CHARACTERIZING DYNAMIC MRI USING OBJECTIVE IMAGE QUALITY METRICS

by

Nathan Murtha

Submitted in partial fulfillment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
July 2017

For all my family, who supported me throughout this thesis.

Table of Contents

List of Tables

List of Figures

Abstract

Quantitative mapping in dynamic magnetic resonance imaging (MRI) aims to recover quantitative parameters describing properties of patient tissues. However, quantitative mapping in dynamic MRI is difficult due to limitations imposed on temporal resolution to preserve image quality. Compressed sensing (CS) is a technique allowing increased temporal resolution while maintaining higher image quality than was traditionally possible. However, the compromise between quantitative mapping performance and image quality in dynamic CS-MRI remains to be explored. In this thesis we present three studies, building up towards the application of objective image quality metrics (IQMs) to characterize the performance of quantitative mapping in dynamic CS-MRI. We chose five IQMs for study: the root mean squared error (RMSE), the gradient magnitude similarity deviation (GMSD), the structural similarity index (SSIM), the multi-scale SSIM, and the information-weighted SSIM.

We first explored the design of a k-space acquisition scheme designed specifically for dynamic CS-MRI, utilizing the IQMs to assess the performance of each design. It was observed that designs did not differ significantly in performance within a fixed k-space undersampling factor.

Incorporating knowledge from the k-space acquisition scheme design study, we next investigated the properties of each IQM in static CS-MRI of the pelvis. Using simulated k-space acquisitions, the properties of the IQMs in static CS-MRI of 15 pelvic image data sets were investigated. It was found that the GMSD had the largest inter-image variation, followed by the RMSE and then the SSIM family of IQMs. Furthermore, the ability of each IQM to select favourable CS reconstruction parameters under different reconstruction conditions as undersampling of k-space increased was observed and characterized.

With knowledge of the behaviour of the IQMs in static CS-MRI, a first-generation simulation framework for quantitative mapping in dynamic CS-MRI was developed. We present preliminary results to validate the performance of the simulation framework. The ability to investigate changing CS reconstruction parameter needs for both

an emphasis on image quality and an emphasis on quantitative mapping performance in dynamic CS-MRI was demonstrated. Finally, evidence of correlations between IQM scores and quantitative mapping performance in dynamic MRI applications was demonstrated.

List of Abbreviations and Symbols Used

$C(\vec{r})$	Coil sensitivity profile
$G(t)$	Gradient magnetic field
GSM	Gaussian scale mixture
J	Side length of nested square in CIRCUS generation
K	Number of possible indices on square of side length J
P_C	Point spread function coherence
P_I	Point spread function incoherence
T_1	Longitudinal recovery time
T_2	Transverse relaxation time
Φ	Sensing basis.
Ψ	Sparse transform
γ	Gyromagnetic ratio
λ	Sparse regularization weight
$\mathcal{F}$	Fourier transform operator
$\phi(\vec{r}, t)$	Magnetic vector phase
ρ	Effective proton density (i.e. MRI image)
φ	The golden ratio
$\vec{B}_0$	External magnetic field in MRI scanner
$\vec{B}_1$	Perturbing magnetic field
$\vec{M}$	Net magnetization
$\vec{M}_z$	Parallel magnetization
$\vec{M}_{xy}$	Transverse magnetization
$\vec{\mu}$	Magnetic moment
$\vec{\omega}_0$	Larmor frequency
$\vec{c}$	Sparse expansion of $\vec{v}$ in Ψ
$\vec{k}$	Spatial frequency vector
$\vec{r}$	Spatial position
$\vec{s}$	Signal in $\mathbb{R}^M$, measurements of $\vec{v}$ with $M < N$

$\vec{v}$	Signal in $\mathbb{R}^N$
b	CIRCUS radial parameter
c	CIRCUS spiral parameter
q	CIRCUS quanta index
BART	Berkeley Advanced Reconstruction Toolbox
CIRCUS	CIRcular Cartesian UnderSampling
CS	Compressed sensing
DCE	Dynamic contrast enhanced
EES	Extravascular extracellular space
GMS	Gradient magnitude similarity
GMSD	Gradient magnitude similarity deviation
HVS	Human visual system
IQM	Image quality metric
IW-SSIM	Information-weighted structural similarity index
LLR	Locally low rank
MRI	Magnetic resonance imaging
MS-SSIM	Multi-scale structural similarity index
NSS	Natural scene statistics
PI	Parallel imaging

PSF	Point spread function
RF	Radiofrequency
RMSE	Root mean squared error
SSIM	Structural similarity index
TV	Total variation

Acknowledgements

There are many people who I would like to thank for making this thesis possible. My supervisors, Steven Beyea and Sharon Clarke, and my committee members, James Rioux and Chris Bowen, all of whom provided valuable guidance and answered many questions throughout the course of this work.

Acknowledgements must be paid to all of my family, for supporting me throughout during all of this work and throughout the writing of this thesis. To my friends as well, for many insightful conversations.

Finally, a special thanks to Peter Lee for giving me access to his personal computer at a time when I was in desperate need for computational power.

Chapter 1

Introduction

Magnetic resonance imaging (MRI) is a powerful imaging technique, capable of producing high quality diagnostic images of patient anatomy non-invasively and with no exposure to ionizing radiation. Since its introduction to the clinical scene in the early 1980's[1], MRI has experienced a boom in innovation and application. As part of this continued growth, objective image quality metrics (IQMs) developed in the field of image processing as an alternative to the root mean squared error (RMSE) have begun to be of interest to MRI researchers. These IQMs show better correlation with observer subjective scoring than traditional error-based IQMs, such as the RMSE, and may have utility not only in assessing the quality of static image reconstructions but also for assessing the accuracy of quantitative dynamic compressed sensing (CS) MRI applications.

In this thesis we investigate the proper implementation conditions of a k-space sampling scheme designed for CS-MRI. We then study five objective IQMs in static CS-MRI reconstructions, four of which are alternatives to the RMSE, utilizing the aforementioned k-space sampling scheme and incorporating the results of our implementation study. Finally, we design and validate the initial version of a simulation framework for the assessment of quantitative dynamic CS-MRI, which incorporates the previously investigated k-space sampling scheme and objective IQMs. Although we refer to quantitative dynamic contrast enhanced MRI as an exemplar in this thesis, the results presented in this thesis will motivate future work in many dynamic CS-MRI applications.

1.1 Quantitative Dynamic Magnetic Resonance Imaging (MRI)

MRI is a medical imaging modality with a wide number of uses. MRI can be used to obtain images of static anatomy, or it can be used to acquire a series of images to represent a dynamic process. A dynamic process is one where some aspect of the

images change over time. Examples of dynamic MRI applications include cardiac cine MRI and dynamic contrast enhanced (DCE) MRI. We utilize DCE-MRI as a motivating example of a dynamic MRI application. This was motivated by the existence of a many clinical applications for quantitative mapping utilizing DCE-MRI. We stress that the work presented in this thesis may be extended to other dynamic MRI applications.

DCE-MRI involves administering a bolus injection of paramagnetic contrast agent to the patient and characterizing its flow through the tissues. This is done by acquiring images prior to, during, and after injection of the contrast agent. The contrast enhancement profiles of tissues can inform radiologists as to the underlying properties of the tissue (e.g. perfusion, permeability) which in turn can inform them in making a diagnosis. While DCE-MRI has often been analysed in a non-quantitative method, whereby radiologists simply observe relative enhancement of a region of interest to a control region, or a semi-quantitative method, whereby non-physiological measures such as slope of enhancement profile or time to peak enhancement are taken, recent advancement in the field of MRI have been trending towards a quantitative analysis of DCE-MRI data.

Quantitative analyses of DCE-MRI involve modelling the tissues in a patient as a set of compartments through which the contrast agent flows, a process also known as pharmacokinetic mapping. These compartments represent the blood plasma and the extracellular extravascular space (EES). Model parameters describe factors such as the volume fraction of each compartment per unit volume of tissue, the influx of contrast agent into the blood plasma compartment, leakage of contrast agent from the blood plasma compartment to the EES compartment, and back leakage from the EES compartment into the blood plasma compartment. An example of a widely used pharmacokinetic model is the extended Tofts model[2;3]. The end result of pharmacokinetic mapping is a set of quantitative parameter maps for each parameter in the model.

Quantitative mapping with DCE-MRI has traditionally been challenging in MRI. High temporal resolution is needed to characterize the rapid uptake of contrast agent through tissues. To get high temporal resolution, the data must be undersampled. This will quickly degrade image quality. Compressed sensing, introduced in section

1.2, permits maintenance of image quality at higher undersampling factors. Combined with golden ratio sampling, which permits flexible retrospective decision in the compromise between temporal resolution and image quality (and is introduced in section 1.3), quantitative mapping with DCE-MRI has gained traction in clinical studies. However, a method for characterizing the compromise between image quality and temporal resolution is required. Quantitative DCE-MRI involves voxel-wise extraction of enhancement profiles to which the chosen signal evolution model is fit; poor image quality will degrade the enhancement profiles, decreasing the accuracy of recovered quantitative parameters. We propose to use objective IQMs, introduced in section 1.4, to characterize quantitative dynamic MRI.

An exemplar anatomy for the application of quantitative dynamic MRI is the prostate. Prostate cancer exhibits rapid enhancement in DCE-MRI, necessitating high temporal resolution[4;5]. Image quality is also paramount for prostate imaging since the prostate is a relatively small organ with a nested structure[6;7], such that deteriorating image quality may drastically affect the accuracy of quantitative mapping of the prostate. For this reason, we utilized clinical pelvic MRI images as part of our study; it is emphasized that the scope of the work presented in this thesis extends to other anatomical applications however.

1.2 Compressed Sensing (CS) and MRI

While MRI provides excellent soft tissue contrast and does not subject patients to ionizing radiation, the acquisition of high quality images often takes a long time compared to the rate of dynamic processes in the body (e.g. contrast agent flow). Data acquisition in MRI involves sampling the spatial frequency content of an image, also known as the *k-space* of the image. Sampling k-space necessitates repeated perturbation of the protons in a tissue sample to generate an MRI signal, which in turn causes the inherent "slowness" of MRI. Producing quality diagnostic MR images through sufficient sampling of k-space can take on the order of minutes. The resulting limitations in temporal resolution pose difficulties for quantitative dynamic techniques like quantitative DCE-MRI. To increase temporal resolution k-space may be undersampled. Undersampling has traditionally been performed by uniformly skipping portions of the k-space, but this rapidly leads to severe losses in image

quality.

Compressed sensing (CS) is a strategy that allows significant undersampling of a signal while still achieving high fidelity reconstructions of that signal, thereby speeding up the acquisition process. There are three conditions that must be satisfied for CS to provide reasonable results[8–10]:

1. The signal must be sampled (pseudo-)randomly, so that any aliasing artifacts introduced by undersampling the signal are incoherent (i.e. "noise-like"). By the Shannon-Nyquist sampling theorem, uniform undersampling of a signal causes coherent aliasing artifacts. It's impossible to recover the original signal from beneath a series of coherent aliases, since the original signal cannot be distinguished. However, incoherent aliasing allows the original signal to be distinguished from beneath the aliasing "noise". In a sense CS-MRI then becomes a denoising problem, suppressing the incoherent aliasing "noise" to recover the image. Figure 1.1 compares a uniformly undersampled reconstruction of a pelvic MRI image to a pseudo-randomly sampled reconstruction of a pelvic MRI image; the advantage of pseudo-random undersampling is immediately clear.

2. The data must be sparse, or at least compressible, in an appropriate transform domain. Sparsity means that the majority of a signal is composed of zeros, whereas compressibility means that the majority of a signal is *nearly* zero. For example, many natural images and medical images are compressible in the wavelet or discrete cosine transform domains[9]. One may get an intuitive feeling for sparsity/compressibility from Figure 1.2. By retaining only the largest 50% of the wavelet coefficients of a pelvic image, a decent reconstruction is still possible.

3. A suitable non-linear reconstruction must be used. At its core, CS is simply a minimization problem where data consistency is exchanged for a sparse representation in a chosen transform basis using a sparse regularization. The simplest implementation of CS is as follows:

$$\arg\min_{\rho} \left(\|\mathcal{F}[\rho] - k\|_2^2 + \lambda \|\Psi\rho\|_1 \right) \tag{1.1}$$

Where $\mathcal{F}$ is the Fourier transform operator, ρ is the image reconstruction based on the k-space measurements k, and λ is the regularization weight given to

enforcing sparsity in a transform domain Ψ. The minimization problem becomes a balancing act, requiring proper choice of a value of Ψ and of λ that ensures sufficient data consistency while still removing the incoherent aliasing artifacts introduced through undersampling k-space. Equation 1.1 may in practise have several regularizations added onto the data consistency term, enforcing sparsity in several different transform domains.

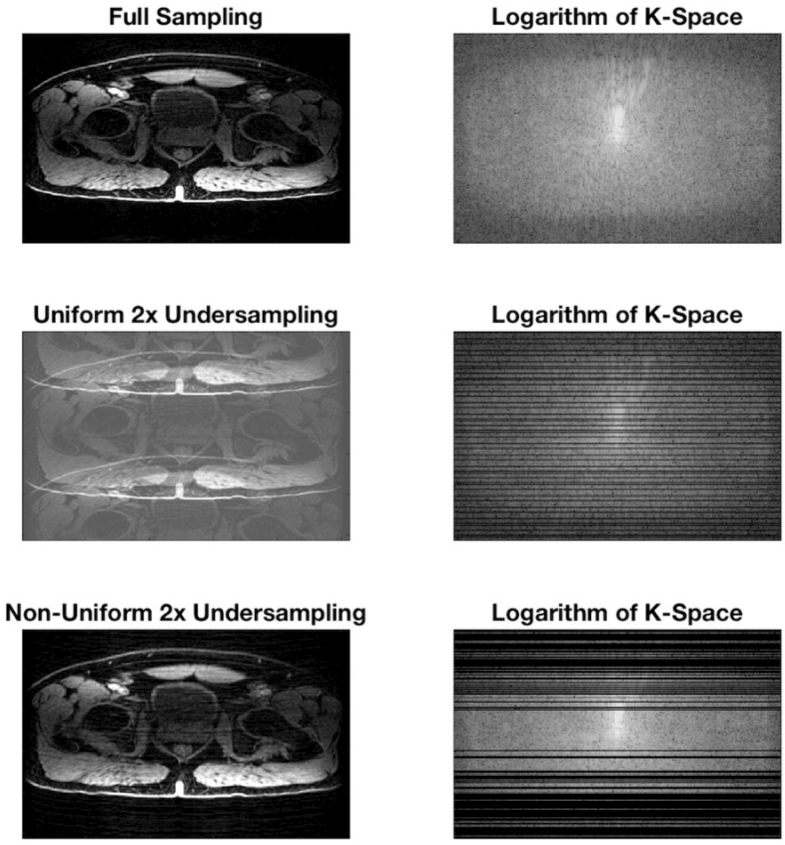

Figure 1.1: Uniform and pseudo-random undersampling of a pelvic MRI image, both retaining 50% of the original k-space data. While uniform undersampling results in coherent aliasing, pseudo-random undersampling creates noise-like incoherent aliasing artifacts below which the original image is easily visible. Pseudo-random sampling can be tailored, as in this example, to concentrate k-space sampling in regions that carry more image contrast (i.e. low spatial frequencies).

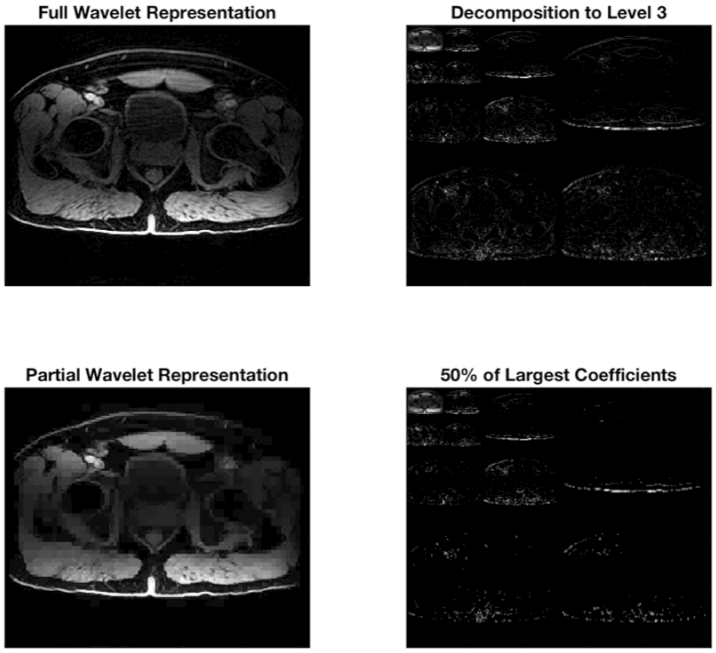

Figure 1.2: Medical images are often compressible in the wavelet domain. By retaining 50% of the wavelet coefficients in a pelvic MRI image, a decent reconstruction is still possible.

CS was in part motivated by lossy compression schemes for data storage (e.g. JPEG compression of images), where it was known that the majority of acquired data could be discarded once it was represented in an appropriate transform domain while still allowing very good recovery of the original data. The philosophy of CS is to acquire the data directly in such a compressed form. CS was originally proposed in a purely mathematical framework[11;12], and in short time its utility for reducing MRI acquisition time by undersampling k-space was realized[8]. Because of the freedom allowed in MRI to the user in acquiring k-space data, pseudo-random sampling of k-space is easily achieved (as demonstrated in Figure 1.1). Additionally, MR images are very often sparse in some appropriate transform domain (as demonstrated in Figure 1.2), and a non-linear reconstruction scheme is easily implemented in MRI to reconstruct the image. Thus, MRI acquisitions satisfy all of the criteria for CS listed above.

CS is beneficial to MRI because acquiring fewer points in k-space results in faster acquisition times (i.e. better temporal resolution). This is of particular importance to quantitative DCE-MRI, where high temporal resolution is needed without a detrimental loss of image quality.

1.3 Golden Ratio Sampling of K-Space

There exist several ways that k-space may be sampled, examples of which are shown in Figure 1.3. Cartesian trajectories acquire one line of k-space at a time through "frequency encoding", restarting at the beginning of the next line via "phase encoding" before the next acquisition begins. Zig-zag trajectories will run a continuous path through k-space, running across one frequency encoding line before being incremented to another line of k-space by a phase encoding gradient and running parallel-and-opposite to the original frequency encoding line, and so on until the end of the acquisition. Spiral trajectories either spiral towards or away from the center of k-space. Radial trajectories trace diameters through the centre of k-space, with equal angles between each successive diameter. With any of these sampling techniques, the temporal resolution of the scan is largely limited by the number of k-space points acquired.

To achieve the flexibility of retrospectively compromising between temporal resolution and image quality, small amounts of k-space data can be considered as "base-units" that are interleaved after acquisition in appropriate amounts to mediate the trade-off. For example during radial k-space acquisitions one diameter in k-space may be considered as a base-unit of k-space data, and diameters may be interleaved after k-space acquisition to create k-space volumes for image reconstruction. Interleaving many diameters would allow a high image quality at the expense of lower temporal resolution, whereas interleaving fewer diameters would allow high temporal resolution at the expense of reduced image quality. The number of base-units that are interleaved may easily be chosen retrospectively for the desired application (e.g. high quality images for viewing by a radiologist or high temporal resolution images for quantitative dynamic MRI applications).

Winkelmann et. at.[13] devised a k-space acquisition scheme that allows for flexible choice between reconstructed image quality and temporal resolution utilizing an

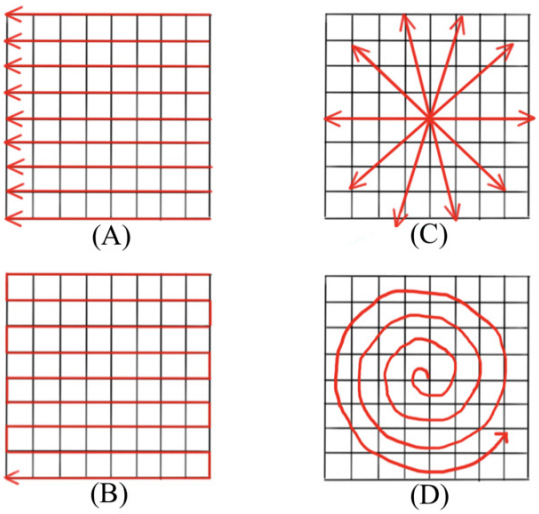

Figure 1.3: Some trajectories sample on a uniform grid. Examples include (A) Cartesian and (B) zig-zag patterns. Other trajectories trace paths in k-space that do not conform to a uniform grid. Examples include (C) radial and (D) spiral patterns/

extension of the simple example presented above. To achieve uniform k-space coverage for an arbitrary number of interleaved diameters, they proposed incrementing the angle between diameters in radial sampling using the golden ratio ($\varphi \approx 1.618$) to modulate the incrementation angle rather than using an incrementation angle calculated for a fixed number of diameters that were a priori chosen for interleaving. That is, rather than incrementing the angle by $180°/N_d$ for a pattern with N_d intended diameters, they proposed incrementing the angle between two successive diameters by $180°/\varphi \approx 111.248°$ to achieve a relatively uniform sampling of k-space for an arbitrary number of combined diameters. This is demonstrated in Figure 1.4, where standard radial sampling with five diameters is compared to the flexibility given by golden ratio sampling. Additionally, because the golden ratio is an irrational number, the k-space diameters will never overlap one another except at the centre of k-space. This ensures that for any number of combined diameters there will always be distinct points of k-space represented with a uniform coverage, allowing complete retrospective selection of the k-space interleaving.

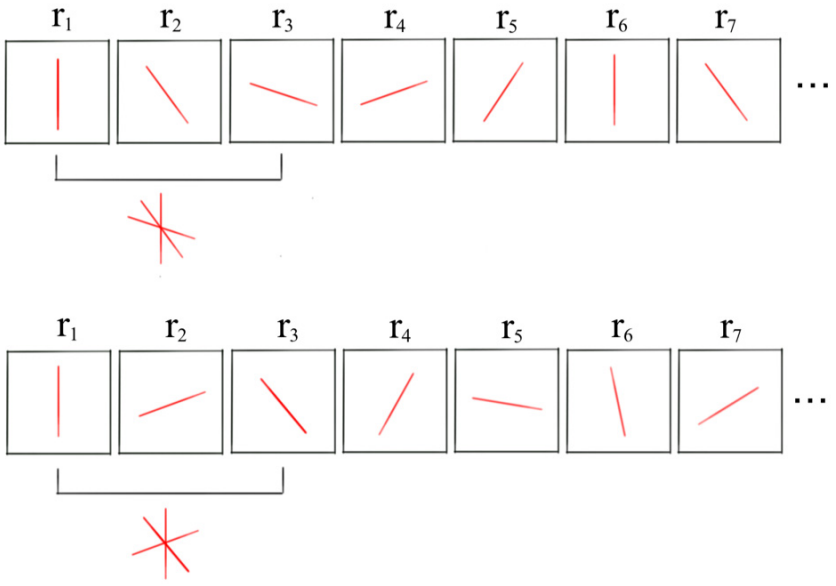

Figure 1.4: Interleaving three diameters with standard radial sampling designed with $N_d = 5$ results in an asymmetric coverage of k-space (top row), whereas interleaving more than five will result in k-space overlaps and no new k-space data. However, golden ratio sampling allows for interleaving of an arbitrary number of diameters while still maintaining relatively uniform coverage of k-space (bottom row).

An extension of the golden ratio sampling scheme has recently been proposed to take advantage of the flexible resolution capabilities offered by a golden ratio radial sampling of k-space and improve its suitability for CS, while adapting the theory for 3D Cartesian sampling. CIRcular Cartesian UnderSampling (CIRCUS)[14] combines aspects of both golden ratio radial sampling and spiral sampling k-space techniques, producing pseudo-random sampling patterns that can be interleaved arbitrarily with few overlapping points. The pseudo-random sampling patterns produced by CIRCUS are more amenable to CS-MRI than are traditional diameters in radial sampling, due to the incoherent aliasing they produce (see section 1.2 for the requirements of CS). By adjusting two parameters b and c, roughly corresponding to the "shear" and "twist" of the CIRCUS quanta, the operator can change the properties of the CIRCUS

trajectory. In this thesis we explore the b and c parameter space, with the intention of determining if favourable choices of b and c exist with regards to the ability of the CIRCUS patterns to inherently promote good CS-MRI performance.

1.4 Objective Image Quality Metrics (IQMs)

The golden standard of image quality assessment in any MRI application is the opinion of radiologists. However the surge of innovations in MRI hardware, k-space sampling, and image reconstruction techniques make it impractical to have radiologists extensively assess the quality of all of the resulting images. This is particularly true in applications where tuning of parameters is necessary for optimal image reconstruction quality, such as CS-MRI. To expedite the validation of new image reconstruction techniques, objective image quality metrics (IQMs) can be used. In MRI it is common to measure the root mean squared error (RMSE) between a reference image and an image reconstructed using a new experimental technique, where it is assumed that small RMSE implies high subjective quality in the reconstructed image. Studies in the field of image processing have long since established that the RMSE *is not* indicative of subjective image quality as perceived by human observers[15]. This is a result of the fact that the RMSE relies on pixel-wise error, and incorporates no knowledge of the properties of the human visual system (HVS). The need for rigorous investigation of the use of alternative objective IQMs incorporating properties of the HVS in the field of MRI is evident.

There are three broad categories of advanced IQMs. Full-reference IQMs compare a degraded image to a reference image, where the reference image is assumed to be ideal. No-reference IQMs make no comparison to a reference image, and assume that certain statistical properties should exist in high quality images; if the statistics of the input image differ, a no-reference IQM would thus indicate that the image is degraded from ideal in some way. A reduced-reference IQM sits between the previous two extremes, assuming the use of partial information from a reference image (e.g. a set of salient features). While reduced-reference IQMs haven't been given much attention in MRI, there has been a slowly growing body of interest in the field for both no-reference and full-reference IQMs. In this thesis we focus on the application of full-reference objective IQMs.

Full-reference IQMs model properties of the HVS in some way in order to assess how the degraded image will perceptually compare to the reference image (a detailed discussion of how this is done can be found in section 2.4.1). The structural similarity index (SSIM) developed by Wang et. al.[16] has achieved enormous success as a full-reference IQM in the field of image processing, and has increasingly been studied in the field of MRI. Kumar et. al.[17] used the SSIM to measure the quality of brain phantom images at increasing levels of JPEG compression, though made no reference to radiologist subjective scoring. A separate study by another group found that the SSIM was responsive to JPEG compression, additive white Gaussian noise, blurring, and contrast stretching in an MRI image of the brain[18]. More advanced studies have used the SSIM to quantify the performance of denoising techniques in MRI[19], or to quantify the performance of MR tractography[20]. Studies beginning to correlate the SSIM with radiologist subjective scoring in static image observation have also been undertaken. Miao et. al.[21] presented brain, cardiac, and abdominal MRI images that had Gaussian noise, Gaussian blur, or undersampling artifacts to two radiologists and found that the SSIM was reflective of their subjective scoring of image quality. Recent work by Akasaka et. al.[22] investigated the use of the SSIM in optimizing CS parameters for image quality in time-of-flight MR angiography of the brain, in which they found that the SSIM correlated well with subjective scoring of two radiologists and allowed for optimization of regularization weights in CS reconstructions. However, no investigations of IQMs in the assessment of quantitative mapping accuracy had been undertaken at the time of writing of this thesis. We believe that image quality may be a valuable predictor of the performance of quantitative dynamic MRI techniques.

A recent review of the assessment of medical image quality by Chow et. al.[23] asserted that the use of full-reference IQMs in medical imaging was inappropriate, due to the lack of a "perfect" reference image in practise[i]. While we agree that this is true for some imaging modalities in practise, such as in computed tomography where the production of a reference image would necessitate unnecessary dose to the patient, we disagree in the context of MRI. One method by which a reference image can safely be obtained in practical MRI is via a high resolution baseline scan, at the expense

[i]We wish to emphasize that Chow et. al. made this statement with broad reference to all medical imaging modalities. Their review was not limited solely to MRI.

of a few extra minutes spent scanning the patient. Image reconstruction quality in dynamic MRI applications can then be measured against the baseline image.

Given the prior discussions of this section, we recognize the need for a comprehensive study of full-reference IQMs in many MRI applications, particularly quantitative dynamic MRI. We believe that full-reference IQMs can serve as a valuable tool in the optimization of CS-MRI reconstruction parameters for high quality image reconstruction, and further that full-reference IQMs could be a valuable tool in predicting the accuracy of quantitative dynamic MRI applications (e.g. quantitative DCE-MRI). Towards this end we study the behaviour of four alternative IQMs in static CS-MRI images before developing a dynamic simulation framework to assess their utility in predicting the accuracy of quantitative mapping, and include the RMSE for comparison. The four alternative IQMs investigated are:

1. The SSIM, due to its recognition in the field of image processing and its increasing interest in MRI applications.

2. The multi-scale SSIM (MS-SSIM), an improvement on the SSIM intended to include the sensitivity of the HVS to distortions at different scales.

3. The information-weighted SSIM (IW-SSIM), which takes the MS-SSIM and includes statistical information weighting techniques in an attempt to highlight distortions that may be salient to the HVS.

4. The gradient magnitude similarity deviation (GMSD), which uses image gradients to detect and quantify the distortions and is designed to be computationally cheaper compared to IQMs like the SSIM.

1.5 Contributions of this Thesis

This thesis presents preliminary investigations of the use of objective IQMs in quantitative dynamic CS-MRI applications. To study the use of objective IQMs in dynamic CS-MRI applications, it is first necessary to find suitable methods of k-space acquisition and to characterize the behaviour of the IQMs in static CS-MRI image reconstructions. Towards this end we present three studies, the first two of which build

towards the development of a simulation framework for the assessment of quantitative dynamic CS-MRI.

We first explored the parameter range of the CIRCUS k-space sampling scheme discussed in section 1.3. Our goal was to determine if there were preferential CIRCUS trajectory designs that inherently allowed for better objective CS-MRI performance. This will assist in the compromise between image quality and temporal resolution that must occur in dynamic MRI. Utilizing a pelvic MRI image as an exemplar for an anatomy in which quantitative dynamic MRI applications exist and for which high image quality would be desired (see section 1.1), we simulated k-space acquisitions with many CIRCUS patterns by varying the design parameters b and c and utilized the five IQMs discussed in section 1.4 to characterize image reconstruction performance.

We next investigated quantitative behaviours of the five aforementioned IQMs in static CS-MRI reconstructions of 15 clinical pelvic MRI images, using results from our study of CIRCUS k-space acquisitions to obtain simulated k-space data. Understanding the properties of each IQM in static CS-MRI is essential before they be applied to quantitative dynamic MRI. Our goal was to obtain the quantitative background necessary to move forward with confidence in the development of a simulation framework for quantitative dynamic MRI, and furthermore to inform future studies utilizing IQMs to assess quantitative dynamic MRI applications in real world experiments.

Finally we developed a simulation framework for quantitative dynamic MRI using DCE-MRI as a motivating example, and present preliminary results validating its performance in a simple temporally evolving synthetic phantom. Our goal was to develop a simulation framework such that correlations between image quality and the accuracy of quantitative mapping in CS-MRI may be studied, and to present proof of concept results that objective IQMs may characterize quantitative dynamic CS-MRI.

Chapter 2

Background

2.1 Basic MRI Physics

Magnetic resonance imaging (MRI) uses strong magnetic fields, combined with brief applications of perturbing magnetic fields, to manipulate protons in the anatomy of the patient and produce an image. In this section, we summarize the physics of MRI and the methods by which data is acquired and images are formed. The description of MRI begins with reviewing the behaviour of a magnetization vector in a magnetic field. Following this, topics specific to encoding the MRI signal and recovering an image will be described. The section is concluded with a description of parallel imaging in MRI.

2.1.1 Classical Description of MRI

When placed in an external magnetic field $\vec{B}_0$, quantum mechanics predicts that protons in a sample may exist in discrete energy levels ϵ. Because protons possess a spin of one half, there exist two possible energy states:

$$\epsilon_{+1/2} = -\frac{1}{2}\gamma\hbar B_0 \tag{2.1}$$

$$\epsilon_{-1/2} = \frac{1}{2}\gamma\hbar B_0 \tag{2.2}$$

Where γ is known as the gyromagnetic ratio and is equal to 267.513 rad s^{-1} T^{-1} for protons and $\hbar$ is Planck's constant. Within a collection of protons at temperature T, the probability of finding a proton in an energy state ϵ is given by the Boltzmann distribution:

$$P(\epsilon) = \frac{e^{-\epsilon/kT}}{\sum_\epsilon e^{-\epsilon/kT}} \tag{2.3}$$

At typical human body temperatures of 37 °C = 310.15 K and at magnetic field strength of 3 T, the balance between thermal energy due to proton motions compared to the quantum energy levels gives rise to a relative excess of roughly 1 proton in

100,000 in the lower energy state. Approximating tissue as water, this means that each kilogram of tissue contains 6.69×10^{20} excess protons in the low energy state state. The enormous quantity of excess protons in the low energy state gives rise to a net magnetization vector that we may treat with classical mechanics. The resultant net magnetization per unit volume, $\vec{M}$, is given by the vector sum of all magnetic moments $\vec{\mu}$ of the protons in the sample of volume $\vec{V}$:

$$\vec{M} = \frac{1}{V} \sum_{\substack{\text{protons} \\ \text{in V}}} \vec{\mu} \tag{2.4}$$

Consider an external magnetic field which is aligned along the z-axis, $\vec{B}_0 = B_0\hat{z}$. When left at equilibrium, $\vec{M}$ forms parallel to $\vec{B}_0$ with equilibrium magnitude M_0. Because $M_0 \ll B_z$, detection of components of $\vec{M}$ parallel to $\vec{B}_0$ is impossible. However, perturbing $\vec{M}$ through the momentary application of a perpendicular magnetic field delivered via a radiofrequency pulse generates transverse magnetization components that *can* be detected. When perturbed from equilibrium in this fashion, $\vec{M}$ will precess about $\vec{B}_0$ as it returns to the equilibrium condition of $\vec{M} = M_0\hat{z}$. This motion is demonstrated in Figure 2.1 and described by the Bloch equation[24]:

$$\frac{d\vec{M}}{dt} = \gamma\vec{M} \times \vec{B}_0 + \frac{(M_0 - M_z)\hat{z}}{T_1} - \frac{\vec{M}_{xy}}{T_2} \tag{2.5}$$

Where $\vec{M}$ can be described in terms of parallel magnetization $\vec{M}_z$ and transverse magnetization $\vec{M}_{xy}$ via:

$$\begin{aligned} \vec{M} &= \vec{M}_z + \vec{M}_{xy} \\ \vec{M}_z &= M_z\hat{z} \\ \vec{M}_{xy} &= M_x\hat{x} + M_y\hat{y} \end{aligned} \tag{2.6}$$

The values T_1 and T_2 are empirically determined constants that describe the rate of exponential regrowth of $\vec{M}_z$ towards $\vec{M}_z = M_0\hat{z}$ and the rate of exponential decay of $\vec{M}_{xy}$ towards $\vec{M}_{xy} = 0$ respectively.

Equation 2.5 implies that the frequency of precession of $\vec{M}$ about $\vec{B}_0$ is linearly related to the magnitude of $\vec{B}_0$. This frequency is called the Larmor frequency, ω_0, and is given in general vector form by:

$$\vec{\omega}_0 = \gamma\vec{B}_0 \tag{2.7}$$

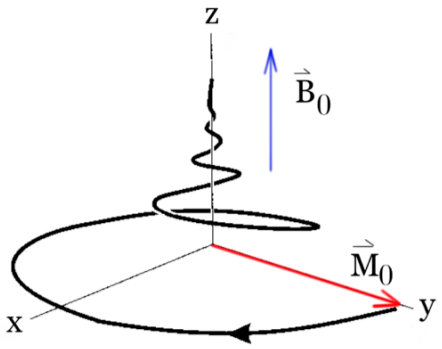

Figure 2.1: After the magnetization vector (shown in red) is perturbed away from equilibrium along the external magnetic field (shown in blue), it follows a helical path back towards equilibrium as the parallel magnetization regrows exponentially and the transverse magnetization decays exponentially. Figure adapted from Haacke et. al.[24]

2.1.2 Signal Detection and the K-Space Formalism

The source of the signal in MRI is the transverse component of magnetization, as discussed in section 2.1.1. As the transverse magnetization precesses at the Larmor frequency, the varying magnetic field lines induce an electromotive force (EMF) in nearby receive coils according to Faraday's law of induction:

$$EMF = \frac{d\Phi}{dt} \tag{2.8}$$

Where Φ is the magnetic flux (i.e. the component of the magnetic field $\vec{B}$ passing perpendicular to the surface of the coil at each spatial location):

$$\Phi = \int_{\substack{\text{coil} \\ \text{area}}} \vec{B} \cdot d\vec{A} \tag{2.9}$$

The vector $d\vec{A}$ is normal to the surface area of the coil and represents a differential area component. For the discussion of this section, we will assume a single receive coil with uniform sensitivity across the sample[i]. In section 2.1.5 we introduce the concept of individual coils with their own sensitivity profiles.

[i]In reality there may be an array of coils whose signals are combined prior to processing, such that no distinguishing information remains from the individual coils. In this case, approximation by a single uniform coil suffices for our purposes

Taken together, equation 2.8 and equation 2.9 can be used to show that the measured signal $s(t)$ is an exponentially decaying, rapidly oscillating function that depends on the Larmor frequency and on the initial magnitude of the transverse magnetization $\vec{M}_{xy}(\vec{r}, t = 0)$ following the application of a perturbing magnetic field. This is known as the *free induction decay*[24].

$$s(t) \propto \omega_0 e^{-t/T_2(\vec{r})} \int M_{xy}(\vec{r}, 0) e^{i(\omega_0 t + \phi(\vec{r},t))} d^3r \tag{2.10}$$

Where $\phi(\vec{r}, t)$ is term describing the relative phase of the precessing magnetization vectors in the sample. This term is paramount to the k-space formulation of MRI, and we will return to it in greater detail later in this section.

If we make the assumption that signal acquisition takes place over a period that is much shorter than $T_2(\vec{r})$ at all locations in the sample, $e^{-t/T_2(\vec{r})} \approx 1$ and we may ignore the effects of transverse magnetization decay in our subsequent analysis. To make the proportionality an equivalence, we introduce a constant η that accounts for the gain factors in the detector coils and in the electronics processing the signal. We then obtain a simplified equation describing the signal evolution which we may analyze:

$$s(t) = \eta\omega_0 \int M_{xy}(\vec{r}, 0) e^{i(\omega_0 t + \phi(\vec{r},t))} d^3r \tag{2.11}$$

If we now translate to a reference frame rotating at angular frequency ω_0[ii] and combine η, ω_0, and $M_{xy}(\vec{r}, 0)$ into an effective proton density $\rho(\vec{r})$, the signal in equation 2.11 is conveniently expressed as[24]:

$$s(t) = \int \rho(\vec{r}) e^{\phi(\vec{r},t)} d^3r \tag{2.12}$$

The importance of the phase term $\phi(\vec{r}, t)$ alluded to at the beginning of this section now becomes more apparent. The phase of the magnetization vectors in the sample can be used to spatially distinguish them. To see this, recall that phase is given by the amount of rotation incurred by the precessing magnetization over time:

$$\phi(\vec{r}, t) = \int_0^t \omega(\vec{r}, \tau) d\tau \tag{2.13}$$

In the rotating frame of reference that we have assumed, the transverse magnetization vectors across the sample appear stationary (since they too rotate at ω_0), and equation

[ii]The change of perspective is known as *demodulation*, and is performed by components of the signal detection system in practice.

