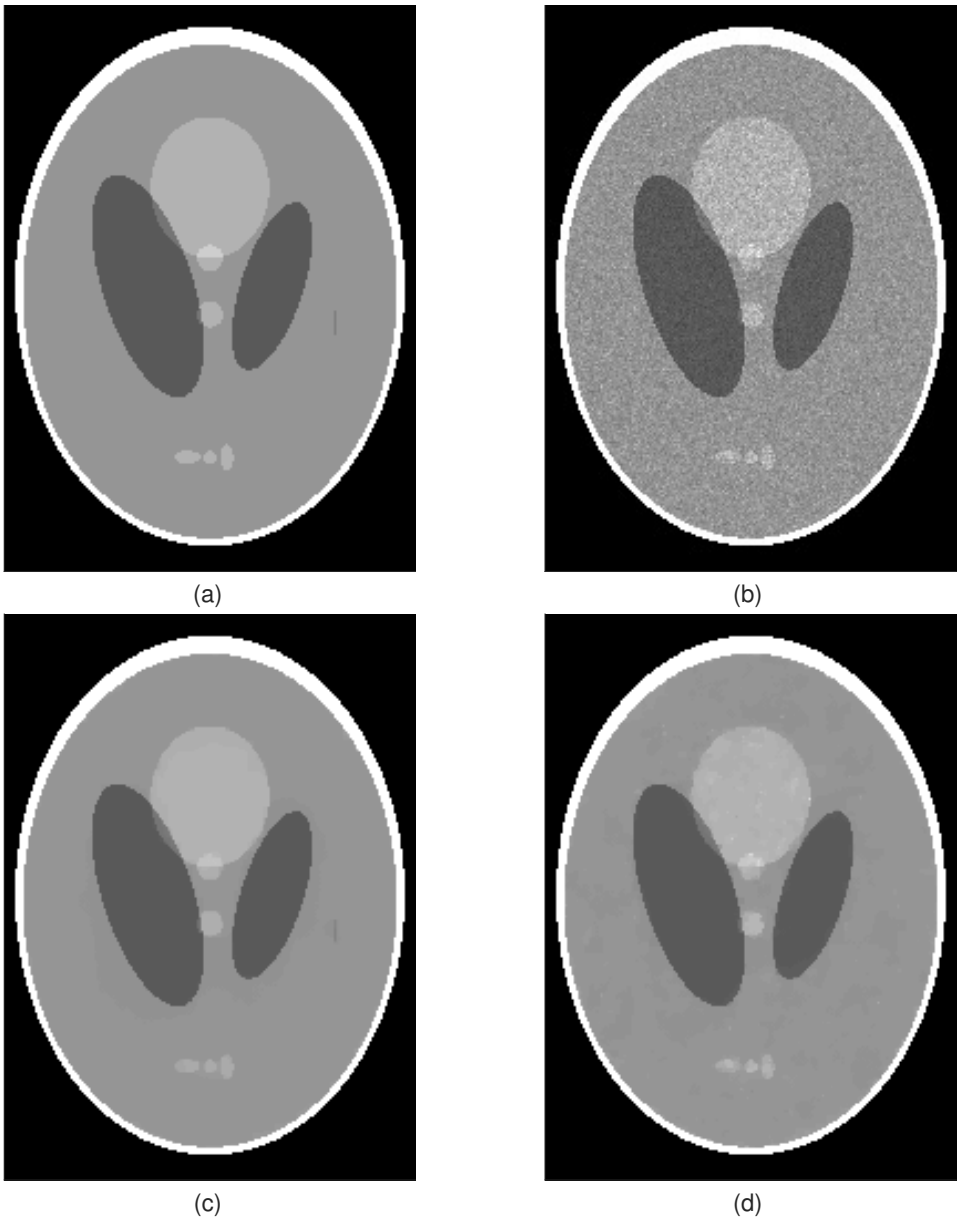


Figure 5.5: $T_1$ maps when the TV regularization used (a) Noise Free (b) NLLS (c) *mid* (d) *post*. The noise used is additive Gaussian noise. *mid* recovers the lesion while *post* does not.

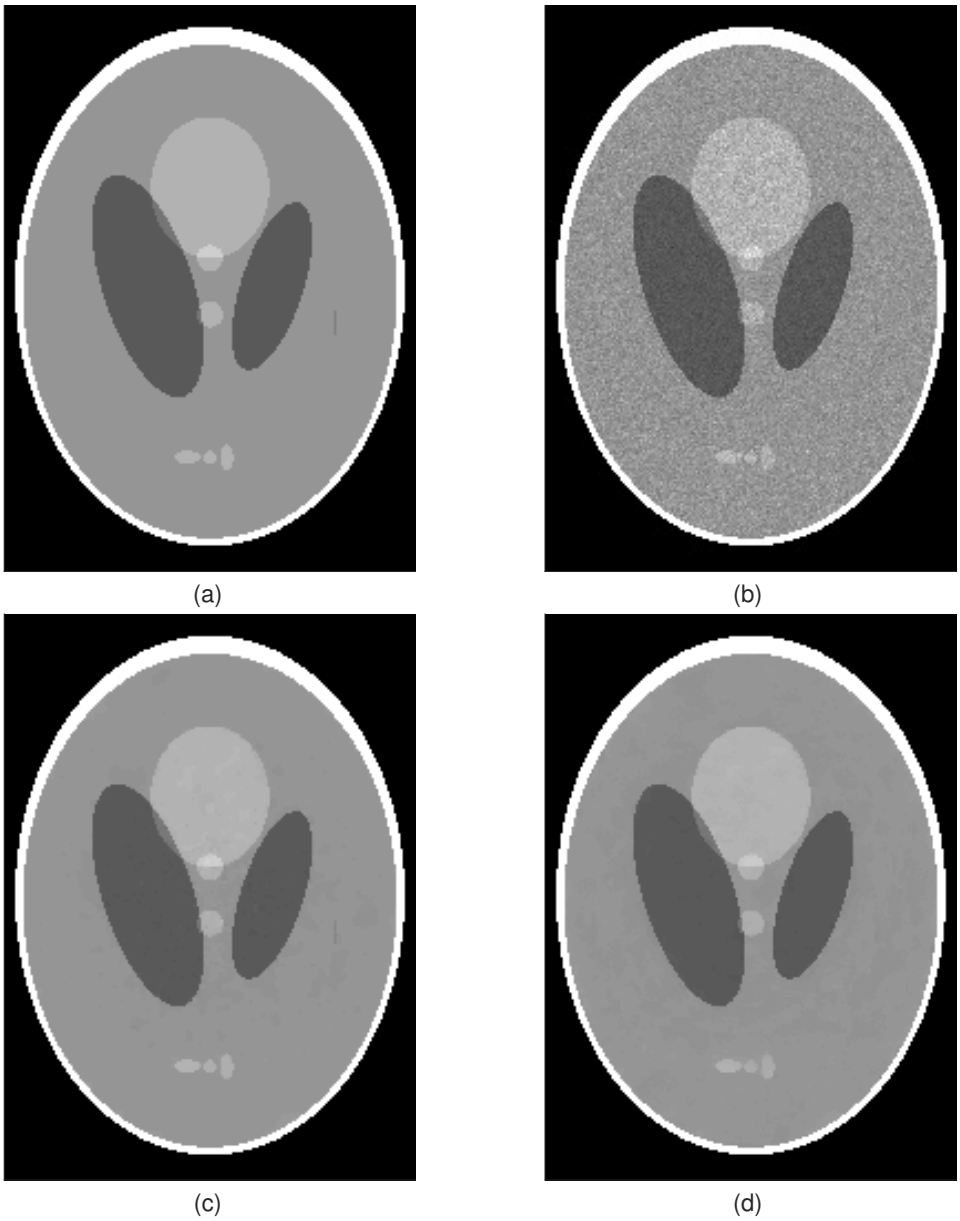


Figure 5.6: $T_1$ maps when the TV regularization used (a) Noise Free (b) NLLS (c) *mid* (d) *pre*. The noise used is additive Gaussian noise, and the FA at all voxels are modulated by the $B_1$ map shown in Figure 5.7.

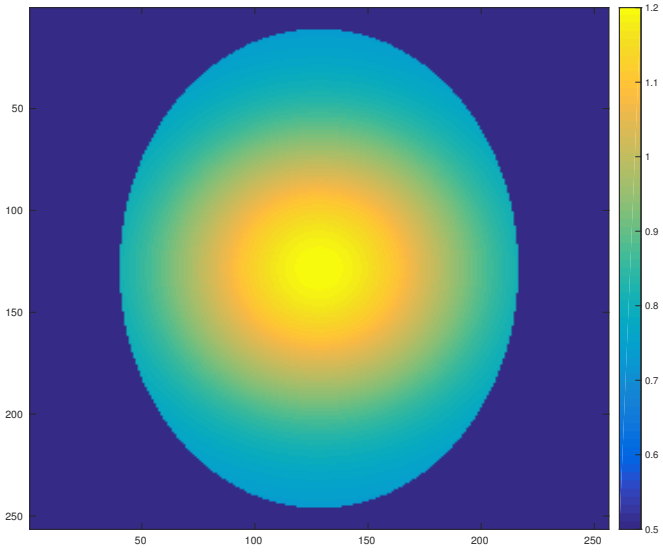


Figure 5.7: $B_1$ map used to modulate the flip angles at different voxels for the *mid* vs. *pre* results in Figure 5.6.

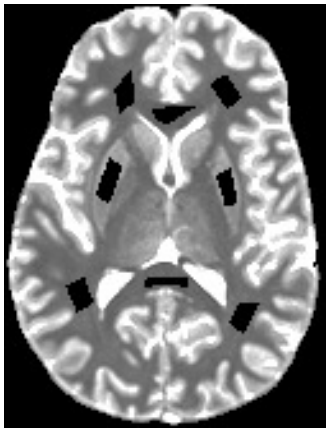

Figure 5.8: The different ROIs used for estimating the mean and standard deviation of the $T_1$. 8 ROIs are used and are the black patches. These ROIs are chosen in both white matter and gray matter tissues.

## 5.3 Discussion and Conclusions

A total variation (TV) denoising procedure was used to reduce the amount of noise in VFA $T_1$ estimation. The results indicate that in all simulations, TV based $T_1$ estimation outperforms

the usual NLLS $T_1$ estimation since TV uses the additional hypothesis that the $T_1$ of adjacent voxels are similar. Since the ROI chosen were small in size, $T_1$ values of all voxels in these ROI are similar and so the standard deviation of $T_1$ is a reasonable measure of estimation performance. Mean of the TV based $T_1$ estimate in different ROI was similar to the ground truth which implies that TV does not give a biased estimate, while the standard deviation of TV was lower than NLLS but not dropping significantly below the standard deviation of the ground truth which means that TV was improving the $T_1$ estimate without over smoothing the $T_1$ map, and thereby retaining the same resolution as the ground truth $T_1$ image. Nevertheless, poor $T_1$ estimates can be obtained if the regularization parameter $\lambda$ is not chosen optimally. The amount of spatial smoothing of the $T_1$ maps is controlled by $\lambda$. If $\lambda$ is too large, the estimated $T_1$ maps become blurry while if $\lambda$ is small, the noise does not get eliminated. The approach described in [14] to choose $\lambda$ maintains a good balance between eliminating the noise while at the same time maintaining spatial resolution. This balance was not achieved using the L-curve [20] approach to choosing $\lambda$, possibly because L-curve was designed for Tikhonov regularization [20] and not for TV.

It is not surprising that TV denoising works in the Gaussian noise case since TV has previously successfully denoised images corrupted by Gaussian noise [17], albeit in those cases, the objective was a linear function of the quantity to be estimated, whereas in our case, the objective is a nonlinear function of $T_1$. However, for the undersampled $k$-space case, it was unclear apriori if TV could actually reduce these undersampling artifacts. Here, TV regularization only works for those cases in which $k$-space sampling patterns result in undersampling artifacts that are like Gaussian noise. A radial sampling scheme was chosen since it results in undersampling artifacts that are approximately Gaussian noise. We also tried a Cartesian sampling scheme in which the center of $k$-space was heavily sampled and the edges were sparsely sampled. This did not result in undersampling artifacts that were like Gaussian noise and so

the TV based $T_1$ estimation did not result in any improvement over the NLLS $T_1$ estimation procedure. While TV did perform well for radial by reducing the undersampling streak artifacts, small amounts of streak artifacts still remain. Increasing $\lambda$ does eliminate these but it oversmoothes the $T_1$ maps and the resulting standard deviation of $T_1$ in the different ROIs become smaller than noise free $T_1$ and this indicates a loss of resolution. Removing the streak artifacts while retaining the resolution will be a topic of future work. Other *k*-space sampling procedures like spiral and Poisson also result in artifacts that are like Gaussian noise. Future work will involve using these patterns to potentially get better results. Additionally this TV based quantitative $T_1$ estimation can be used in other approaches like inversion recovery which is described in Section 4. At higher resolutions, this TV based $T_1$ estimation will have an even higher impact both because the noise is higher at higher resolutions and so will require denoising of the $T_1$ maps, and because the spatial redundancy of $T_1$ increases at higher resolutions and so TV will be better suited to eliminate artifacts.

For most cases, TV done using *pre*, *mid*, and *post* gives similar results. However, there are a few cases in which *mid* outperforms both *pre* and *post*.