2.13 implies that there is net-zero phase across the entire sample[iii]. By superimposing linear gradient magnetic fields $\vec{G}(t)$ on top of the static magnetic field across the sample, a spatially dependent change in precessional frequency occurs as a result of the dependence of the Larmor frequency on the magnetic field (equation 2.7):

$$\omega(\vec{r},t) = \omega_0 + \gamma\vec{r}\cdot\vec{G}(t) \tag{2.14}$$

Thus, in the rotating reference frame, the phase accrued by the magnetization vectors across the sample is determined by the difference in their precession frequency under the presence of external gradients from the Larmor frequency in the original static magnetic field. The resulting phase is:

$$\phi_G(\vec{r},t) = -\gamma\int_0^t \vec{r}\cdot\vec{G}(\tau)d\tau \tag{2.15}$$

Where the negative factor is introduced to account for the clockwise precession of the magnetization having positive gyromagnetic ratio. The application of spatially dependent magnetic gradient fields is thus seen to spatially encode the relative phases of the protons. If we make the following substitution:

$$\vec{k} = \frac{\gamma}{2\pi}\int_0^t \vec{G}(\tau)d\tau \tag{2.16}$$

Then the combination of equations 2.12, 2.15, and 2.16 gives:

$$s(\vec{k}) = \int \rho(\vec{r})e^{-i2\pi\vec{k}\cdot\vec{r}}d^3r \tag{2.17}$$

Equation 2.17 is immediately recognizable as the Fourier transform of the effective proton density. Recovery of the effective proton density, and thus the MRI image, is achieved by taking the inverse Fourier transform:

$$\rho(\vec{r}) = \int s(\vec{k})e^{i2\pi\vec{k}\cdot\vec{r}}d^3k \tag{2.18}$$

Equations 2.17 and 2.18 represent the k-space formalism of MRI. The application of magnetic field gradients on top of the static magnetic field in an MRI scanner allows for the reconstruction of spatially encoded proton densities from the measured free induction decay signal. Because the magnetic field gradient $\vec{G}$ has units of Tesla

[iii] We have assumed that our rotating frame of reference was chosen to align the axis of phase measurement with the initial transverse magnetization.

per meter, the units of $\vec{k}$ are in spatial frequency. MRI may thus be thought of as a tool that measures spatial frequencies in an image, and reconstructs the image by taking the inverse Fourier transform of the acquired frequencies.

While $\vec{k}$ is continuous and spans an infinite domain in theory, sampling of k-space necessitates a discrete sampling process. Equation 2.17 and equation 2.18 are then discretized by the use of the discrete Fourier transform. For an $N_x \times N_y \times N_z$ image, the discrete Fourier relationship between the image and the k-space is:

$$s(k_x, k_y, k_z) = \sum_{x=0}^{N_x-1} \sum_{y=0}^{N_y-1} \sum_{x=0}^{N_z-1} \rho(x, y, z) e^{-i2\pi\left(\frac{(\Delta k)_x x}{N_x} + \frac{(\Delta k)_y y}{N_y} + \frac{(\Delta k)_z z}{N_z}\right)} \tag{2.19}$$

$$\rho(x, y, z) = \frac{1}{N_x N_y N_z} \sum_{x=0}^{N_x-1} \sum_{y=0}^{N_y-1} \sum_{x=0}^{N_z-1} s(k_x, k_y, k_z) e^{i2\pi\left(\frac{(\Delta k)_x x}{N_x} + \frac{(\Delta k)_y y}{N_y} + \frac{(\Delta k)_z z}{N_z}\right)} \tag{2.20}$$

Where k_r in each dimension is now discretized with a stepsize $(\Delta k)_r$. The choice of $(\Delta k)_r$ is important in MRI, and will determine whether or not aliasing artifacts occur.

2.1.3 Sampling the K-Space

A k-space trajectory refers to the path traced in the spatial frequency domain as magnetic field gradients are applied to a sample during imaging, as per equation 2.16. While there are virtually infinite k-space trajectories in theory, in practise patient comfort/safety concerns and hardware limitations mean that the development of k-space trajectories requires careful planning and implementation.

There are two classes of k-space sampling trajectories: those that sample on a uniform grid and those that do not. A variety of k-space trajectories are illustrated in Figure 1.3. For the purposes of this thesis, we are concerned only with Cartesian sampling of k-space. As such, we restrict our focus to the introduction of terminology and concepts relevant to Cartesian sampling trajectories.

Equation 2.16 indicates that movement through k-space is achieved through a balance of gradient magnitude and gradient duration; it is *gradient area* that matters. We utilize this fact during *phase encoding* and *frequency encoding* in standard Cartesian imaging. For simplicity, we illustrate these concepts for 2D imaging with

perfect boxcar gradients of magnitude G[iv]. Under this assumption, equation 2.16 may be simplified:

$$k = \frac{\gamma}{2\pi} G \Delta t \tag{2.21}$$

The relationship between gradient area and position in k-space is now made exceedingly obvious. For example, equivalent distances are covered in k-space through the application of a gradient with a magnitude of 9 mT/m for 3 msec or for the application of a gradient with a magnitude of 3 mT/m for 9 msec.

We now briefly detail movement through a 2D plane in k-space. Let a sample of protons reside in a static magnetic field $\vec{B}_0 = B_z \hat{z}$, to which constant gradient magnetic fields $\vec{G}_y = G_y \hat{y}$ and $\vec{G}_x = G_x \hat{x}$ are intermittently applied after the application of a perturbing magnetic field $\vec{B}_1$ delivered by a radiofrequency (RF) pulse. Without loss of generality, assume that phase encoding takes place along the y-axis and that frequency encoding takes place along the x-axis. Phase encoding involves the relatively brief application of a gradient to achieve a step through k-space in one dimension using the gradient magnitude to obtain gradient area. Frequency encoding, which is performed after phase encoding and in the perpendicular dimension, entails the application of a constant gradient as time evolves to build up gradient area and will achieve many steps in k-space along the line specified by the phase encoding step. The differences are illustrated in Figure 2.2. To move through k-space in Cartesian imaging, phase encodes using gradients of varying magnitude are performed following the application of a perturbing RF pulse to generate transverse magnetization. The time between RF pulses is called the repetition time (TR). The need to regenerate transverse magnetization through the application of an RF pulse between each phase encode makes phase encoding the limiting factor in fast acquisition times using Cartesian imaging.

K-space traversal in 3D imaging is identical in concept, simply involving an extra phase encoding step for movement in the third dimension.

[iv] In reality, the gradient fields require a finite period of time to reach their intended magnitude, and must have their profiles adjusted for to account for the production of eddy currents. This is accounted for during an MRI scan, and further discussion is not relevant to the content of this thesis.

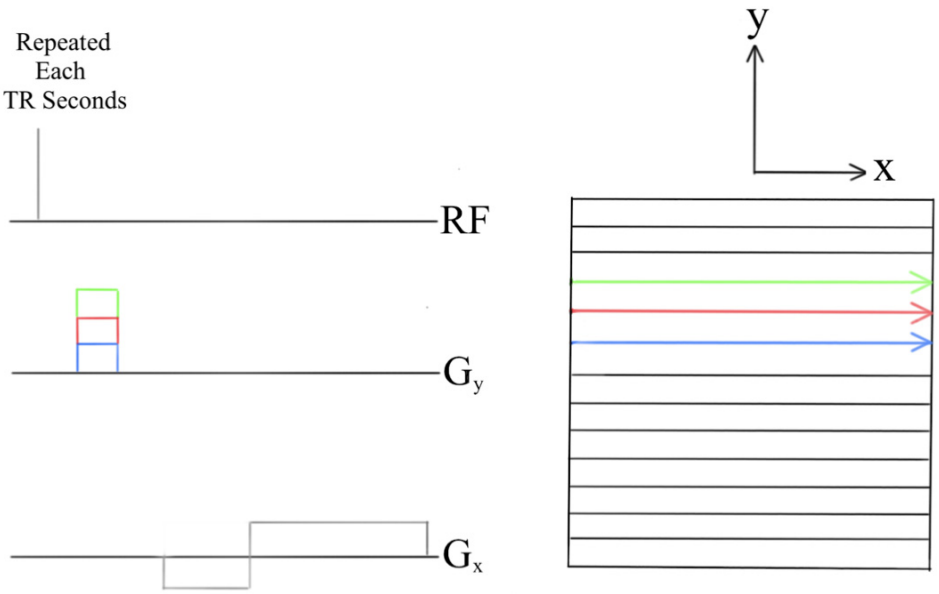

Figure 2.2: Phase encoding in the y-direction results in the selection of a particular row of k-space, after which frequency encoding will traverse the selected row to acquire data along the x-direction. Each new phase encoding step requires the application of an RF pulse to generate transverse magnetization, with TR seconds elapsing between the application of the RF pulses.

2.1.4 Relationship Between K-Space and the Image

In this section, we review the relationships between k-space and the image. We assume Cartesian sampling of k-space, whereby k-space samples are taken on a uniform grid via the frequency encoding and phase encoding processes summarized in section 2.1.3. A discussion of non-Cartesian k-space sampling techniques is beyond the scope of this thesis, but it is worth noting that non-Cartesian k-space techniques often employ a technique known as gridding that resamples the data onto a uniform Cartesian grid in order to facilitate image reconstruction[25]. Thus, while not entirely representative, a discussion of Cartesian k-space sampling properties will loosely inform the properties of non-Cartesian k-space sampling as well.

It is instructive to understand the pieces of k-space. The center of k-space contains low spatial frequencies, which carry the contrast information in the image. The periphery of k-space contains high spatial frequencies, which define the spatial resolution of the image. This is demonstrated in Figure 2.3. While there are no hard boundaries defining "low" from "high" spatial frequencies, in general there will be more detail and less contrast contributed to the final image as k-space coefficients are taken further from the center of the k-space.

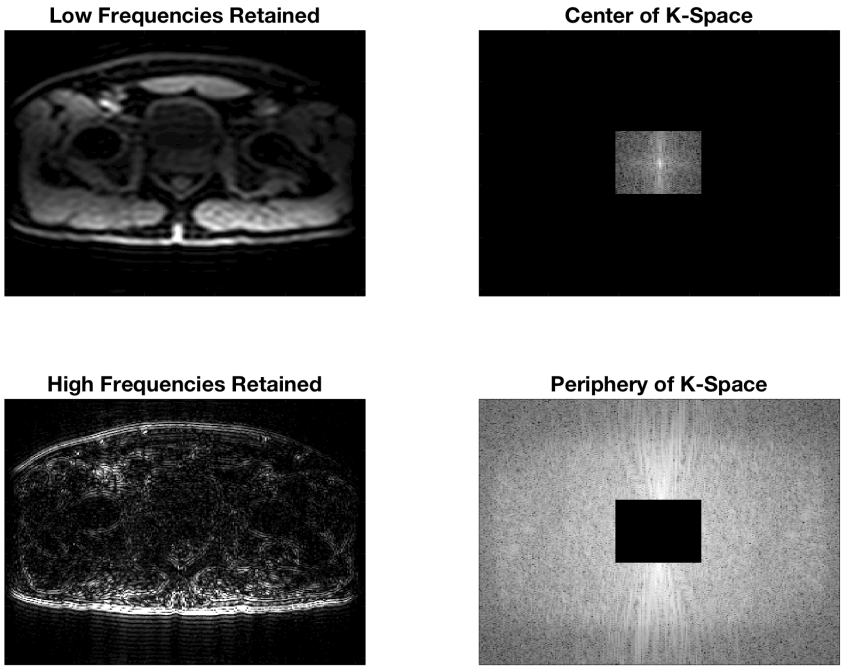

Figure 2.3: The center of k-space contains low spatial frequency data defining the contrast of the final image. The periphery of k-space contains high spatial frequency data defining the resolution and detail of the final image..

From equation 2.16, it is understood that traversal of k-space is accomplished through the application of magnetic field gradients for a specified duration of time. For k-space sampling with spacing $(\Delta k)_r$ between samples, where r represents any of

the three Cartesian axes, the field of view (FOV) is given by:

$$(FOV)_r = \frac{1}{(\Delta k)_r} \tag{2.22}$$

When the highest spatial frequency sampled in k-space[v] along Cartesian axis r is $(k_{max})_r$, the maximum image spatial resolution in that direction is given by:

$$\Delta r = \frac{1}{(k_{max})_r} \tag{2.23}$$

The sampling density of k-space therefore determines the FOV of the image, while the extent to which high spatial frequencies are sampled determines the image resolution.

Conventional MRI acquisitions must consider the relationships in equation 2.22 and equation 2.23 when acquiring k-space data. In an effort to speed the data acquisition process, a common technique in Cartesian imaging is to maintain the k_{max} required for acceptable maximum image resolution while uniformly skipping phase encodings and replacing the missing data with zeros. Undersampling k-space in this way will lead to aliasing artifacts as the FOV decreases and repetitions of the image begin to come into frame[vi]. To mitigate the introduction of these aliases, k-space sampling must follow the Nyquist criteria[24]

$$\Delta k_r < \frac{1}{D_r} \tag{2.24}$$

Where D_r is the spatial dimension of the object to be imaged along direction r. That is, to maintain a FOV free of aliasing artifacts when uniformly undersampling the k-space data, the k-space acquisition matrix must have steps no larger than the inverse of the spatial extent of the object being imaged.

The concept of aliasing is important to CS-MRI, which is discussed in detail section 2.2. The Nyquist criteria applies only to *uniformly* undersampled data and describes conditions for which aliasing artifacts will be *coherent*. If undersampling is performed in a non-uniform fashion, *incoherent* aliasing artifacts are introduced. The notion of coherence and incoherence is quantified by the point spread function (PSF). The PSF characterizes the blurring that will be experienced by a hypothetical point

[v] Here we assume that the k-space is sampled symmetrically for negative and positive frequencies in the direction of r.

[vi] This is sometimes called a "wrap-around" artifact, though in truth there is no wrapping occurring at all. Instead, indentical copies of the image are brought closer as the FOV decreases until their superposition makes it *appear* as if a wrap-around phenomena has occurred.

source when imaged by the MRI scanner. For a k-space trajectory design $H(k)$, the PSF $h(x)$ is given by the inverse Fourier transform[24]:

$$h(x) = \mathcal{F}^{-1}[H(k)] \tag{2.25}$$

Previous authors have proposed quantifying the coherence P_C or incoherence P_I of the PSF via the sidelobe-to-peak ratio[9] and the peak-to-noise ratio[26] respectively:

$$P_C = \frac{L_{side}}{L_{main}} \tag{2.26}$$

$$P_I = \frac{L_{main}}{\sigma_{side}} \tag{2.27}$$

Where L_{main} and L_{side} are the magnitude of the main lobe and maximum side lobe in the PSF and σ_{side} is the standard deviation of the side-lobe peaks. Coherent and incoherent aliases have very different properties, which are demonstrated in Figure 2.4 for a k-space in which 50% of the coefficients are sampled. Figure 2.4 makes clear that it is not just *how much* of the k-space that is sampled, but also *how* it is sampled that will affect image quality.

Finally, there is a relationship between the energy of the image and the energy contained in k-space. By Parseval's theorem, the energy in the image should be exactly equal to the energy in the sampled k-space.

$$\sum_{x=0}^{N_x-1}\sum_{y=0}^{N_y-1}\sum_{z=0}^{N_z-1}|\rho(x,y,z)|^2 = \frac{1}{N_xN_yN_z}\sum_{x=0}^{N_x-1}\sum_{y=0}^{N_y-1}\sum_{z=0}^{N_z-1}|s(k_x,k_y,k_z)|^2 \tag{2.28}$$

Where $\rho(x,y,z)$ and $s(k_x,k_y,k_z)$ are the image and k-space values, respectively (see equations 2.20 and 2.19).

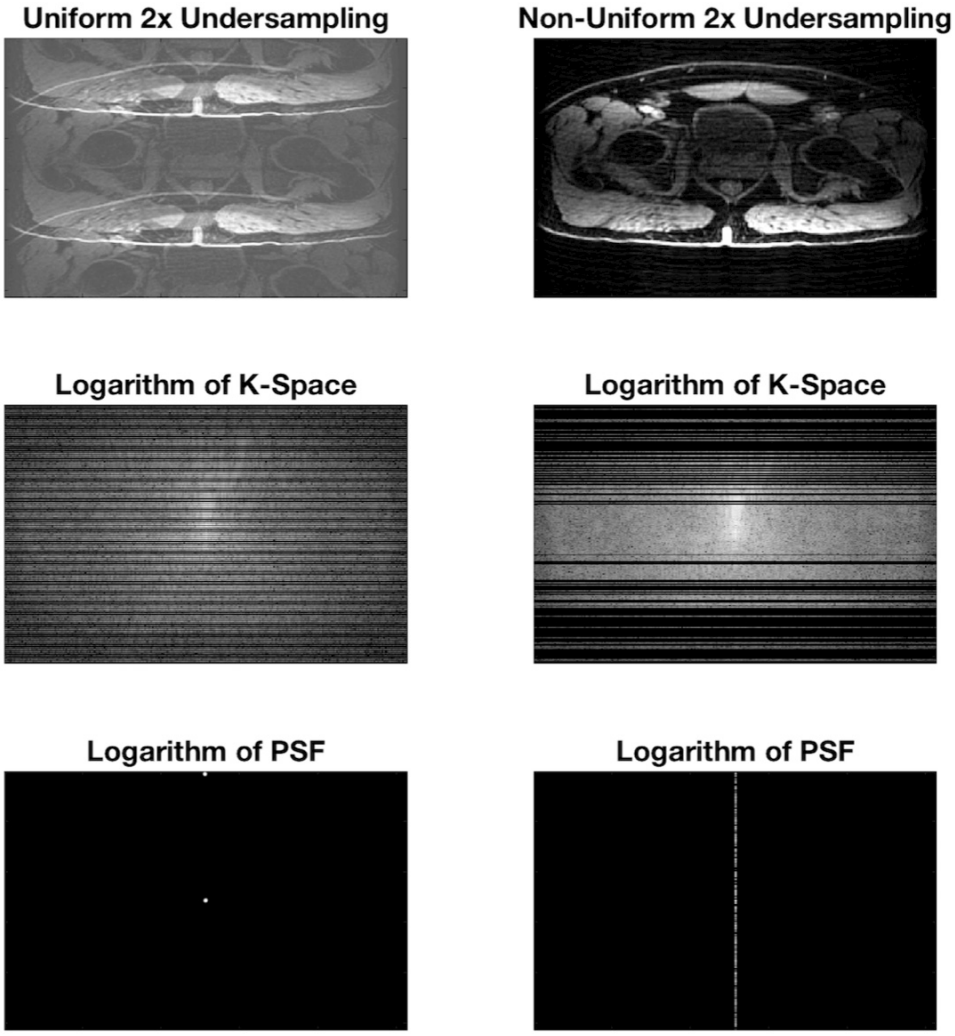

Figure 2.4: Uniform undersampling of k-space (left column) demonstrates coherent aliasing. The PSF of the k-space trajectory is highly coherent, consisting of repeated spikes spaced by half the field of view. Non-uniform undersampling of k-space (right column) demonstrates incoherent aliasing. The PSF is now incoherent, with small side lobes relative to the main lobe. The logarithm of the PSFs has been taken for visibility.

2.1.5 Parallel Imaging

In the derivation of equation 2.17, we assumed that there was a single receive coil with uniform sensitivity across the sample. In parallel imaging (PI) multiple receive coils are used in parallel, each with their own spatial sensitivity profile and with

their own channel in the data processing system. With knowledge of the sensitivity profiles of each coil, the spatial information encoded with each coil may be leveraged to allow the acquisition of fewer phase encoding lines while retaining image quality. The resulting decrease in acquisition time has made PI a standard in clinical MRI.

For a collection of coils, let the coil sensitivity profile of coil l be $C_l(\vec{r})$. Equation 2.17 is then modified in the following way:

$$s_l(\vec{k}) = \int C_l(\vec{r})\rho(\vec{r})e^{-i2\pi\vec{k}\cdot\vec{r}}d^3r \tag{2.29}$$

For a single uniform coil with sensitivity $C(\vec{r}) = 1$, equation 2.29 is observed to reduce to equation 2.17. Comparison of equation 2.17 to equation 2.29 reveals that with one uniform coil encoding of the signal is accomplished entirely through the application of gradients via the term $e^{-i2\pi\vec{k}\cdot\vec{r}}$, whereas the use of multiple coils in parallel allows encoding of the signal through the use of gradients *and* spatial coil sensitivities via $C_l(\vec{r})e^{-i2\pi\vec{k}\cdot\vec{r}}$. It is the additional encoding information contained in the coil sensitivities that permits reduced data acquisition time through the acquisition of fewer phase encoding steps, while still maintaining high image quality.

There are two main classes of PI reconstruction techniques, differing in the way that they reconstruct the images and not in the way that coils are employed in parallel. SENSitivity Encoding (SENSE)[27] first reconstructs aliased images obtained by each of the individual coils, and then uses knowledge of individual coil sensitivities to remove the aliases in the combination of the individual coil images. GeneRalized Autocalibrating Partial Parallel Acquisition (GRAPPA)[28] utilizes fully sampled lines in the center of k-space to calibrate weighting factors for each coil, which are used to make local kernels in k-space that interpolate missing phase encode lines *before* reconstruction of the final unaliased image. The reader is referred to the literature for a detailed discussion of the methodology of SENSE or GRAPPA. We illustrate the concept of spatial encoding within PI and the possibility for increased undersampling, and thus faster acquisition time, that follows as a result.

As mentioned in section 2.1.2, sampling k-space is a discrete procedure in practise. Let us define a discrete encoding function for an $N_x \times N_y \times N_z$ image:

$$E_l(x, y, z, k_x, k_y, k_z) = C_l(x, y, z)e^{-i2\pi\left(\frac{(\Delta k)_x x}{N_x} + \frac{(\Delta k)_y y}{N_y} + \frac{(\Delta k)_z z}{N_z}\right)} \tag{2.30}$$

In light of equations 2.19, 2.29, and 2.30, the discretized sampling of k-space in PI

may be viewed as[29]:

$$s_l(k_x, k_y, k_z) = \sum_{x=0}^{N_x-1} \sum_{y=0}^{N_y-1} \sum_{z=0}^{N_z-1} \rho(x, y, z) E_l(x, y, z, k_x, k_y, k_z) \tag{2.31}$$

The user may choose, through the process of phase encoding and frequency encoding, the number N_k of total k-space samples taken. The signal encoding process is conveniently expressed in matrix notation by concatenating the N_k k-space samples for N_c coils into a column vector and concatenating the 3D image into a column vector:

$$\vec{s} = E\vec{\rho} \tag{2.32}$$

Where $\vec{s}$ measures $N_c N_k \times 1$, E measures $N_c N_k \times N_x N_y N_z$, and $\vec{\rho}$ measures $N_x N_y N_z \times 1$. For the solution to equation 2.32 to be uniquely determined (i.e. for E to be invertible), we must have $N_c N_k = N_x N_y N_z$. Thus the greater that N_c becomes, the smaller N_k may be and thus that k-space can be undersampled in theory. We define the undersampling factor R to be the ratio of the number of acquired k-space points to the number of k-space points $(N_k)_{max}$ required to fully sample the discretized k-space:

$$R = \frac{N_k}{(N_k)_{max}} \tag{2.33}$$

Noting that $(N_k)_{max} = N_x N_y N_z$ as a result of the discrete Fourier transform relating k-space to the image, N_k is then given in terms of the undersampling factor via:

$$N_k = \frac{N_x N_y N_z}{R} \tag{2.34}$$

In light of equation 2.34, the maximum undersampling factor R that is theoretically achievable with PI is equal to the number of coils N_c that are used[vii] (this will keep E invertible).

As previously mentioned, undersampling in Cartesian sampling is most practically achieved by the omission of phase encoding lines. PI in modern MRI often uses 32 or more coils, theoretically allowing significant undersampling of k-space and a correspondingly significant reduction in data acquisition time. When combined with CS, which discussed in detail in section 2.2, even more significant reduction in acquisition time is possible.

[vii] In practise, the maximum undersampling factor will be less than N_c due to noise contamination and coil arrangement factors. Relatively significant undersampling is still achievable in practical PI however.

2.2 Compressed Sensing (CS)

CS techniques aim to acquire data directly in a compressed state. This differs from conventional sampling techniques which acquire a complete representation of the data in accordance with the Nyquist sampling theorem, following which compression techniques may be applied. The first mathematical methods to loosely venture in this direction were developed as far back as 1795, when Gaspard de Prony proposed an algorithm for estimating parameters of complex exponentials from a set of noisey samples[30]. The theory did not emerge as an established framework until the early 21st century however, when Candès, Romberg, Tao, and Donoho established conditions by which a finite-dimensional signal with a sparse or compressible representation in an appropriate basis could be recovered from a small (i.e. sub-Nyquist) set of non-adaptive measurements, provided the measurement scheme obeyed certain conditions.

In this section, we review the basics of CS theory and demonstrate the compatibility of CS with MRI. We conclude this section with a brief review of the Berkeley Advanced Reconstruction Toolbox (BART) for CS-MRI.

2.2.1 The CS Formalism

We shall examine the sensing formalism for discrete signals, since realistic sampling is necessarily a discrete process. In the most general sense, signals are represented as vectors in an N-dimensional signal space, where the coefficients in the signal vector represent the value of the signal at that location (e.g. measured voltage in time or image pixel luminance). Sampling a signal $\vec{v}$ in $\mathbb{R}^N$ can be formalized as taking projections of $\vec{v}$ onto the members of a sampling basis[10;29;30]. Suppose that we choose as a sampling basis the collection of M orthonormal vectors $[\phi_i \in \mathbb{R}^N : i = 1, 2, \ldots, M]$, and from this collection we form an $M \times N$ sensing matrix Φ formed by taking ϕ_i as its rows. Then a sampling $\vec{s} \in \mathbb{R}^M$ of the signal $\vec{v} \in \mathbb{R}^N$ is given by:

$$\vec{s} = \Phi\vec{v} \tag{2.35}$$

In CS, an undersampled set of measurements is taken and $M < N$. This results in equation 2.35 being underdetermined, and no unique solution for $\vec{v}$ can be found from knowledge of $\vec{s}$ without making assumptions about the nature of the signal $\vec{v}$.

According to CS theory, the appropriate assumption to make when trying to recover $\vec{v}$ from an undersampled set of measurements $\vec{s}$ is that $\vec{v}$ is sparse in some transform basis[10]. Let Ψ be the $N \times N$ transform basis formed by taking orthonormal basis elements $[\psi_i \in \mathbb{R}^N : i = 1, 2, \ldots, N]$ as its columns. Then there is some collection of coefficients $\vec{c} \in \mathbb{R}^N$ such that $\vec{v}$ can be expanded in Ψ:

$$\vec{v} = \Psi \vec{c} \tag{2.36}$$

The assumption is then that $\vec{c}$ is sparse, a notion which is discussed in section 2.2.3. As will discussed in section 2.2.3, real signals are very unlikely to be sparse, though they are often compressible. We will move forward in our discussion with the understanding that, when we refer to $\vec{c}$ as sparse, we often mean that $\vec{c}$ is sufficiently compressible to be accurately approximated as a sparse.

Equations 2.35 and 2.36 can be combined to reveal the full machinery of the sensing process in CS:

$$\vec{s} = \Phi \Psi \vec{c} \tag{2.37}$$

In words the meaning of equation 2.37 is, "through the basis Ψ our signal $\vec{v}$ exists as a sparse signal $\vec{c}$, and we obtain samples $\vec{s}$ of $\vec{c}$ through the sampling basis Φ". The challenge of CS is to recover $\vec{c} \in \mathbb{R}^N$ from a lower dimensional set of samples $\vec{s} \in \mathbb{R}^M$, where $M < N$.

To provide some intuition for equation 2.37, consider the simple analogy in Figure 2.5. A 3D shape is defined by a collection of vertices[viii] in $\mathbb{R}^3$. However, if an observer briefly views the shape edge-on from different orientations, they obtain 2D cross-sectional samples of the shape in $\mathbb{R}^2$. With appropriate assumptions about the nature of the full 3D object (e.g. the assumption that there are no divots in the surface), the observer could reconstruct the full 3D shape from a small collection of simplified 2D observations. That is, they can move from a lower dimensional vector space to a higher dimensional vector space.

[viii] In more complicated examples, we may have chosen a shape with divets and bumps on its surface as well, in which case we would also have to make note of the location and extent of those surface features.

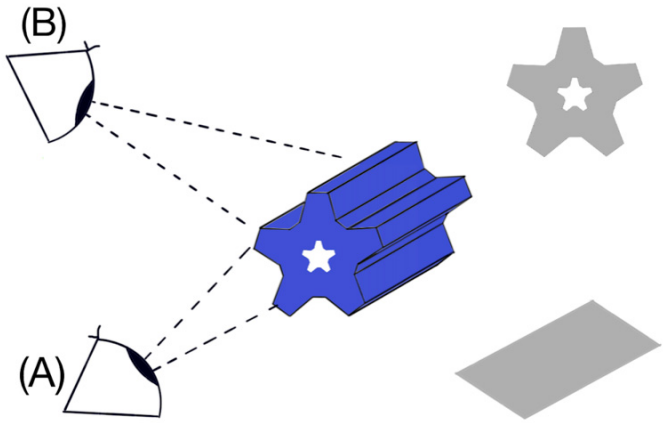

Figure 2.5: A simple analogy to demonstrate the sensing process of CS. An observer samples the cross-section profile of a 3D object from two positions, (A) and (B). From the small set of 2D cross-sectional samples, with appropriate assumptions regarding the nature of the full 3D object (e.g. smooth surfaces), the observer may reconstruct the full 3D object. CS allows recovery of information in $\mathbb{R}^N$ from information in $\mathbb{R}^M$ when $M < N$.

2.2.2 Conditions for Successful CS Reconstruction

There are three requirements for the successful application of CS[9]:

1. The signal must be sparse in some transform basis.

2. The signal must be sampled incoherently.

3. The signal must be reconstructed in a nonlinear manner.

We detail each of these requirements in the following sections.

2.2.3 Sparsity and Compressibility of Signals and Images

A signal $\vec{v}$ is said to be $p-$sparse if at most p of its coefficients are nonzero. For example, the following three signals in $\mathbb{R}^5$ are all 3-sparse:

$$[1,\ 1,\ 1,\ 0,\ 0]$$
$$[1,\ 1,\ 0,\ 0,\ 0]$$
$$[1,\ 0,\ 0,\ 0,\ 0]$$

Compressible signals are signals which are not truly *sparse*, but can be approximated as sparse due to the presence of many small but nonzero coefficients. If we sort the magnitude of the coefficients v_n of a compressible signal $\vec{v} \in \mathbb{R}^N$, such that $|v_1| \geq |v_2| \geq \ldots \geq |v_n|$, then the magnitudes would decay according to a power law[30]. For example, the following signal in $\mathbb{R}^5$ is compressible and might be approximated as a 2-sparse signal:

$$[1, 1, 0.0005, 0.0002, 0.0001] \approx [1, 1, 0, 0, 0]$$

The majority of signals are not sparse or compressible in their representative basis (e.g. an image directly in image space). Rather, an appropriate transform basis can often be found in which the signal has a compressible or even a sparse set of coefficients. This is known as *transform sparsity*, and is central to CS theory. The idea is conveniently illustrated by the Fourier transform of a sinusoid. Figure 2.6 shows a 1 Hz sinusoidal signal in time, and the corresponding Fourier transform of the signal. The sinusoid is clearly not sparse in the temporal domain, but requires only one coefficient to be represented with 100% accuracy in the Fourier domain; it is 1-sparse in the Fourier basis.

However, no data sampling in the real world comes without noise. Figure 2.7 shows a 1 Hz sinusoidal signal sampled in the presence of noise. While no longer sparse in the Fourier domain, the signal is *compressible* in the Fourier domain. Indeed, the sinusoidal behaviour of the signal can easily be captured by setting a threshold on the Fourier coefficients and reconstructing using only the coefficients with magnitude larger than the threshold. For example, setting a threshold at 10% of the magnitude of the largest Fourier coefficient is found to perfectly recover a denoised sinusoidal signal. Thus, the signal is compressible in the Fourier domain because it can be

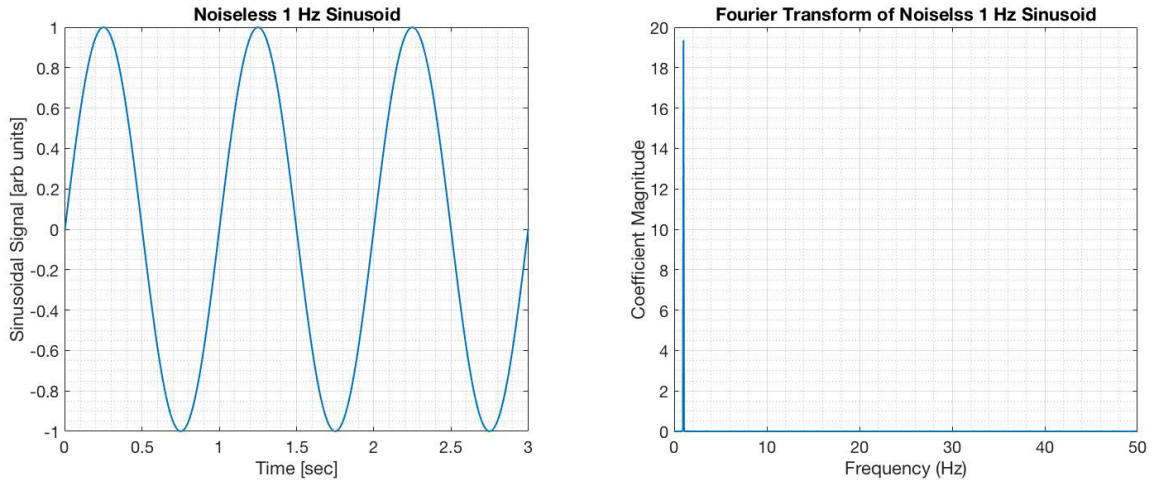

Figure 2.6: A 1 Hz sinusoid is not sparse in the time domain (left). The sinusoidal signal *is* sparse in an appropriate transform domain, in this case through the Fourier transform (right).

approximated very accurately by a sparse collection of the largest coefficients. This example suggests an alternative viewpoint for CS as a denoising problem. We shall return to this point of view in the context of applying CS to MRI.

Extending our discussion to the compressibility of images, it has been observed that the majority of images are not sparse or compressible directly in the image domain but can be made sparse under appropriate transform[9]. As an example, Figure 1.2 shows the effect of taking a wavelet transform of a pelvic MRI image. While not sparse or compressible in the image domain, the image is revealed to be compressible in the wavelet domain. The transform sparsity of MRI images satisfies the sparsity condition required for successful CS.

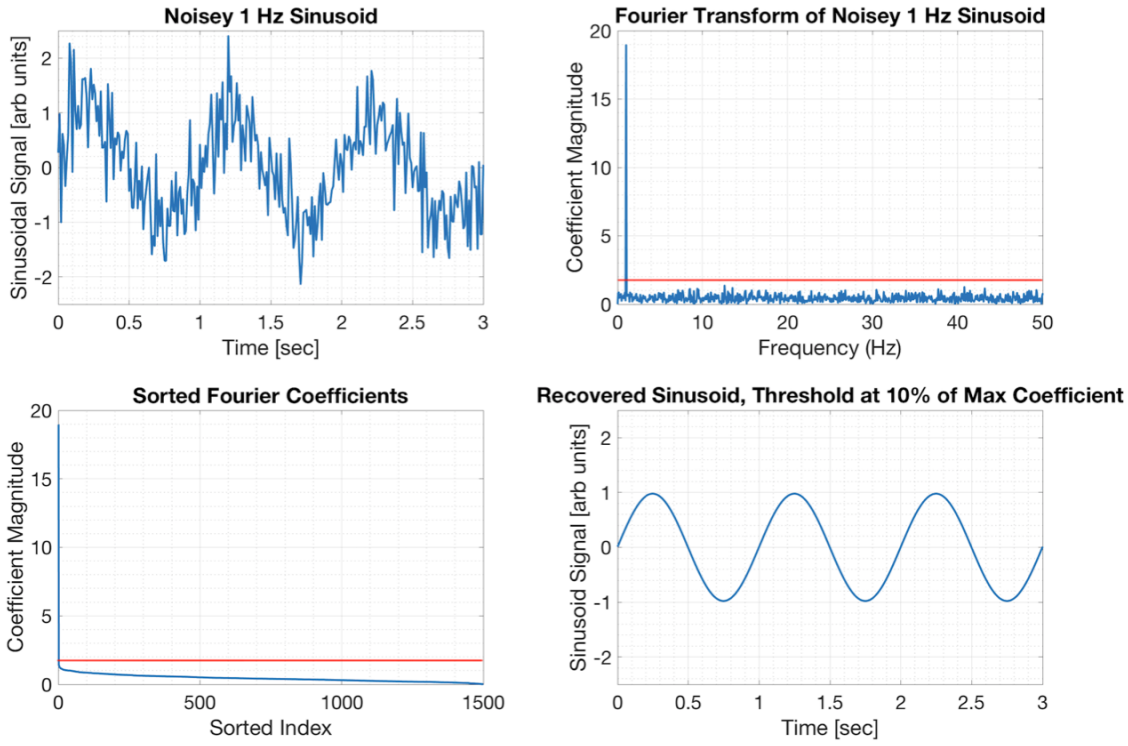

Figure 2.7: A noisey 1 Hz sinusoid (top left) and its Fourier transform (top right). The signal is not strictly sparse due to the presence of high frequencies from noise, though the power of the higher frequencies is relatively weak. Sorting the Fourier coefficients by their magnitude (bottom left) reveals the compressible nature of the noisey sinusoid in the Fourier basis. Setting a threshold on the Fourier coefficients to retain only the coefficients larger than 10% of the magnitude of the largest coefficient reveals the sinusoidal nature of the signal (bottom right). Threshold is displayed via red horizontal line.

2.2.4 Incoherent Sampling

CS requires that the relationship between Φ and Ψ be incoherent[ix], that is, minimally coherent. Coherence plays a role in determining the number of measurements, $\vec{s} \in \mathbb{R}^M$, required to recover a p-sparse signal $\vec{c} \in \mathbb{R}^N$. Lower coherence means fewer measurements are required (i.e. M may be smaller)[10].

[ix] When discussing incoherence, we *must* make reference to both a sampling and a transform basis. The notion cannot be defined without knowledge of both.

Defining $A = \Phi\Psi$, equation 2.37 reduces to:

$$\vec{s} = A\vec{c} \tag{2.38}$$

The coherence is then quantified by the mutual coherence μ of A[29]:

$$\mu(A) = \max_{i \neq j} \frac{|\langle A_i, A_j \rangle|}{\|A_i\|_2 \|A_j\|_2} \tag{2.39}$$

When the mutual coherence of A is low, the incoherence of A is high and the members of Φ and Ψ are largely uncorrelated. A convenient method for estimating $\mu(A)$ is to calculate the Gramian matrix G_A of A.

$$G_A = A^* A \tag{2.40}$$

Where the notation A^* is the complex transpose of A. When the off-diagonal entries of G_A are small, $\mu(A)$ is small, and the sampling basis Φ is incoherent with respect to the sparsifying transform Ψ[29].

In MRI, samples are taken in the Fourier basis. Figure 2.8 shows examples of G_A for Ψ taken as the Haar wavelet transform and Φ taken as complete Fourier sampling, uniform Fourier undersampling, and random Fourier undersampling. The coherence is lower when undersampling of the Fourier domain is random as opposed to uniform. Figure 2.8 makes clear the compatibility between MRI the incoherence requirements of CS. In MRI, there is a large amount of freedom in designing k-space trajectories, and random sampling of the k-space can be relatively easily achieved. However, pseudo-random k-space trajectories that have higher sampling probability at the center are commonplace in CS-MRI because of the role that low spatial frequencies have in determining image contrast[9].