If there is a small lesion in the $T_1$ map that has a different $T_1$ compared to its neighboring $T_1$ values, *mid* outperforms *post* by retaining this lesion while *post* eliminates this lesion. The reason why *post* does not perform well in this case is because when the noise is large enough as shown in Figure 5.6, the lesion after the NLLS-$T_1$ fitting gets buried in surrounding noise, and becomes indistinguishable from its neighborhood. Since in *post*, TV denoising is done on this NLLS-$T_1$ map, it cannot recover this lesion. *Mid* on the other hand can recover this lesion because of the better modelling procedure.

For the case in which the flip angles are modulated by the $B_1$ map, *pre* does not perform as well as *mid* because the $B_1$ map modulates the source images and makes them less spatially redundant, and so a TV on the source images (as done in *pre*) does not perform as well as

the case in which $B_1$ was constant across the brain. However, the *mid* performs well because the underlying $T_1$, $M_0$ maps are unaffected by the $B_1$ modulation, and since *mid* performs a regularization on just $T_1$ and $M_0$, it gives results that are comparable to the case in which $B_1$ is constant across the brain.

# Chapter 6

# Summary and Future Directions

In this thesis, techniques to speed up the acquisition of Quantitative MRI (Q-MRI) are discussed. The acquisition speed of Q-MRI is determined by the number of $k$-space (i.e. the space of spatial-frequencies) samples that are acquired. We reduced the acquisition time of Q-MRI by acquiring fewer $k$-space samples, followed by using more sophisticated algorithms like compressed sensing (CS) to reconstruct the desired quantitative $T_1$ maps.

In the first project, we compared the performance of two different $k$-space sampling techniques, Vastly Undersampled Isotropic Projection (VIPR) and Stack-of-Stars (SOS), in CS algorithms to estimate $T_1$ using the Variable Flip Angle (VFA) acquisition paradigm. For most scan times, SOS outperformed VIPR. This is despite the fact that VIPR sampling is more incoherent than SOS, and thereby more likely to give improved results in CS algorithms. SOS outperformed VIPR because our current $k$-space sampling techniques enables us to acquire samples in SOS tailored to the shape of the brain whereas in VIPR, we are currently unable to tailor the sampling pattern to the shape of the brain. In the future it might be possible to modify the VIPR

*k*-space sampling pattern such that it can take advantage of the shape of the brain like was done in SOS and this might result in VIPR outperforming SOS at all scan times.

In the second project, we developed a new temporal CS regularizer, Total Generalized Variation (TGV) and used this to speed up the acquisition of quantitative $T_1$ using the MPnRAGE inversion recovery paradigm. TGV outperformed the Zero Filled (ZF) approach currently used for MPnRAGE, and it also outperformed the classical Total Variation (TV) regularizer by eliminating its staircase artifacts. These different approaches were tested on both phantom and in-vivo data, and TGV always performed the best. However, a shortcoming of our simulation approach was that the data used was a single 2D slice of the brain. This was due to the difficulty in acquiring a fully sampled in-vivo 3D brain data in realistic scan time, and also due to the computational complexity associated with reconstructing 3D volumes using our algorithm on our currently available hardware. A compliant subject who can remain still for an extremely long period of time in the MRI scanner can help acquire 3D data, and improved computational hardware units like Graphics Processing Unit (GPU) can help in reconstructing image-space from *k*-space data in a reasonable amount of time.

In the third project, we developed a new $T_1$ estimation approach for Variable Flip Angle (VFA) using spatial Total Variation (TV) regularization. In noisy datasets, this approach outperformed the nonlinear least-squares (NLLS) based $T_1$ estimation that is currently used. TV based $T_1$ estimation could be performed in three different ways, and among these *mid* outperformed both *pre* and *post*. All these different approaches estimated $T_1$ from image-space data. We hypothesize that estimating $T_1$ directly from *k*-space using a spatial TV regularizer without going through the intermediate step of estimating the source images can potentially give better results.

# Chapter 7

# Appendix

We show that the expected value of the $T_1$ estimate in the presence of noise is larger than than the true $T_1$ for all choices of acquisition parameters like flip angles ($\alpha$), repetition time ($T_R$), noise standard deviation ($\sigma$).

The VFA signal model in the noise free case is

$$x(\alpha) = \frac{M_0(1-E_1)\sin\alpha}{1-E_1\cos\alpha},$$

where $E_1 = \exp(-\frac{T_R}{T_1})$. Rearranging the terms in the above equation, we get

$$\frac{x(\alpha)}{\sin\alpha} = M_0(1-E_1) + \frac{x(\alpha)}{\tan\alpha}E_1.$$

Since data are acquired at multiple flip angles $\alpha$, we can write the above equation for multiple

$\alpha$ as,

$$\begin{bmatrix} \frac{x(\alpha_1)}{\sin(\alpha_1)} \\ \frac{x(\alpha_2)}{\sin(\alpha_2)} \\ \frac{x(\alpha_3)}{\sin(\alpha_3)} \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 & \frac{x(\alpha_1)}{\tan(\alpha_1)} \\ 1 & \frac{x(\alpha_2)}{\tan(\alpha_2)} \\ 1 & \frac{x(\alpha_3)}{\tan(\alpha_3)} \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} M_0(1-E_1) \\ E_1 \end{bmatrix}$$

We use the notation $s(\alpha) := \sin(\alpha)$ and $t(\alpha) := \tan(\alpha)$.

In the noisy case, $x(\alpha)$ above is replaced by $x(\alpha) + n(\alpha)$, where $n(\alpha_i) \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$.

In noisy case, we write the above equation as $y = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \theta = A\theta$,
where $y = \begin{bmatrix} \frac{x(\alpha_i)}{\sin(\alpha_i)} \end{bmatrix}$, $a_1 = \begin{bmatrix} 1 \end{bmatrix}$, $a_2 = \begin{bmatrix} \frac{x(\alpha_i)}{\tan(\alpha_i)} \end{bmatrix}$.

The least-square estimate of $\theta$, $\hat{\theta}$ is given by,

$$\hat{\theta} = (A^T A)^{-1} A^T y$$

$$= \left( \begin{bmatrix} a_1^T \\ \hdashline \\ a_2^T \end{bmatrix} \begin{bmatrix} a_1 & \vdots & a_2 \end{bmatrix} \right)^{-1} \begin{bmatrix} a_1^T \\ \hdashline \\ a_2^T \end{bmatrix} y$$

$$= \begin{bmatrix} < a_1,\, a_1 > & < a_1,\, a_2 > \\ < a_2,\, a_1 > & < a_2,\, a_2 > \end{bmatrix}^{-1} \begin{bmatrix} < a_1,\, y > \\ < a_2,\, y > \end{bmatrix}$$

$$\hat{\theta} = \begin{bmatrix} \hat{\theta}_1 \\ \\ \hat{\theta}_2 \end{bmatrix} = \begin{bmatrix} < a_1,\, a_1 > & - < a_1,\, a_2 > \\ - < a_2,\, a_1 > & < a_2,\, a_2 > \end{bmatrix} \begin{bmatrix} < a_1,\, y > \\ < a_2,\, y > \end{bmatrix} / \det(A^T A)$$

We are interested only in $\hat{\theta}_2 = \hat{E}_1$

$$\hat{\theta}_2 = \frac{- < a_2,\, a_1 > \; < a_1,\, y > + < a_1,\, a_1 > \; < a_2,\, y >}{\det(A^T A)}$$

$$\hat{\theta}_2 = \frac{- < a_2,\, a_1 > \; < a_1,\, y > + < a_1,\, a_1 > \; < a_2,\, y >}{< a_2,\, a_2 >< a_1,\, a_1 > -(< a_2,\, a_1 >)^2} := \frac{\hat{D}}{\hat{N}}$$

Let $\hat{F}_1 = \dfrac{1}{\hat{E}_1} = \dfrac{1}{\hat{\theta}_2} = \dfrac{\hat{N}}{\hat{D}}$.

To compute $\mathbb{E}\left[\hat{F}_1\right]$, we need to compute $\mathbb{E}\left[\hat{N}\right]$, $\mathbb{E}\left[\hat{D}\right]$. To do this we need to compute $\mathbb{E}$ of the following four terms,

1. $\mathbb{E}\left[ < a_1, a_1 >< a_2, a_2 > \right]$

$$\begin{aligned}
\mathbb{E}\left[ < a_2, a_2 > \right] &= \mathbb{E}\left[ \sum_{\alpha} \left( \frac{x(\alpha) + n(\alpha)}{t(\alpha)} \right)^2 \right] \\
&= \mathbb{E}\left[ \sum_{\alpha} \frac{x^2(\alpha) + n^2(\alpha) + 2x(\alpha)n(\alpha)}{t^2(\alpha)} \right] \\
&= \mathbb{E}\left[ \sum_{\alpha} \frac{x^2(\alpha) + \alpha^2}{t^2(\alpha)} \right] \\
\Rightarrow \mathbb{E}\left[ < a_1, a_1 >< a_2, a_2 > \right] &= < a_1, a_1 > \mathbb{E}\left[ < a_2, a_2 > \right]
\end{aligned}$$

2. $\mathbb{E}\left[ (< a_2, a_1 >)^2 \right]$

$$\begin{aligned}
\mathbb{E}\left[ (< a_2, a_1 >)^2 \right] &= \mathbb{E}\left[ \left( \sum_{\alpha} \frac{x(\alpha) + n(\alpha)}{t(\alpha)} \right)^2 \right] \\
&= \mathbb{E}\left[ \left( \sum_{\alpha} \frac{x(\alpha) + n(\alpha)}{t(\alpha)} \right) \left( \sum_{\beta} \frac{x(\beta) + n(\beta)}{t(\beta)} \right) \right] \\
&= \mathbb{E}\left[ \sum_{\alpha\beta} \frac{x(\alpha)x(\beta) + n(\alpha)x(\beta) + x(\alpha)n(\beta) + n(\alpha)n(\beta)}{t(\alpha)t(\beta)} \right] \\
&= \mathbb{E}\left[ \sum_{\alpha\beta} \frac{x(\alpha)x(\beta) + n(\alpha)n(\beta)}{t(\alpha)t(\beta)} \right] \\
&= \sum_{\alpha\beta} \frac{x(\alpha)x(\beta)}{t(\alpha)t(\beta)} + \mathbb{E}\left[ \sum_{\alpha} \frac{n^2(\alpha)}{t^2(\alpha)} \right] \\
&= \left( \sum_{\alpha} \frac{x(\alpha)}{t(\alpha)} \right)^2 + \sum_{\alpha} \frac{\sigma^2}{t^2(\alpha)}
\end{aligned}$$

3. $\mathbb{E}\left[ < a_2, a_1 >< a_1, y > \right]$

$$\begin{aligned}
\mathbb{E}\left[ < a_2, a_1 >< a_1, y > \right] &= \mathbb{E}\left[ \left( \sum_{\alpha} \frac{x(\alpha)+n(\alpha)}{t(\alpha)} \right) \left( \sum_{\beta} \frac{x(\beta)+n(\beta)}{s(\beta)} \right) \right] \\
&= \mathbb{E}\left[ \sum_{\alpha\beta} \frac{x(\alpha)x(\beta)+n(\alpha)x(\beta)+n(\beta)x(\alpha)+n(\alpha)n(\beta)}{t(\alpha)s(\beta)} \right] \\
&= \sum_{\alpha\beta} \frac{x(\alpha)x(\beta)}{t(\alpha)s(\beta)} + \mathbb{E}\left[ \sum_{\alpha\beta} \frac{n(\alpha)n(\beta)}{t(\alpha)s(\beta)} \right] \\
&= \left( \sum_{\alpha} \frac{x(\alpha)}{t(\alpha)} \right) \left( \sum_{\beta} \frac{x(\beta)}{t(\beta)} \right) + \sum_{\alpha} \frac{\sigma^2}{t(\alpha)s(\alpha)}
\end{aligned}$$

4. $\mathbb{E}\left[ < a_2,\ y > \right]$

$$\begin{aligned}
\mathbb{E}\left[ < a_2,\ y > \right] &= \mathbb{E}\left[ \sum_{\alpha} \frac{(x(\alpha)+n(\alpha))^2}{t(\alpha)s(\alpha)} \right] \\
&= \mathbb{E}\left[ \sum_{\alpha} \frac{x^2(\alpha)+n^2(\alpha)+2x(\alpha)n(\alpha)}{t(\alpha)s(\alpha)} \right] \\
&= \sum_{\alpha} \frac{x^2(\alpha)}{t(\alpha)s(\alpha)} + \sigma^2 \sum_{\alpha} \frac{1}{t(\alpha)s(\alpha)}
\end{aligned}$$