Calculating the incoherence between the sampling basis and the transform basis can be difficult at times, such as when defining a matrix Ψ is not obvious. Calculating the coherence or the incoherence of the PSF of a k-space sampling trajectory, via equations 2.26 and 2.27 respectively, is often used as a convenient surrogate in CS-MRI[8;9;29], where it is assumed that incoherent PSFs correspond to incoherent CS pairs. An incoherent PSF (or conversely a PSF with low coherence) will induce incoherent aliasing artifacts[x] in the final image that appear similar to noise, from which

[x]The dual use of the terms coherence and incoherence with reference to the relationship between Φ and Ψ and with reference to the aliasing properties of the PSF is unfortunate, but represents standard terminology across the mathematical field of CS and the applied field of CS-MRI. The reader is reminded to keep the distinction clear in their mind.

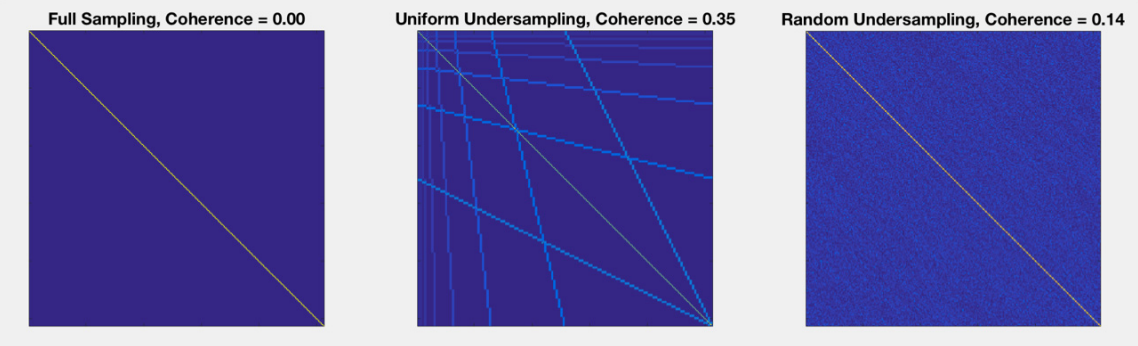

Figure 2.8: Full Fourier sampling (left) produces the lowest coherence. When the Fourier basis is uniformly undersampled (center) the coherence is higher than when it is randomly undersampled (right). This suggests random undersampling techniques should be used in CS-MRI.

CS can uncover the true image. The motivation for using PSF coherence/incoherence as a predictor of CS-MRI quality is made clear by Figure 2.4, where uniform undersampling has a very coherent PSF whereas pseudo-random undersampling has a very incoherent PSF. When the PSF gives rise to incoherent aliasing, the resulting noise-like aliases may be suppressed as CS recovers the underlying image.

2.2.5 CS Reconstruction

Recovery of a signal $\vec{v} \in \mathbb{R}^N$ from the samples $\vec{s} \in \mathbb{R}^M$ is generally not possible when $M < N$, since equation 2.35 is underdetermined. However, making the assumption that $\vec{v}$ is sparse in a transform domain (equation 2.36) provides us with extra information to make accurate recovery of $\vec{v}$ plausible. We seek an approximation $\vec{c}_s$ to $\vec{c}$ such that the acquired data is equivalent to the sparse estimation to within a noise tolerance ϵ[30]:

$$\min \|\vec{c}_s\|_0 \text{ s.t. } \|\vec{s} - A\vec{c}_s\|_2^2 < \epsilon \tag{2.41}$$

Where $\|\vec{c}_s\|_0 = \sum_i^N |c_i|^0$ counts the number of nonzero entries and is known as the ℓ_0-"norm"[xi]. CS theory would predict that $\vec{c}_s$ and $\vec{c}$ are *identical*, with arbitrarily

[xi]The quotation marks are essential, because the ℓ_0-"norm" is not a true mathematical norm. For example it is not absolutely scalable, since $\|\alpha\vec{c}\|_0 \neq |\alpha|\|\vec{c}\|_0$ for $\alpha \in \mathbb{C}$.

high probability, under noiseless sampling for appropriate choice of A. With noise contamination equivalence is guaranteed, with arbitrarily high probability, to within the specified tolerance for appropriate choice of A.

Unfortunately, minimiziation of the ℓ_0-"norm" is an NP-hard problem and is not practical in application. Fortunately, a convex relaxation of the minimization problem using the ℓ_1-norm (defined by $\|\vec{c}_*\|_1 = \sum_i^N |c_i|$) has been shown to provide equivalent results to the use of the ℓ_0 norm, provided that previously established incoherence conditions are met[29]. With this relaxation, the constrained minimization problem in 2.41 becomes:

$$\min \|\vec{c}_s\|_1 \text{ s.t. } \|\vec{s} - A\vec{c}_s\|_2^2 < \epsilon \tag{2.42}$$

To recast the recovery algorithm in 2.42 in a fashion compatible with standard minization algorithms, Lagrange multipliers are used to transform the recovery into an unconstrained problem:

$$\arg\min_{\vec{c}_s} \|\vec{s} - A\vec{c}_s\|_2^2 + \lambda\|\vec{c}_s\|_1 \tag{2.43}$$

Where the parameter λ is known as the sparse regularization weight, and balances the trade-off between the data consistency term and the sparsity enforcement term.

2.2.6 BART and CS-MRI

The recovery algorithm of equation 2.43 is readily formulated for image recovery in MRI. The Berkely Advanced Reconstruction Toolbox (BART) employs CS for MRI via the following equation:

$$\arg\min_{\rho} \left(\|\mathcal{F}[C\rho] - k\|_2^2 + \lambda\|\Psi\rho\|_1\right) \tag{2.44}$$

Where F is the Fourier transform, ρ represents the image, C represents the spatial sensitivity profiles of any coils used in PI, k are the measured k-space data, and Ψ is a sparsifying transform. CS-MRI is often viewed as a denoising problem[9], where the coefficients of the reconstructed image are regularized using the ℓ_1-norm in a sparse transform domain in an effort to suppress incoherent aliasing "noise". The regularization weight λ determines the extent to which coefficients in the reconstructed image are enforced to comply with sparsity in the transform domain specified by Ψ at the expense of data consistency.

In the sections below, we briefly cover the three sparse transforms included with BART.

Total Variation Transform

The total variation (TV) transform of an $N_x \times N_y \times N_z$ image ρ is defined as the sum of the gradients of the image along each dimension[31]:

$$TV = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \sum_{k=1}^{N_z} \left(|D_x(\rho_i)| + |D_y(\rho_j)| + |D_z(\rho_k)| \right) \tag{2.45}$$

Where D are the numerical derivative operators, applied along each dimension, and can be defined by:

$$\begin{aligned} D_x(i,j,k) &= x(i,j,k) - x(i-1,j,k) \\ D_y(i,j,k) &= x(i,j,k) - x(i,j-1,k) \\ D_z(i,j,k) &= x(i,j,k) - x(i,j,k-1) \end{aligned} \tag{2.46}$$

The assumption made by enforcing sparsity with the TV is that the image is piecewise constant, which is a good approximation of some anatomical imaging applications in MRI. Under the assumption of a piecewise constant image, the best CS reconstruction will be the one with the lowest derivative values across the entire image; this is what is quantified by the TV.

Wavelet Transform

The wavelet transform of an image decomposes an image in the *wavelet basis*, the results of which contain information about both the position and spatial frequency content of an image. For the purposes of instruction, we consider the one dimensional wavelet transform of a signal $\rho(x)$. The principles presented below extend to higher dimensional applications.

Similar to the Fourier transform, which correlates the signal of interest with sinusoids, the continuous wavelet transform correlates the signal with a set of wavelets[32]:

$$W(a,b) = \int_{-\infty}^{\infty} \rho(x) \psi_{a,b}^{*} dx \tag{2.47}$$

Where $\psi_{a,b}$ are child wavelets generated by translating and scaling the mother wavelet ψ:

$$\psi_{a,b} = \frac{1}{\sqrt{a}} \psi \left(\frac{x-b}{a} \right) \tag{2.48}$$

Where $a \in \mathbb{R} > 0$ and $b \in \mathbb{R}$. The choice of ψ is restricted to functions which have net-zero area[32]:

$$\int_{-\infty}^{\infty} \psi(x)dx = 0 \tag{2.49}$$

For example, the Haar wavelet is composed of two rectangular lobes of equal height but opposite magnitude.

Equation 2.47 demonstrates that a wavelet transform involves scaling the mother wavelet to a desired scale by fixing a, and then translating the child wavelet across the signal by varying b. At each stage of the translation projections of $\rho(x)$ onto $\psi_{a,b}$ are taken, quantifying the "amount" of $\psi_{a,b}$ in the signal at scale a and location b. The wavelet transform is thus a multi-parametric transformation, with coefficients W indexed by the scale and location at which they were obtained.

In practical numerical applications of the wavelet transform, discretization of a and b is required. In general, selection of $a = a_0^m$ for $m \in \mathbb{Z}$ and $b = nb_0a_0^m$ for $n \in \mathbb{Z}$, with $a_0 > 1$ and $b_0 > 0$, discretizes the production of children wavelets in a way that provides equivalent coverage across scales[32]. This yields the discrete wavelet transform for a signal of size N:

$$W(m,n) = \sum_{i=1}^{N} \rho(x_i)\psi_{m,n}^* \tag{2.50}$$

Where the production of children wavelets is now obtained by the discretized parameters m and n:

$$\begin{aligned} \psi_{m,n}(x) &= \frac{1}{\sqrt{a_0^m}}\psi\left(\frac{x - nb_0a_0^m}{a_0^m}\right) \\ &= \frac{1}{\sqrt{a_0^m}}\psi\left(\frac{x}{a_0^m} - nb_0\right) \end{aligned} \tag{2.51}$$

In virtually all applications, the choices $a_0 = 2$ and $b_0 = 1$ are made.

The assumption made by enforcing sparsity in a wavelet transform is that the majority of the coefficients in W at each scale are zero. This is a very common property of natural images[xii], and is often true for medical images as well.

[xii] The JPEG 2000 lossy compression scheme takes advantage of this fact to compress fully sampled image data.

Locally Low Rank Sparsity

Suppose that we have an $N_x \times N_y \times N_z$ image ρ, and that we partition it into L smaller blocks of size $n_x \times n_y \times n_z$. These 3D blocks can then be reshaped into 2D matrices by concatenating one dimension along either the rows or the columns. The locally low rank (LLR) sparse enforcement attempts to minimize the *rank* of the local block matrices b_l (i.e. the dimension of the vector space spanned by the columns of b_l). Direct minimization of the rank of a matrix is an NP-hard problem however. It has been shown that minimizing the sum of the singular values of a matrix provides a convex relaxation to the rank minimization problem[33]. We define the nuclear norm $\|b_l\|_*$ of a block b_l as the norm returning the sum of the singular values of b_l. An LLR sparsity enforcement then amounts to minimizing the sum of the nuclear norms of each block b_l, and equation 2.44 takes the form of equation 2.52.

$$\arg\min_{\rho} \left(\|\mathcal{F}[C\rho] - k\|_2^2 + \lambda \sum_{l=1}^{L} \|b_l\|_* \right) \tag{2.52}$$

The ℓ_1-norm of equation 2.44 has been replaced by the nuclear norm because sparsity is enforced in the singular value decomposition of ρ, not in an explicit transform basis.

The assumption made when enforcing sparsity with the LLR method is that nearby patches of an image tend to be more similar to one another than to the global image. This is certainly the case in many imaging applications, including medical images.

2.3 CIRcular Cartesian Undersampling (CIRCUS)

In dynamic MRI applications, temporal resolution and image quality are determined by the k-space content chosen for each reconstructed image. Golden ratio sampling, first proposed by Winkelmann et. al.[13] in the context of radial sampling, allows for retrospective selection of the temporal resolution and image quality while maintaining relatively uniform coverage of k-space. In this section we introduce golden ratio sampling in its original radial k-space sampling context, and then describe a recent k-space sampling scheme that incorporates golden ratio sampling properties for Cartesian sampling with CS-MRI.

2.3.1 Golden Ratio Sampling of K-Space

Traditional radial sampling involves the selection of a fixed number N_d of diameters that are used to sample k-space per image reconstruction, such that two temporally adjacent diameters are at an angle of $\theta = {}^{180^\circ}/_{N_d}$ relative to each other. By specifying that each reconstructed image contain N_d diameters, the temporal and spatial resolutions of the results are fixed a priori to maintain uniform k-space coverage and cannot be changed following the scan should the results be found to be suboptimal.

To permit the retrospective selection of temporal resolution and image quality while obtaining very uniform coverage of k-space, it has been proposed that the succesive diameters should be incremented by the following angle[13]:

$$\theta_{GR} = \frac{180^\circ}{\varphi} \approx 111.25^\circ \tag{2.53}$$

Where φ is the golden ratio[xiii] defined by:

$$\phi = \frac{(\sqrt{5}+1)}{2} \tag{2.54}$$

It has been shown that incrementation by θ_{GR} provides a very uniform coverage of k-space for an *arbitrary* number of combined diameters compared to standard radial sampling[13]. This is demonstrated in Figure 1.4. A further advantage is that θ_{GR} is an irrational number, so in theory the sampled diameters will never overlap as more of them are interleaved. This permits interleaving of very few diameters if high temporal resolution is desired, and of many diameters if high image resolution is desired.

2.3.2 CIRCUS Base Pattern

Liu et. al. recently proposed a 3D Cartesian sampling scheme that incorporates properties of golden ratio sampling, while incorporating influences from radial and spiral sampling to increase the incoherence of the aliasing artifacts in an effort to boost CS-MRI performance[14]. They named their sampling scheme CIRcular Cartesian UnderSampling (CIRCUS). We outline the construction of CIRCUS patterns below.

We begin with the assumption that phase encoding dimensions are equal; rectangular phase encoding planes will be discussed in section 2.3.4. The CIRCUS base

[xiii] The golden ratio is known by a variety of other names, such as the golden section and the divine proportion. The *golden angle* is commonly confused for the *golden ratio*, but in fact the golden angle is constructed from the golden ratio and the two are not the same thing.

pattern is generated by combining a number of *quanta*. Quanta are generated by the selection of sampling indices on the periphery of a series of nested squares extending outwards from the center of k-space in the direction of the two phase encoding dimensions. The side length of each nested square is $J \in 2\mathbb{Z}$, such that for a square of side length J there are $K = 4(J-1)$ possible sampling indices $i(q, J)$ for the generation of quanta q . The indices $i(q, J)$ are given by equation 2.55.

$$i(q, J) = \left\lfloor \mathrm{mod}\left(\frac{q}{\varphi}, 1\right) \cdot K \right\rfloor \tag{2.55}$$

Where φ is the golden ratio defined in equation 2.54, and mod is the modulo function, defined by:

$$\mathrm{mod}\ (a, n) = a - n \cdot \left\lfloor \frac{a}{n} \right\rfloor \tag{2.56}$$

The index $i(q, J)$ can take a value from 0 to $K - 1$, where the first index $i = 0$ is chosen as the bottom left corner of each nested square for consistency.

2.3.3 CIRCUS Radial and Spiral

One technique the authors implement to increase the incoherence of CIRCUS pattern PSFs is to add a linear shift to the sampling indices generated for each quanta in the CIRCUS base pattern. This shift is controlled by a parameter $b \in \mathbb{Z}$, and is implemented by altering equation 2.55 in the following manner:

$$i'(b, q, J) = \left\lfloor \mathrm{mod}\left(\frac{q + bJ}{\phi}, 1\right) \cdot K \right\rfloor \tag{2.57}$$

The authors name the patterns resulting from the production of sampling indices via equation 2.57 "CIRCUS radial" patterns. The effect of the shift introduced in equation 2.57 is to "shear" the straight quanta produced by equation 2.55.

A second technique implemented by the authors in an effort to increase the incoherence of CIRCUS pattern PSFs is to add a nonlinear shift to the sampling indices produced by the CIRCUS base. This nonlinear shift, resulting in the "CIRCUS spiral" patterns, is controlled by a parameter $c \in 0 \cup (1, 2)$ and obtained by adding a shift to equation 2.55 in the following way:

$$i''(c, q, J) = \quad \mathrm{mod}\ \left(i(q, J) + \lceil J^c \rceil - 1, K\right) \tag{2.58}$$

The effect of the spiral shift is to "twist" the base quanta.

For maximally incoherent sampling patterns, the radial and spiral shifts may be combined. The combination of equations 2.55, 2.57, and 2.58 lead to the general form of the equation for CIRCUS sampling indices on each nested square:

$$i(b, c, q, J) = \mod\left(\left\lfloor \mod\left(\frac{q + bJ}{\phi}, 1\right) \cdot K \right\rfloor + \lceil J^c \rceil - 1, K\right) \tag{2.59}$$

Representative CIRCUS patterns for various b and c parameters generated with an undersampling factor R of 10 are shown in Figure 2.9.

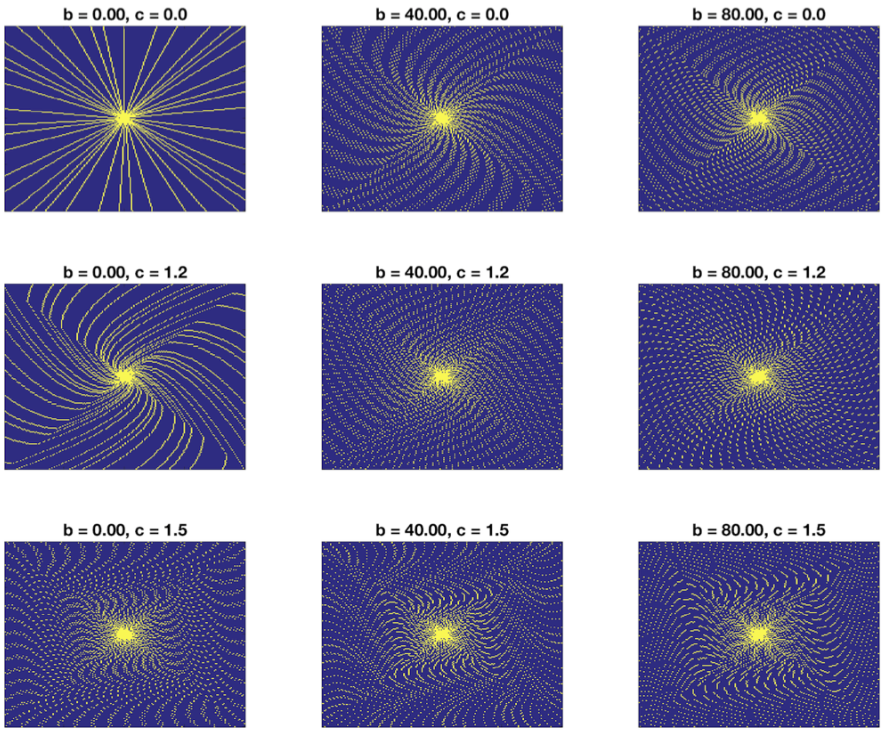

Figure 2.9: Examples of CIRCUS trajectories obtained with various b and c combinations with an undersampling factor of $R = 10$. The CIRCUS base pattern is shown at the top left.

2.3.4 CIRCUS with Rectangular Phase Encoding Plane

In practise, the number of phase encode locations along one dimension is often significantly smaller than in the other dimension. For example, in 3D pelvic MRI, there

may be on the order of 256 phase encode locations in the y-direction and on the order of 32 in the z-direction. Generation of rectangular CIRCUS patterns is thus necessary.

Let N_y and N_z represent the number of phase encode points along the y- and z-direction respectively, where $N_z < N_y$. Generation of an $N_y \times N_z$ CIRCUS pattern begins with the construction of an $N_y \times N_y$ CIRCUS pattern, as described in sections 2.3.2 and 2.3.3. N_z lines are then selected along the short phase encode dimension according to the golden ratio profile described by equation 2.55. The selected lines are assembled, maintaining their relative positions, to obtain a rectangular CIRCUS pattern measuring $N_y \times N_z$.

2.4 Image Quality Assessment

The gold standard of image quality assessment is the subjective qualification of image quality via the perception of a human observer. However, this process is time consuming and expensive. This is especially true of diagnostic radiology, where radiologists are faced with tight schedules and cannot afford to regularly assess large quantities of MRI images to determine optimum reconstruction quality. The field of image processing has seen extensive work in the development of image quality metrics (IQMs) which aim to objectively quantify image quality by considering properties of the human visual system (HVS) to return scores that are indicative of observer subjective scoring. In this section we summarize five objective IQMs: the root mean squared error (RMSE), the gradient magnitude similarity deviation (GMSD)[34], the structural similarity index (SSIM)[35], the multi-scale structural similarity index (MS-SSIM)[36], and the information weighted structural similarity index (IW-SSIM)[37]. The reader is referred to the original publication for each IQM for specific calculation details. The latter four IQMs have been verified to correlate better with subjective human scoring, compared to the RMSE, when viewing natural images with a variety of distortions. In the following descriptions we assume that an $N_x \times N_y$ reference image R is being compared to a distorted representation D of equivalent size.

2.4.1 Objective IQM Design

There are three broad classes of IQMs[38]. Full-reference IQMs compare a distorted image to a readily available reference image. Reduced-IQMs make use of partial information obtained from the reference image. No-reference IQMs make no comparison a reference image, instead assuming statistical properties that would be contained in a theoretical reference image or properties of a specific family of distortions that may be present in the image. In this thesis we investigate full-reference IQMs. The acquisition of a high quality baseline image prior to clinical imaging protocols is feasible in MRI, permitting the use of full-reference IQMs in practise.

The design of an IQM may be classified as either bottom-up or top-down[38]. The bottom-up approach uses the results of psychophysical studies in an attempt to simulate the functionality of component stages of the HVS. These might include estimates of an optical point spread function (PSF), estimates of a contrast-sensitivity function (CSF), estimates of the relative importance of spatial scales and orientations (subband decompositions), and so on. These stages are then strung together as a series of building blocks in an attempt to create an IQM that is directly reflective of many component stages of the HVS. The top-down approach makes general, higher level hypotheses regarding the functionality of the HVS, and devises an IQM based on the hypothesized HVS behaviour. The HVS is treated is a black box in such an approach, with the guiding hypothesis often taken that the HVS is especially sensitive to structural changes in an observed scene[xiv]. The difference between bottom-up and top-down IQMs are illustrated in Figure 2.10. While one method of IQM design is not "the right way", the introduction of the SSIM by Wang et. al. spurred the top-down movement that many IQM designs use today. In addition to being computationally simpler to implement, the design of top-down metrics is not heavily reliant on the results of psychovisual studies, many of which are performed under limited artificial conditions that don't accurately capture the full complexity of the HVS[35].

The black box treatment of the HVS in the top-down framework typically results in a quality map (the HVS step in the top-down IQM flowchart in Figure 2.10 generates such a map), which is pooled in some fashion to give a final score that is reflective of

[xiv]This hypothesis originates from the observation that natural scenes comprise a small and very structured subset of all possible images. Since the HVS has evolved over millions of years in the presence of these structured scenes, it should be sensitive to changes in their structure[38].

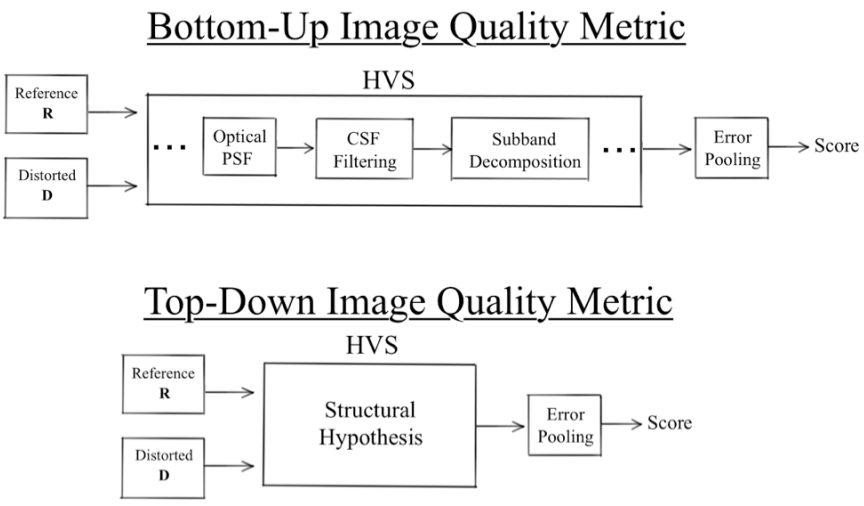

Figure 2.10: Bottom-up approaches attempt to simulate the functionality of stages of the HVS, building a computational model of the HVS. Top-down approaches assume a high level functionality of the HVS in terms of structural perception, and treat the details of the HVS as a black box.

the quality of the distorted image. A common approach is to take the mean of the image quality map, though top-down metrics utilizing different pooling techniques have been proposed.

2.4.2 Root Mean Squared Error (RMSE)

The RMSE quantifies the pixel-wise errors between R and D:

$$RMSE = \sqrt{\frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} [R(x,y) - D(x,y)]^2} \tag{2.60}$$

The RMSE is independent of the signs of the error in each pixel due to the squared error term in equation 2.60. Furthermore, the RMSE makes no consideration of the local image properties during calculation of the local error map.

Under the framework of a top-down image quality metric, pictured in Figure 2.10, the RMSE can be thought of as assuming that the HVS is sensitive to absolute errors at each pixel of the image, such that the black box treatment of the HVS results in an error map. Pooling of the error map then involves squaring the errors, taking

their mean, and reporting the square root of this mean to arrive at the final RMSE score. The HVS is not highly sensitive to errors on a pixel-wise basis, and it has been found that the RMSE does not correlate particularly well with human perception of distorted images[15].

For an image measuring 256×256, the RMSE takes on the order of 0.0005 seconds to calculate.

2.4.3 Structural Similarity Index (SSIM)

The overall hypothesis of the SSIM is that the HVS is primarily sensitive to relative changes in luminance, contrast, and structure in an image. The SSIM compares R and D using a sliding window technique to compare local image properties (i.e. windows are passed over both R and D, with calculations made at equivalent pixels in each image). The authors chose to use an 11×11 circularly symmetric Gaussian window (with standard deviation of 1.5) to avoid "blocking" artifacts introduced at the boundary of sliding windows with equivalently weighted elements[35]. The weights of the window are normalized to unit sum. Within each local window, calculations of the luminance (quantified by the mean), contrast (quantified by the standard deviation), and structure (quantified by the correlation) of each image are made. The calculations of each of these three qualities between the individual images are then combined to make three terms that reflect the perceivable change in luminance, contrast, and structure between R and D in each of the local regions of the images. A weighted product of these three terms at each pixel in the images gives an SSIM quality map, where the weighting is typically taken as equivalent for each term. The final SSIM score is obtained by taking the mean of the SSIM quality map. The range of the SSIM is from -1 to 1, where negative scores are hardly ever achieved in practise and a score of 1 is optimal.

The SSIM is a single scale IQM (the concept of image scale is introduced in section 2.4.4), meaning that it is intended to quantify the effects of distortions to image features at a specific scale. Images must be brought to that scale in order for the SSIM to optimally correlate with subjective scoring. The authors propose that the images be downsampled, if necessary, such that the smallest dimension measure 256

pixels across[xv]. Downsampling by a factor of d_s involves convolving the images with a $d_s \times d_s$ averaging kernel and selecting every d_s^{th} pixel in each dimension. For example, images measuring 512×768 would be downsampled by a factor of 2 using a 2×2 averaging kernel and retaining every second pixel in the images to form downsampled copies measuring 256×384.

For an image measuring 256×256, the SSIM takes on the order of 0.003 seconds to calculate.

2.4.4 Multi-Scale Structural Similiarity Index (MS-SSIM)

The scale for which the SSIM should be applied is dependent on image size and image resolution, and in practise the optimal scale will not accurately be reached. The use of multiple image scales makes the assessment of image quality robust to such factors, and further improves the objective image quality assessment by considering the sensitivity of the HVS to distortions at different scales[36].

The concept of *scale* in image processing may be confused with the use of the word scale in the colloquial sense. To be clear, the image scale in image processing refers not to the size of the image, but to the relative size of features made available for observation in the image by sampling the image at different resolutions. To understand this, consider the fine scale image shown in the left of Figure 2.11. We recognize that there are features of several different scales in the image; the cameraman is a very large scale feature in the image, whereas the handle of his camera and the blades of grass are small scale features. The HVS easily recognizes and distinguishes between these features, assessing each scale with a different sensitivity. To include the concept of feature scale into an image processing application, the processing algorithm must be made sensitive to features at each scale. One method of achieving this is to adjust the resolution of the image. By resampling an image at decreasing resolutions, the large scale features of an image will be emphasized. This is demonstrated in the right of Figure 2.11, where a 10 times downsampling of the cameraman image emphasizes coarse scaled features.

[xv] In the original paper, the authors recommended downsampling by a fixed factor of 2[35]. This recommendation was later updated on the website of the author (`https://ece.uwaterloo.ca/~z70wang/research/ssim/`), and in the code written and distributed by the author, to the recommendation presented here.

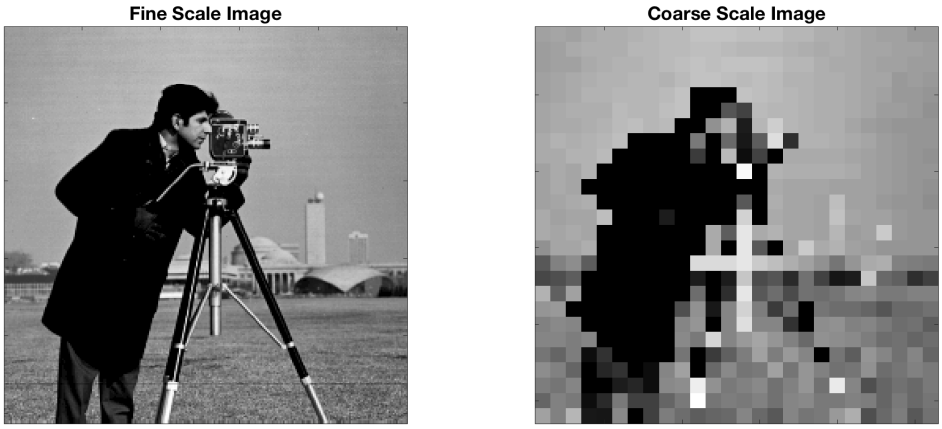

Figure 2.11: Fine scale features are easily resolved at the highest resolution expression of the image (left), for example the handle of the camera stand. Decimating the resolution of the image gives a coarse scale representation (right). In the coarse scale image, large features remain evident while small features are lost; we can discern the cameraman, but few details of his equipment.

To calculate changes in luminance, contrast, and structure at progressively coarser scales for a maximum of five scales, the authors implement a 2x downsampling procedure. At each of the scales, a contrast comparison and a structural comparison are calculated via the same approach as for the SSIM to generate a "constrast-structure" map; luminance comparison is done only for the coarsest scale[36]. At each scale the quality maps are pooled by taking the mean of to obtain an intermediate quality score for that scale. Each intermediate score is weighted according to the sensitivity of the HVS to distortions at that scale, and the final MS-SSIM score is obtained by taking the product of all weighted intermediate scale scores.

For an image measuring 256×256, the MS-SSIM takes on the order of 0.008 seconds to calculate.

2.4.5 Information-Weighted Structural Similarity Index (IW-SSIM)

Sophisticated methods of pooling that attempt to weight the contributions of salient image regions have been found to improve the correlation between subjective observer scores and objective IQM scores[37]. The IW-SSIM builds upon the MS-SSIM by weighting the local image quality maps at each scale according to the perceived statistical information in each of R and D; the calculation of the IW-SSIM is otherwise the same as that of the MS-SSIM. The hypothesis underlying the use of weighting local quality maps with statistical information is that the HVS is an optimal extractor of information content; this hypothesis is widely held across the field of computational vision[37].

To develop a statistical information weighting technique, a statistical model describing images must be developed. Extensive studies of natural scene statistics (NSS)[xvi] have shown that natural scenes are well modelled in the wavelet domain by Gaussian scale mixture (GSM) models of wavelet coefficients[39]. A GSM may intuitively be thought of as a mixture of many zero-mean Gaussians with differing variance; the wavelet coefficients of natural scenes display a sharp peak near zero and rapid decay for coefficients of increasing magnitude.

To simplify the calculation process, the authors proposed to use a GSM to model the coefficients of a Laplacian pyramid[40] decomposition[xvii]. A Laplacian pyramid is formed by subtracting blurred copies of an image from itself at each scale and retaining the difference image; the exception is at the coarsest scale of the pyramid, where the blurred image is retained. The formation of a Laplacian pyramid is demonstrated in Figure 2.12. Because the coarsest scale in a Laplacian pyramid is simply a blurred image, information weighting is not performed at that scale[37].

In the process of statistically modelling the coefficients at each level of the Laplacian pyramid, it is assumed that a coefficient is only related to coefficients in a nearby 3×3 neighbourhood. This neighbourhood may additionally include a single coefficient

[xvi] A *natural scene* is one which can be observed by the naked human eye and that can be encountered in nature. Medical images are not a subset of the natural scenes.

[xvii] While wavelets are an *oriented* transform, the Laplacian pyramid is *non-oriented*. We contacted Zhou Wang via email in May of 2017 for justification. He states that he believed the GSM model holds equally well for non-oriented transforms, given the highly empirical nature of the study of NSS.

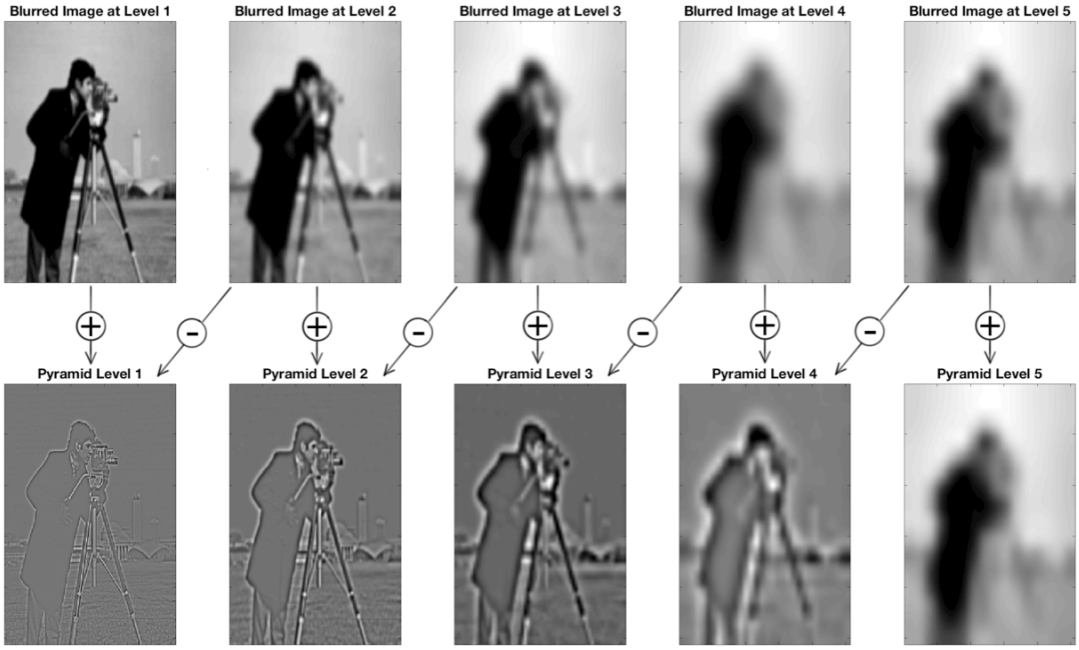

Figure 2.12: Each scale of the Laplacian pyramid is formed by subtracting a blurred image from the base image, with the exception of the coarsest scale. The scale of the image increases for higher levels of the pyramid, incorporating a multi-scale element into the IW-SSIM.

from the next coarsest scale in the pyramid[xviii]. The end result is either a $K = 3 \times 3$ or a $K = 3 \times 3 + 1$ neighbourhood, with the later choice being the default used by the authors. Borrowing from the results of NSS, the coefficients of the Laplacian pyramid in the neighbourhood for the reference image are then modelled as a GSM, yielding a collection of transform coefficients R_K. The authors then model the coefficients of the Laplacian pyramid in the neighbourhood for the distorted image by assuming that a distortion channel generated D_K from R_K through scaling of the GSM coefficients via a gain factor and with the addition of additive noise. Finally the perception of each set of coefficients by the HVS is modelled via a perception channel which adds independent white noise that is reflective of overall internal neural noise[41], yielding

[xviii] Wang et. al. refer to this as a "parent" coefficient, though the terminology may be confusing given that the "parent" scale in the pyramid is in fact generated from the "child" scale under their terminology... nevertheless, the reader should be aware of this when consulting the literature.

perceived models P_{R_K} of the coefficients of the reference image and P_{D_K} of the coefficients of the distorted image. The information w_K extracted by the HVS from each of the reference and distorted image coefficients in the neighbourhood is then quantified via the sum of mutual informations I between each set of image coefficients and their output from the perception channels, less the mutual information between the perception channel outputs of each set of coefficients. The final information weighting of the coefficient at which the neighbourhood was centered is given by:

$$w_K = I(R_K; P_{R_K}) + I(D_K; P_{D_K}) - I(P_{R_K}; P_{D_K}) \tag{2.61}$$

Figure 2.13 schematically relates all the contributions towards the calculation of w_K.

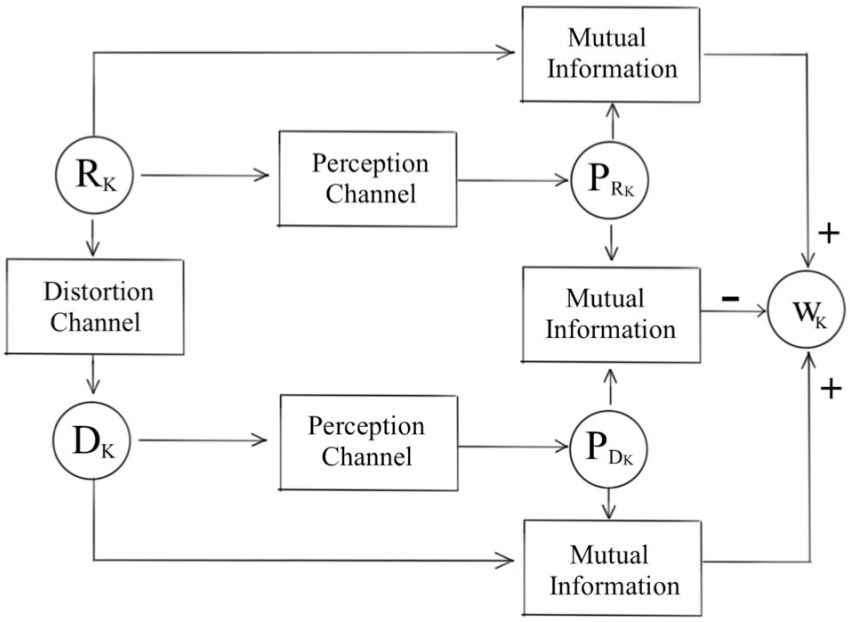

Figure 2.13: The final weight allocated for the Laplacian pyramid coefficients during the pooling stage is given by the sum of mutual informations between the original and perceived coefficients for both reference and distorted images, less the mutual information between the perception model outputs.

For an image measuring 256×256, the IW-SSIM takes on the order of 0.08 seconds to calculate.

2.4.6 Gradient Magnitude Similarity Deviation (GMSD)

The GMSD compares the gradients of R and D, since it has been found that common distortions affecting image structure (e.g. noise, compression artifacts) are very prominent in the gradient domain[34]. The primary goal of the GMSD was computational expediency, and this is achieved by pixel-wise comparison of the gradient magnitudes of R and D as opposed to several local window calculations as in the SSIM family of IQMs. For an image measuring 256×256, the GMSD takes on the order of 0.001 seconds to calculate, marginally slower than the RMSE and faster than any of the SSIM family of IQMs.