$$
\begin{aligned}
\Rightarrow \mathbb{E}\left[\hat{N}\right] &= \mathbb{E}\left[ < a_2,\, a_2 > \;\; < a_1, a_1 > \;-\; (< a_2,\, a_1 >)^2) \right] \\
&= < a_1,\, a_1 > \sum_{\alpha} \frac{x^2(\alpha)+\sigma^2}{t^2(\alpha)} - \left( \sum_{\alpha} \frac{x(\alpha)}{t(\alpha)} \right)^2 - \sigma^2 \sum_{\alpha} \frac{1}{t^2(\alpha)} \\
&= \underbrace{< a_1,\, a_1 > \sum_{\alpha} \frac{x^2(\alpha)}{t^2(\alpha)} - \left( \sum_{\alpha} \frac{x(\alpha)}{t(\alpha)} \right)^2}_{\overset{\cdot\cdot}{N}} \\
&\quad + \sigma^2 \left( < a_1,\, a_1 > \sum_{\alpha} \frac{1}{t^2(\alpha)} - \sum_{\alpha} \frac{1}{t^2(\alpha)} \right) \\
\mathbb{E}\left[\hat{N}\right] &= N + \sigma^2(< a_1,\, a_1 > -1)U, \text{where } U = \sum_{\alpha} \frac{1}{t^2(\alpha)} \\
\mathbb{E}\left[\hat{D}\right] &= \mathbb{E}\left[ < a_2,\, y >< a_1,\, a_1 > - < a_2,\, a_1 >< a_1,\, y > \right] \\
&= < a_1,\, a_1 > \sum_{\alpha} \frac{x^2(\alpha)}{t(\alpha)s(\alpha)} + < a_1,\, a_1 > \sigma^2 \sum_{\alpha} \frac{1}{t(\alpha)s(\alpha)} \\
&\quad - \left( \sum_{\alpha} \frac{x(\alpha)}{t(\alpha)} \right) \left( \sum_{\beta} \frac{x(\beta)}{s(\beta)} \right) - \sigma^2 \sum_{\alpha} \frac{1}{t(\alpha)s(\alpha)} \\
&= \underbrace{< a_1,\, a_1 > \sum_{\alpha} \frac{x^2(\alpha)}{t(\alpha)s(\alpha)} - \left( \sum_{\alpha} \frac{x(\alpha)}{t(\alpha)} \right) \left( \sum_{\beta} \frac{x(\beta)}{t(\beta)} \right)}_{\overset{\cdot\cdot}{D}} \\
&\quad + \sigma^2 \left( < a_1,\, a_1 > \sum_{\alpha} \frac{1}{t(\alpha)s(\alpha)} - \sum_{\alpha} \frac{1}{t(\alpha)s(\alpha)} \right) \\
&= D + \sigma^2(< a_1,\, a_1 > -1)V,
\end{aligned}
$$

where $V = \sum_{\alpha} \frac{1}{t(\alpha)s(\alpha)}$.

$$
\mathbb{E}\left[\hat{F}_1\right] = \mathbb{E}\left[\frac{\hat{N}}{\hat{D}}\right]
$$

We now show that $\mathbb{E}\left[\frac{\hat{N}}{\hat{D}}\right] \approx \frac{\mathbb{E}\left[\hat{N}\right]}{\mathbb{E}\left[\hat{D}\right]}$.

Taylor series expansion 1st order of $\mathbb{E}\left[f(X,Y)\right] = \mathbb{E}\left[\frac{X}{Y}\right]$ around the point $\left(\mathbb{E}\left[X\right],\mathbb{E}\left[Y\right]\right) \coloneqq (X_0\ Y_0)$.

$$
\begin{aligned}
\mathbb{E}\left[f(X,Y)\right] &\approx \mathbb{E}\left[f(X_0,Y_0)\right] + \mathbb{E}\left[f_x(X_0,Y_0)(X-X_0)\right] + \mathbb{E}\left[f_y(X_0,Y_0)(Y-Y_0)\right] \\
&= \mathbb{E}\left[f\big(\mathbb{E}\left[X\right],\mathbb{E}\left[Y\right]\big)\right] \\
&\quad + \mathbb{E}\left[f_x\Big(\mathbb{E}\left[X\right],\mathbb{E}\left[Y\right]\Big)\Big(X-\mathbb{E}\left[X\right]\Big)\right] \\
&\quad + \mathbb{E}\left[f_y\Big(\mathbb{E}\left[X\right],\mathbb{E}\left[Y\right]\Big)\Big(Y-\mathbb{E}\left[Y\right]\Big)\right] \\
&= f\Big(\mathbb{E}\left[X\right],\mathbb{E}\left[Y\right]\Big) + f_x\Big(\mathbb{E}\left[X\right],\mathbb{E}\left[Y\right]\Big)\cdot 0 + f_y\Big(\mathbb{E}\left[X\right],\mathbb{E}\left[Y\right]\Big)\cdot 0 \\
&= \frac{\mathbb{E}\left[X\right]}{\mathbb{E}\left[Y\right]} \\
\Rightarrow \mathbb{E}\left[\frac{\hat{N}}{\hat{D}}\right] &\approx \frac{\mathbb{E}\left[\hat{N}\right]}{\mathbb{E}\left[\hat{D}\right]} \\
\Rightarrow \mathbb{E}\left[\frac{\hat{N}}{\hat{D}}\right] &\approx \frac{N+\sigma^2\big(<a_1,\ a_1>-1\big)U}{D+\sigma^2\big(<a_1,\ a_1>-1\big)V} \\
\Rightarrow \mathbb{E}\left[\frac{\hat{N}}{\hat{D}}\right] - \frac{N}{D} &= \frac{N+\sigma^2\big(<a_1,\ a_1>-1\big)U}{D+\sigma^2\big(<a_1,\ a_1>-1\big)V} - \frac{N}{D} \\
&= \frac{ND+\sigma^2(<a_1,\ a_1>-1)UD - ND - \sigma^2(<a_1,\ a_1>-1)UN}{(D+\sigma^2(<a_1,\ a_1>-1)V)D} \\
&= \frac{\sigma^2(<a_1,\ a_1>-1)(UD-VN)}{(D+\sigma^2(<a_1,\ a_1>-1)V)D}
\end{aligned}
$$

We will now show that this $\leq 0$.

<u>First $U < V$</u>

Since $s(\alpha) = \sin(\alpha)$ and $t(\alpha) = \tan(\alpha)$ and $0 < \alpha < 20^\circ \Rightarrow \alpha$ is in the 1st quadrant and so

$$0 < \cos\alpha < 1 \Rightarrow \tan\alpha = \frac{\sin\alpha}{\cos\alpha} > \sin\alpha \Rightarrow \frac{1}{\tan\alpha} < \frac{1}{\tan\alpha} \Rightarrow \frac{1}{\tan^2\alpha} < \frac{1}{\tan\alpha\sin\alpha}$$

$$\sum_{\alpha} \frac{1}{t^2(\alpha)} < \sum_{\alpha} \frac{1}{t(\alpha)s(\alpha)}$$

$$U < V$$

Next $D < N$

In the noise free case, $n = 0, \sigma = 0$ and so the estimate $\hat{\theta}$ is

$$\hat{\theta} = (A^T A)^{-1} A^T y = (A^T A)^{-1} A\theta = \theta = \begin{bmatrix} M_0(1-E_1) \\ \\ E_1 \end{bmatrix}$$

$$\Rightarrow \hat{\theta}_2 = E_1 \Rightarrow \hat{F}_1 = \frac{1}{\hat{\theta}_2} = \frac{1}{E_1} = \frac{1}{\exp(-\frac{T_R}{T_1})} = \exp\left(\frac{T_R}{T_1}\right)$$

Since $T_R, T_1 > 0 \Rightarrow \exp(T_R/T_1) > 1 \Rightarrow \hat{F}_1 > 1$. Now from the earlier computations, if $\sigma = 0$, then,

$\hat{F}_1 = \frac{N}{D}$, $\quad N, D$ defined before

Now since $\hat{F}_1 > 1 \Rightarrow \frac{N}{D} > 1 \Rightarrow N > D$

Now since $U < V$ and $D < N \Rightarrow UD < VN \Rightarrow UD - VN < 0$

Next $(D+\sigma^2(<a_1,\ a_1>-1)V)>0$

$$V=\sum_{\alpha}\frac{1}{t(\alpha)s(\alpha)}>0 \text{ since } 0^\circ<\alpha<20^\circ$$

$$<a_1,\ a_1>=\sum_{\alpha}1$$

$$\Rightarrow <a_1,\ a_1>-1=\sum_{\alpha}1-1>0 \text{ if the number of } \alpha>1$$

$$\Rightarrow \text{since D}>0\ (D+\sigma^2(<a_1,\ a_1>-1)V)>0$$

$$\text{All that together } \Rightarrow \mathbb{E}\frac{\hat{N}}{\hat{D}}-\frac{N}{D}<0$$

$$\text{Since } F_1=\frac{\hat{N}}{\hat{D}}\Rightarrow \mathbb{E}\hat{F_1}-\frac{N}{D}<0$$

$$\text{Now}\hat{F_1}=\exp\left(\frac{T_R}{\hat{T}_1}\right) \text{ and } \frac{N}{D}=\exp\left(\frac{T_R}{T_1}\right)$$

($\hat{T}_1$ is the estimate of $T_1$ in noise)

$$\Rightarrow\ \exp\left(\frac{T_R}{\hat{T}_1}\right)-\exp\left(\frac{T_R}{T_1}\right)<0$$

$$\Rightarrow \mathbb{E}\left[\exp\left(\frac{T_R}{\hat{T}_1}\right)\right]>\exp\left(\frac{T_R}{T_1}\right)$$

$$\Rightarrow \mathbb{E}\left[\exp\left(\frac{T_R}{\hat{T}_1}-\frac{T_R}{T_1}\right)\right]<1$$

Now $\phi(x) = e^x$ is convex in $x$ and so by Jensen's inequality,

$$\begin{aligned}
&\phi(\mathbb{E}[X]) \leq \mathbb{E}[\phi(X)] \\
&\Rightarrow e^{\mathbb{E}[X]} \leq \mathbb{E}[e^X] \\
&\Rightarrow e^{\mathbb{E}\left[\frac{T_R}{\hat{T}_1} - \frac{T_R}{T_1}\right]} \leq \mathbb{E}\left[\exp\left(\frac{T_R}{\hat{T}_1} - \frac{T_R}{T_1}\right)\right] < 1 \\
&\text{if } e^x < 1 \Rightarrow x < 0 \\
&\Rightarrow \mathbb{E}\left[\frac{T_R}{\hat{T}_1} - \frac{T_R}{T_1}\right] < 0 \\
&\mathbb{E}\left[\frac{1}{\hat{T}_1}\right] < \mathbb{E}\left[\frac{1}{T_1}\right] = \frac{1}{T_1}
\end{aligned}$$

Take $\hat{\hat{T}}_1 = \max(\delta, \hat{T}_1)$ where $\delta > 0$ and $\delta$ is very small

$$\begin{aligned}
&\Rightarrow \hat{\hat{T}}_1 > 0 \text{ and } \hat{\hat{T}}_1 \geq \hat{T}_1 \Rightarrow \frac{1}{\hat{\hat{T}}_1} \leq \frac{1}{\hat{T}_1} \\
&\Rightarrow \mathbb{E}\left[\frac{1}{\hat{\hat{T}}_1}\right] \leq \mathbb{E}\left[\frac{1}{\hat{T}_1}\right] < \frac{1}{T_1}
\end{aligned}$$

Now $\phi(x) = \frac{1}{x}$ is convex in $x$ for $x > 0$ since $\phi''(x) = \frac{2}{x^3}$. So by Jensen's inequality,

$$\begin{aligned}
\phi(\mathbb{E}[X]) &\leq \mathbb{E}[\phi(X)] \\
\Rightarrow \frac{1}{\mathbb{E}[X]} &\leq \mathbb{E}\left[\frac{1}{X}\right] \\
\text{Since } \hat{\hat{T}}_1 > 0 \Rightarrow \frac{1}{E\left[\hat{\hat{T}}_1\right]} &\leq \mathbb{E}\left[\frac{1}{\hat{\hat{T}}_1}\right] < \frac{1}{T_1} \\
\Rightarrow \mathbb{E}\left[\hat{\hat{T}}_1\right] &> T_1.
\end{aligned}$$

# Bibliography

[1] Andrew L Alexander, Samuel A Hurley, Alexey A Samsonov, Nagesh Adluru, Ameer Pasha Hosseinbor, Pouria Mossahebi, Do PM Tromp, Elizabeth Zakszewski, and Aaron S Field. Characterization of cerebral white matter properties using quantitative magnetic resonance imaging stains. *Brain connectivity*, 1(6):423–446, 2011.

[2] Alex Mackay, Kenneth Whittall, Julian Adler, David Li, Donald Paty, and Douglas Graeb. In vivo visualization of myelin water in brain by magnetic resonance. *Magnetic Resonance in Medicine*, 31(6):673–677, 1994.

[3] Mark D Does and John C Gore. Compartmental study of t1 and t2 in rat brain and trigeminal nerve in vivo. *Magnetic resonance in medicine*, 47(2):274–283, 2002.

[4] Sean CL Deoni, Brian K Rutt, Tarunya Arun, Carlo Pierpaoli, and Derek K Jones. Gleaning multicomponent t1 and t2 information from steady-state imaging data. *Magnetic Resonance in Medicine*, 60(6):1372–1387, 2008.

[5] Robert W Brown, Y-C Norman Cheng, E Mark Haacke, Michael R Thompson, and Ramesh Venkatesan. *Magnetic resonance imaging: physical principles and sequence design*. John Wiley & Sons, 2014.

[6] Matt A Bernstein, Kevin F King, and Xiaohong Joe Zhou. *Handbook of MRI pulse sequences*. Elsevier, 2004.

[7] Klaas P Pruessmann, Markus Weiger, Markus B Scheidegger, Peter Boesiger, et al. Sense: sensitivity encoding for fast mri. *Magnetic Resonance in Medicine*, 42(5):952–962, 1999.