The GMSD is a single scale IQM, and the authors downsample the images by a factor of 2 *regardless* of the initial image sizes[34][xix]. Gradients of R and D are then calculated using 3×3 Prewitt filters (though the authors acknowledge that specific choice of gradient filter will not significantly affect the performance of the GMSD). A gradient magnitude similarity (GMS) map is made through a pixel-wise comparison of the gradient maps of R and D; the assumption is that sufficient local structural information in the images has been considered during the generation of each gradient map. The different structures in an image will suffer different degradation to their gradient magnitude as the image is distorted, and the GMSD capitalizes on changes in the gradient magnitude from R to D to quantify the presence of visible distortions. Pooling the GMS into a GMSD score involves taking the standard deviation of the entire GMS map, such that low GMSD score means better objective image quality. The justification for this pooling strategy is the assumption that global variation in the local quality map is reflective of overall image quality[34].

[xix] The authors of the GMSD proposed this downsampling based on a similar recommendation made in the original SSIM paper. The authors of the SSIM later changed their downsampling recommendation, but the authors of the GMSD appeared to have missed the update.

Chapter 3

Common Methodology

3.1 Acquisition of Pelvic MRI Data

This study was research ethics board approved. Informed consent was obtained from 15 clinical prostate cancer patients for use of their image data. Imaging was performed on a 3T GE MR750 MRI system (GE Healthcare, Waukesha, WI) using a free breathing DIfferential subsampling with cartesian ordering (DISCO)[42] 3D spoiled gradient-echo acquisition to acquire axial images of the pelvis. Acquisition parameters were:

- TR/TE: 3.7 ms/1.1 ms
- Flip angle: 12°
- Field of view: 34×34 cm^2
- Axial spacing: 3 mm
- K-space matrix size: 224×192
- PI acceleration: 2×1

Fat-water seperation was performed via a Dixon method[43], and image reconstructions performed using the GE Orchestra SDK image reconstruction suite.

3.2 Berkeley Advanced Reconstruction Toolbox (BART)

The Berkeley Advanced Reconstruction Toolbox (BART)[44] is a current CS reconstruction toolbox that has become ubiquitous in the field of CS-MRI. BART possesses a wide variety of image reconstruction algorithms, including CS with wavelet, total variation, and locally low rank sparse regularizations (or any combination thereof). BART offers support for combining parallel imaging with CS, further increasing the

ability to accelerate data acquisition. It is an open source toolbox[i] built in the C language for UNIX operating systems, with MATLAB interfacing capabilities for convenience. We utilized version 3.00 of BART to reconstruct the simulated CS-MRI images.

3.3 Software Implementation

All computational aspects of this thesis were implemented in MATLAB 2016b (MathWorks, Natick MA). CS-MRI reconstructions with optional PI utilized BART (see section 3.2), called via a supported MATLAB interface.

3.4 Simulated Coil Sensitivities

We utilized artificial coil sensitivity profiles when implementing parallel imaging (PI) in our calculations. The sensitivity profiles were chosen such that coils were most sensitive to nearby regions, with a decline in spatial sensitivity as distance from the coil increased. The coils themselves were modelled as "point coils", having no spatial extent. While these assumptions are not entirely realistic, they allow for the inclusion of spatial encoding in our calculations and the results will be broadly representative of realistic PI acquisitions that leverage spatial encoding. Future work may wish to accurately measure the coil sensitivity profiles for a specific PI set-up, allowing for specific investigation of that system.

The coil sensitivity profile generation algorithm supported the use of either one uniformly sensitive coil, or the placement of any even number of individual coil with sensitivity profiles C_l given by:

$$C_l(x, y, z) = \cos\left(\frac{2\pi}{P_x}(x - \Delta x_l)\right)\cos\left(\frac{2\pi}{P_y}(y - \Delta y_l)\right)\cos\left(\frac{2\pi}{P_z}(z - \Delta z_l)\right) \qquad (3.1)$$

Where l denotes the coil number, P_r is the period of the sinusoid in the r^{th} dimension and is chosen such that the first "zero-crossing" occurs at the boundary of the image (i.e. $P_r = 4 * L_r$ for L_r the voxel-extent of the image object in the r^{th} dimension), $r \in [x, y, z]$ represent the row, column and slice coordinates and Δr_l are the coordinates for the placement of coil l. Following the generation of coil sensitivity profiles via

[i]Available from `https://mrirecon.github.io/bart/`

equation 3.1, the coil sensitivity profiles are normalized such that the sum of squares of sensitivities within a single voxel across all sensitivity profiles is 1 (as per BART requirements).

Figure 3.1 demonstrates examples of coil placements for 8 and 10 total coils. Coil placements Δr_l were made with symmetric placement of coils across from one another in the y-direction (e.g. placement of 8 coils in total results in 4 coils in the anterior plane and 4 coils in the posterior plane of the pelvic image, with symmetric placement of the 4 coils between planes). Δx_l and Δz_l were then chosen dependent on whether the number of coils placed in each of the anterior and posterior planes was even or odd. Even numbers of coils per plane positioned coils such that two lines of coils divided each plane into thirds; odd numbers of coils per plane placed one coil centrally in each plane and the remaining coils along the same lines used for even placement of coils.

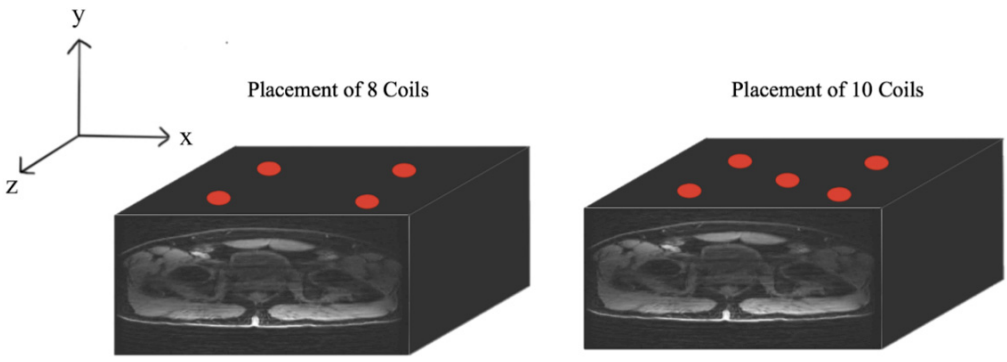

Figure 3.1: Coils (represented by red ellipses) were placed symmetrically in the anterior and posterior planes. Even numbers of coils per plane (e.g. 8 coils total, 4 per plane, left) placed coils along two lines dividing the slice direction into thirds. Odd numbers of coils per plane (e.g. 10 coils total, 5 per plane, right) placed coils as in the even case, with the extra coil added to the central region of each plane.

3.5 3D Extensions of Objective IQMs

3.5.1 RMSE

The RMSE was calculated by extending equation 2.60 to 3D:

$$RMSE = \sqrt{\frac{1}{N_x N_y N_z} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} \sum_{z=1}^{N_z} [R(x,y,z) - D(x,y,z)]^2} \tag{3.2}$$

For an image measuring $256 \times 256 \times 30$, the RMSE takes on the order of 0.006 seconds to calculate.

3.5.2 SSIM

As an addition to the image processing toolbox in MATLAB version 2014a, MATLAB includes a 3D implementation of the SSIM[ii]. We implemented the SSIM using the function included with MATLAB.

The SSIM function included with MATLAB differs from the SSIM code written by the authors of the SSIM in two ways. First, while the authors handle edge cases during local window calculations by entirely avoiding the edge pixels of an image, the included MATLAB implementation pads the image array with replicates of the border voxels. We preferred the latter technique because it included all voxels in the 3D images. Ultimately, the difference in handling of edge-cases was not significant. This is demonstrated in Figure 3.2, where the handling of border pixels is not a significant factor in the calculation of the 2D SSIM and is not expected to be significant in the 3D SSIM either.

The second difference between the included MATLAB function and the SSIM code originally written by the authors is that the MATLAB function does not downsample the image such that the smallest dimension measures 256 voxels across. While the images utilized in our study did not exceed 256 voxels in any dimension, this difference in SSIM implementations may be relevant to future work so we briefly address it here. The purpose of such a downsampling step is to bring the images into a scale similar to which the SSIM was originally correlated against non-expert subjective scoring of natural images. The logic of this downsampling step does not necessarily translate to

[ii]Documentation for the function ssim.m included with MATLAB can be found at `https://www.mathworks.com/help/images/ref/ssim.html`.

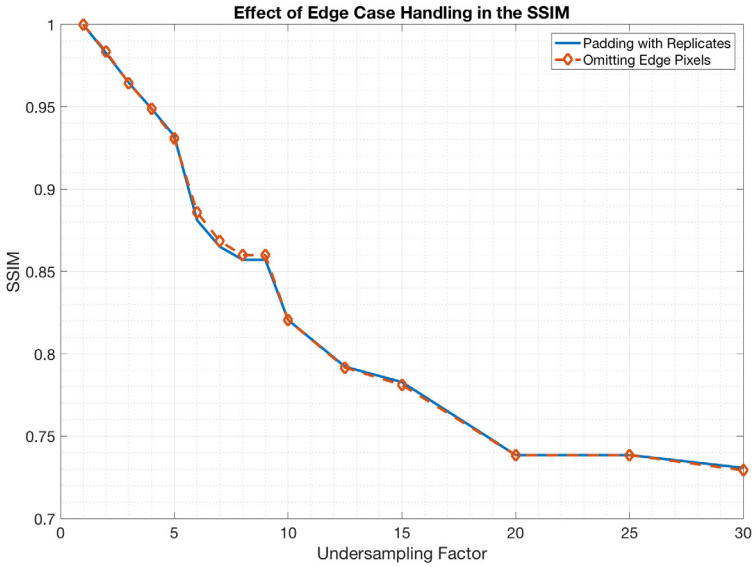

Figure 3.2: The method by which border pixels were handled during sliding window calculations made an insignificant impact on the SSIM.

3D medical images where the slice direction dimension may be significantly smaller than the in-plane direction. Furthermore advanced generations of the SSIM have attempted to include the response of the HVS to distortions at different scales, so we believe it most appropriate to evaluate the SSIM on MRI images at their original scale.

In selecting the dimensions of the circularly symmetric Gaussian window for calculation of the local luminance, contrast, and structure changes, we chose in-plane dimensions of 11×11 voxels with standard deviation of 1.5[iii] as suggested by the authors of the SSIM. We found that the extent of the window along the slice direction did not make a significant difference in the performance of the SSIM when choosing a slice extent between 1 and 11 slices (see Figure 3.3). The insignificant change in SSIM response to the addition of extra slices is partially a result of the fact that the

[iii] A standard deviation of 1.5 means that over 99% of the weights are contained within 11 voxels. Changing window size requires changing the standard deviation accordingly.

SSIM utilizes a Gaussian window to calculate local luminance, contrast, and structure terms. Voxels in additional slices are weighted less than voxels that are in the plane of the center of the window, and thus emphasis is placed primarily on the "in-plane" voxels during local window calculation. This is particularly true of relatively small window sizes (i.e. between 1 and 11 voxels), where the voxel weights relative to one another will decline quickly compared to larger window sizes. In addition to this nearby slices in a 3D MRI image contain very similar structural features, and the incoherent aliasing artifacts appear relatively consistent in behaviour across the image when inspected visually. Thus, nearby slices would be expected to have similar structural degradation quality. The inclusion of additional nearby slices in the SSIM calculation would not be expected to significantly change the result since they will all be of similar structural quality. This effect will be dependent upon how quickly the structure of the images changes, though for the pelvic images we studied the relative change in overall visual structure between slices tends to be slow (e.g. if a chosen slice contains views of the femoral head, prostate, and so on, the adjacent slices will contain broadly similar views of the same anatomy). This, combined with the fact that slices further from the center of the Gaussian window are weighted less, explains why the SSIM was not observed to differ greatly when choosing between 1 and 11 slices for inclusion in the Gaussian window.

Given these results, we chose to implement the SSIM with a sliding window that had equal voxel-extent in each dimension as opposed to scaling the window to have equal spatial-extent in each dimension. This decision also facilitates implementation of the SSIM, since the spatial extent of the image in each direction need not be considered at the time of calculation. Thus, we implemented the SSIM with an $11 \times 11 \times 11$ voxel circularly symmetric Gaussian window.

The relative weight of luminance, contrast, and structure related distortions were assigned to be equivalent based on the recommendations of the SSIM authors[35].

Implementing the SSIM as described above, it takes on the order of 16.6 seconds to calculate a score for a $256 \times 256 \times 30$ image.

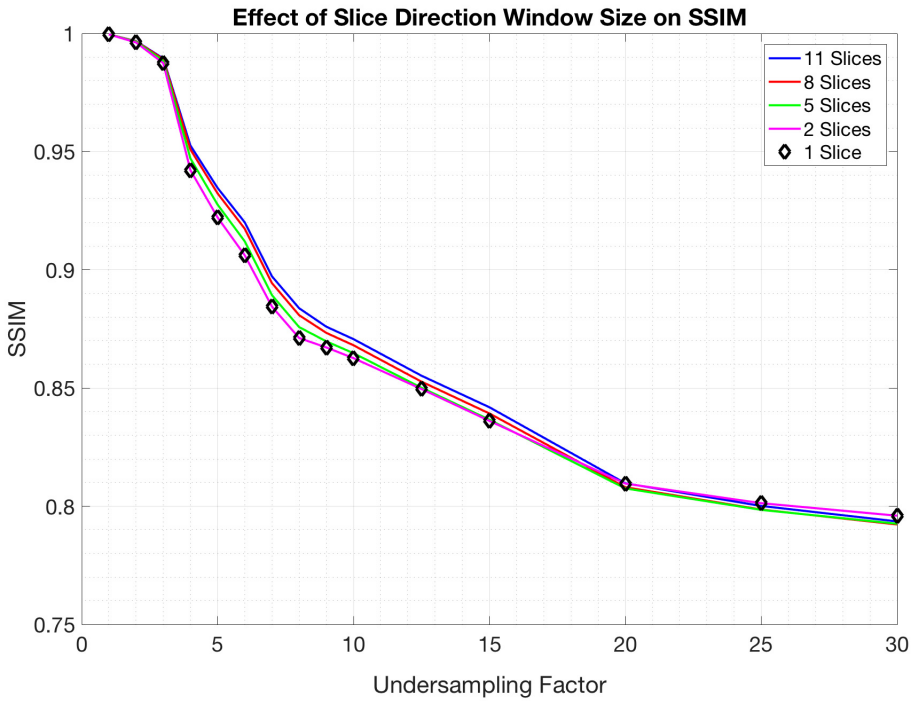

Figure 3.3: Changing the extent of the sliding window along the slice direction was found to give insignificant changes to the performance of the SSIM.

3.5.3 MS-SSIM

We extended the MS-SSIM to 3D by extending the 2×2 voxel averaging window used in the image re-scaling process for the original 2D implementation of the MS-SSIM to a $2 \times 2 \times 2$ voxel averaging window. Calculation of the MS-SSIM score for the 3D extension was then similar to the 2D MS-SSIM as originally put forward by its authors[36]. In line with our implementation of the SSIM (section 3.5.2), we chose to handle edge cases during luminance, contrast, and structure term calculations by padding with replicates of border voxels, as opposed to the technique of ignoring border voxels employed by the authors of the MS-SSIM.

Because of the inherent need for downsampling in the process of calculating the MS-SSIM, we chose to evaluate the 3D MS-SSIM with 3 scales as opposed to the 5 scales that were studied by the authors of the MS-SSIM. For 3D images with on the

order of 32 slices, undersampling twice to achieve two additional scales (three scales in total) will reduce the image extent along the slice direction to 8 slices; beyond this we considered downsampling as meaningless due to the excessive averaging of image features along the slice direction. Figure 3.4 shows the change in MS-SSIM as the number of scales considered increases. As more scales were added, the MS-SSIM reported higher scores. This is a result of the fact that the MS-SSIM is composed of a weighted product of intermediate scores calculated at each scale. At each intermediate scale, the image is lowpass filtered to accentuate coarser scale features, the result being that coarse scales show higher intermediate scores than finer scales (e.g. incoherent aliases will more detrimentally affect the perceived structure of fine scale image features in the pelvic MRI image, resulting in lower intermediate score when examining fine features, while the effect on the coarse features is less severe). As more weight is allocated to intermediate scales, the inclusion of extra scales results in a MS-SSIM score calculated from larger intermediate scores that have been allocated more weight, raising the MS-SSIM score. Due to the fact that radiologists are frequently concerned with small details in an image, we consider the inclusion of three scales with the MS-SSIM as an appropriate compromise between scale sensitivity and representation of scales that are most prominently desired in medical images. Future work must determine the effect of image scale when medical images are viewed by radiologists, since it is almost certain that radiologist preference will lie in images with greater preservation of fine detail (e.g. blood vessels) relative to large detail (e.g. outline of the pelvis).

We chose to implement luminance, contrast, and structural term calculations with an $11 \times 11 \times 5$ Gaussian window, though the MS-SSIM was not found to differ greatly as the extent of the window in the slice direction changed (Figure 3.5), the explanation for which is similar to that given for the SSIM in section 3.5.2. Because an image with on the order of 32 slices analysed at 3 scales will result in the smallest image having 8 slices, we felt that a Gaussian window extending 5 voxels in the slice direction permitted the window to capture sufficient unique detail in the image during structural term calculation (as opposed to a window with 11 voxel extent, for example, which will already exceed the dimension of the slice direction at the third scale).

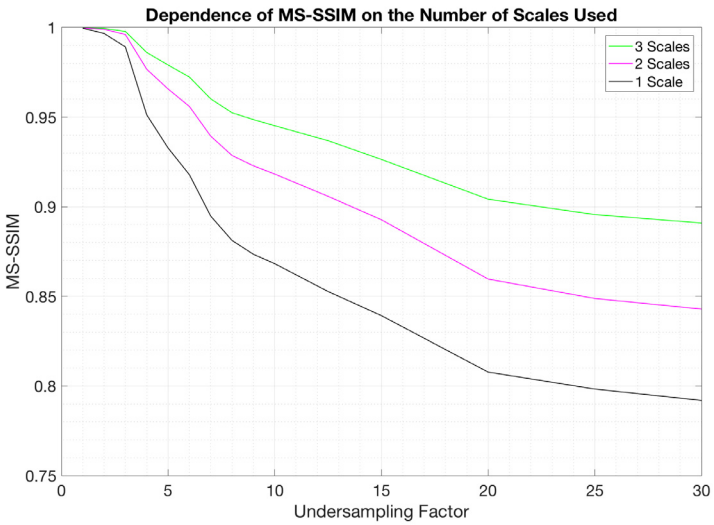

Figure 3.4: The MS-SSIM scores increased as the number of scales used increased. Future studies relating radiologist perception to MS-SSIM scores may wish to consider these differences.

Implementing the MS-SSIM as described above, it takes on the order of 5.2 seconds to calculate a score for a $256 \times 256 \times 30$ image. The reason that the MS-SSIM takes less time than our implementation of the SSIM is that we have used a smaller Gaussian window in calculating the MS-SSIM (i.e. $11 \times 11 \times 11$ for the SSIM vs $11 \times 11 \times 5$ for the MS-SSIM), and thus less padding of the images is required.

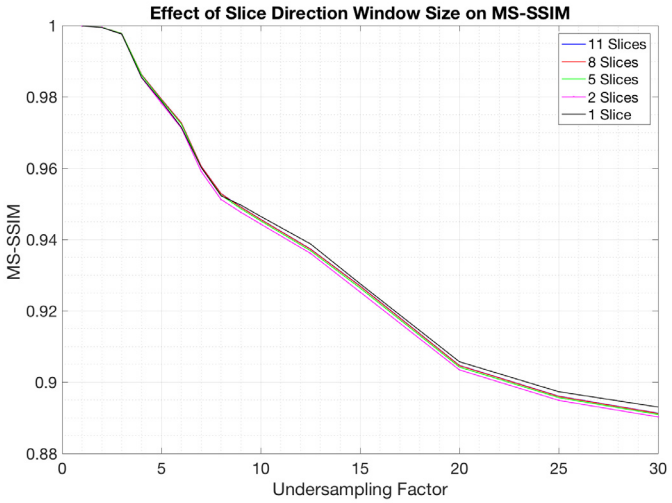

Figure 3.5: The MS-SSIM scores did not vary greatly as the extent of the Gaussian window used for local structural term calculations along the slice direction changed.

3.5.4 IW-SSIM

The information weighting scheme used in the IW-SSIM assumes that the Laplacian pyramid coefficients are distributed according to a Gaussian scale mixture (GSM) model, an assumption which was borrowed from detailed studies on the statistics of natural images showing that the marginal distribution of wavelet coefficients are modelled by a GSM[39]. While an in-depth study of the statistics of medical images is beyond the scope of this thesis, we conducted a very brief investigation of the marginal statistics of both Laplacian pyramid and wavelet coefficients in pelvic MRI images to provide motivation for the use of the IW-SSIM in the context of MRl images[iv]. We wished only to observe that the statistics of the pelvic MRI images we studied were broadly similar to those of natural images, for which the assumptions of the IW-SSIM information weighting scheme were made. Figure 3.6 shows exemplar results of the statistics observed for pelvic MRI images compared to a natural image. The distribution of coefficients were found to be qualitatively similar; both classes of

[iv]The reader is reminded that the subset of natural images *do not* include medical images, and so assumptions made for natural images do not immediately transfer to MRI images.

images showed peaks at coefficients of zero with heavy tails as coefficient magnitude increased. Thus, the assumption that a GSM can model the transform coefficients of a pelvic MRI image is valid, and the IW-SSIM can be studied in the context of medical imaging with some measure of confidence.

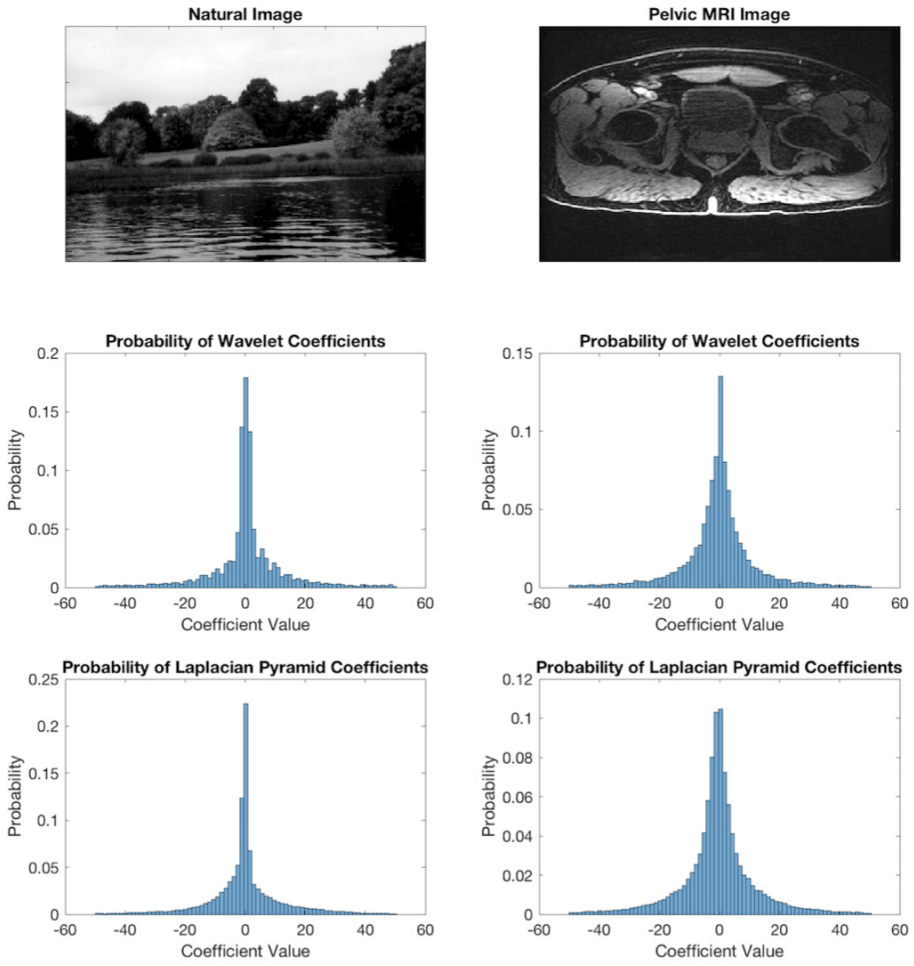

Figure 3.6: The 15 pelvic images studied showed GSM-like marginal distributions of both the Laplacian pyramid coefficients and the wavelet coefficients. The GSM distributions modelling natural scene coefficients are characterized by heavy tails and sharp peaks near zero. The pelvic MRI images we studied showed GSM-like marginal distributions for transform coefficients, motivating the study of the IW-SSIM in pelvic MRI images.

We implemented a 3D Laplacian pyramid generation by employing a $5 \times 5 \times 5$ 3D

extension of the 5×5 binomial kernel used in the re-scaling step for the 2D Laplacian pyramid. The 3D Laplacian pyramid generation then follows the 2D pyramid generation steps detailed in section 2.4.5.

As in our 3D implementations for the SSIM and the MS-SSIM, we chose to handle edge cases during the local window calculations of luminance, contrast, and structure terms by padding with replicates of the border values as opposed to ignoring the border values. We implemented the IW-SSIM local window calculations using an $11 \times 11 \times 5$ Gaussian window, evaluated at three scales of the Laplacian pyramid. We chose to implement three scales in the IW-SSIM for the same reasons discussed for the MS-SSIM (see section 3.5.3). The effect to the IW-SSIM of changing window size along the slice direction was similar to the effect in the MS-SSIM (see Figure 3.5), the explanation for which is similar to the discussion given in section 3.5.2 for the SSIM. The IW-SSIM responded less significantly than the MS-SSIM to different numbers of included scales (Figure 3.7). This is a result of the fact that statistical information weighting has been included at each intermediate scale during calculation of the IW-SSIM, tailoring the intermediate scores according to the presence of salient distortions. For reasons similar to those stated in section 3.5.3, we felt it appropriate to use three scales with a Gaussian window slice-extent of 5 voxels.

As opposed to the default use of a neighbourhood that includes a parent coefficient in the original IW-SSIM, we chose to omit a parent coefficient when calculating information weighting maps. This was done primarily for coding convenience when extending the IW-SSIM to 3D images. The omission of a parent coefficient in calculating the information weighting does not significantly change the results of the 2D IW-SSIM, as shown in Figure 3.8, and would not be expected to change the results of the 3D IW-SSIM either.

Implementing the IW-SSIM as described above, it takes on the order of 15.7 seconds to calculate a score for a $256 \times 256 \times 30$ image. The reason that the IW-SSIM takes less time than our implementation of the SSIM is that we have used a smaller Gaussian window in calculating the IW-SSIM (i.e. $11 \times 11 \times 11$ for the SSIM vs $11 \times 11 \times 5$ for the IW-SSIM), and thus less padding of the images is required. However, the calculation of the information weighting terms causes the IW-SSIM to be slower in implementation than the MS-SSIM.

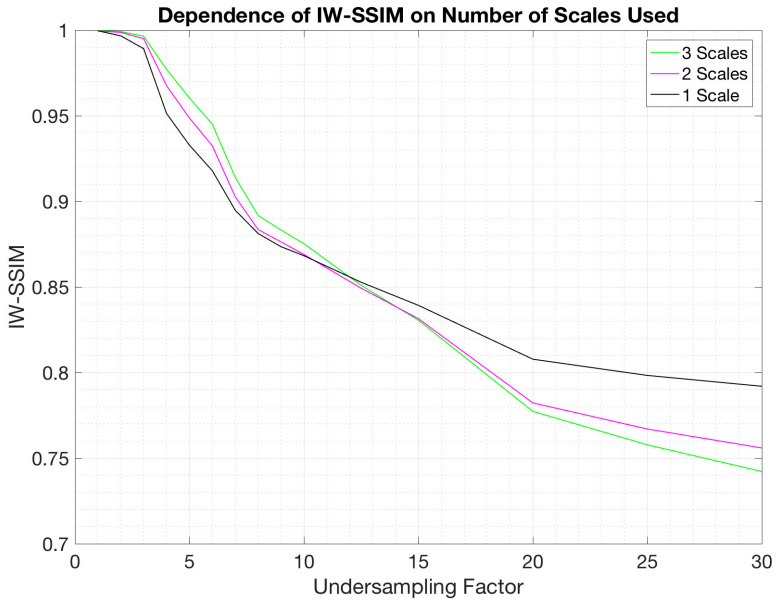

Figure 3.7: The change in IW-SSIM was not drastic as the number of scales considered increased.

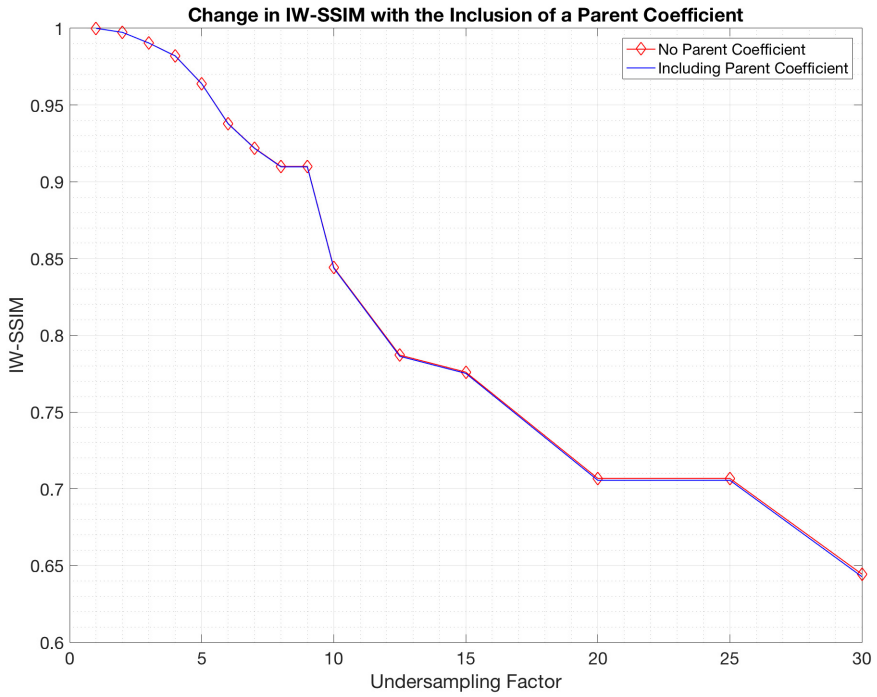

Figure 3.8: The inclusion of a parent coefficient in the information weighting calculation does not make a significant impact on the IW-SSIM.

3.5.5 GMSD

The GMSD, described in section 3.5.5, was extended for use in 3D images. The 3×3 Prewitt filters used in the 2D implementation of the GMSD were extended to $3 \times 3 \times 3$ gradient filters, normalizing the entries appropriately. Figure 3.9 demonstrates how the horizontal gradient filter was extended to 3D; vertical and slice-directed gradient filters were designed similarly. Gradient magnitude similarity map (GMS)

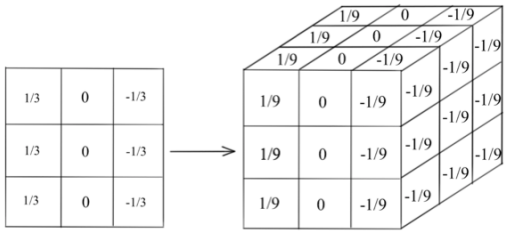

Figure 3.9: The 3×3 Prewitt filters were extended to $3 \times 3 \times 3$ filters, renormalizing the entries appropriately (i.e. by number of active elements in each directional component of the filters). Shown here is the extension of a horizontal filter to 3D.

computation was then performed analogously to the 2D GMSD scheme, filtering both the reference image R and reconstructed image D with the three directional gradient filters. The standard deviation of the resulting 3D GMS map was taken to report a GMSD score.

Contrary to the original 2D GMSD implementation, which uses zero-padding to return convolution results at edge cases, we handle convolution at edge voxels by padding with replicates of border values. We chose this approach to maintain similar voxel intensities when padding of the image is necessary, though the difference in padding techniques is more a matter of preference than performance; because both R and D are subjected to the same filtering process, the difference in performance between padding techniques is not significant.

A notable difference in our implementation of the GMSD is that we do not perform the 2-times downsampling step suggested by the authors. We chose not to include 2-times downsampling largely because the slice direction of 3D MRI images is small

compared to the in-plane dimensions of the images, and so the images would suffer unnecessary loss of detail in the slice direction. It should also be noted that the authors of the GMSD recommended a 2-times undersampling step by citing the use of such a step in the calculation of the SSIM, however the authors of the SSIM only recommended an undersampling step be used under certain conditions[v] that were not met in our implementation. Finally, given that the purpose of such a downsampling step is to bring the images into a scale for which the GMSD was purportedly validated with the use of natural images, its necessity in medical imaging applications does not necessarily follow. Regardless, the relative response of the GMSD with and without downsampling is similar (Figure 3.10), and no preference can be determined in the context of medical imaging until future work correlates the GMSD with radiologist opinion.

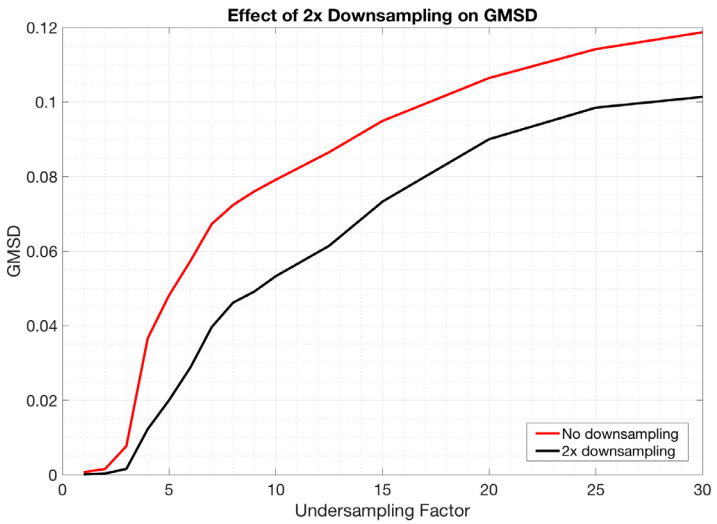

Figure 3.10: The trends in the GMSD with and without 2-times downsampling are very similar.

Implementing the GMSD as described above, it takes on the order of 0.4 seconds to calculate a score for a $256 \times 256 \times 30$ image.

[v] Wang et. al. initially recommended 2x undersampling of the luminance channel of images [35], though later empirically suggest downsampling the image such that the minimum dimension of the image measures 256 (see `https://ece.uwaterloo.ca/~z70wang/research/ssim/`).

Chapter 4

Implementing CIRCUS in CS-MRI

4.1 Purpose of Investigation

Although CIRCUS k-space trajectories allow flexible retrospective compromise between temporal resolution and image quality, the design of the CIRCUS quanta used during k-space sampling is determined prior to data acquisition through the choice of b and c and cannot be changed retrospectively. Selection of some values of b and c may provide inherently better performance of CS-MRI using CIRCUS. CIRCUS trajectories with inherently better performance for a fixed undersampling factor, as quantified by objective image quality metrics (IQM), will allow more leniency in the compromise between temporal resolution and image quality. An informed decision regarding the parameters b and c is thus paramount in ensuring favourable performance using CIRCUS k-space sampling, both for static and dynamic MRI applications.

Though the authors of CIRCUS originally specified that $b \in \mathbb{Z}$ and $c \in 0 \cup (1, 2)$, no justification was given for these limitations[14]. Choosing any $b, c \in \mathbb{R}$ is permitted by equation 2.59 in theory. By extending the range of possible b and c choices, we extend the possible range of good CIRCUS parameter choices with regards to CS-MRI image reconstruction performance.

In this section we investigated the properties of simulated CIRCUS acquisitions using various b and c parameters in a static pelvic CS-MRI image reconstruction. Our intent was twofold:

1. Determine whether choices of b and c exist that consistently give improved performance in CS-MRI, as quantified by objective image quality metrics (IQMs), at a fixed undersampling factor.

2. Determine if the coherence or the incoherence of the CIRCUS pattern point spread function (PSF) can be used as a surrogate for choosing the best CIRCUS parameters for CIRCUS trajectories generated for a fixed undersampling factor.

We began by seeking choices of b and c that produced unique CIRCUS trajectories, to establish what choices of b and c were truly valid. We then extensively investigated the PSF coherence and incoherence measures of many CIRCUS trajectories at increasing undersampling factors, to determine if any unusual trajectories existed across undersampling factors. Following this, with knowledge of the behaviour of CIRCUS trajectory PSFs at many undersampling factors, we characterized the performance of CIRCUS trajectories for a single fixed undersampling factor using objective IQMs and sought correlations between IQM scores and PSF measures.

Then end goal of this chapter was to provide confidence in CIRCUS trajectory implementations for use with CS-MRI. Our aim was to provide suggestions as to which values of b and c are favourable for implementation, and to seek whether or not PSF measures can serve as a tool for predicting CIRCUS trajectory performance. By utilizing CIRCUS trajectories that inherently allow CS-MRI to produce better quality images, as quantified by objective IQMs, dynamic MRI applications are afforded a measure of increased leniency in the compromise between temporal resolution and image quality.

4.2 Hypotheses

We hypothesize the following:

1. With regards to valid b and c values for unique CIRCUS trajectories:
   (a) No two distinct b_1 and b_2 will produce identical CIRCUS trajectories for fixed c.
   (b) There will exist a subset $C \subset \mathbb{R}$ such that $c_1, c_2 \in C$ produce identical CIRCUS trajectories. For all choices $c_1, c_2 \in \mathbb{R} \setminus C$, no two distinct c_1 and c_2 will produce identical CIRCUS trajectories for fixed b.
2. With regards to the coherence and incoherence of CIRCUS trajectory PSFs (equations 2.26 and 2.27 respectively) with varying b and c both at increasing undersampling factors and within a fixed undersampling factor:
   (a) There will exist a region of the b-c parameter space for which consistent PSF measure behaviours are observed.

(b) There will exist a region of the b-c parameter space for which PSF coherence will reach an average minimum behaviour and for which PSF incoherence will reach an average maximum behaviour.

3. With regards to the performance, as quantified by objective IQMs, of CIRCUS trajectories with fixed undersampling factor as b and c vary:

   (a) There will exist a region of the b-c parameter space for which objective image quality will be consistent on average.

   (b) Objective IQM scores will correlate with both the coherence and incoherence measures of the CIRCUS PSFs.

4.3 Methods

4.3.1 Point Spread Function (PSF) Coherence and Incoherence

CIRCUS trajectories measuring of $256 \times 256 \times 34$ voxels were generated, equivalent to the size of the pelvic MRI image used for the CIRCUS trajectory performance study (see section 4.3.2). Varying the sizes moderately was not observed to significantly change the final trends in the PSF coherence and incoherence. Undersampling occurred along the z and y phase encoding axes, with full sampling along the x frequency encode axis. CIRCUS pattern generation used values of b ranging from 0 to 90 in steps of 0.5 (181 values) and values of c ranging from 0 to 5 in steps of 0.05 (101 values), resulting in 18281 CIRCUS trajectories. These values were chosen as a representative range of parameter values, and it is assumed that any trends observed will hold as the range of parameters is reasonably extended. The PSF coherence and incoherence were calculated for each CIRCUS trajectory using equation 2.26 and 2.27 respectively. This was performed for undersampling factors of 2, 4, 6, 8, 10, and 12. Two-way analysis of variance (ANOVA) testing was performed to determine if changes in b or c produced significant differences in mean PSF measures. Tukey honest significant difference (HSD) post hoc analysis (described in appendix A.2), was used to further investigate the effect of b and c on PSF measures.