[8] Paul M Margosian, Gordon DeMeester, and Haiying Liu. Partial fourier acquisition in mri. *eMagRes*, 2007.

[9] Douglas C Noll, Dwight G Nishimura, and Albert Macovski. Homodyne detection in magnetic resonance imaging. *IEEE Transactions on Medical Imaging*, 10(2):154–163, 1991.

[10] Jan Cuppen and Andre van Est. Reducing mr imaging time by one-sided reconstruction. *Magnetic Resonance Imaging*, 5(6):526–527, 1987.

[11] Michael Lustig, David L Donoho, Juan M Santos, and John M Pauly. Compressed sensing mri. *IEEE Signal Processing Magazine*, 25(2):72–82, 2008.

[12] Julia V. Velikina, Andrew L. Alexander, and Alexey Samsonov. Accelerating mr parameter mapping using sparsity-promoting regularization in parametric dimension. *Magnetic Resonance in Medicine*, 70(5):1263–1273, 2013.

[13] Florian Knoll, Kristian Bredies, Thomas Pock, and Rudolf Stollberger. Second order total generalized variation (tgv) for mri. *Magnetic Resonance in Medicine*, 65(2):480–491, 2011.

[14] M Green. Statistics of images, the tv algorithm of rudin-osher-fatemi for image denoising and an improved denoising algorithm. *CAM Report*, 2:55, 2002.

[15] Emmanuel J. Candes, Justin K. Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223, 2006.

[16] Steven Kecskemeti, Alexey Samsonov, Samuel A. Hurley, Douglas C. Dean, Aaron Field, and Andrew L. Alexander. Mpnrage: A technique to simultaneously acquire hundreds of differently contrasted mprage images with applications to quantitative t1 mapping. *Magnetic Resonance in Medicine*, 75(3):1040–1053, 2016.

[17] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(14):259 – 268, 1992.

[18] David Strong and Tony Chan. Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Problems*, 19(6):S165, 2003.

[19] Tom Goldstein and Stanley Osher. The split bregman method for l1-regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343, 2009.

[20] Per Christian Hansen. *Discrete inverse problems: insight and algorithms*. SIAM, 2010.

[21] David L Donoho. For most large underdetermined systems of linear equations the minimal l1-norm solution is also the sparsest solution. *Communications on pure and applied mathematics*, 59(6):797–829, 2006.

[22] Dana C Peters, Pratik Rohatgi, René M Botnar, Susan B Yeon, Kraig V Kissinger, and Warren J Manning. Characterizing radial undersampling artifacts for cardiac applications. *Magnetic resonance in medicine*, 55(2):396–403, 2006.

[23] Kai Tobias Block, Martin Uecker, and Jens Frahm. Undersampled radial mri with multiple coils. iterative image reconstruction using a total variation constraint. *Magnetic resonance in medicine*, 57(6):1086–1098, 2007.

[24] Mariya Doneva, Peter Börnert, Holger Eggers, Christian Stehning, Julien Sénégas, and Alfred Mertins. Compressed sensing reconstruction for magnetic resonance parameter mapping. *Magnetic Resonance in Medicine*, 64(4):1114–1120, 2010.

[25] Frederike H Petzschner, Irene P Ponce, Martin Blaimer, Peter M Jakob, and Felix A Breuer. Fast mr parameter mapping using k-t principal component analysis. *Magnetic resonance in medicine*, 66(3):706–716, 2011.

[26] Henry Z Wang, Stephen J Riederer, and James N Lee. Optimizing the precision in t1 relaxation estimation using limited flip angles. *Magnetic Resonance in Medicine*, 5(5):399–416, 1987.

[27] Klaas P Pruessmann, Markus Weiger, Peter Börnert, and Peter Boesiger. Advances in sensitivity encoding with arbitrary k-space trajectories. *Magnetic Resonance in Medicine*, 46(4):638–651, 2001.

[28] Vasily L Yarnykh. Actual flip-angle imaging in the pulsed steady state: a method for rapid three-dimensional mapping of the transmitted radiofrequency field. *Magnetic resonance in Medicine*, 57(1):192–200, 2007.

[29] Gilles Puy, Pierre Vandergheynst, and Yves Wiaux. On variable density compressive sampling. *IEEE Signal Processing Letters*, 18(10):595–598, 2011.

[30] Rachel W Chan, Elizabeth A Ramsay, Edward Y Cheung, and Donald B Plewes. The influence of radial undersampling schemes on compressed sensing reconstruction in breast mri. *Magnetic resonance in medicine*, 67(2):363–377, 2012.

[31] Sean CL Deoni, Terry M Peters, and Brian K Rutt. Determination of optimal angles for variable nutation proton magnetic spin-lattice, t1, and spin-spin, t2, relaxation times measurement. *Magnetic resonance in medicine*, 51(1):194–199, 2004.

[32] Lin-Ching Chang, Cheng Guan Koay, Peter J Basser, and Carlo Pierpaoli. Linear least-squares method for unbiased estimation of t1 from spgr signals. *Magnetic resonance in medicine*, 60(2):496–501, 2008.

[33] Joëlle K Barral, Erik Gudmundson, Nikola Stikov, Maryam Etezadi-Amoli, Petre Stoica, and Dwight G Nishimura. A robust methodology for in vivo t1 mapping. *Magnetic resonance in medicine*, 64(4):1057–1067, 2010.

[34] Matthias C Schabel and Glen R Morrell. Uncertainty in t1 mapping using the variable flip angle method with two flip angles. *Physics in medicine and biology*, 54(1):N1, 2008.

[35] Mingqiang Zhu. *Fast numerical algorithms for total variation based image restoration*. ProQuest, 2008.

[36] Florian Knoll. *Constrained MR image reconstruction of undersampled data from multiple coils*. Citeseer, 2011.

[37] Kristian Bredies, Karl Kunisch, and Thomas Pock. Total generalized variation. *SIAM Journal on Imaging Sciences*, 3(3):492–526, 2010.

[38] Alessandro Sbrizzi, Hans Hoogduin, Jan J Lagendijk, Peter Luijten, and Cornelis AT den Berg. Robust reconstruction of b1+ maps by projection into a spherical functions space. *Magnetic resonance in medicine*, 71(1):394–401, 2014.

[39] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.

[40] Urs Gamper, Peter Boesiger, and Sebastian Kozerke. Compressed sensing in dynamic mri. *Magnetic resonance in medicine*, 59(2):365–373, 2008.

[41] Hong Jung, Kyunghyun Sung, Krishna S Nayak, Eung Yeop Kim, and Jong Chul Ye. k-t focuss: A general compressed sensing framework for high resolution dynamic mri. *Magnetic Resonance in Medicine*, 61(1):103–116, 2009.