4.3.2 Effect of b and c on IQM Scores

The root mean squared error (RMSE), gradient magnitude similarity deviation (GMSD), structural similarity index (SSIM), multi-scale SSIM (MS-SSIM), and information-weighted SSIM (IW-SSIM) were calculated for 3D CS-MRI reconstructions of a single pelvic MRI image ($256 \times 256 \times 34$ voxels). RMSE results were normalized to the mean reference image intensity to yield a unitless normalized RMSE (nRMSE). K-space acquisition used CIRCUS patterns as generated in section 4.3.1 with a simulated 8 coil PI acquisition. Acquisitions were made at an undersampling factor of $R = 10$, and CS-MRI reconstructions were performed using total variation (TV), wavelet, and locally low rank (LLR) sparsity enforcements, all with regularization weight chosen as $\lambda = 0.02$. A single undersampling factor was used because the purpose of the IQM investigation was to determine how IQMs quantified CIRCUS trajectories at a fixed R, and whether those scores correlated with PSF measures. The choice of λ was arbitrary, since the goal was to characterize the relative performance of each CIRCUS trajectory and not to determine the best regularization weight for a given undersampling factor.

4.4 Results

4.4.1 Unique CIRCUS Trajectory Parameter Choices

Repetitions in b with Fixed c

The full derivation of these results may be found in appendix B.2.

For two choices $b_1, b_2 \in \mathbb{R}$ of the CIRCUS radial parameter b, identical CIRCUS trajectories occur for the trivial condition when b_2 is chosen to be very near in value to b_1. Assuming that a reasonable user does not specify b to more than two decimal places, this effect will not be observed in practise. The results describing this effect are therefore presented in appendix B.2 for the interested reader. Aside from trivial choices of b_1 and b_2 made exceedingly near to one another, no two distinct b_1 and b_2 generate identical CIRCUS trajectories.

Repetitions in c with Fixed b

The full derivation of these results may be found in appendix B.3.

For two choices $c_1, c_2 \in \mathbb{R}$ of the CIRCUS radial parameters c, identical CIRCUS trajectories occur for the trivial condition when c_2 is chosen to be very near in value to c_1. This effect will not be observed in practise, assuming that a reasonable user doesn't specify c to beyond two decimal places. The results describing this effect are therefore presented in appendix B.3 for the interested reader. It was found that for two choices of $c_1, c_2 \in \mathbb{Z} \geq 2$, the CIRCUS trajectories produced will be identical. For all other choices of c_1 and c_2, CIRCUS trajectories will be unique.

4.4.2 Effect of b and c on CIRCUS PSF Measures

As exemplar results of the PSF measure behaviours, Figure 4.1 shows the coherence and incoherence of CIRCUS PSFs with undersampling factors of 2, 6, and 10 for various combinations of b and c. These results are representative of results for $R = 4, 8$, and 12, which may be found in appendix B.4.

Two-way ANOVA tests of the effects of b and c on PSF coherence revealed a significant ($p < 0.05$) difference between mean PSF coherence with fixed b and varying c and the mean PSF coherence with fixed c and varying b the majority of the time; only for $R = 2$ and 4 did the PSF coherence not show significant difference in mean coherences at fixed b as c varied and only for $R = 4$ did the PSF coherence not show significant difference in mean coherences at fixed c as b varied. In all cases, the PSF incoherence showed significant difference in mean values as b and c varied.

Tukey HSD post hoc analysis revealed evidence of regimes in CIRCUS trajectory PSF incoherences at higher undersampling factors, though the same was not observed for the PSF coherences. Figure 4.2 gives an exemplar visualization of the effect of b and c on PSF coherence and incoherence for CIRCUS trajectories with $R = 10$. Shown in the subfigures are the mean PSF coherence and incoherence values for fixed b and for fixed c, plus-or-minus Tukey HSD comparison intervals for each mean. It was found that no statistically significant trend towards decreased PSF coherence existed with respect to either b or c (i.e. non-overlapping Tukey comparison intervals occurred sporadically). There was a statistically insignificant trend towards decreased mean PSF coherence for $c > 1$ observed for $R \geq 4$. There were no consistent trends

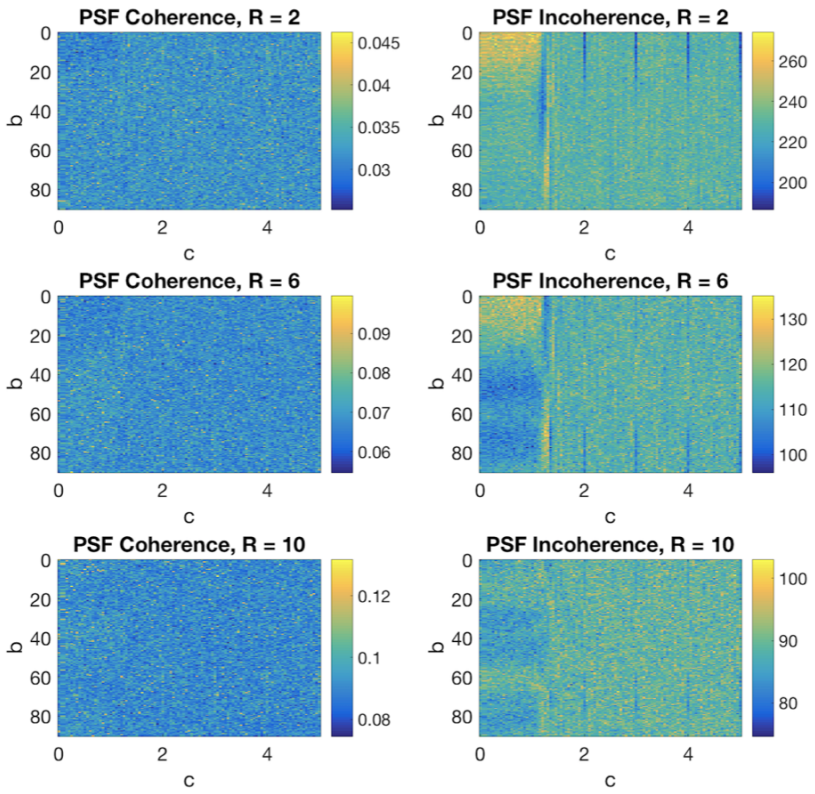

Figure 4.1: Coherence (left column) and incoherence (right column) of CIRCUS pattern PSFs generated for a variety of b and c at increasing undersampling factor.

in the mean PSF incoherences with changes in b. Finally, a statistically significant trend towards higher PSF incoherence for $c \in \mathbb{R} \setminus \mathbb{Z} > 2$ relative to $c < 1$ was observed for $R \geq 6$, with a similar trend observed without statistical significance for $R = 4$. The relative difference in mean PSF incoherences with this statistically significant increase in mean PSF was typically on the order of 5%.

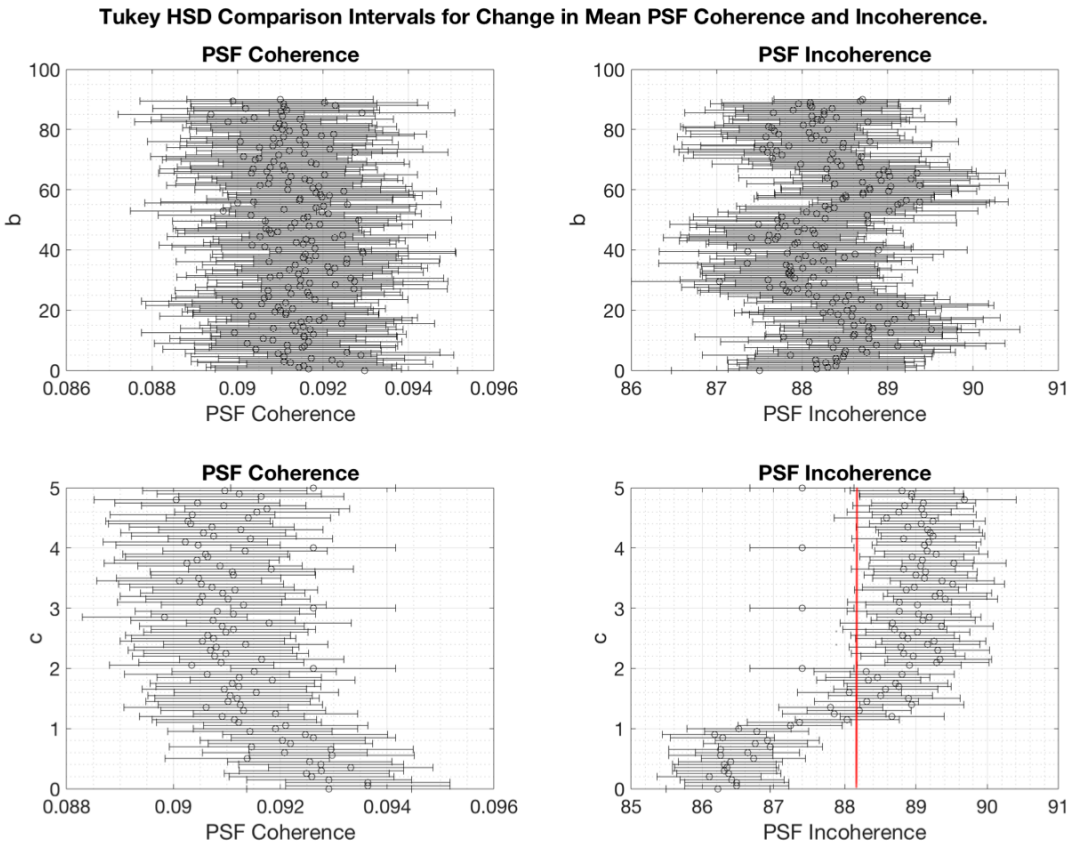

Figure 4.2: Visualization of Tukey HSD post hoc analysis for mean PSF coherence and incoherence as b changes (top row) and as c changes (bottom row) at $R = 10$. No significant trends are observed in the PSF coherence measure, nor with the PSF incoherence measure as b changes. For $c \in \mathbb{R} \setminus \mathbb{Z} > 2$, there is a higher occurance of statistically significant gains in mean PSF incoherence relative to both $c < 1$ and to $c \in \mathbb{Z} > 2$ (red line demarcates the maximum of the comparison intervals for $c \in \mathbb{Z} > 2$. This was observed for $R \geq 6$.

4.4.3 Effect of b and c on IQM Scores

Results for the nRMSE, GMSD, and SSIM are shown in Figure 4.3. These are representative of the MS-SSIM and IW-SSIM, which may be found in appendix B.5.

For approximately $c < 1$, noticeably different IQM responses to b and c were visible for CS-MRI reconstructions using all three sparsity enforcements. For approximately

$c > 1$ the IQMs all appeared to reach an average behaviour, with no distinct regions emerging as superior or inferior in terms of objective image quality.

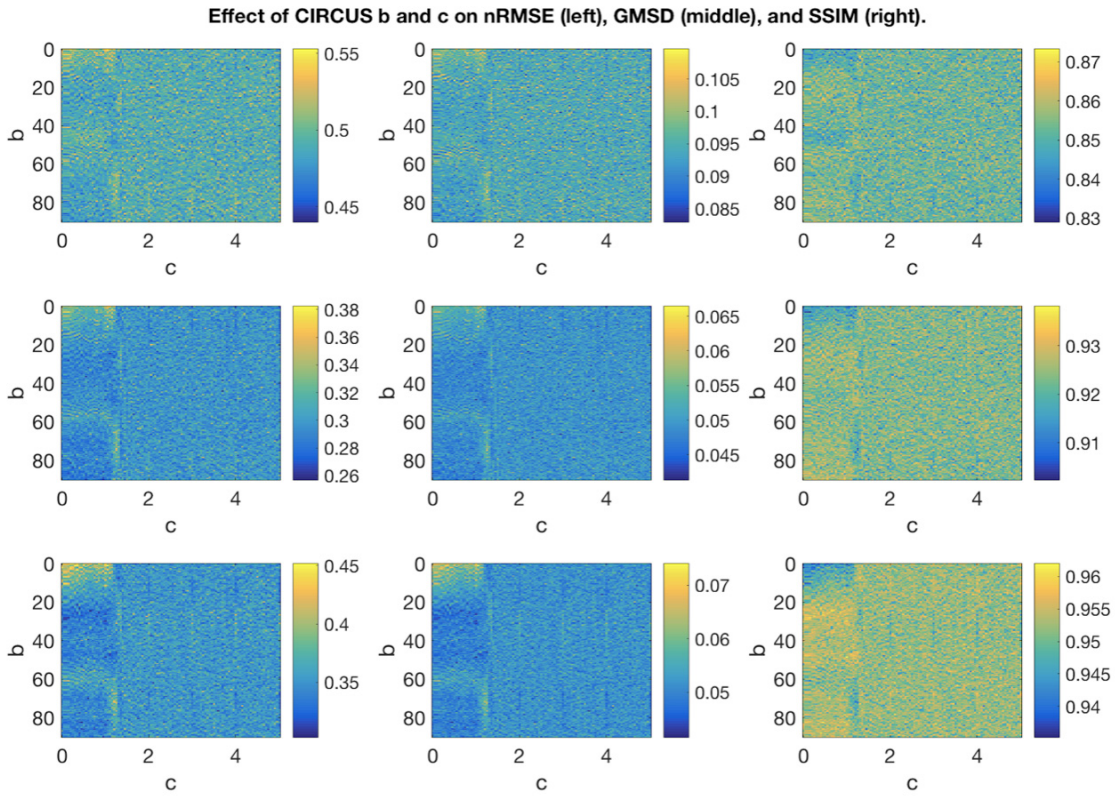

Figure 4.3: Response of the nRMSE (left column), GMSD (middle column), and SSIM (right column) to changes in b and c for TV sparsity enforcement (top row), wavelet sparsity enforcement (middle row), and LLR sparsity enforcement (bottom row).

Two-way ANOVA testing of the effect of b and c on each of the IQMs indicated that for all IQMs and all sparsity enforcements there was a significant ($p < 0.05$) change in the mean IQM scores as b changed. The difference in mean IQM scores with respect to varying c was significant only for the nRMSE and the GMSD with a wavelet sparsity enforcements; for all other IQMs and sparsity enforcements the differences in mean IQM scores were not significant as c changed.

Tukey HSD post hoc analysis was performed to determine whether domains of b and c values existed that consistently gave statistically significant improvements in

IQM score relative to the other values of b and c. Figure 4.4 shows representative results of the Tukey HSD post hoc analysis of the SSIM results. No statistically significant trends were observed for any of the IQMs utilizing any of the sparsity enforcements. Statistically insignificant trends towards improved IQM scores existed for all IQMs and all sparsity enforcements for approximately $b > 20$. Statistically insignificant trends towards poorer IQM scores existed for all IQMs and all sparsity enforcements for approximately $c > 1$.

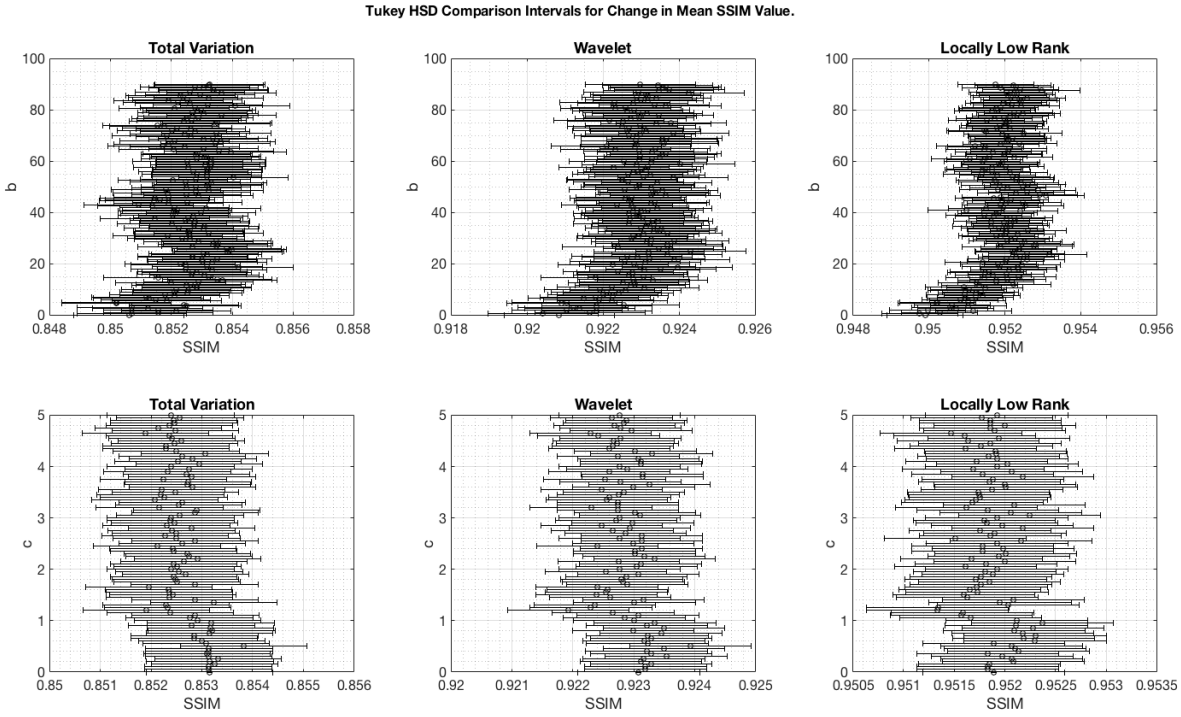

Figure 4.4: Visualization of overlap in Tukey HSD comparison intervals for mean SSIM score as b changes (top row) and as c changes (bottom row), for each of the three sparsity enforcements.

4.4.4 Correlations Between IQM Scores and PSF Measures

Tables 4.1 to 4.4 show the coefficients of the Pearson linear correlation and Spearman rank correlation between each of the objective IQM scores and the PSF coherence

and incoherence measures calculated for the CIRCUS trajectories with $R = 10$. In all cases, there was not strong correlation between PSF coherence and incoherence measures and the resulting objective IQM scores. Figures representing the relationship between each IQM and the PSF coherence and incoherence may be found in the appendix.

Table 4.1: Pearson linear correlation coefficients for correlation between each IQM and PSF coherence, for each of the three sparsity enforcements studied. Bolded entries were statistically significant ($p < 0.05$).

	nRMSE	GMSD	SSIM	MS-SSIM	IW-SSIM
TV	**0.0400**	**0.0298**	**-0.0425**	**-0.0428**	**-0.0440**
W	**0.0173**	**0.0155**	**-0.0367**	**-0.0430**	**-0.0442**
LLR	**0.0226**	**0.0183**	**-0.0339**	**-0.0403**	**-0.0473**

Table 4.2: Spearman rank correlation coefficients for correlation between each IQM and PSF coherence, for each of the three sparsity enforcements studied. Bolded entries were statistically significant ($p < 0.05$).

	nRMSE	GMSD	SSIM	MS-SSIM	IW-SSIM
TV	**0.0308**	**0.0168**	**-0.0372**	**-0.0393**	**-0.0409**
W	0.0050	-0.0019	**-0.0265**	**-0.0371**	**-0.0417**
LLR	0.0048	-0.0027	**-0.0209**	**-0.0349**	**-0.0391**

Table 4.3: Pearson linear correlation coefficients for correlation between each IQM and PSF incoherence, for each of the three sparsity enforcements studied. Bolded entries were statistically significant ($p < 0.05$).

	nRMSE	GMSD	SSIM	MS-SSIM	IW-SSIM
TV	0.0077	**0.0300**	**0.0282**	**0.0373**	**0.0451**
W	**0.0592**	**0.0636**	-0.0089	**0.0176**	**0.0304**
LLR	**0.0866**	**0.0980**	**-0.0506**	**-0.0145**	-0.0129

Table 4.4: Spearman rank correlation coefficients for correlation between each IQM and PSF incoherence, for each of the three sparsity enforcements studied. Bolded entries were statistically significant ($p < 0.05$).

	nRMSE	GMSD	SSIM	MS-SSIM	IW-SSIM
TV	0.0110	**0.0302**	**0.0264**	**0.0339**	**0.0413**
W	**0.0676**	**0.0689**	-0.0129	0.0139	**0.0269**
LLR	**0.1048**	**0.1162**	**-0.0587**	**-0.0198**	**-0.0165**

4.5 Discussion

4.5.1 Unique CIRCUS Trajectory Parameter Choices

We investigated when choices of two different radial parameters b_1 or b_2 (at fixed c) or two different spiral parameters c_1 or c_2 (at fixed b) lead to equivalent CIRCUS patterns. This was done in an effort to determine if redundant subsets of b and c existed, putting natural limits on the feasible ranges of b and c that could be selected. With regards to b, it was found that CIRCUS patterns repeated for two choices of b_1 and b_2 only if b_1 and b_2 were chosen exceedingly near in value. This effect is negligible in practise, assuming that a user doesn't specify b to beyond 2 decimal places, and so we conclude that CIRCUS trajectories do not repeat with b in practical implementations. This confirms hypothesis 1a.

It was found that CIRCUS patterns will repeat for c taken as any integer greater than or equal to two for fixed b. Though two choices c_1 and c_2 may give identical CIRCUS patterns if they are chosen exceedingly near in value, this effect is negligible in practise assuming that the user is not specifying c to more than two decimal places. We may conclude that CIRCUS patterns will not repeat with two choices of c in general, except at integer values of c greater than or equal to two. This confirms hypothesis 1b, where the hypothesized subset C is $C = \mathbb{Z} \setminus [0, 1]$.

Given the golden ratio design elements incorporated in CIRCUS, it is not unexpected that CIRCUS trajectories are largely unique. Knowledge of the theoretical freedom in CIRCUS parameter choices allows for an unrestricted investigation of CIRCUS trajectory designs. Furthermore, it opens future possibilities in the implementation of CIRCUS trajectories. For example, rather than designing a series of CIRCUS quanta at fixed b and c for dynamic acquisition of k-space data, the uniqueness of CIRCUS trajectories at different b and c allows the user to generate a series of different CIRCUS quanta at many different b and c for k-space sampling and to retrospectively interleave the quanta with knowledge that acquired k-space coefficients will not entirely overlap between CIRCUS quanta designed with different b and c. Future work may investigate if such a CIRCUS implementation is beneficial.

4.5.2 Effect of b and c on CIRCUS PSF Measures

Figure 4.1 demonstrates exemplar behaviour of the PSF coherence and incoherence as b and c change. Noticeable trends between all PSF coherence and incoherence results existed as undersampling factor increased. Namely, for approximately $c < 1$ both the coherence and incoherence showed noticeably different response to b and c than for approximately $c > 1$. For higher undersampling factors, the variations in the PSF coherence and incoherence for $c < 1$ were more consistent. In all cases, for $c > 1$ the PSF coherence and incoherence measures achieved relatively uniform response to b and c. This confirms hypothesis 2a.

The PSF incoherence was a more responsive metric than the PSF coherence, which showed no consistent statistically significant trends as b or c changed. The bottom right panel in Figure 4.2 suggests that three domains of PSF incoherence may exist with regards to the effect of c on the behaviour of the resulting CIRCUS pattern; relatively low incoherence for $c < 1$ (the "sub-CIRCUS regime" of c), comparatively higher incoherence for $c \in \mathbb{R} \setminus \mathbb{Z} > 2$ (the "super-CIRCUS regime" of c), and a transitionary region between $c = 1$ and $c = 2$ within which the authors of CIRCUS originally proposed the scheme follow (the "CIRCUS regime" of c). These regimes were observed, with statistical significance, for undersampling factors greater than or equal to six. Figure 4.5 demonstrates the change in CIRCUS pattern structure for values of c taken in each of these three domains prior to the generation of rectangular phase encoding tables as described in section 2.3.4. While the rectangular CIRCUS patterns are the one actually used in practise, it is instructive to view the underlying square CIRCUS patterns from which they are generated. In the sub-CIRCUS regime of c, the change in CIRCUS pattern structure is dominated by changes in b, with changes in the value of c causing minimal change in the sampled phase encode coordinates. Within the CIRCUS regime of c, obvious structure characteristic to the CIRCUS scheme in the sampled phase encode coordinates exists; this structure might be described as "boxy spirals". For non-integer values of c chosen in the super-CIRCUS regime, the CIRCUS patterns are visibly less structured. Indeed, the reader may *visualize* the increase in incoherence of the resulting CIRCUS patterns as c changes through these three regimes, consistent with the results of Figure 4.2. No such regimes were observed for varying b, consistent with the Tukey HSD post

hoc analysis results of Figure 4.2 (top right panel). Figure 4.6 show the resultant

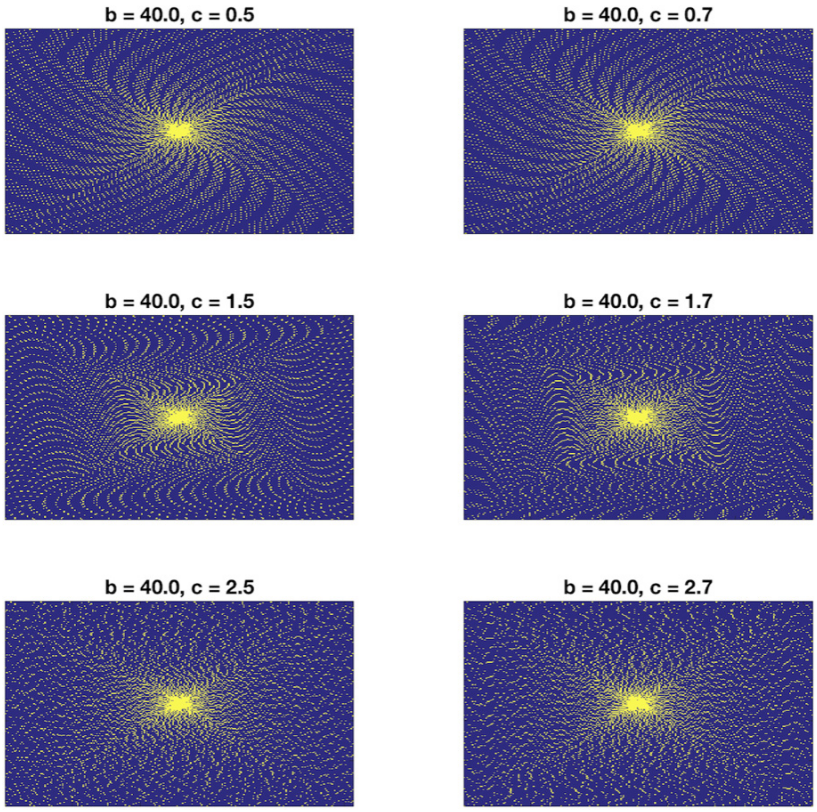

Figure 4.5: Change in structure of CIRCUS patterns ($R = 10$) prior to rectangular pattern generation when c is taken in the sub-CIRCUS regime (top row), the CIRCUS regime (middle row), and the super-CIRCUS regime (bottom row). The values of c shown for the super-CIRCUS regime are representative of the entire super-CIRCUS regime. The relative behaviour of each regime is consistent for all values of R. Similar regimes in b were not observed.

CIRCUS structures for each of the three CIRCUS domains following the generation of rectangular patterns, as described in section 2.3.4, from the square patterns shown in Figure 4.5. While the distinct structures aren't as obvious in the rectangular patterns, the rectangular patterns in Figure 4.6 inherit aspects of the square patterns shown in Figure 4.5 – hence the existence of the regimes we observed in the PSF incoherence. However it also becomes clear that through the generation of rectangular

CIRCUS patterns, the difference in CIRCUS structures is reduced. This explains the low range in PSF measures that we observed.

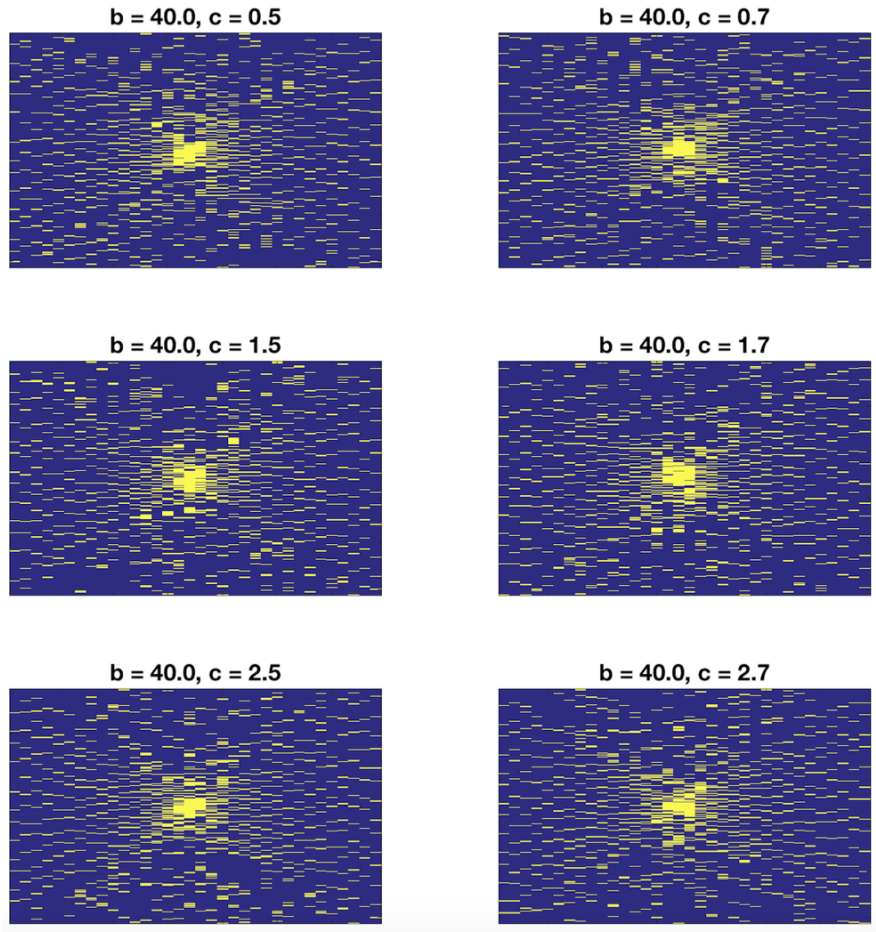

Figure 4.6: Change in structure of CIRCUS patterns ($R = 10$) following the rectangular pattern generation when c is taken in the sub-CIRCUS regime (top row), the CIRCUS regime (middle row), and the super-CIRCUS regime (bottom row). The values of c shown for the super-CIRCUS regime are representative of the entire super-CIRCUS regime. The relative behaviour of each regime is consistent for all values of R. Similar regimes in b were not observed.

These results only partially support hypothesis 2b. With regards to both the coherence and incoherence measures of the CIRCUS PSFs, there were no preferential regimes of b. With regards to the coherence of the PSF, a trend towards lower mean

coherence with $c > 1$ is observed, but it is generally not statistically significant. With regards to the PSF incoherence, maximal incoherence is achieved when c is selected as a non-integral value in the super-CIRCUS regime of values (i.e. $c \in \mathbb{R} \setminus \mathbb{Z} > 2$). This increase in PSF incoherence was not very large however, and was observed only for $R \geq 6$ with statistical significance. The results of Figure 4.1 and supporting results in the form of Figure 4.2 would suggest that for $c > 1$ the resulting PSF measures reach an overall average response to b and c compared to $c < 1$ across all undersampling factors. To fully incorporate the effect of the spiral offset introduced by c, and maintain uniform CIRCUS properties across different R, it is suggested that c be chosen above the sub-CIRCUS regime of values, i.e. $c > 1$ (the reader is again referred to Figure 4.5). While a marginal gain in PSF incoherence was observed for $c \in \mathbb{R} \setminus \mathbb{Z} > 2$, this was not found to translate to significantly better IQM scores (see section 4.4.4) and thus we make no recommendation to specifically operate in the super-CIRCUS regime of c values.

4.5.3 Effect of b and c on IQM Scores

Similar to the PSF coherence and incoherence, the IQM scores all showed noticeably different behaviour in the sub-CIRCUS regime of c values (e.g. Figure 4.3) and reached a stable average response to b and c above the sub-CIRCUS regime of c values. Regions of generally poorer IQM scores were often observed near $b = 0$ and $c = 0$, though the difference in mean IQM scores relative to larger values of b and c were typically not statistically significant (e.g. Figure 4.4). Unlike the PSF incoherence results, no regimes of consistently improved IQM scores existed for c taken as a non-integral value in the super-CIRCUS regime. In fact, all of the IQMS demonstrated a slightly higher sensitivity to changes in b, though even with changes in b there are no statistically significant regimes of consistently improved IQM scores. The only observations of note are a statistically insignificant trend towards improved mean IQM scores for approximately $b > 20$ and a statistically insignificant trend towards *degraded* mean IQM scores for approximately $c > 1$. The relative difference in mean IQM scores under these trends is exceedingly small however. An explanation of these trends can be obtained upon inspection of the IQM score results (e.g. Figure 4.3). The worst IQM scores occurred for b and c near zero, and some of the best IQM

scores occurred for $b > 0$ with $c < 1$ (e.g. between $b = 30$ and $b = 40$ in Figure 4.3). These extremes shift the *mean* IQM scores at fixed b or at fixed c in just such a way as to demonstrate the statistically insignificant trends suggested in Figure 4.4. These results only partially support hypothesis 3a. While the IQMs did achieve the most uniform behaviour for approximately $c > 1$, the same as for the PSF measures, the IQM scores were sporadically superior in the sub-CIRCUS regime.

An interesting case occurs when $b = 0$ and $c = 0$; the CIRCUS pattern generated with this combination of parameters is similar to a radial sampling pattern. Our results indicated that the worst objective image quality tended to occur when CIRCUS parameters were chosen for such "radial-like" CIRCUS patterns (e.g. Figure 4.3, where the nRMSE and GMSD were highest near $b = 0$ and $c = 0$ and the SSIM was lowest for the same parameter selections). An explanation for this may be discerned by examining Figure 4.7. Radial-like CIRCUS patterns are distinctly less random than CIRCUS patterns incorporating the effects of non-zero b and c. Because CS-MRI benefits from random undersampling of the k-space coefficients, it is expected that non radial-like CIRCUS patterns will impart the most benefit to CS-MRI. In terms of the PSF, it's seen that the radial-like CIRCUS pattern exhibits lots of energy in the side-lobes compared to non radial-like CIRCUS patterns. The non radial-like CIRCUS patterns all exhibited relatively similar PSFs compared to the radial-like CIRCUS pattern PSF, offering an explanation for why the radial-like CIRCUS pattern consistently resulted in lower objective image quality.

The results of this section would suggest that all CIRCUS trajectories perform more or less equivalently when c is chosen above the sub-CIRCUS regime. A "safe zone" in parameter selection may be considered for $c > 1$, where CIRCUS performed uniformly for all parameter choices and would be expected to perform uniformly across different undersampling factors based on the results discussed in section 4.5.2. Considering all these factors, our conclusion is that users should operate above the sub-CIRCUS regime of c values (i.e. $c > 1$) in order to take advantage of the relative stability of these regimes.

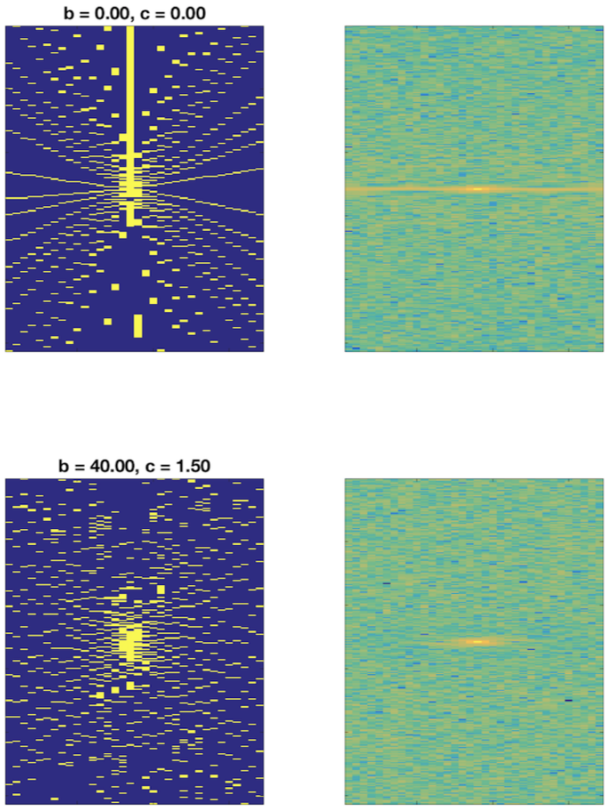

Figure 4.7: "Radial-like" CIRCUS patterns (top left) exhibit a pattern that is distinctly less random than CIRCUS patterns with non-zero b and c (bottom left). CS theory prefers random sampling of k-space, thus it is expected that the radial-like pattern will not perform as well. An alternative viewpoint comes by inspecting the PSF, where a radial-like CIRCUS pattern PSF (top right) exhibits more energy in the side lobes compared to a CIRCUS pattern PSF with non-zero b and c (bottom right, representative of many CIRCUS patterns).

4.5.4 Correlations Between IQM Scores and PSF Measures

No strong correlation, either linear correlation or rank correlation, was found between any of the objective IQMs and the PSF coherence or incoherence of CIRCUS patterns generated at a fixed undersampling factor. Hypothesis 3b is therefore rejected, and neither the PSF coherence nor the PSF incoherence are recommended as quantitative tools for selection of one CIRCUS pattern over another when considered at the same

undersampling factor.

The lack of correlation between the PSF coherence and incoherence and each of the five IQMs is likely a result of the fact that few significant differences existed in the PSF coherences and incoherences and in the IQM scores; there was very little variation in all of the values. It may therefore reasonably be concluded that all values of b and c studied resulted in CIRCUS trajectories that did not differ appreciably in terms of their objective CS-MRI performance, and so no correlations could emerge between the objective performance of the CIRCUS trajectories in CS-MRI reconstructions and their PSF measures.

4.6 Conclusions

In this section, we have investigated what range of CIRCUS b and c parameters resulted in unique implementations of CIRCUS. We further investigated the coherence and incoherence of CIRCUS trajectory PSFs at undersampling factors ranging from 2 to 12. Finally, we investigated CIRCUS trajectory performance as quantified by objective IQMs at an undersampling factor of 10 and investigated the link between PSF measures and the resulting objective IQM scores.

For typical selections of b and c, CIRCUS trajectories repeat only for two values of c chosen in the subset $c \in \mathbb{Z} \setminus [0, 1]$. This not only provides confidence that redundant choices in CIRCUS trajectories will not be made, but opens avenues for future work to investigate the implementation of several unique CIRCUS trajectories in dynamic CS-MRI applications.

The mean PSF incoherence demonstrated statistically significant increases for $c \in \mathbb{R} \setminus \mathbb{Z} > 2$ for $R \geq 6$, though the relative increase in the PSF incoherence was small. Both the PSF coherence and incoherence achieved an overall uniform response to b and c for approximately $c > 1$ at all R studied, indicating that the most uniform behaviour of the CIRCUS trajectories across undersampling factors will be expected when $c > 1$ is chosen.

The objective IQMs did not demonstrate any statistically significant trends towards improvement for any regimes of b and c, though were often observed to report lowest image reconstruction quality for b and c chosen near zero. For parameters chosen such that $c > 1$, the overall response of the IQMs were uniform. No significant

correlation was found between the IQMs and the PSF coherence/incoherence measures at an undersampling factor of 10 due to the small range of the PSF measures and the small range of the IQM scores. Thus the PSF is not recommended as a method of selecting CIRCUS trajectory parameters for use at a given undersampling factor. Given the uniform overall nature of the CIRCUS and super-CIRCUS regimes of c with regards to objective IQM scores and the PSF measures observed at many undersampling factors, it is recommended that CIRCUS parameters be chosen such that $c > 1$.

The outcome of this chapter is a knowledge of CIRCUS trajectory implementations, with regards to the uniqueness of CIRCUS trajectories and selections of b and c that will perform favourably on average across undersampling factors. Moving forward with this knowledge, we are confident that any CIRCUS trajectory designed with any selection of b and any $c > 1$ will not inherently limit the capabilities of CS-MRI reconstructions, and will provide the most uniform performance as undersampling factor is adjusted.

Chapter 5

Quantitative Behaviour of IQMs in Static CS-MRI

5.1 Purpose of Investigation

The unofficial standard for quantifying image reconstruction quality in the field of MRI is the root mean squared error (RMSE). This is in part due to the simple interpretation of RMSE values, as well as the rapid speed and ease of implementation of RMSE calculation. However, the RMSE has been found to correlate poorly with human perception of image quality[15]. We investigated the quantitative behaviours of four alternative objective image quality metrics (IQMs) in static CS-MRI reconstructions, using pelvic image reconstructions as an exemplar anatomy, and included the normalized RMSE (nRMSE) for comparison. The four alternative IQMs investigated were the gradient magnitude similarity deviation (GMSD), the structural similarity index (SSIM), the multi-scale SSIM (MS-SSIM), and the information-weighted SSIM (IW-SSIM).

Study of the quantitative properties of the objective IQMs in static CS-MRI is necessary to obtain foundational knowledge upon which their investigation in dynamic CS-MRI may begin. Towards this goal, we studied how the IQM scores varied for CS-MRI reconstructions across 15 pelvic CS-MRI images, we investigated the change in favoured sparse regularization weight for each IQM as R increased, and we studied whether objective IQM scores implied unique image degradation properties. In addition to informing the use of objective IQMs in dynamic CS-MRI applications, the results of this chapter will serve as valuable knowledge for future work correlating the IQM scores to radiologist subjective scoring and future work aiming to optimize CS-MRI reconstruction parameters.

5.2 Hypotheses

We hypothesize the following:

1. Objective IQM scores will be more consistent (i.e. less variation from the sample mean) across patients for the four alternative IQMs than for the nRMSE, for all sparsity regularizations.

2. For all five of the IQMs and for each sparsity regularization, the regularization weight λ giving the best IQM score will increase monotonically (i.e. never decrease, but may remain constant) as the undersampling factor R increases.

3. For any of the IQMs, image reconstructions with the same objective IQM score will not possess degradations with the same visual properties when reconstructed with different sparse regularizations.

5.3 Methods

5.3.1 Inter-Patient IQM Score Variability

The k-space of images from 15 pelvic MRI images were retrospectively undersampled using CIRCUS patterns designed with $b = 40$ and $c = 1.5$ using a simulated 8 coil acquisition at undersampling factors R ranging from 1 to 12 in steps of 0.5. The choice of $b = 40$ and $c = 1.5$ were made with the knowledge obtained via the results of chapter 4. Separate CS reconstructions were performed with total variation (TV), wavelet, and locally low rank (LLR) sparse regularizations using regularization weights λ ranging from 0 to 0.05 in steps of 0.02. The quality of each reconstruction was quantified by the five objective IQMs described in section 3.5. For each of the sparse regularizations this generated 5 sets of 15 IQM score matrices, one matrix per patient, which will be referred to as R-λ matrices.

The variation between the R-λ matrices calculated for a given sparsity regularization and a given IQM was quantified by the normalized root mean squared distance (RMSD) of the matrices from their sample mean, where normalization was done relative to the norm of the mean R-λ matrix. Larger RMSD implies larger variation of the R-λ matrices from their mean, and hence larger inter-patient variability in the IQM score results. The distance between matrices was quantified by the Frobenius

norm. For a matrix A of size $M \times N$, the Frobenius norm $\|A\|_F$ is given by:

$$\|A\|_F = \sqrt{\sum_{m=1}^{M}\sum_{n=1}^{N}|a_{m,n}|^2} \tag{5.1}$$

5.3.2 Favourable Sparse Regularization Weights at Increased Undersampling

For each undersampling factor R, corresponding to the rows of the R-λ matrices described in section 5.3.1, the value of λ that gave the minimum RMSE and GMSD were recorded and the value of λ that gave the maximum SSIM, MS-SSIM, and IW-SSIM were recorded. This was repeated for all 15 patient R-λ matrices, for each IQM, for each of the three sparse regularizations investigated.

5.3.3 Perceptual Equality of Degradations at Similar IQM Score

For each IQM, a metric score Q was selected. The R-λ matrices for a single patient across each of the three sparse enforcements were then searched to determine which combination of R and λ resulted in an image reconstruction with objective IQM score nearest to Q. The value of R was not constrained in the search, and was allowed to differ between the results for each sparsity enforcement as image reconstructions with IQM score nearest to Q were sought. The goal was to investigate the visual properties of CS-MRI image reconstructions possessing similar IQM scores, not to determine if one sparsity enforcement was better able to maintain high IQM scores for a given undersampling factor. The images were then visually compared by one graduate student, looking only for high level perceptual differences.

5.4 Results

5.4.1 Inter-Patient IQM Score Variability

The R-λ matrices may be found in appendix C.1. For each combination of sparsity enforcement and objective IQM, there were 15 R-λ matrices, resulting in 225 total R-λ matrices.

Table 5.1 shows the nRMSD of the collection of R-λ matrices from their sample means for each sparsity enforcement and each IQM. The GMSD was observed to have

the highest inter-patient variation, followed by the nRMSE, and then the SSIM family of metrics.

Table 5.1: nRMSD between image R-λ matrices and their sample mean, for CS-MRI reconstructions with each of the three sparsity enforcements as evaluated by each of the five objective IQMs.

	nRMSE	GMSD	SSIM	MS-SSIM	IW-SSIM
TV	0.0796	0.1325	0.0194	0.0075	0.0115
W	0.1196	0.1817	0.0121	0.0038	0.0057
LLR	0.0767	0.1407	0.0167	0.0069	0.0107

5.4.2 Favourable Sparse Regularization Weights at Increased Undersampling

The favourable λ value for CS-MRI image reconstruction with respect to each IQM did not always increase monotonically with R. The R-λ matrices in appendix C.1 have favourable λ at each R marked with red diamonds for convenient visualization. Table 5.2 summarizes instances when optimal λ for each IQM increased monotonically with increasing R for each patient using a total variation (TV) sparsity enforcement for the CS reconstruction. Table 5.3 and Table 5.4 summarize the same results for wavelet and locally low rank (LLR) sparsity enforcements.

Table 5.2: Indication of whether the favourable λ for each IQM increased monotonically with increasing R for each CS-MRI image reconstruction using a TV sparsity enforcement. Of the 75 cases, 23 saw favourable λ increase monotonically with R.

	nRMSE	GMSD	SSIM	MS-SSIM	IW-SSIM
Patient 1	True	False	False	False	False
Patient 2	False	False	False	False	False
Patient 3	False	False	False	False	False
Patient 4	True	False	False	False	False
Patient 5	True	False	False	False	False
Patient 6	False	False	True	True	False
Patient 7	False	False	False	False	False
Patient 8	False	False	False	False	False
Patient 9	True	True	True	True	True
Patient 10	True	False	True	True	True
Patient 11	True	False	True	True	False
Patient 12	False	False	False	False	False
Patient 13	True	False	True	True	False
Patient 14	False	False	True	True	True
Patient 15	False	False	False	False	False

Table 5.3: Indication of whether the favourable λ for each image quality metric increased monotonically with increasing R for each CS-MRI reconstruction using a wavelet sparsity enforcement. Of the 75 cases, 50 saw favourable λ increase monotonically with R.

	nRMSE	GMSD	SSIM	MS-SSIM	IW-SSIM
Patient 1	True	False	True	True	True
Patient 2	True	False	True	True	False
Patient 3	False	False	False	False	False
Patient 4	True	False	False	True	True
Patient 5	True	False	True	True	True
Patient 6	True	True	True	True	True
Patient 7	True	False	False	False	False
Patient 8	True	False	True	True	True
Patient 9	True	False	True	True	True
Patient 10	True	False	True	True	False
Patient 11	True	False	True	True	True
Patient 12	True	False	False	False	False
Patient 13	True	True	True	True	True
Patient 14	True	True	True	True	True
Patient 15	True	False	True	True	True

Table 5.4: Indication of whether the favourable λ for each IQM increased monotonically with increasing R for each CS-MRI reconstruction using a LLR sparsity enforcement. Of the 75 cases, 50 saw favourable λ increase monotonically with R.

	nRMSE	GMSD	SSIM	MS-SSIM	IW-SSIM
Patient 1	True	True	True	True	True
Patient 2	True	False	True	True	True
Patient 3	False	False	False	False	False
Patient 4	True	False	True	True	True
Patient 5	True	False	True	True	True
Patient 6	True	False	True	True	True
Patient 7	False	False	True	True	True
Patient 8	False	False	False	False	False
Patient 9	True	True	True	True	True
Patient 10	True	False	True	True	True
Patient 11	True	True	True	True	True
Patient 12	False	False	True	True	True
Patient 13	True	False	True	True	True
Patient 14	True	True	True	True	True
Patient 15	False	False	False	False	False

5.4.3 Similar IQM Scores and Perceptual Equality

Visual properties of image reconstructions using different sparsity enforcements, though possessing similar objective IQM scores, were noticeably different even to a non-expert observer. Figure 5.1 shows exemplar results of reconstructions using each of the three sparsity transformations that have similar IW-SSIM scores. Though all of the reconstructed images shown in Figure 5.1 have similar IW-SSIM score, the visual properties between the images are noticeably different. Reconstruction using a wavelet sparsity enforcement showed more of a "grainy" degradation than reconstructions using a TV or an LLR sparsity enforcement, which produced images that appeared to have more of a "blurred" degradation. This trend was common to other patient image reconstructions and to the other IQMs.

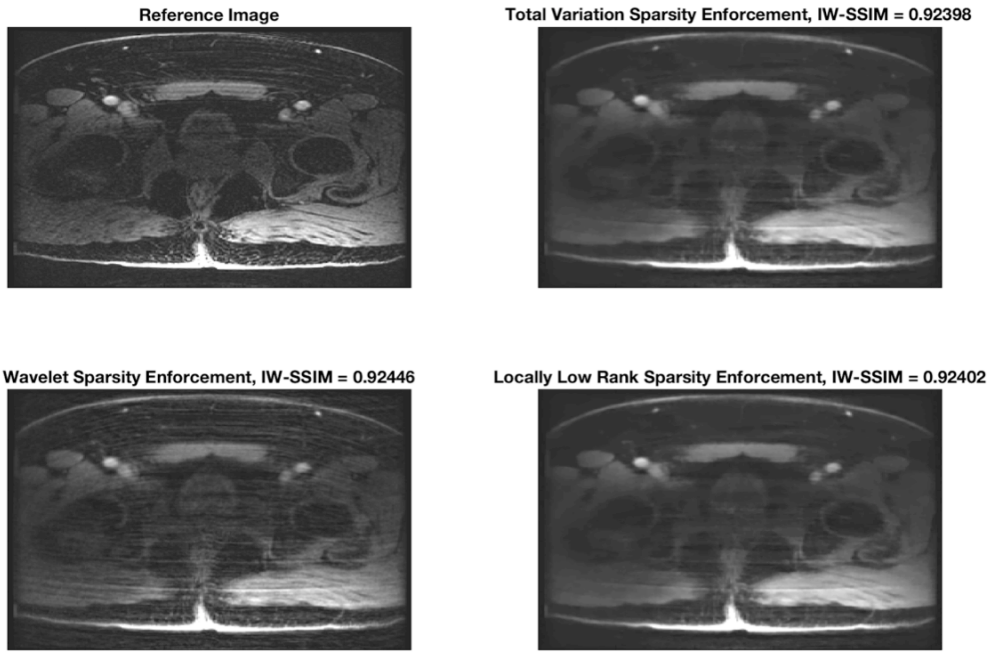

Figure 5.1: Example CS-MRI image reconstructions with similar IW-SSIM score using each of the three sparsity enforcements. Reconstruction using a wavelet sparsity enforcement showed noticeably different artifacts than TV or LLR transforms, despite having similar IW-SSIM score.

5.5 Discussion

5.5.1 Inter-Patient IQM Score Variability

It was hypothesized that IQM scores produced by the GMSD, SSIM, MS-SSIM and IW-SSIM would have less variance between patient image reconstructions than nRMSE scores. This hypothesis was motivated by the design of each of the alternative IQMs. The GMSD quantifies loss of image quality based on findings that the gradient of an image can capture local image structure, and on the assumption that the global variation of local image structure degradation can reflect the degradation of overall image quality. The SSIM family of IQMs all assume that the human visual system (HVS) is highly adapted for extraction of structural features in the image domain; the MS-SSIM includes extra consideration of the sensitive of the HVS to distortions

of different scales, while the IW-SSIM further includes consideration of the ability of the HVS to extract statistical information from the image. The RMSE does not consider any of these factors, and has been found to be sensitive to properties of image artifacts that might otherwise be imperceptible to a human observer[15]. As the undersampling of k-space increased it was anticipated that the structural degradation in the image reconstructions through incoherent aliases would be uniform across patients compared to the pixel-wise introduction of error across patients, and thus that the four alternative IQMs would show more uniform scoring between patient image reconstructions than the nRMSE at each combination of R and λ.

Analyzing the distribution of R-λ matrices produced with each sparsity enforcement and each IQM via the nRMSD from the corresponding sample mean showed that the GMSD had highest relative variance between patient images, followed by the nRMSE and then the SSIM IQM family. The relative similarity between R-λ matrices for each IQM may also be inferred through careful observation of Figures C.1 through C.15 in appendix C.1. These results partially support hypothesis 1, with the notable exception that the GMSD showed higher inter-patient variability than the RMSE.

The low variance between SSIM, MS-SSIM and IW-SSIM metric scores across each of the 15 patient images studied at different combinations of undersampling, sparsity enforcement, and regularization weight suggests that these IQMs may have utility in predicting image properties in a larger setting. Assuming that the underlying hypothesis of the SSIM family of IQMs is accurate (i.e. that they all quantify degradation to image structure that is salient to the HVS), low variability of these IQMs patient image reconstructions implies that a quantifiable loss of structure can potentially be predicted for a given set of image reconstruction conditions. This implication can be taken even further and a hypothesis made that CS-MRI introduces a common class of structural distortions to all patient images that may be statistically modelled. The statistics of this structural distortion model could be studied and used to design a CS-MRI specific no-reference IQM. There is therefore two avenues for future work that may be taken with regards to the SSIM, MS-SSIM and IW-SSIM:

1. Correlate the SSIM, MS-SSIM, and IW-SSIM scores with radiologist subjective scoring of the quality of CS-MRI image reconstructions. Only then can the

assumption that each of these three IQMs is responding to perceptually relevant structural degradations be validated.

2. Motivated by the low inter-patient variance in IQM scores, a statistical model of CS-MRI structural distortion may be proposed and investigated. Knowledge of the statistical changes in image properties under the effect of CS-MRI could permit the construction of a CS-MRI specific no-reference IQM.

Contrary to hypothesis 1, the GMSD scores showed higher variance between patients than the nRMSE. Inspection of Figures C.4 to C.6 show that much of the variance is attributable to differences in GMSD scores across patients at higher values of R and λ for all three sparsity enforcements. The introduction of incoherent aliasing artifacts is dependent on the content of the sampled k-space data and can appear very much like an additive noise structure, especially at high undersampling factors. Increased regularization weight further manipulates the properties of the reconstructed image and the appearance of the incoherent aliasing artifacts. The image gradient domain has been found to be very sensitive to the introduction of noise[34]. Highly incoherent aliasing artifacts approximately appear like an additive noise structure; it is thus not unreasonable that image-dependent artifacts which are known to invoke a sensitive response in the gradient of an image cause wider variation in GMSD scores across different images. Furthermore the GMSD evaluates gradient similarity on a voxel-by-voxel basis and considers nearby image structure only during generation of the gradient maps. In this respect, the GMSD is similar to the RMSE in that it is effectively blind to the local structure of the domain within which it operates. The sensitive nature of the GMSD with respect to noiselike artifacts, combined with its lack of awareness of the nearby voxels during local quality calculation in the gradient similarity map, make the GMSD less consistent between patient image reconstructions.

The fact that the SSIM, MS-SSIM and IW-SSIM show lower inter-patient variability than the GMSD suggests that the effect of CS-MRI on the voxel-wise gradient magnitude of the image is more dependent on the content of the image than is the effect on perceived image structural content. Future work must still determine the ability of the GMSD to correlate with radiologist perception, though the GMSD will not possess the same reliability as the SSIM family of IQMs in predicting a common

loss of image structure *between* patient images. If the GMSD is found to correlate well with radiologist perception, however, its advantage over the SSIM family of IQMs comes through significantly faster execution time due to its relative computational simplicity (i.e. fewer local window calculations). For example, whereas the SSIM takes on the order of 16 seconds to evaluate a 3D image, the GMSD takes on the order of 0.4 seconds. The gradient-based SSIM[45] may also be investigated in the future as an alternative to the GMSD for a gradient-based image quality metric with low inter-patient variability, although presumably at a higher computational cost compared to the GMSD.

With regards to quantitative dynamic CS-MRI applications, objective IQMs with lower inter-patient variability will allow for more consistent predictive capabilities. If correlations between quantitative mapping accuracy and objective IQM scores are found to exist, determining objective IQM score thresholds for a desired quantitative mapping accuracy will be more tractable if the IQM scores are relatively consistent across patient image reconstructions. The results of this section would thus suggest that the SSIM family of objective IQMs may provide better performance in predicting quantitative mapping performance in dynamic applications across a wide range of patients.

In the pursuit of expediency, the objective IQMs are intended to be applied in analysing images a whole. This saves the time of having to manually select a region of the images for objective quality analysis. However, a radiologist may desire to assess the objective quality of a smaller subset of the image (e.g. the prostate in a pelvic MRI acquisition), since they may be concerned primarily with the quality of the image within the specified subset. We predict that the relative inter-patient variabilities of the objective IQMs we've observed when analysing an entire MRI image would similarly exist if only a subset of the image were analysed. In a broad sense, the pelvic MRI images we studied are all quite similar; patients may roughly be approximated as equivalent, all containing very similar anatomical features. This is true of a 3D pelvic image, as well as for a 3D prostate image. Additionally, the incoherent distortions present in CS-MRI are roughly uniformly distributed through the reconstructed images. Given that patient anatomies are roughly consistent when

analysis similar portions of the body, and that the distortions introduced by CS-MRI are roughly uniformly distributed throughout the reconstructed images, we do not anticipate significant changes in the observed inter-patient variability of the the objective IQMs if future work investigated the effects of selecting only a subset of an image for objective image quality assessment.

Limitations of our analysis must be addressed. In assessing the inter-patient variability of the objective IQMs under similar reconstruction conditions, we assumed that some semblance of the MRI images were discernible (i.e. that we were not including what amounted to pure incoherent aliasing artifacts in our analysis) and that no unexpected distortions were present in the image reconstructions. In the former case, the accidental inclusion of image reconstructions that had no obvious anatomical image content would bias the results, since objective IQM scores would be calculated by comparing the reference image to what is effectively random noise. Visual inspection of a random selection of image reconstructions confirmed that this did not occur. In the latter case, the presence of consistent erroneous artifacts (i.e. artifacts that arising from mistakes during data acquisition or image reconstruction, as opposed to the artifacts expected to remain from reconstruction of undersampled data) may affect the inter-patient variability of the objective IQM scores. For example, in the trivial case that all image reconstructions were made exactly equivalent to one another the objective IQM scores would have been identical and the RMSD results reported in section 5.4.1 would indicate no difference in the inter-patient variability of the IQMs. Visual inspection of a random selection of image reconstructions confirmed that no trivial errors in image reconstruction occurred.

Our analysis may also be susceptible to consistent erroneous artifacts that aren't reflective of true CS-MRI image reconstruction artifacts. For example, by artificially zeroing a subsection measuring $40 \times 40 \times 6$ voxels in the center of each of the 15 patient MRI images, we similarly find the the SSIM family of IQMs demonstrate lower relative inter-patient variability than either the RMSE or the GMSD. This may be problematic given that a loss of the central features of the image may render the data diagnostically useless, and hence knowledge of the relative variability of the IQMs useless. Thus observer intervention is still required intermittently to ensure that the image reconstruction pipeline is performing as expected, and that the image artifacts

being assessed with objective IQMs are reflective of the anticipated reconstruction process.

5.5.2 Favourable Sparse Regularization Weights at Increased Undersampling

Favourable regularization weights in past studies have typically been chosen empirically, usually by method of a radiologist observing several image reconstructions and choosing the value of λ that resulted in the best perceptual image quality at a single undersampling factor. It has often been assumed that the regularization weight should increase as undersampling increases to offset the introduction of increasingly severe incoherent aliasing artifacts. For example, Feng et. al.[26] have suggested that after a reference regularization weight λ_{ref} is established at a reference undersampling factor R_{ref} by an experienced radiologist, the value of λ to be used for other undersampling factors should vary proportionally with the change in R:

$$\lambda \propto \left(\frac{R}{R_{ref}}\right) \lambda_{ref} \tag{5.2}$$

It should be noted that Feng et. al. do acknowledge that equation 5.2 is only a rough guideline. Nonetheless, it is demonstrative of the assumptions that must be made in CS-MRI without an objective method for tuning the CS reconstruction parameters.

Contrary to hypothesis 2, and to assumptions frequently made in CS-MRI, our results indicated that the favourable λ with respect to each IQM does not necessarily increase monotonically with increasing R. This holds true for all five of the IQMs studied and for each of the three sparsity enforcements (see Tables 5.2 to 5.4). For some patients a monotonic increase in preferential λ is observed as R increases, whereas for others it is not. We chose this strict analysis as a means of verifying the assumptions that are commonly made in CS-MRI, though an argument can be made that a "tolerance" should be allowed when discussing whether or not a given patient saw monotonic increase in preferential λ as R increased. For example, patient 4 in Figure C.8 did not technically exhibit a monotonic increase in preferential λ with increasing R, though the increase in λ was *roughly* monotonic with R. However, we still observed that some patient image reconstructions distinctly broke from the assumption of monotonically increasing λ as R increased (e.g. patient 3 in Figure C.8),

providing confidence in our rejection of hypothesis 2. It should also be noted that some patient results in the figures shown in section C.1 demonstrated a "saturation" in preferential λ at higher R values (i.e. the red diamonds all stick to the maximum value of λ studied past a certain value of R, such as in patient 12 in Figure C.8). This is likely an artifact of the discrete selection of λ values we were forced to make, and finer sampling of λ values or extended range in λ values would aid in further distinguishing the favourable λ for the affected patient $R - \lambda$ matrices. Due to the fact that many patients do not exhibit this saturation in favourable λ, we expect similar results regarding monotonicity to hold.

These results underline the need for an objective predictor of image quality that has been well characterized. It is unlikely that a value of λ found to be favourable for the reconstruction of CS-MRI images for one patient should be used for the CS-MRI image reconstruction of all other patient images, especially if the sparse enforcement used is changed. However, the time required to bring in radiologists for the perception-based determination of an appropriate λ on a patient-by-patient basis and for difference sparse enforcement methods is infeasible. Future work correlating these IQMs with radiologist subjective scoring could, if good correlation is found, provide tools that allow for the expedient selection of favourable CS reconstruction parameters on a patient-by-patient basis.

Knowledge that favourable regularization weight does not necessarily increase monotonically with increased R also benefits quantitative dynamic MRI, where assumptions along the lines of equation 5.2 would otherwise similarly be applied in reconstructing the dynamic image series. The compromise between image quality and temporal resolution is made more efficient through the selection of favourable CS reconstruction parameters. With knowledge that larger R does not necessarily require larger λ, dynamic image series may be reconstructed with higher objective quality at a given temporal resolution, on a patient-by-patient basis. Future work that correlates the objective IQM scores with radiologist subjective scores with thus aid quantitative dynamic CS-MRI.

5.5.3 Similar IQM Scores and Perceptual Equality

One of the main advantages of each of the GMSD, SSIM, MS-SSIM, and IW-SSIM is that they are distortion nonspecific; they are designed to capture a wide range of common distortions and quantify the perceptual impact of those distortions. However, this means that similar IQM scores can be obtained for distorted images that are noticeably different in visual properties.

The visual properties of CS-MRI image reconstructions are dependent upon the nature of the sparse enforcement used. Figure 5.1 shows representative results obtained with each of the three sparsifying transforms having similar IW-SSIM score. For all five IQMs it was observed that although the image reconstructions had similar IQM score, image reconstructions using a wavelet sparsity enforcement appeared to possess a "grainy" degradation not present when reconstructions were made with either the TV or the LLR sparsity enforcement. The wavelet sparsity enforcement seemed to result in images with less of a blurred quality to them however. These results demonstrate that future studies of the correlation of objective IQM scores and subjective radiologist scores must consider the sparsifying transform(s) used for CS-MRI reconstructions. The IQMs may be found to correlate better with radiologist perception under certain classes of sparsifying transforms than under others, and certain combinations of sparsifying transforms may be found to complement each other in a way that drastically improves perceptual quality of the image reconstructions. For example, Figure 5.1 might suggest pairing a wavelet sparsity enforcement with a mild TV sparsity enforcement to remove the "grainy" aspect of the reconstruction while taking advantage of any potentially desirable qualities of a wavelet transform. Indeed, this combination is frequently seen in the CS-MRI literature. Characterization of objective IQMs can provide a quantitative means of justifying such sparse enforcement combinations.

It is worth considering whether this phenomena would be observed if a subset of an MRI image is selected to be analysed for objective image quality. The artifacts present in CS-MRI images with equivalent IQM scores are characteristic of the sparsity enforcement used, and thus a similar behaviour would be observed even if assessing only a subset of the images. For example, if objective IQM scores were

calculated only for the prostate portion of a pelvic MRI image using the total variation, wavelet, and locally low rank sparsity enforcements in a CS-MRI reconstruction, prostate images that had the same objective IQM scores would still demonstrate artifacts with qualities characteristic of the sparsity enforcement used (e.g. wavelet sparsity enforcements would still impart a "grainy" quality to the image). Thus future work must be concerned with the potential impacts of CS-MRI based artifacts that might remain in a region of interest.

These results may also pertain to quantitative dynamic MRI outcomes. Although the enforcement of different sparsity enforcements in CS-MRI reconstructions may result in images with equivalent IQM scores, the degradations present in the images may prove worse for some sparsity enforcements than for others. For example, referring again to Figure 5.1, enforcement of only a spatial wavelet enforcement during CS-MRI reconstruction of dynamic MRI data may introduce too much noise to the extracted time series and degrade the accuracy of the quantitative mapping relative to a combination of sparsity enforcements. The development of a no-reference CS-MRI specific IQM, as discussed in section 5.5.1, may aid in distinguishing the presence of distortions that would otherwise be deemed equivalent by IQMs such as the SSIM family.

5.6 Conclusions

Our goal in this section was to study the quantitative properties of the RMSE and four alternative objective IQMs in order to obtain foundational knowledge that can inform the application of objective IQMs to quantitative dynamic CS-MRI. Additionally, we intended to provide a base from which future work investigating correlations between objective IQM scores with subjective radiologist scores may draw. To accomplish this we have characterized the inter-patient variability of scores from each IQM, assessed the change in favourable sparse regularization weight of each IQM, and studied whether similar objective IQM scores imply unique image degradations

We demonstrated that the SSIM, MS-SSIM and IW-SSIM show lower relative inter-patient variability across a wide range of undersampling factors and sparsity regularization weights compared to the RMSE, and that the GMSD shows *higher* inter-patient variability under the same circumstances. This motivates future work

seeking common structural degradations induced by CS-MRI, the properties of which may allow development of a CS-MRI specific no-reference IQM. The low inter-patient variability of the SSIM family of IQMs would also suggest that the SSIM family of IQMs may provide better performance in predicting quantitative mapping performance in dynamic applications across many patients, due to less variability in any thresholds that future work may establish.

We demonstrated that, contrary to assumptions often made in the field of CS-MRI, it may not be appropriate to increase the sparse regularization weight monotonically with increasing undersampling factor. This is demonstrates the need for future work correlating objective IQM scores with radiologist subjective scoring, so that an objective tool for choosing favourable CS reconstruction parameters may be implemented. Such knowledge would also benefit quantitative dynamic MRI. The compromise between image quality and temporal resolution is made more efficient through the selection of favourable CS reconstruction parameters, obtaining higher image quality for a given undersampling factor.

Finally, we found that the properties of the degradations of images with similar objective IQM scores are not necessarily equivalent, and depend on the sparsity enforcement used during CS reconstruction. Although the IQMs may report equivalent scores, the degradations present under the enforcement of different sparsity enforcements may be found to affect the accuracy of quantitative dynamic MRI techniques in different ways. The development of a no-reference CS-MRI specific IQM may aid in distinguishing the presence of distortions that would otherwise be deemed equivalent by IQMs such as the SSIM family, which in turn may assist in improving the accuracy of quantitative mapping techniques.

The outcome of this chapter has been the development of foundational knowledge necessary to motivate future studies of objective IQMs in CS-MRI, as well as to assist in the implementation of objective IQMs in characterizing quantitative dynamic MRI techniques.

Chapter 6

Simulations of a Temporally Dynamic Synthetic Phantom

6.1 Purpose of Investigation

As advancements in MRI image reconstruction techniques, such as compressed sensing (CS), begin to permit greater temporal resolutions while maintaining higher image quality compared to conventional techniques, the potential for accurate quantitative dynamic MRI applications is growing. Methods for characterizing the accuracy of these quantitative dynamic techniques are currently lacking however. We propose that objective image quality metrics (IQMs) have utility in characterizing the accuracy of quantitative dynamic MRI.

As an exemplar of a dynamic application with quantitative mapping capabilities, we chose to design a simulation investigating a simple dynamic contrast enhancing synthetic phantom. The synthetic phantom contains features that evolve dynamically in intensity according to simple, closed-form models. With knowledge of "ground truth" quantitative parameters describing the dynamic evolution of the synthetic phantom features, we may investigate the correlations between objective IQM scores and quantitative parameter recovery accuracy. In this chapter we present a series of results from simple simulations that validate the expected performance of our simulation framework, allowing it to serve as a valuable tool for future work. We also demonstrate evidence of correlations between objective IQM scores and quantitative parameter recovery accuracy.

In this chapter we present three simple experiments. We begin by investigating the changing regularization needs for obtaining the most favourable objective image quality at increasing undersampling factors, comparing the results broadly to trends observed in chapter 5 to validate the behaviour of the simulation framework against data obtained from actual MRI image data. We then assess how changes in undersampling factor affect the recovery accuracy for quantitative parameters, demonstrate that the simulation framework may assess the sparse regularization needs for most

favourable quantitative mapping accuracy, and assess whether feature size affects the accuracy of quantitative parameter recovery accuracy. Finally, we investigate the correlations between objective IQM scores and the accuracy of recovered quantitative parameters under simple, well understood conditions. These simple conditions are meant to emphasize the affect of image quality, as opposed to temporal resolution, on the accuracy of quantitative mapping.

The goals of this chapter were to validate the simulation framework as a functioning tool for future work and to demonstrate proof of concept evidence of correlations between objective IQM scores and the accuracy of quantitative mapping in dynamic CS-MRI. These results are intended to serve as a proof of concept to motivate novel research into characterizing the accuracy of quantitative mapping in dynamic CS-MRI. Additionally, our intention was to validate the performance of our simulation framework under simple conditions, such that future work may use it as a tool for extensive investigation of correlations between quantitative mapping accuracy and objective IQM scores.

6.2 Hypotheses

We hypothesize the following results from our simple set of simulations:

1. Sparse regularization of fully sampled k-space data will always degrade image quality. Evidence of changing CS-MRI regularization needs for favourable image reconstruction performance, as quantified by objective IQMs, will exist as undersampling factor increases.

2. With regards to the accuracy of recovered quantitative parameters:

   (a) Recovered quantitative parameters related to the scale of the temporal evolution (e.g. the initial magnitude of an exponential decay) will show larger percent error than recovered quantitative parameters related to the rate of temporal evolution (e.g. the decay constant of an exponential decay) as undersampling factor increases.

   (b) The accuracy of recovered quantitative parameters in smaller features will not be significantly worse than the accuracy of quantitative parameters recovered in larger features.

3. There will evidence of correlations between quantitative parameter recovery accuracy and image quality as quantified by objective IQMs.

6.3 Methods

To validate the performance of our newly developed simulation framework, as well as to seek preliminary evidence of correlations between objective IQM scores and quantitative mapping accuracy, we have designed and carried out three experiments. All of the following simulations were performed with noiseless "snapshot sampling" of the simulated k-space data of a temporally evolving synthetic phantom using a single uniform coil (see section 6.3.1), enforcing a spatial wavelet sparsity regularization during the CS reconstructions. While the simulation framework is capable of a more detailed set of simulations, these simple conditions allow a confident validation of the simulation framework performance while still allowing proof of concept results that objective IQMs can characterize quantitative dynamic MRI performance. Use of snapshot sampling at fixed temporal resolution will emphasize the effect of image quality on quantitative mapping performance.

6.3.1 Simulation Summary

A more detailed overview of the simulation framework may be found in appendix D.1. We present a brief summary of the relevant pieces of the simulation framework in this section.

The simulation framework incorporates a synthetic phantom containing a series of embedded features that evolve dynamically in time. The features, pictured in Figure 6.1, include three planes of embedded cylindrical features and two planes of "pin grids". For the purposes of this thesis, we made use of two sets of embedded cylindrical features with decreasing diameters, measuring 19, 15, 11 and 7 voxels across. One set of embedded cylindrical features decayed exponentially in intensity while another oscillated sinusoidally over time. The models describing the change in feature intensity over time for the exponentially decaying features and the sinusoidally evolving features are given by equations 6.1 and 6.2 respectively, where $I(t)$ is the

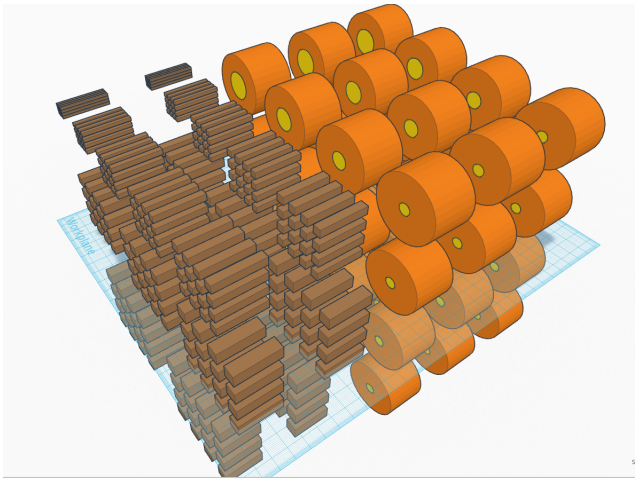

Figure 6.1: A 3D rendering of the features in the synthetic phantom. These features sit inside a rectangular phantom body.

signal intensity at time t and A, B, C, and D are user specified model parameters.

$$I(t) = Ae^{-t/B} + C \tag{6.1}$$

$$I(t) = A \sin\left(\frac{2\pi}{B}\left[t - C\right]\right) + D \tag{6.2}$$

We chose to implement "snaphot sampling" of k-space. As opposed to a realistic k-space acquisition which samples a new frequency encoding line every TR seconds, snapshot sampling assumes that the entire k-space sampling trajectory is sampled instantaneously (i.e. TR = 0 seconds) at each point in time and allows for a user specified lag between snapshot samples to allow the synthetic phantom to evolve in intensity. We implemented snapshot sampling of CIRCUS trajectories designed with $b = 40$ and $c = 1.5$, and a lag between snapshot samples of 4 seconds. The choice of $b = 40$ and $c = 1.5$ was motivated by the results of chapter 4. Although it is not entirely reflective of the compromise that must occur between temporal resolution and the amount of k-space data acquired, we chose to maintain a 4 second temporal lag at all R in order to demonstrate the fundamental relationship between image quality (as quantified by objective IQMs) and quantitative parameter recovery accuracy, as well as to validate the simulation framework under simple conditions. CS-MRI reconstructions were performed using each of the acquired CIRCUS trajectories.

At each simulated image reconstruction time, the IQM scores between the CS-MRI reconstructions and the corresponding true phantom image were calculated. The mean and standard deviation of each IQM score over time in each simulation were calculated for reporting of results in this chapter.

To perform quantitative mapping, time series were extracted on a voxel-by-voxel basis for each set of features across all of the reconstructed images. Least-squares fitting of the models in equations 6.1 and 6.2 to each extracted time course were performed for the appropriate features, recovering quantitative model parameters whose accuracy could be assessed against knowledge of "ground truth".

6.3.2 Effect of Undersampling on Favourable Sparse Regularization Weight

The first experiment demonstrates the ability of the simulation framework to study the changing favourable sparse regularization weight required to obtain relatively high objective image quality as undersampling factor increases. By utilizing snapshot sampling of k-space, we may temporarily ignore the temporal evolution of the phantom features and effectively approximate the snapshot sampled k-space data as a series of static k-space acquisitions for a sample population of phantoms. In this case, we may study the variance in the observed IQM scores and the behaviour of the mean IQM scores as regularization weight and k-space undersampling both vary. This allows broad comparison between any observed simulation trends and the results presented in chapter 5.

A temporally evolving synthetic phantom measuring $200 \times 200 \times 32$ voxels was sampled via snapshot sampling of k-space, as described in section 6.3.1. This was performed for undersampling factors of $R = 1, 1.5, 2, 3, 5, 8$ and 12. CS-MRI reconstructions enforced a spatial wavelet regularization with regularization weights of λ = 0, 0.00125, 0.0025, 0.00375, 0.005, 0.01, 0.0150, 0.02, 0.025, 0.03, 0.04, and 0.05. Objective IQM scores were calculated as described in section 6.3.1.

6.3.3 Accuracy of Recovered Quantitative Parameters

Simulations were carried out as described in section 6.3.2. By considering the temporal evolution of the features of the synthetic phantom, as opposed to approximating

them as a series of static phantom images, we may perform quantitative mapping to recover the model parameters describing the temporal evolution of the features. Quantitative parameters were recovered via voxel-by-voxel least squares fitting, as described in section 6.3.1. Results are presented for the magnitude and time related parameters of each model (i.e. initial magnitude and decay constant for the exponentially decaying features, and amplitude and period for the sinusoidally oscillating features), since the offset parameters C and D in equations 6.1 and 6.2 serve primarily to afford the fitting algorithm a measure of freedom. A set of exponentially decaying embedded cylindrical features with an initial intensity of $0.9(2^{15})$ arbitrary units and a decay constant of 80 seconds, and a set of sinusoidally evolving features with an amplitude of $0.8(2^{15})$ and a period of 80 seconds, were chosen for analysis. The diameters of the embedded cylindrical features in each set were 19, 15, 11, and 7 voxels. Total simulated time length was 400 seconds.

6.3.4 Correlations Between Objective IQMs and Quantitative Mapping Accuracy

Correlations between the objective IQM scores calculated from the results of section 6.3.2 and the error in recovered quantitative parameters from the results of section 6.3.3 were calculated. Pearson linear correlation coefficients and Spearman rank correlation coefficients were calculated.

6.4 Results

6.4.1 Effect of Undersampling on Favourable Regularization Weight

Figure 6.2 shows the effect of increased undersampling factor on the favourable regularization weights for best objective image reconstruction quality, as quantified by each of the objective IQMs. Plotted are the mean IQM scores for an entire simulation, with standard deviations (i.e. each plotted point in Figure 6.2 represents the average IQM score across a single dynamic simulation).