[42] Sajan Goud Lingala and Mathews Jacob. Blind compressive sensing dynamic mri. *IEEE transactions on medical imaging*, 32(6):1132–1145, 2013.

[43] Frederike H. Petzschner, Irene P. Ponce, Martin Blaimer, Peter M. Jakob, and Felix A. Breuer. Fast mr parameter mapping using k-t principal component analysis. *Magnetic Resonance in Medicine*, 66(3):706–716, 2011.

[44] Dana C Peters, Frank R Korosec, Thomas M Grist, Walter F Block, James E Holden, Karl K Vigen, and Charles A Mistretta. Undersampled projection reconstruction applied to mr angiography. *Magnetic Resonance in Medicine*, 43(1):91–101, 2000.

[45] Jianing V Shi, Wotao Yin, Aswin C Sankaranarayanan, and Richard G Baraniuk. Video compressive sensing for dynamic mri. *BMC Neurosci*, 13:183, 2012.

[46] Emmanuel J Candes and Terence Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE transactions on information theory*, 52(12):5406–5425, 2006.

[47] Scott Shaobing Chen, David L Donoho, and Michael A Saunders. Atomic decomposition by basis pursuit. *SIAM review*, 43(1):129–159, 2001.

[48] David L Donoho and Xiaoming Huo. Uncertainty principles and ideal atomic decomposition. *IEEE Transactions on Information Theory*, 47(7):2845–2862, 2001.

[49] David L Donoho and Philip B Stark. Uncertainty principles and signal recovery. *SIAM Journal on Applied Mathematics*, 49(3):906–931, 1989.

[50] Thierry Blu, Pier-Luigi Dragotti, Martin Vetterli, Pina Marziliano, and Lionel Coulot. Sparse sampling of signal innovations. *IEEE Signal Processing Magazine*, 25(2):31–40, 2008.

[51] Emmanuel J Candes and Michael B Wakin. An introduction to compressive sampling. *IEEE signal processing magazine*, 25(2):21–30, 2008.

[52] Michael Unser. Sampling-50 years after shannon. *Proceedings of the IEEE*, 88(4):569–587, 2000.

[53] David L Donoho, Yaakov Tsaig, Iddo Drori, and Jean-Luc Starck. Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 58(2):1094–1121, 2012.

[54] Zhongmin Wang and Gonzalo R Arce. Variable density compressed image sampling. *IEEE Transactions on image processing*, 19(1):264–270, 2010.

[55] Hong Jung, Jong Chul Ye, and Eung Yeop Kim. Improved k–t blast and k–t sense using focuss. *Physics in medicine and biology*, 52(11):3201, 2007.

[56] PP Vaidyanathan. Generalizations of the sampling theorem: Seven decades after nyquist. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 48(9):1094–1109, 2001.

[57] Albert Macovski. Noise in mri. *Magnetic Resonance in Medicine*, 36(3):494–497, 1996.

[58] John I Jackson, Craig H Meyer, Dwight G Nishimura, and Albert Macovski. Selection of a convolution function for fourier inversion using gridding (computerised tomography application). *IEEE transactions on medical imaging*, 10(3):473–478, 1991.

[59] Daniel Rosenfeld. An optimal and efficient new gridding algorithm using singular value decomposition. *Magnetic Resonance in Medicine*, 40(1):14–23, 1998.

[60] Philip J Beatty, Dwight G Nishimura, and John M Pauly. Rapid gridding reconstruction with a minimal oversampling ratio. *IEEE transactions on medical imaging*, 24(6):799–808, 2005.

[61] David L Donoho and Jared Tanner. Precise undersampling theorems. *Proceedings of the IEEE*, 98(6):913–924, 2010.

[62] Hossein Sedarat and Dwight G Nishimura. On the optimality of the gridding reconstruction algorithm. *IEEE Transactions on Medical Imaging*, 19(4):306–317, 2000.

[63] Deanna Needell and Rachel Ward. Stable image reconstruction using total variation minimization. *SIAM Journal on Imaging Sciences*, 6(2):1035–1058, 2013.

[64] Deanna Needell and Rachel Ward. Near-optimal compressed sensing guarantees for total variation minimization. *IEEE Transactions on Image Processing*, 22(10):3941–3949, 2013.

[65] Clarice Poon. On the role of total variation in compressed sensing. *SIAM Journal on Imaging Sciences*, 8(1):682–720, 2015.

[66] Dana C Peters, Rene M Botnar, Kraig V Kissinger, Susan B Yeon, Evan A Appelbaum, and Warren J Manning. Inversion recovery radial mri with interleaved projection sets. *Magnetic resonance in medicine*, 55(5):1150–1156, 2006.

[67] Jeffrey D Blanchard, Coralia Cartis, and Jared Tanner. Compressed sensing: How sharp is the restricted isometry property? *SIAM review*, 53(1):105–125, 2011.

[68] Tony Chan, Antonio Marquina, and Pep Mulet. High-order total variation-based image restoration. *SIAM Journal on Scientific Computing*, 22(2):503–516, 2000.

[69] Kai Tobias Block. Advanced methods for radial data sampling in magnetic resonance imaging. *SUB University of Goettingen*, 2008.

[70] Kristian Bredies and Martin Holler. A tgv-based framework for variational image decompression, zooming, and reconstruction. part i: Analytics. *SIAM Journal on Imaging Sciences*, 8(4):2814–2850, 2015.

[71] Kristian Bredies and Martin Holler. A tgv-based framework for variational image decompression, zooming, and reconstruction. part ii: Numerics. *SIAM Journal on Imaging Sciences*, 8(4):2851–2886, 2015.

[72] Antonin Chambolle, Vicent Caselles, Daniel Cremers, Matteo Novaga, and Thomas Pock. An introduction to total variation for image analysis. *Theoretical foundations and numerical methods for sparse recovery*, 9(263-340):227, 2010.

[73] Tony Chan, Selim Esedoglu, Frederick Park, and A Yip. Recent developments in total variation image restoration. *Mathematical Models of Computer Vision*, 17(2), 2005.

[74] Yiqiu Dong, Michael Hintermüller, and M Monserrat Rincon-Camacho. Automated regularization parameter selection in multi-scale total variation models for image restoration. *Journal of Mathematical Imaging and Vision*, 40(1):82–104, 2011.

[75] David Moroni Strong and Tony F Chan. *Relation of regularization parameter and scale in total variation based image denoising*. Department of Mathematics, University of California, Los Angeles, 1996.

[76] Tony F Chan and Selim Esedoglu. Aspects of total variation regularized l 1 function approximation. *SIAM Journal on Applied Mathematics*, 65(5):1817–1837, 2005.

[77] Jian-Feng Cai, Bin Dong, Stanley Osher, and Zuowei Shen. Image restoration: total variation, wavelet frames, and beyond. *Journal of the American Mathematical Society*, 25(4):1033–1089, 2012.