All IQMs indicated a very slight reduction in image quality for regularization of fully sampled data. As undersampling factor increased, all of the IQMs began to demonstrate a preference for nonzero regularization weight; within the standard

deviation of the IQM scores, the preference towards increased regularization weight is not always significant. At the highest undersampling factors studied, all IQMs show less sensitive response to changes in regularization weight. Noteably, the GMSD indicates that low regularization weight may be preferential at high undersampling factors.

The SSIM family of IQMs all exhibited lower variance relative to mean scores than did the nRMSE or the GMSD.

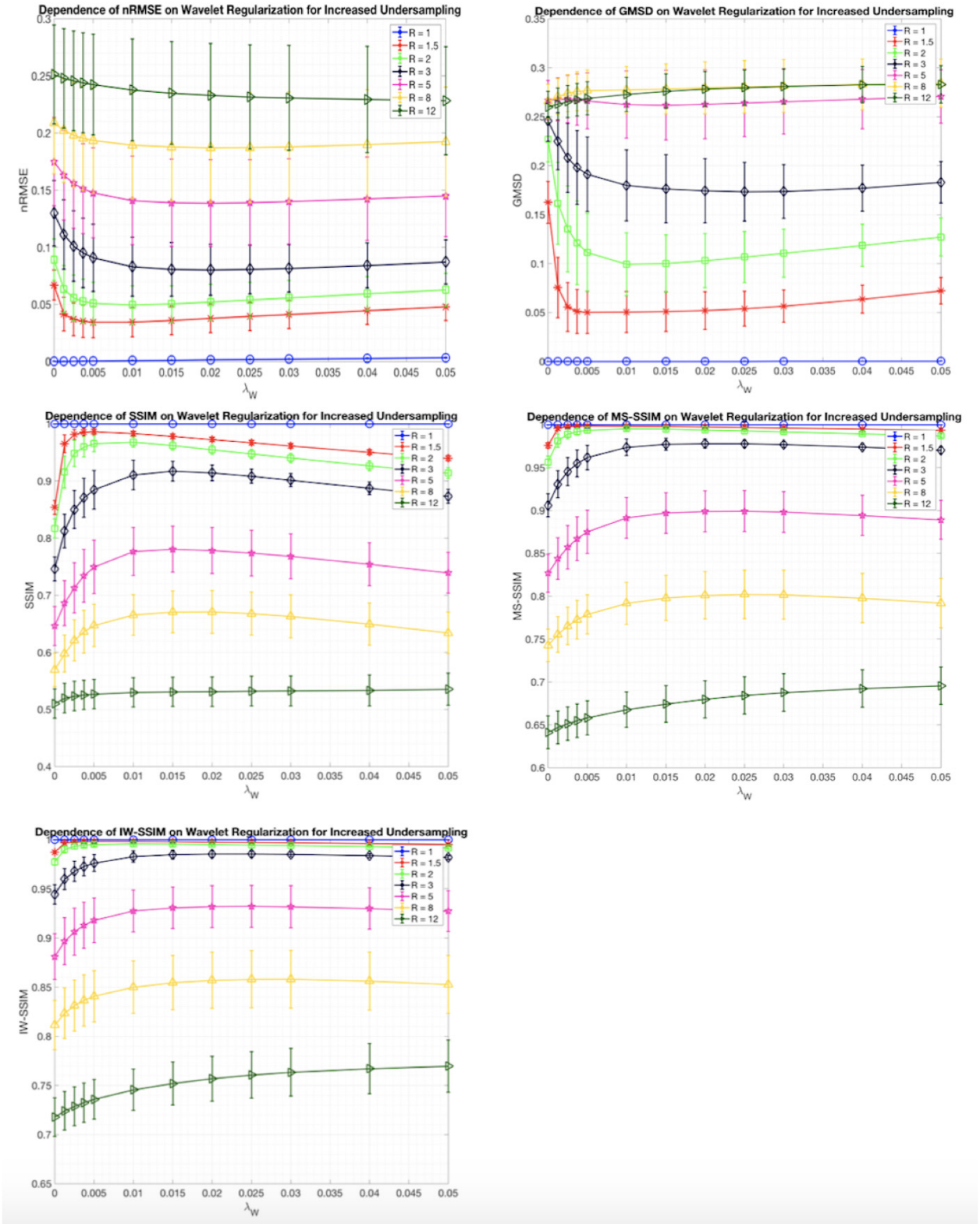

Figure 6.2: Changing regularization needs for favourable objective image quality at increasing undersampling factors.

6.4.2 Accuracy of Recovered Quantitative Parameters

Figure 6.3 shows the accuracy in recovered quantitative parameters for an exponentially decaying set of embedded cylindrical features. Figure 6.4 shows the accuracy in recovered quantitative parameters for a sinusoidally evolving set of embedded cylindrical features. Both plots show mean quantitative parameter accuracies across each feature, with standard deviations. For the 7, 9, 11, and 19 voxel diameter cylindrical features studied, there were 111, 267, 531, and 879 constituent voxels per feature respectively.

For both the exponentially decaying and sinusoidally oscillating features, scale-related parameters showed tendency towards larger error as undersampling factor increased. Evidence of favourable regularization weight for best mean quantitative mapping performance was observed for both scale and temporal evolution parameters, in both the exponentially decaying and sinusoidally evolving models.

Within the bounds of variance, the size of the features was not observed to make a significant difference in the accuracy of the recovered quantitative parameters.

A full representation of the change in quantitative parameter accuracy across all undersampling factors studied may be found in appendix D.2. Figure 6.3 and Figure 6.4 drew example results from the matrices presented in appendix D.2.

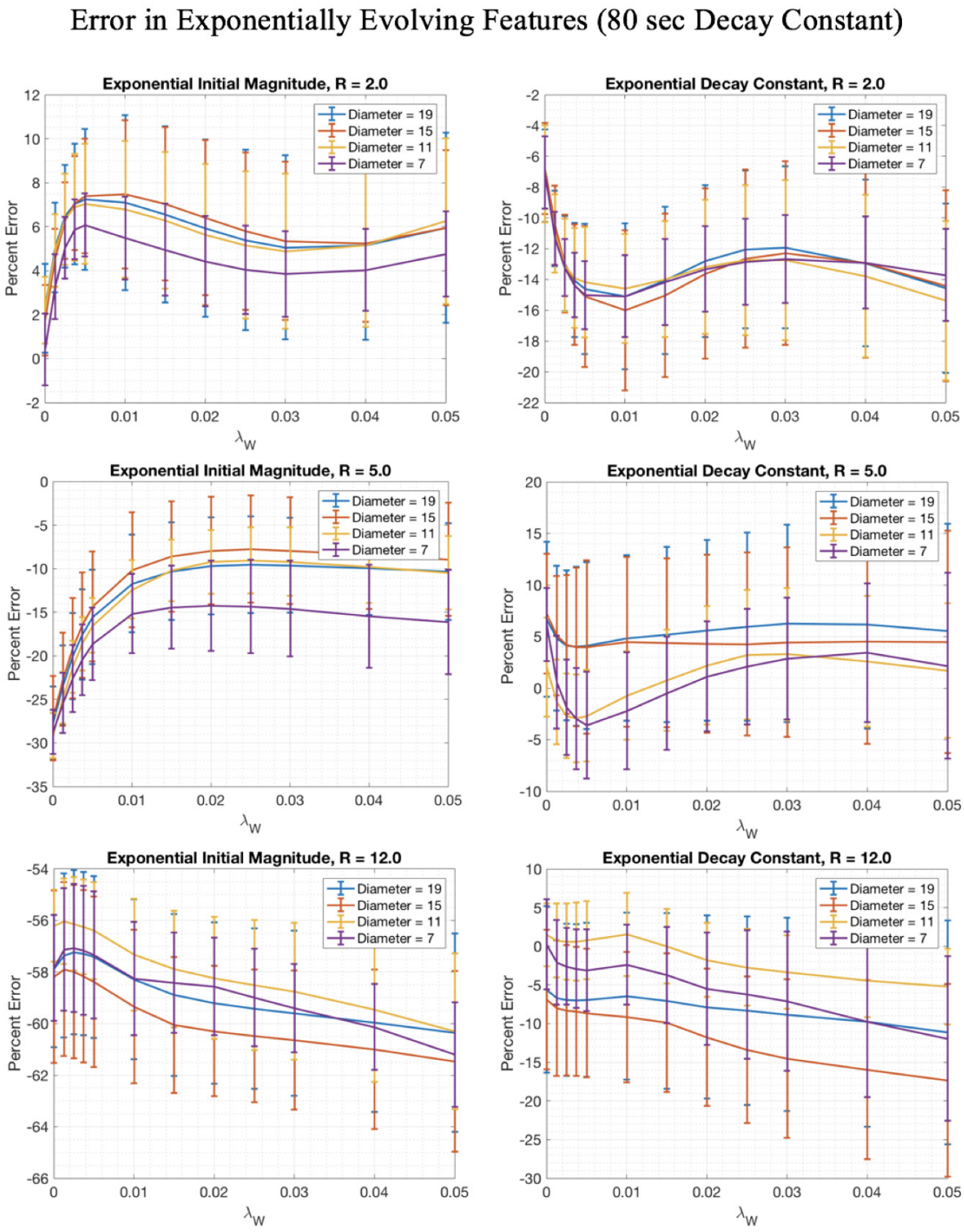

Figure 6.3: Effect of regularization weight on the error in recovered exponential parameters for embedded cylindrical features with decreasing diameter (measured in voxels) as undersampling factor and regularization weight increases.

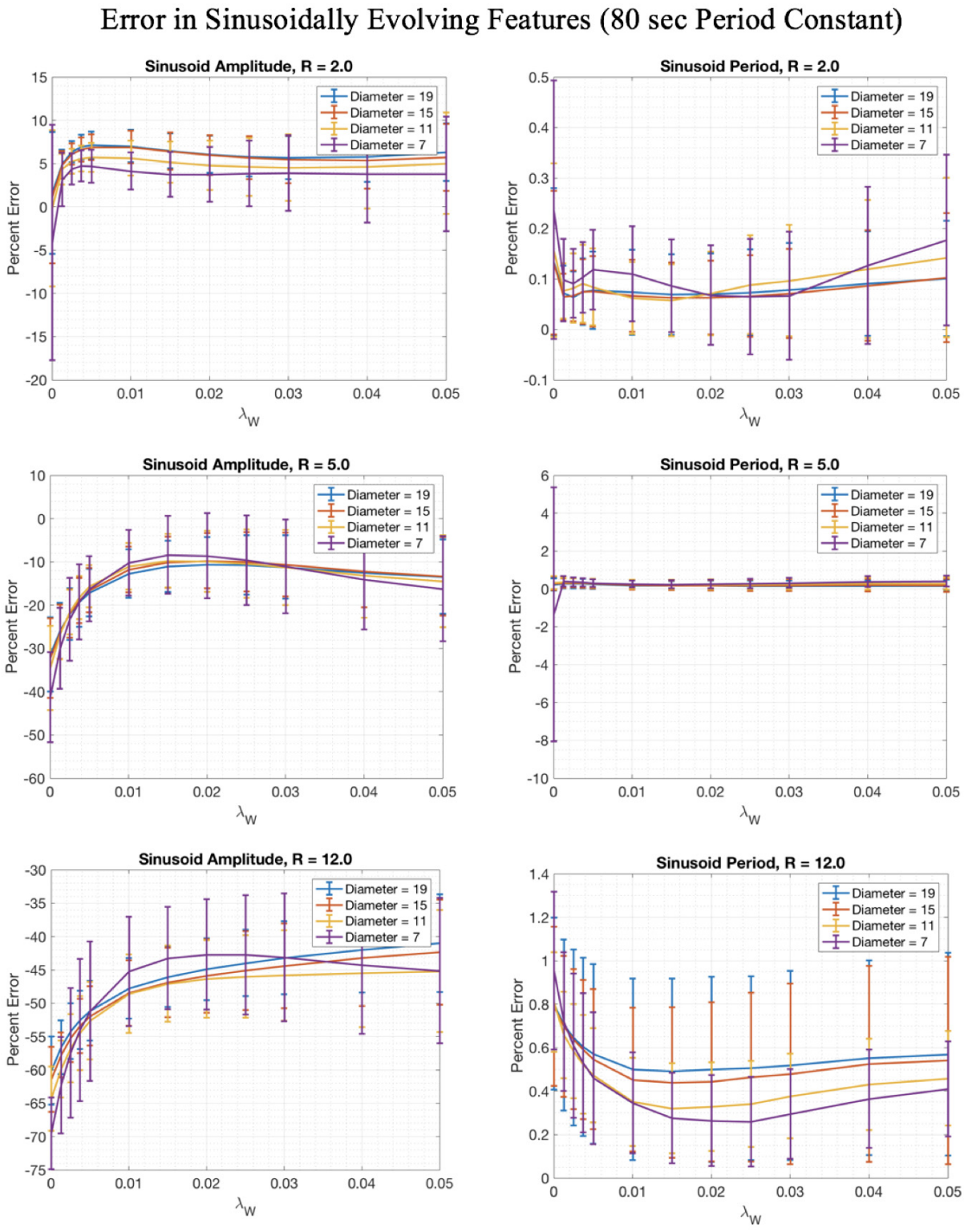

Figure 6.4: Effect of regularization weight on the error in recovered sinusoidal parameters for embedded cylindrical features with decreasing diameter (measured in voxels) as undersampling factor and regularization weight increases.

6.4.3 Correlations Between Objective IQMs and Quantitative Mapping Accuracy

Figure 6.5 demonstrates the relationship between the accuracy of recovered exponential model parameters and the RMSE, GMSD and SSIM. Figure 6.6 demonstrates the same for the sinusoidal model parameters. Results for the MS-SSIM and the IW-SSIM may be found in the appendix.

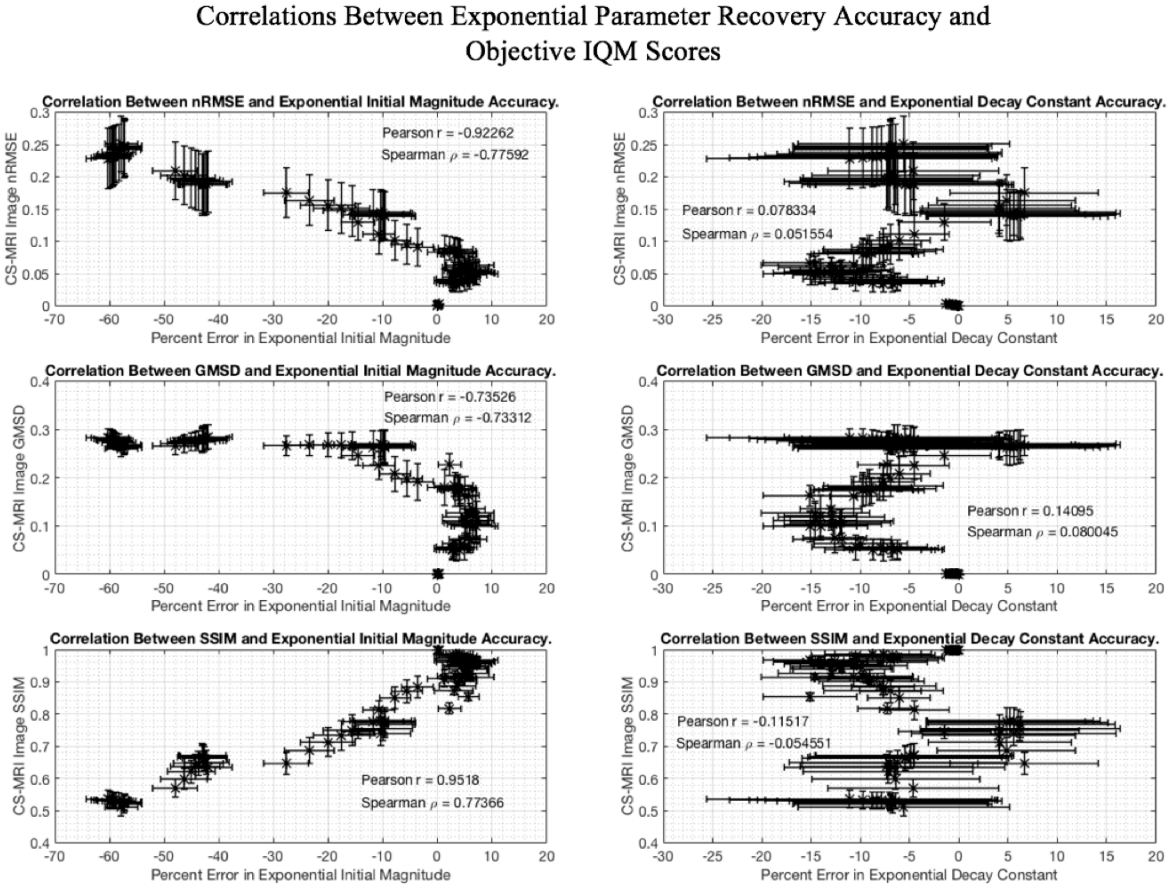

Figure 6.5: Correlation between RMSE, GMSD, and SSIM scores and the accuracy of recovered parameters from exponentially decaying features.

For both the exponentially decaying and the sinusoidally varying features, statistically significant correlations ($p < 0.05$ for both Pearson and Spearman correlations) were observed between the accuracy in recovered scale related parameters and the

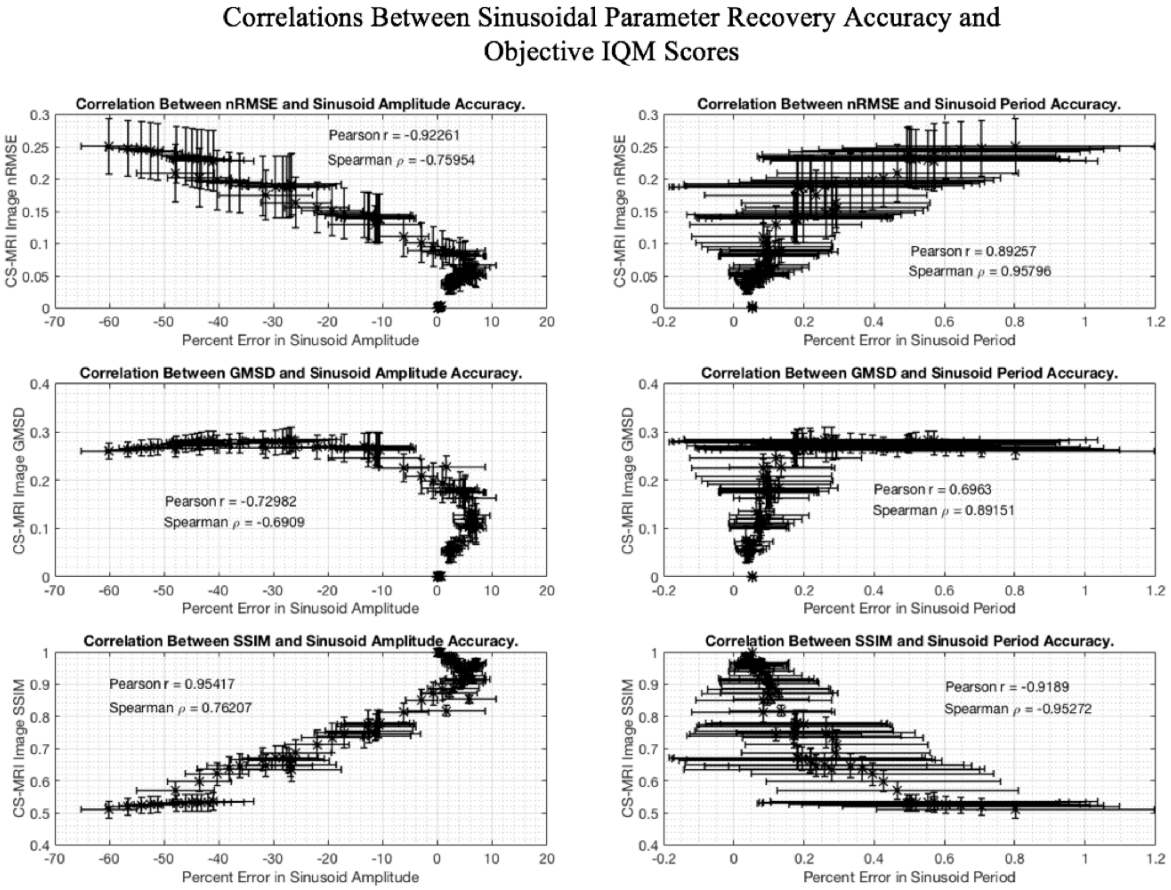

Figure 6.6: Correlation between RMSE, GMSD, and SSIM scores and the accuracy of recovered parameters from sinusoidally varying features.

objective IQM scores. Correlations between accuracy in parameters related to rate of temporal evolution and the IQM scores were statistically significant only for the sinusoidally varying features.

6.5 Discussion

6.5.1 Effect of Undersampling on Favourable Sparse Regularization Weight

Figure 6.2 demonstrates the capability of the simulation framework in studying the effects of undersampling on image quality in dynamic CS-MRI applications, and lend

support for hypothesis 1. As a performance check, simulations with fully sampled k-space (i.e. an undersampling factor of 1) were acquired and subjected to increasing sparse regularization weight. As would be expected, there was a slight degradation in image quality as regularization weight increased, indicated by marginal increases in the RMSE and the GMSD and marginal decreases in the SSIM, MS-SSIM and IW-SSIM. Enforcing sparsity on otherwise fully sampled data results in minor artifacts characteristic to the sparsifying transform; Figure 6.7 shows the effects of wavelet sparsity enforcement on an image with fully sampled k-space data. Through loss of high detail wavelet coefficients there are distortions at feature boundaries, and through an overall loss of power in the wavelet representation there are intensity reductions in the smooth regions of the image. This is a result of CS trying to "de-noise" the fully sampled data in the absence of any incoherent aliasing. In particular, over-regularization of the data in this way causes a reduction of some of the wavelet coefficients through the use of the ℓ_1-norm in the CS-MRI reconstruction algorithm (equation 2.44), which in turn causes the observed compression artifacts.

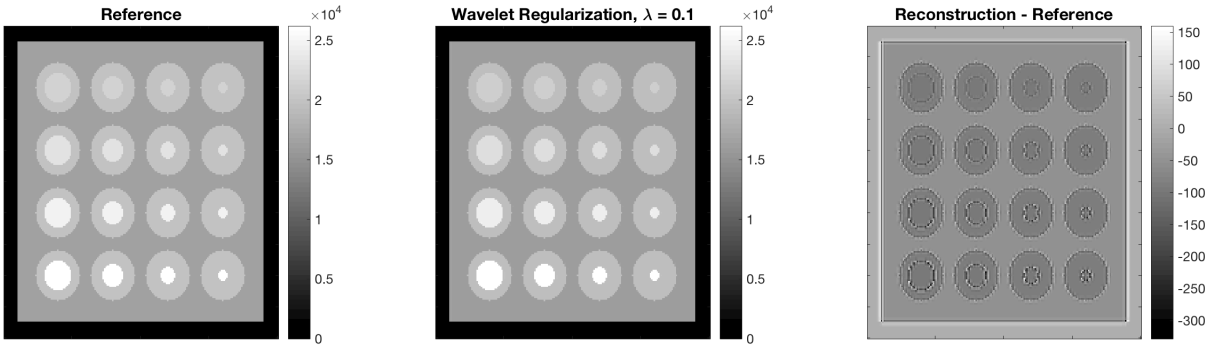

Figure 6.7: Enforcing sparsity in the wavelet domain when a complete representation of the image data is acquired results in compression of the image. Loss of high frequency wavelet coefficients results in boundary artefacts, while overall loss of power in the wavelet coefficients results in loss of signal intensity throughout smooth regions of the image.

As undersampling factor increases, the objective IQMs begin to report different trends in objective image quality. In particular, the SSIM family of IQMs suggests

that image quality becomes distinctly worse as undersampling factor and regularization weight increase, whereas the GMSD would suggest that image reconstruction quality reaches the worst performance for roughly five-times undersampling of k-space. Within the SSIM family of IQMs, the SSIM suggests that appropriate choice of sparsity enforcement at a higher undersampling factor may return better image quality on average than no regularization at a lower undersampling factor (e.g. $R = 5$ and $R = 3$), whereas the MS-SSIM and IW-SSIM would suggest that the ability of CS to recover image quality is not as high for these same conditions. Such considerations may be useful when considering which IQM to use in future work.

It is instructive to broadly compare the results shown in section 6.4.1 with the results of chapter 5. While exact equivalence in the results cannot be expected, given the artificial nature of our synthetic phantom, validation of the simulation framework as a tool for assessing CS-MRI requires that similar trends be observed between simulation results and trends observed in real MRI data. It was observed that simulated reconstructions had less relative variance in the SSIM, MS-SSIM, and IW-SSIM compared to the RMSE or the GMSD scores. This is similar to the results observed in chapter 5, where the SSIM family of IQMs had lower inter-patient variability than the GMSD or the RMSE for the same image reconstruction conditions. Additionally, the SSIM family of IQMs showed a distinct loss of objective image quality as the undersampling factor increased (i.e. no overlap of the resultant IQM score curves within the bounds of variance). This may suggest that a distinct type of structural distortion is being imposed on the image space using CS-MRI, to which the SSIM family of IQMs is more sensitive. This result parallels our conclusion made in chapter 5. Thus, the simulation framework exhibits trends that are similar to those observed in real MRI images in chapter 5.

As undersampling of the k-space data increases, regularization of the image reconstruction in a sparse domain begins to show benefit in terms of image reconstruction quality in many cases. Considering the RMSE, SSIM, MS-SSIM and IW-SSIM, favourable regularization weight for highest mean objective image quality increased monotonically with R. Favourable regularization weight for the highest mean objective image quality did not increase monotonically with regards to the GMSD, with an affinity for lower values of λ demonstrated at undersampling factors above 5. These

results demonstrate the utility of the simulation framework as a tool that future work may use to assess the changing CS reconstruction parameter needs for objective image quality in dynamic CS-MRI. They further demonstrate that correlation of objective IQMs with radiologist subjective scoring is essential for future work, in order to select an appropriate set of objective IQMs. Whereas the SSIM family of IQMs predicts distinctly worse image quality as undersampling factor increases, and an average increase in image quality when enforcing sparsity in a transform domain at these undersampling factors, the GMSD would imply that the worst image quality is rapidly achieved by undersampling k-space and that CS would potentially *hinder* the image quality on average.

6.5.2 Accuracy of Recovered Quantitative Parameters

The strength of a simulation framework is that "ground truth" knowledge of quantitative parameters describing an intensity evolution model makes it possible to investigate the effects of different factors on the recovery accuracy of the quantitative parameters. To validate the performance of our temporally dynamic simulator against known theory, we sought to investigate the accuracy of parameters as they relate to different properties of the signal evolution and to investigate the effect of feature size on recovered quantitative parameter accuracy.

Figures 6.3 and 6.4 demonstrate that the scale-related parameters for both the exponentially decaying features and the sinusoidally oscillating features suffered greater error compared to their respective temporal evolution parameters, validating hypothesis 2a. This is the result of a loss of image energy as the k-space is undersampled, according to Parsevals theorem (equation 2.28), which results in reduced image magnitudes as the undersampling factor increases. Thus is should be expected that scale related parameters suffer inherently larger error as the undersampling factor increases, validating the performance of our simulations. The comparatively low temporal parameter errors are explained by our choice to implement snapshot sampling of k-space. Although a loss of image energy reduces the image magnitude in general, the relative evolution of features between images in the dynamic series is the important factor for temporal parameter recovery accuracy. Since the relative evolution of feature intensity is not as susceptible to the loss of image intensity, the recovery of the quantitative

parameters describing the rate of temporal evolution shows lower error. There will still be effects on the extracted time series due to the presence of incoherent aliasing in the dynamic images, which will affect the accuracy of the temporal evolution parameters. Thus by suppressing the effects of temporal resolution via the use of snapshot sampling of k-space, we obtain the expected behaviour in quantitative parameter recovery accuracy as undersampling factor increases. Had we chosen to implement a realistic sampling of k-space for these results (i.e. a nonzero repetition time), we would expect the "temporal blurring" that affects the k-space acquisitions to begin to affect the temporal dynamics, reducing the accuracy of the temporal evolution parameters at lower undersampling factors.

These results demonstrate the ability of the simulation framework to study not only the accuracy of the recovered quantitative parameters in a dynamic application, but also the effect on recovery accuracy due to the nature of the parameter in the model. Future implementations of the simulation framework incorporating physiologically realistic pharmacokinetic models, such as the extended Tofts model, may study the loss in quantitative parameter accuracy for each of the physiologically motivated parameters due to the effect of CS-MRI.

The results also reveal the ability of the simulation framework to study the changing regularization needs for quantitative accuracy when CS-MRI reconstructions operate with more or less k-space data. The regularization needs differed across undersampling factors and between the exponentially decaying and sinusoidally evolving features. Using our validated simulation framework, future work may examine detailed simulations in order to guide CS reconstruction parameter choice for quantitative dynamic CS-MRI.

Figures 6.3 and 6.4 show no significant difference in parameter error as a result of feature size across all undersampling factors. Previous results have shown that it is the feature *constrast*, rather than the feature size, that primarily determines the ability of CS-MRI to resolve it in the reconstructed image[8]. Because each of the four embedded cylindrical features for both the exponentially decaying and sinusoidally evolving features followed a common evolution, their relative contrasts did not differ at any point in time and so no significant differences in quantitative parameter recovery accuracy were incurred. This supports hypothesis 2b.

6.5.3 Correlations Between Objective IQMs and Quantitative Mapping Accuracy

Figures 6.5 and 6.6 demonstrate the ability of the simulation framework to investigate correlations between image reconstruction quality, as quantified by objective IQMs, and the accuracy of recovered quantitative parameters. The results partially support hypothesis 3, with correlations observed for the recovery of scale-related parameters in the both the exponentially decaying features and in the sinusoidally evolving features, and with the sinusoidal period.

Because quantitative parameters are recovered through voxel-by-voxel fitting of a model to time series data extracted from reconstructed images, the accuracy of recovered parameters is affected both by the temporal resolution of the data *and* the presence of incoherent aliasing artifacts in each of the voxels. Thus, higher image quality is valuable for the recovery of quantitative parameters, lest the time series be contaminated by the image artifacts. Figure 6.8 demonstrates the deterioration of image quality in a slice of the synthetic phantom as undersampling factor increases. As the incoherent aliases contaminate the time series used for quantitative parameter recovery, the accuracy of the recovered parameters will decrease. This is the cause of correlation between objective IQM score and quantitative parameter recovery accuracy.

The results also demonstrate that some IQMs may perform better in characterizing dynamic MRI. The correlation between quantitative mapping performance and GMSD scores was not as strong as for the SSIM family of IQMs or for the RMSE. This is a result of the saturation of GMSD scores at high undersampling factors. Thus, the results of our simple simulation would suggest that the GMSD is not an appropriate IQM for characterizing dynamic MRI.

It is acknowledged that, through the use of snapshot sampling of k-space data, the results presented for correlations between objective IQM scores and recovered quantitative parameter accuracy represent a simplified scenario. This was done to validate the simulation framework against known theory under simple conditions that are easily understood and analysed, and to demonstrate the fundamental relationship between image quality and quantitative mapping accuracy. With the use of realistic k-space sampling (i.e. nonzero TR), a dynamically evolving phantom will result in a

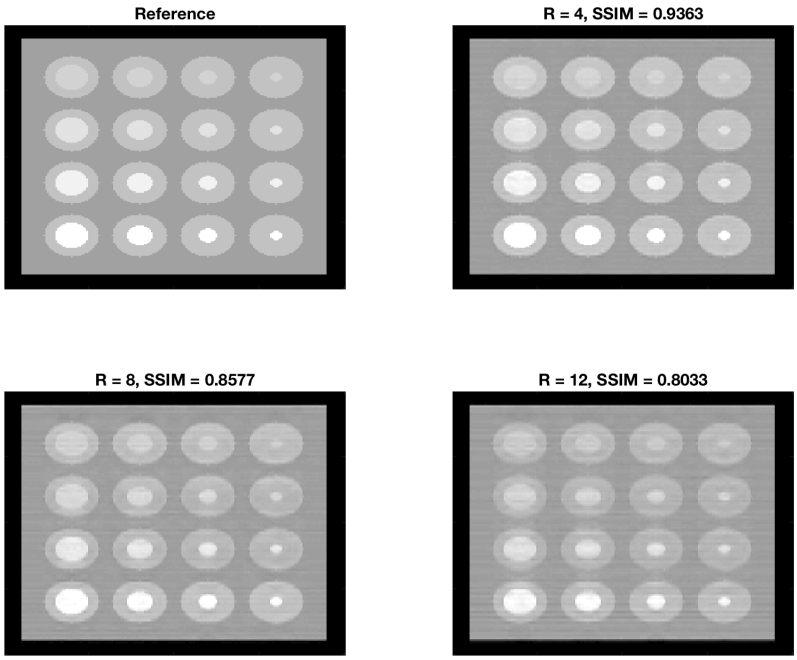

Figure 6.8: As the spatial undersampling increases, the image quality deteriorates (quantified here using the SSIM as an example). Incoherent aliasing artifacts contaminate the extracted timecourse used in model fitting, and so image quality will relate to the accuracy in recovered quantitative parameters.

dynamically evolving k-space during sampling of a single effective CIRCUS quanta. The "temporal blurring" of acquired k-space data will act to degrade reconstructed phantom image quality with respect to a reference phantom image, with the effect being more pronounced for interleaving of larger amounts of k-space data, for faster temporal dynamics, and for longer TR. However we still anticipate a correlation to exist between image quality and quantitative parameter recovery accuracy due to the voxel-by-voxel fitting process, which relies on high quality image data in order to extract accurate time series for quantitative mapping. Future work may extensively investigate correlations between IQM scores and recovered quantitative parameter accuracy using the simulation framework that we have developed. For example, the

balance between the amount of CIRCUS quanta, sampled at a given TR, and interleaved to reconstruct each image versus the temporal resolution required to achieve a threshold in quantitative mapping accuracy may be thoroughly investigated and characterized using our newly validated tool. As realistic sampling of k-space begins to introduced temporal blurring to the k-space data, correlations in quantitative parameters related to temporal evolution are expected to become more pronounced.

Future work with our simulation framework may incorporate physiologically motivated intensity evolution models, and may investigate correlations between the model parameters and objective IQMs. The flexibility of the simulation framework would further allow anatomically specific phantoms to be developed and studied, such that correlations in specific applications may be studied in detail. Beyond just dynamically contrast enhancing applications, future work may be undertaken in characterizing *any* quantitative dynamic MRI technique with objective IQMs. To the best of our knowledge at the time of writing of this thesis, objective IQMs have not been studied in the context of quantitative mapping accuracy, and thus our results serve as a proof of concept which open doors to a wide range of future investigations.

The ability to utilize objective IQMs as a way to predict the accuracy of recovered quantitative parameters would have profound clinical utility; appropriate thresholds in objective image quality may be established that delineate "acceptable" levels of quantitative parameter error from "unacceptable" levels of error. One method to implement the full-reference IQMs could be to acquire a high quality baseline scan prior to the dynamic scan, permitting retrospective adjustment of image quality relative to the baseline scan as a means to control quantitative parameter accuracy.

6.6 Conclusions

In this chapter we have aimed to validate the performance of our simulation framework in preparation for future investigations, and we additionally strived to motivate the use of objective IQMs in characterizing quantitative mapping in dynamic MRI. We began by investigating the ability of the simulation framework to characterize the changing regularization needs for objective image quality as undersampling factor increased, we studied the accuracy of quantitative parameter recovery as undersampling factor increased and feature size decreased, and we sought correlations between

objective IQM scores and the accuracy of recovered quantitative parameters.

The regularization weight required for favourable mean objective image quality increased monotonically for the RMSE, SSIM, MS-SSIM, and IW-SSIM. The GMSD deviated from this trend, showing preference for reduced regularization weight at higher undersampling factors. These results must be considered against the variance in the results, which prevents a hard conclusion being made regarding the overall monotonic nature of favourable regularization weights at each reconstructed image in the dynamic CS-MRI time series. Nonetheless, the results demonstrate the utility of the simulation framework as a means of assessing the changing regularization needs in dynamic CS-MRI. Additionally, trends observed in the simulated data were broadly consistent with the trends observed in static CS-MRI of real MRI image data as discussed in chapter 5. The SSIM family of IQMs demonstrated lower relative variance than the RMSE or the GMSD, reminiscent of the results in chapter 5 that demonstrated lower inter-patient variability for the SSIM family of IQMs. Furthermore, the SSIM family of IQMs indicated a distinct loss of objective image quality as undersampling factor increased, potentially suggesting the presence of a distinct set of structural distortions as a result of CS-MRI. The simulation framework thus exhibited trends that were roughly consistent with those observed in real world MRI data.

The error in recovered quantitative parameters was found to get increasingly larger for scale related parameters as opposed to parameters related to the temporal evolution in both exponentially decaying features and sinusoidally evolving features as the undersampling of k-space increased. This is a result of Parsevals theorem, where a loss of image intensity results from a loss of k-space energy. Additionally, feature size was not found to be a significant factor in determining the quantitative parameter recovery accuracy when features maintained the same relative contrast over time. This is in agreement with results in the CS literature. These results validate that our simulation framework is performing as should be expected, and further demonstrate its utility as a tool for characterizing various effects on the accuracy of recovered quantitative parameters.

Finally, we have proposed the use of objective IQMs for characterizing the accuracy of quantitative mapping in dynamic CS-MRI applications. To the best of our

knowledge, this has not been done in the past. We demonstrated preliminary evidence that objective IQMs are responsive to the inherent relationship between image quality and quantitative mapping accuracy, with statistically significant correlation existing between objective IQM scores and the performance of quantitative mapping.

The outcome of this chapter has been the development of a reliable tool for the assessment of quantitative dynamic CS-MRI, as well as a novel method for characterizing the performance of quantitative mapping techniques. The results we have demonstrated serve as a proof of concept, and motivate wider investigations into the utility of objective IQMs in characterizing dynamic MRI.

Chapter 7

Conclusions

We have presented three studies aimed towards motivating the study and use of objective image quality metrics (IQMs) for the characterization of quantitative dynamic compressed sensing (CS) MRI. To begin, we studied the CIRcular Cartesian UnderSampling (CIRCUS) k-space acquisition scheme designed for CS-MRI in order to obtain knowledge of how best to implement CIRCUS to inherently obtain the most favourable objective image qualities. We then characterized the behaviour of five objective IQMs (the RMSE, GMSD, SSIM, MS-SSIM and IW-SSIM) in static CS-MRI, obtaining foundational knowledge of IQM properties. Our final study utilized simple simulations of a temporally evolving synthetic phantom, taking dynamic contrast enhancing MRI as an exemplar for a quantitative dynamic CS-MRI application, to demonstrate evidence of correlations between IQM scores and quantitative parameter recovery accuracy.

Our first study was of the implementation of the CIRCUS k-space trajectory. In order to obtain maximum efficiency in the compromise between image quality and temporal resolution in dynamic CS-MRI, it is necessary that k-space be sampled in a manner that is agreeable with CS. We explored the CIRCUS parameter space to determine how choices of CIRCUS parameters affected the resulting CIRCUS trajectories, and sought to determine if point spread function (PSF) coherence or incoherence measures would provide a reliable method of determining favourable CIRCUS trajectories. Although there was evidence of regimes of higher PSF incoherence for CIRCUS trajectories with higher undersampling factors designed with $c > 2$, the performance of the CIRCUS patterns (quantified by objective IQMs) did not show similar regimes and the PSF measures were not found to be a reliable predictor of CIRCUS trajectory quality within a fixed undersampling factor. Rather, it was found that most CIRCUS patterns at a fixed undersampling factor perform similarly, and that the most stable performance across undersampling factors occurs provided that

c is chosen to be greater than 1.

Our second study was a characterization of objective IQM properties in static CS-MRI. In order to obtain foundational knowledge required to understand the quantitative behaviour of the objective IQMs, we studied them in retrospective static CS-MRI reconstructions of 15 clinical MRI images. We observed that the SSIM family of IQMs show less inter-patient variability than the RMSE, and that the GMSD shows higher inter-patient variability than the RMSE. This may motivate the description of a structural distortion model of CS-MRI that would allow for future development of a CS-MRI specific no-reference IQM. These results would also suggest that the SSIM family may be better objective IQMs for characterizing the accuracy of quantitative mapping in dynamic CS-MRI, since thresholds could be established with less inter-patient variability affecting the thresholds. In addition to investigating the interpatient variabilities of each IQM, we demonstrated that favourable regularization weight for maintaining high image quality did not necessarily increase monotonically with increasing undersampling factor. This goes against assumptions frequently held in the field of CS-MRI, and motivates the need for a convenient objective assessor of image quality. Finally we showed that CS-MRI images possessing equivalent objective IQM score may not necessarily possess the same visible distortions. This can have implications for quantitative mapping accuracy, where some distortions may affect the accuracy of recovered quantitative parameters differently than others. The development of a CS-MRI specific no-reference IQM may help to address this.

Our third study introduced and validated a simulation framework that we designed for studies of quantitative mapping accuracy in a dynamically enhancing synthetic phantom, with dynamic contrast enhanced MRI taken as an exemplar for dynamic MRI applications. We demonstrated the utility of the simulation framework as a method of characterizing the changing regularization needs for maintaining high objective image quality across a dynamic image series as undersampling factor increased. We then investigated the ability of the simulation framework to study the accuracy of recovered quantitative parameters. Specifically, we demonstrated that CS-MRI will inherently show larger error in scale-related model parameters as undersampling factor increases compared to evolution-related parameters due to expected losses in image

intensity, and we demonstrated that feature size does not significantly affect the accuracy of recovered quantitative parameters for features at equivalent contrasts. These results are consistent with the theory in the literature, validating the performance of our simulation framework and motivating it for future work in physiologically specific applications (e.g. pharmacokinetic modelling). The third study concluded by showing evidence of correlations between objective IQM scores and quantitative parameter recovery accuracy. While the simulations presented were a simplified scenario, the resulting correlations between objective IQM scores and quantitative parameter accuracy are a proof of concept for a novel method of characterizing the accuracy of quantitative mapping in dynamic MRI, and provide motivation for extensive future study of objective IQMs in quantitative dynamic CS-MRI applications.

There are several avenues that future work may take. To begin with, future work may build upon the foundational knowledge of IQM behaviour in static CS-MRI that we have presented in chapter 5. A simple first task would be to repeat the generation of the R-λ matrices with both a finer sampling of λ and an extended range of sampled λ values, to better characterize the change in preferential regularization weights using CS-MRI on a patient-by-patient basis. While we have presented results that characterize the behaviour of the objective IQMs through the independent enforcement of either a total variation (TV), wavelet, or locally low rank (LLR) sparsity models, combinations of these sparsity enforcements must also be investigated. The methods of chapter 5, combined with the suggestions we've just made for future work, should also be investigated in other anatomies (e.g. brain, liver) to determine if similar trends in the objective IQM properties exist across anatomies.

Work must be undertaken in the future to correlate the objective IQMs we have studied with subjective radiologist scoring. If strong correlation between objective IQM scores and subjective radiologist scores are found, one or more objective IQMs can be selected as a quantitative surrogate to tune CS-MRI reconstruction parameters expediently on a patient-by-patient basis. To correlate the objective IQMs with radiologist opinion of image quality, a bank of distorted images must be generated that are reflective of distortions commonly encountered in MRI and CS-MRI. Radiologists must then be asked to rate the diagnostic quality of the images for each

distortion, and the subjective scores compared to objective IQM scores. In particular, the results of chapter 5 demonstrate that care must be taken to seek correlations between objective IQM scores and subjective radiologist scores with knowledge of the sparsity enforcements used in CS-MRI; it was observed that different sparsity enforcements gave rise to image reconstructions with noticeably different visual properties even when the same IQM score was obtained (e.g. wavelet sparsity left a "grainy" quality to the images that wasn't present in TV or LLR sparsity enforcements). It is possible that radiologists will strongly prefer image reconstructions utilizing some sparsity enforcements over others, and that the objective IQMs may be unable to discern the difference between these images. These studies should be repeated for images of several different anatomies, for it is possible that some anatomies may be less amenable to objective IQM analysis than others. The end result of a study that correlates subjective radiologist opinion of diagnostic image quality to objective IQM scores, in several different anatomies, would provide a tool that would allow rapid evaluation of CS-MRI reconstruction capabilities on a patient-by-patient basis.

Future work may endeavour to study no-reference IQMs for medical image analysis. No-reference IQMs can be developed with the intention of quantifying the presence of a single kind of distortion, or to be sensitive to a wide variety of distortions. There have been many no-reference IQMs developed for the assessment of natural scenes[i], which draw on extensive empirical studies of natural scene statistics (NSS). Some examples include Distortion Identification-based Image Integrity and Verity Evaluation (DIIVINE)[46] and Blind Image Notator using Discrete cosine transform Statistics (BLINDS-II)[47]. It is not clear that the statistics of MRI images are equivalent to those of natural scenes, and thus many of these existing no-reference IQMs might need to be adapted for application in MRI. Interestingly, a preliminary study by Woodard et. al.[48] did find that no-reference IQMs incorporating assumptions from NSS were statistically able to differentiate MRI images of medical brain images separately contaminated with additive white Gaussian noise or blurring from those that were not artificially distorted in any way. However the separate additions of additive white Gaussian noise or blurring are very artifical, and correlations between

[i]A natural scene, as the name implies, is essentially one which can be seen by the naked human eye and that is commonly encountered in natural environments. Medical images are not a subset of the natural images.

radiologist subjective scoring and the IQM scores were not studied. This leads us to conclude that further investigation is still required into the use of no-reference IQMs which assume properties from NSS in the field of medical image processing. To address this need a comprehensive study of the statistics of medical images may be undertaken. This may be done for *both* MRI imaging and x-ray based imaging, which may have different statistics due to the different underlying physical processes of image generation (i.e. manipulation of bulk magnetization versus photon interactions). Indeed, a comprehensive study of the statistics of medical images would be a great contribution to the literature in its own right. If the statistics of medical images are found to be very similar to those of NSS, no-reference IQMs developed with assumptions from NSS (such as DIIVINE) may be studied in medical imaging and correlations sought between no-reference objective IQM score and subjective radiologist scoring. This would open the possibility of using a wide range of IQMs that may potentially correlate better with radiologist opinion of diagnostic quality, and may have utility in characterizing dynamic MRI without need of a reference image.

Aside from seeking justification for the use of no-reference IQMs developed for natural images, the specific distortions introduced by CS-MRI may be studied in detail and an attempt made to construct a CS-MRI specific no-reference objective IQM. By tailoring the no-reference IQM to seek distortions specifically introduced by CS-MRI, an IQM may potentially be obtained that better reflects the visual properties of image reconstructions made with various sparsity enforcements (i.e. could differentiate between "grainy" wavelet enforced reconstructions and "blurred" TV enforced reconstructions). Such an objective IQM may correlate much better with subjective radiologist scoring, with the added benefit that no reference image is required.

Much work can be done to improve the simulation framework that we have developed as well. To begin, a physiologically realistic contrast enhancement model may be incorporated to replace the simple closed-form functions we utilized in obtaining our preliminary results. An excellent example would be to have the synthetic phantom features evolve according to the extended Tofts model[2]. Simulations with a finite non-zero repetition time may then be extensively undertaken to assess the balance required between image quality and temporal resolution, and to seek correlations between physiological quantitative parameters and objective IQM score. Additionally,

a set of anatomical phantoms may be developed for accurate investigations of the correlations between objective IQM scores and dynamic MRI performance. Temporal enhancements may be superimposed onto fully sampled, high quality image data in order to fabricate these phantoms.

Utilizing future developments in knowledge of objective IQMs, and in results observed in future generations of our simulation framework, one or more objective IQMs may be selected for characterizing the performance of dynamic MRI applications and thresholds could be established to delineate acceptable loss of dynamic MRI performance from an unacceptable loss. This would have profound utility in clinical practise, providing an expedient and quantitative means of reconstructing a dynamic image series to achieve clinically acceptable performance in quantitative dynamic MRI.

Appendix A

Mathematical Background

A.1 Condition for Modular Congruence

Let $x, y \in \mathbb{R}$ and $n \in \mathbb{Z}$. Taking note of equation 2.56 two modulo operations are equivalent when:

$$\begin{aligned} \text{mod}\ (x, n) &= \quad \text{mod}\ (y, n) \\ x - n \cdot \left\lfloor \frac{x}{n} \right\rfloor &= y - n \cdot \left\lfloor \frac{y}{n} \right\rfloor \\ x - y &= n \cdot \left(\left\lfloor \frac{x}{n} \right\rfloor - \left\lfloor \frac{y}{n} \right\rfloor \right) \\ &\therefore n | (x - y) \end{aligned} \tag{A.1}$$

Where the notation $a|b$ means "a divides b".

A.2 Tukey Honest Significant Difference Post Hoc Analysis

While two-way ANOVA testing will report whether a set of sample means are statistically equivalent or not, the ANOVA test does not report *which* group means are statistically different. Post hoc analyses can be performed to acquire this information. We implemented a Tukey honestly significant difference (HSD) post hoc analysis using the *multcompare* function included with the MATLAB Statistics and Machine Learning Toolbox[i].

Two sample means $\bar{y}_i$ and $\bar{y}_j$ are considered statistically different if they differ by the following range:

$$|\bar{y}_i - \bar{y}_j| > \frac{q_{\alpha,k,N-k}}{\sqrt{2}} \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} \tag{A.2}$$

Where α is the significance level, k is the number of groups studied, n_i are the number of samples in each group, $N = kn$ is the total number of samples, $q_{\alpha,k,N-k}$ is the upper

[i]We obtained the following information from `https://www.mathworks.com/help/stats/multiple-comparisons.html#bum7ugv-1`

$100(1-\alpha)^{\text{th}}$ percentile of the studentized range distribution with parameter k and $N-k$ degrees of freedom, and MSE ratio of the sum of squares from each group to the number of degrees of freedom associated with each source. If the comparison intervals calculated from equation A.2 do not overlap when comparing two sample means, then they are statistically different.

Appendix B

CIRCUS Investigation

B.1 Derivation of Simplified CIRCUS Equation

Equation 2.59 governing sampled indices along each nested square during CIRCUS pattern generation is reprinted below for convenience:

$$i(b,c,q,J) = \mod\left(\left\lfloor \mod\left(\frac{q+bJ}{\varphi},1\right)\cdot K\right\rfloor + \lceil J^c\rceil - 1, K\right) \tag{B.1}$$

Where J is the side length of the current nested square in CIRCUS pattern generation, $K = 4(J-1)$ is the number of possible sampling indices i on each nested square, φ is the golden ratio, and q is the current CIRCUS quanta. For notational convenience, define:

$$\alpha = \left\lfloor \mod\left(\frac{q+bJ}{\varphi},1\right)\cdot K\right\rfloor \tag{B.2}$$

Expanding equation B.1 with the definition of the modulo operator (equation 2.56) and the use of equation B.2:

$$i(b,c,q,J) = (\alpha + \lceil J^c\rceil - 1) - K\cdot\left\lfloor\frac{\alpha+\lceil J^c\rceil-1}{K}\right\rfloor \tag{B.3}$$

To simplify equation B.3 further, we first simplify the expression for α given by equation B.2. Expanding the modulo operator inside the floor function:

$$\alpha = \left\lfloor\left(\frac{q+bJ}{\varphi} - \left\lfloor\frac{q+bJ}{\varphi}\right\rfloor\right)\cdot K\right\rfloor \tag{B.4}$$

$$= \left\lfloor K\left(\frac{q+bJ}{\varphi}\right)\right\rfloor - K\left\lfloor\frac{q+bJ}{\varphi}\right\rfloor \tag{B.5}$$

Where we've used the property $\lfloor x-n\rfloor = \lfloor x\rfloor - n$ for $n\in\mathbb{Z}$. Again moving towards simplification, we simplify the expression $K\left\lfloor\frac{\alpha+\lceil J^c\rceil-1}{K}\right\rfloor$ in equation B.3 using equation

B.4:

$$K\left\lfloor\frac{\alpha+\lceil J^c\rceil-1}{K}\right\rfloor = K\left\lfloor\frac{\left\lfloor K\left(\frac{q+bJ}{\varphi}\right)\right\rfloor - K\left\lfloor\frac{q+bJ}{\varphi}\right\rfloor+\lceil J^c\rceil-1}{K}\right\rfloor \tag{B.6}$$

$$= K\left\lfloor\frac{\left\lfloor K\left(\frac{q+bJ}{\varphi}\right)\right\rfloor+\lceil J^c\rceil-1}{K}\right\rfloor - K\left\lfloor\frac{q+bJ}{\varphi}\right\rfloor \tag{B.7}$$

Combining equations B.4 and B.6 with equation B.3 yields:

$$i(b,c,q,J) = \left\lfloor K\left(\frac{q+bJ}{\varphi}\right)\right\rfloor + \lceil J^c\rceil - 1 - K\left\lfloor\frac{\left\lfloor K\left(\frac{q+bJ}{\varphi}\right)\right\rfloor+\lceil J^c\rceil-1}{K}\right\rfloor \tag{B.8}$$

Which we recognize as the expanded modulo operator (equation 2.56), giving a simplified CIRCUS equation:

$$i(b,c,q,J) = \mod\left(\left\lfloor K\left(\frac{q+bJ}{\varphi}\right)\right\rfloor + \lceil J^c\rceil - 1, K\right) \tag{B.9}$$

B.2 Repetitions with b at fixed c

Let $c=\gamma$ and $q=\eta$ be fixed. For two choices of radial parameters b_1 and b_2, the sampled indices $i(b_1,\gamma,\eta,J)$ and $i(b_2,\gamma,\eta,J)$ on two nested squares are equivalent when:

$$\mod\left(\left\lfloor K\left(\frac{\eta+b_1J}{\varphi}\right)\right\rfloor + \lceil J^\gamma\rceil - 1, K\right) = \mod\left(\left\lfloor K\left(\frac{\eta+b_2J}{\varphi}\right)\right\rfloor + \lceil J^\gamma\rceil - 1, K\right) \tag{B.10}$$

The reader is reminded that J (and consequently K) varies with each nested square during CIRCUS generation. Equation B.10 is true if:[i]

$$K\Big|\left(\left\lfloor K\left(\frac{\eta+b_1J}{\varphi}\right)\right\rfloor + \lceil J^\gamma\rceil - 1\right) - \left(\left\lfloor K\left(\frac{\eta+b_2J}{\varphi}\right)\right\rfloor + \lceil J^\gamma\rceil - 1\right)$$

$$\therefore K\Big|\left\lfloor K\left(\frac{\eta+b_1J}{\varphi}\right)\right\rfloor - \left\lfloor K\left(\frac{\eta+b_2J}{\varphi}\right)\right\rfloor \tag{B.11}$$

Where the notation $a|b$ means "a divides b" or $b=na$ for $n\in\mathbb{Z}$. Condition B.11 must hold true for every J, otherwise there will exist different sampled indices i on at least one nested square between the resulting CIRCUS patterns.

[i] See appendix A.1 for conditions of modular congruence.

Trivial occurrences satisfying condition B.11 occur when the two floor functions are equivalent, and correspond to situations where b_1 and b_2 are very similar. Suppose that $b_2 = b_1 + \delta$ for $\delta \in \mathbb{R}$, and let $z = \left\lfloor K\left(\frac{\eta + b_1 J}{\phi}\right)\right\rfloor$ be the common floor. Then for each J the corresponding tolerance δ_J allowed by the floor function satisfies condition B.11 for a single J is given by:

$$z \leq K\left(\frac{\eta + (b_1 + \delta_J)J}{\phi}\right) < z + 1$$

$$\therefore \frac{1}{J}\left(\frac{z\phi}{K} - b_1 J - q\right) \leq \delta_J < \frac{1}{J}\left(\frac{(z+1)r}{K} - b_1 J - q\right) \tag{B.12}$$

Where it is emphasized that the bounds determining δ_J for each J are not equivalent. To choose a single δ_b that satisfies condition B.11, we must take the most restrictive of the upper and lower boundaries across all δ_J; only then will $b_2 = b_1 + \delta_b$ satisfy condition B.11 for each and every nested square during CIRCUS generation[ii]. δ_b is thus given by:

$$\max_J \left[\frac{1}{J}\left(\frac{z\varphi}{K} - b_1 J - q\right)\right] \leq \delta_b < \min_J \left[\frac{1}{J}\left(\frac{(z+1)\varphi}{K} - b_1 J - q\right)\right] \tag{B.13}$$

The bounds on δ_b are very small for typical values of J_{max}, where J_{max} is the size of the largest nested square used in CIRCUS pattern generation. Figure B.1 shows upper and lower bounds on δ_b as b_1 increases from 0 to 250 with $J_{max} = 200$; the magnitude of the bounds was found to remain consistent even for larger values of b_1, and to become more restrictive for larger values of J_{max}.

Because $\varphi \in \mathbb{R} \setminus \mathbb{Q}$, it follows that $K\left(\frac{q + bJ}{\varphi}\right) \in \mathbb{R} \setminus \mathbb{Q}$ and condition B.11 cannot be further simplified in general. Aside from the cases where b_2 is chosen to be very nearly b_1, as discussed above, condition B.11 is not satisfied for all J in general. Therefore there exist no pairs b_1 and b_2, chosen sufficiently distinct of one another, of b for fixed c in which the generated CIRCUS patterns are equivalent.

[ii] If the most restrictive of the limits on all δ_j are not taken in defining the range of δ_b, there will exist at least one J during CIRCUS generation for which the floor functions in condition B.11 are not equivalent.

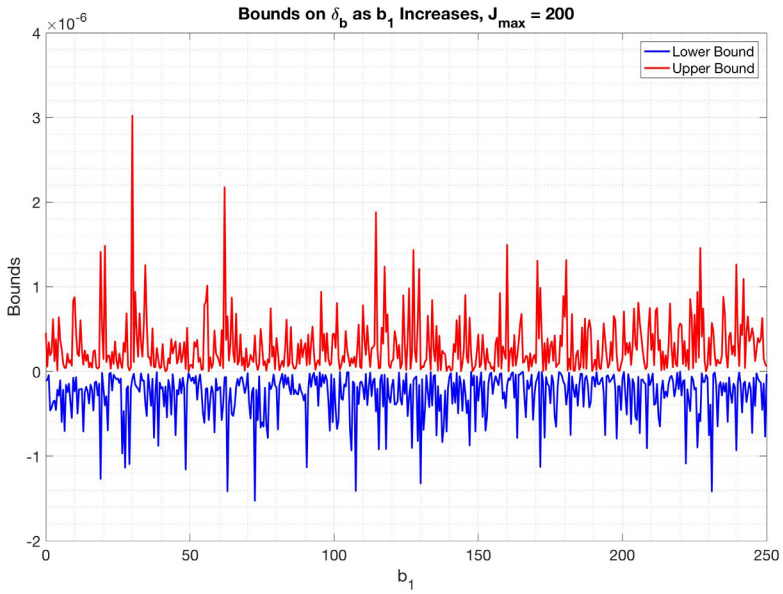

Figure B.1: The upper and lower bounds on δ_b for b_1 and $b_2 = b_1 + \delta_b$ to produce identical CIRCUS patterns, calculated for $J_{max} = 200$. The magnitude of these bounds remains consistent for values of b_1 above 250. As J_{max} increases, the bounds on δ_b become even more restrictive.

B.3 Repetitions with c at fixed b

Let $b = \beta$ and $q = \eta$ be fixed. For two choices of spiral parameters c_1 and c_2, the sampled indices $i(\beta, c_1, \eta, J)$ and $i(\beta, c_2, \eta, J)$ on the two nested squares are equivalent when:

$$K \Big| \left(\left\lfloor K \left(\frac{\eta + \beta J}{\varphi} \right) \right\rfloor + \lceil J^{c_1} \rceil - 1 \right) - \left(\left\lfloor K \left(\frac{\eta + \beta J}{\varphi} \right) \right\rfloor + \lceil J^{c_2} \rceil - 1 \right)$$

$$\therefore K \Big| \lceil J^{c_1} \rceil - \lceil J^{c_2} \rceil \tag{B.14}$$

Trivial pairs of c_1 and c_2 satisfy condition B.14 when $\lceil J^{c_1} \rceil = \lceil J^{c_2} \rceil$, which correspond to situations where c_1 and c_2 were chosen very near to each other. Define $c_2 = c_1 + \delta$ for $\delta \in \mathbb{R}$ and let $z = \lceil J^{c_1} \rceil$ be the common ceiling. Then for each J the corresponding tolerance δ_J allowed by the ceiling function that satisfies condition

B.14 for a single J is given by:

$$z - 1 < J^{c_1+\delta_J} \leq z$$

$$\therefore \log_J(z-1) - c_1 < \delta_J \leq \log_J(z) - c_1 \tag{B.15}$$

The most restrictive limits across all δ_J are chosen for defining the limits on δ such that c_1 and $c_2 = c_1 + \delta$ produce equivalent CIRCUS patterns:

$$\max_J \left[\log_J(z-1) - c_1\right] < \delta_c \leq \min_J \left[\log_J(z) - c_1\right] \tag{B.16}$$

The bounds on δ_c are very small for typical values of J_{max}. Figure B.2 shows upper and lower bounds on δ_c as c_1 increases from 0 to 6 with $J_{max} = 200$. The magnitude of the both the lower and upper bounds converge to zero as c_1 increases, and become more restrictive for larger J_{max}.

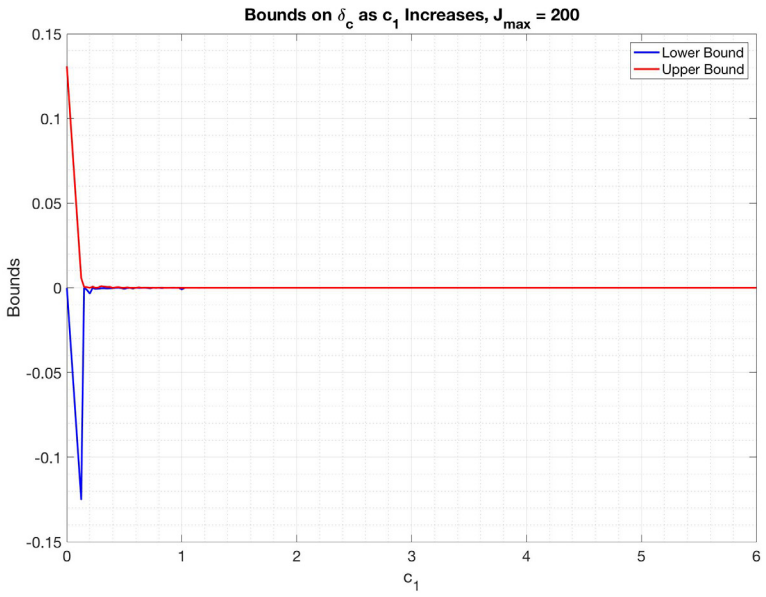

Figure B.2: The upper and lower bounds on δ_c for c_1 and $c_2 = c_1 + \delta_c$ to produce identical CIRCUS patterns, calculated for $J_{max} = 200$. The magnitude of these bounds remains consistent for values of c_1 above 6. As J_{max} increases, the bounds on δ_c gradually become more restrictive.

Condition B.14 cannot be further simplified in general for $c \in \mathbb{R}$. If we restrict the choices of c to $c \in \mathbb{Z} > 0$, condition B.14 may be further simplified due to the

fact that $J \in 2\mathbb{Z}$. Noting that $J^c \in \mathbb{Z}$ when $c \in \mathbb{Z}$, condition B.14 becomes:

$$K \mid J^{c_1} - J^{c_2} \tag{B.17}$$

Without loss of generality, let $c_1 = c_2 + n$ for some $n \in \mathbb{Z} > 0$. Recall also that $K = 4(J-1)$. Condition B.17 becomes:

$$4(J-1) \mid J^{c_2}(J^n - 1) \tag{B.18}$$

Condition B.18 can be further simplified by noting that $(J^n - 1) = (J-1)(1 + J + \ldots + J^{n-1})$. In light of this, condition B.18 requires that the following be true:

$$4 \mid J^{c_2}(1 + J + \ldots + J^{n-1}) \tag{B.19}$$

Since $J \in 2\mathbb{Z}$ it may be equivalently expressed as $J = 2j$ for $j \in \mathbb{Z}$. Thus, condition B.19 may be viewed as:

$$4 \mid (2j)^{c_2} + (2j)^{c_2+1} + \ldots + (2j)^{c_2+n-1} \tag{B.20}$$

Condition B.20 is satisfied for *all* values of j when c_2 is a positive integer greater than or equal than 2.

Thus, two pairs c_1 and c_2 of c will result in identical CIRCUS patterns in only two situations.

1. c_1 and c_2 are chosen as positive integers greater than or equal to 2.

2. c_1 and c_2 are chosen to be very nearly equal in value.

B.4 Additional Point Spread Function Measure Results

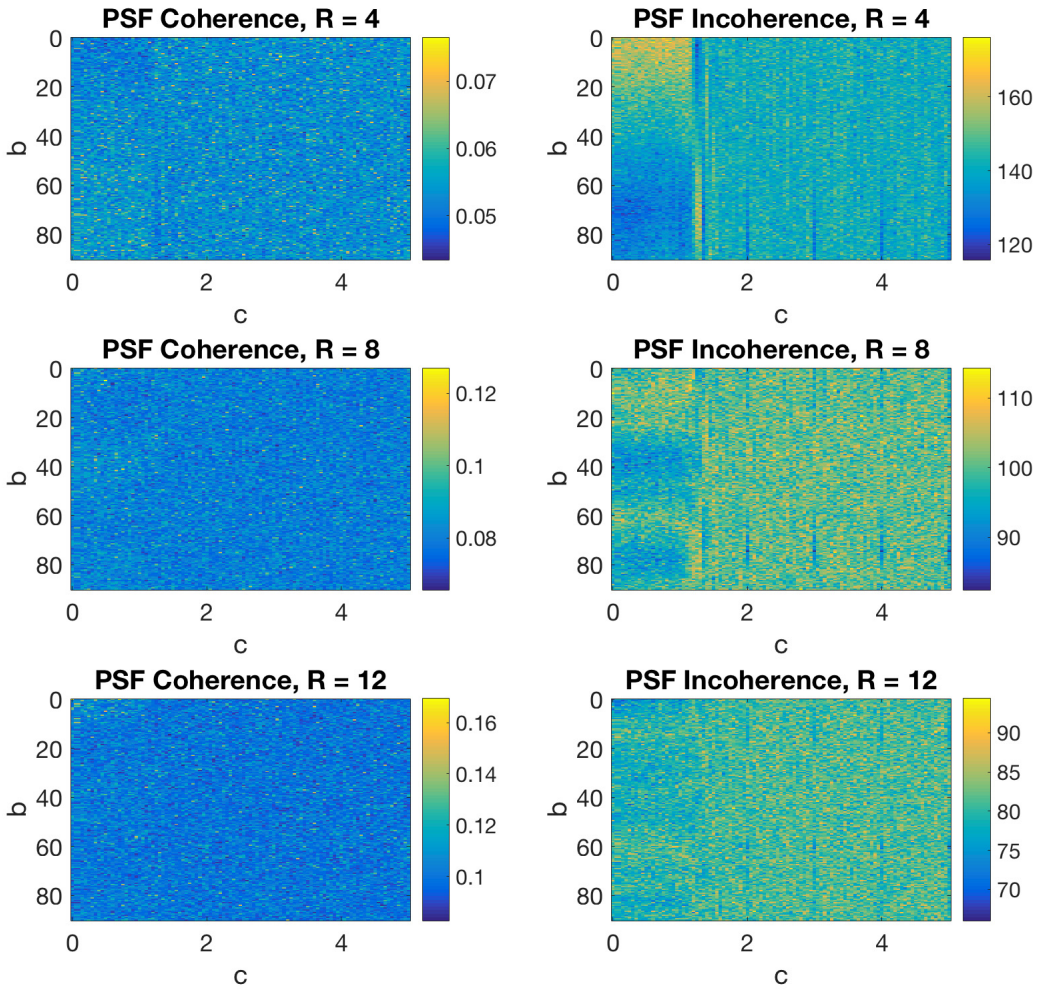

Figure B.3: Coherence (left column) and incoherence (right column) of CIRCUS pattern PSFs generated for a variety of b and c at increasing undersampling factors.

B.5 Additional IQM Results

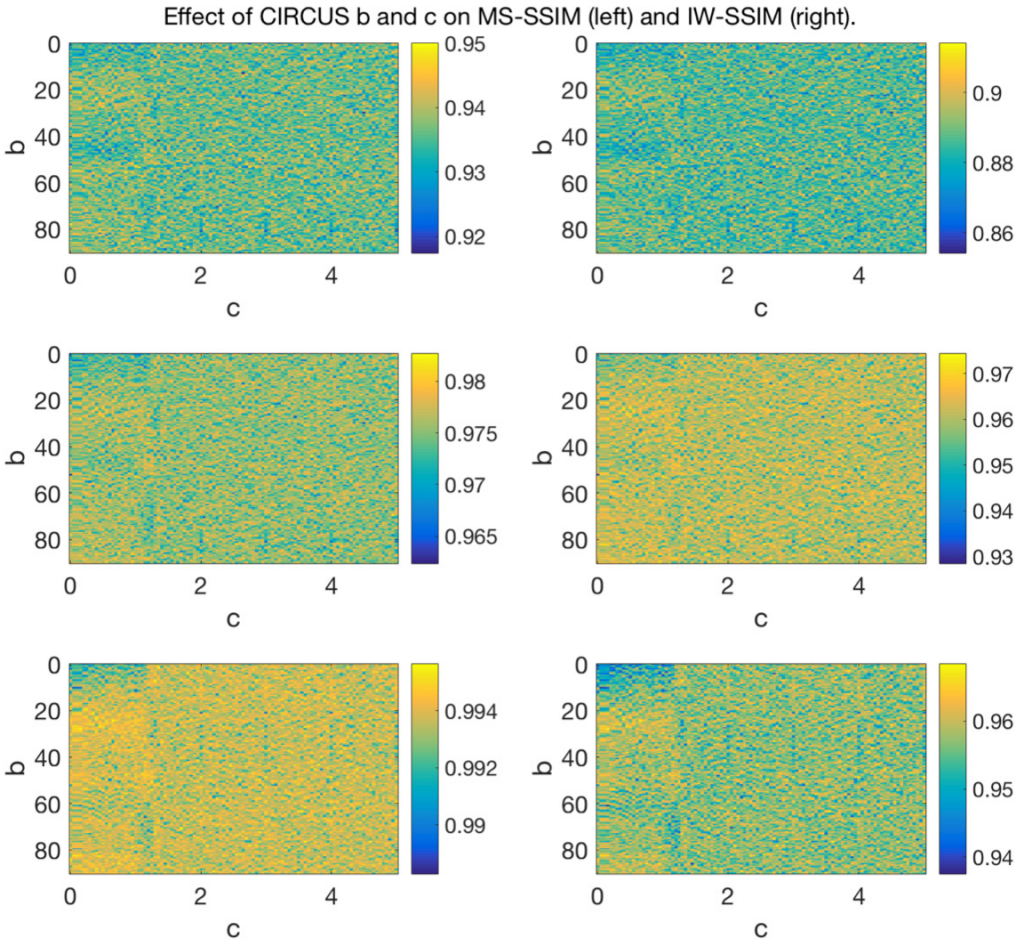

Figure B.4: Response of the MS-SSIM (left column) and IW-SSIM (right column) to changes in b and c for TV regularization (top row), wavelet regularization (middle row), and LLR regularization (bottom row).

B.6 Tukey HSD Intervals for Image Quality Metrics

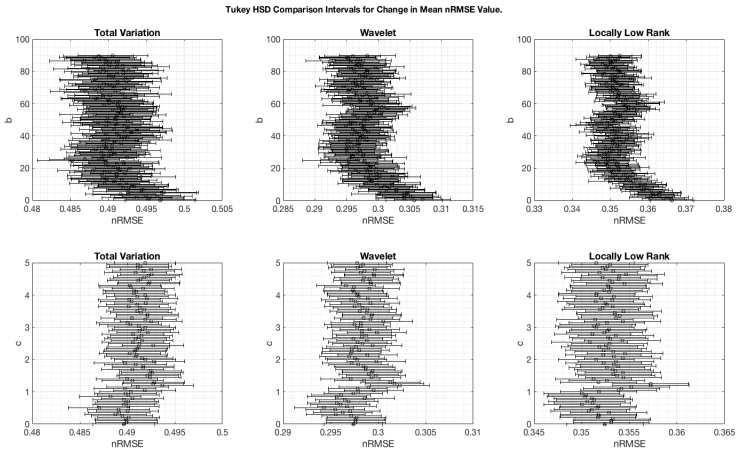

Figure B.5: Visualization of overlap in Tukey HSD comparison intervals for mean normalized RMSE score as b changes (top row) and as c changes (bottom row), for each of the three sparsity transforms.

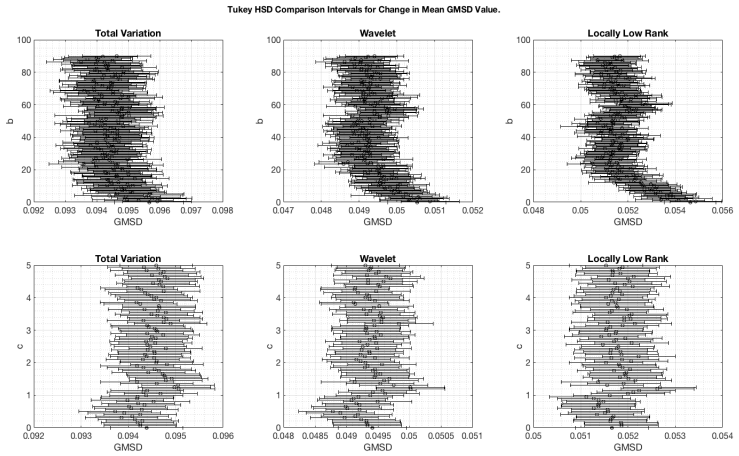

Figure B.6: Visualization of overlap in Tukey HSD comparison intervals for mean GMSD score as b changes (top row) and as c changes (bottom row), for each of the three sparsity transforms.

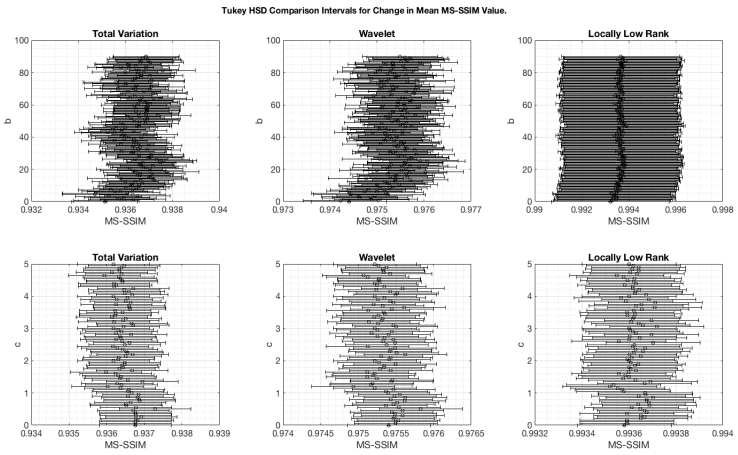

Figure B.7: Visualization of overlap in Tukey HSD comparison intervals for mean MS-SSIM score as b changes (top row) and as c changes (bottom row), for each of the three sparsity transforms.

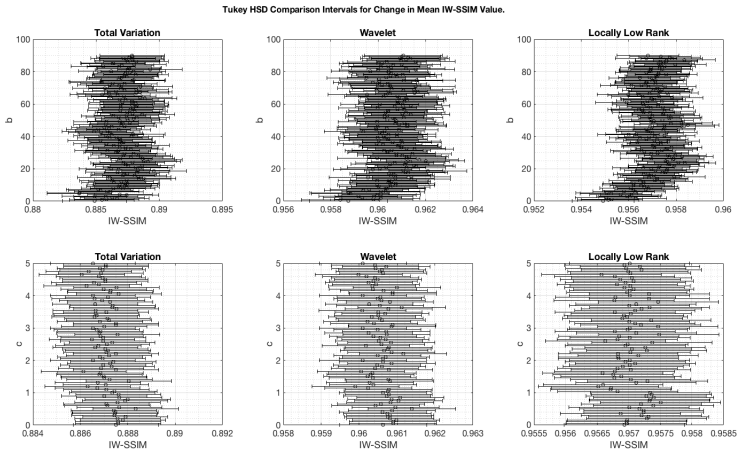

Figure B.8: Visualization of overlap in Tukey HSD comparison intervals for mean IW-SSIM score as b changes (top row) and as c changes (bottom row), for each of the three sparsity transforms.

B.7 Correlations Between Image Quality Metrics and PSF Measures

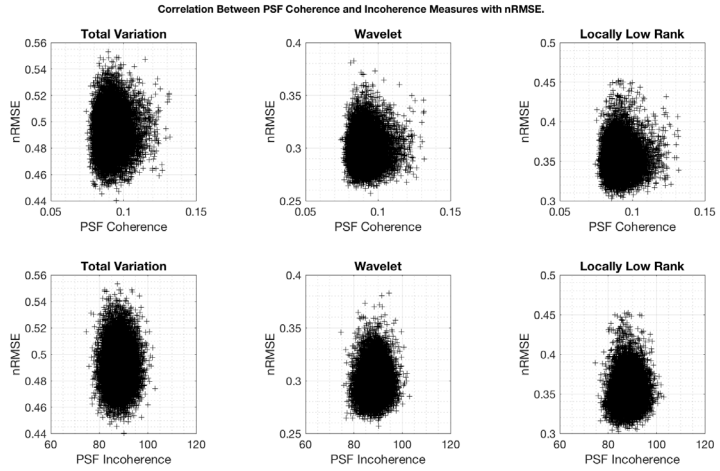

Figure B.9: Correlation between the nRMSE scores and PSF coherence and incoherence, for each of the three sparsity regularizations studied.

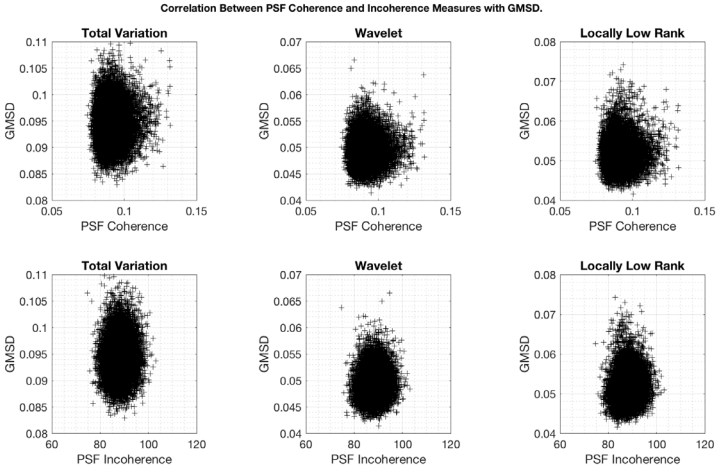

Figure B.10: Correlation between the GMSD scores and PSF coherence and incoherence, for each of the three sparsity regularizations studied.

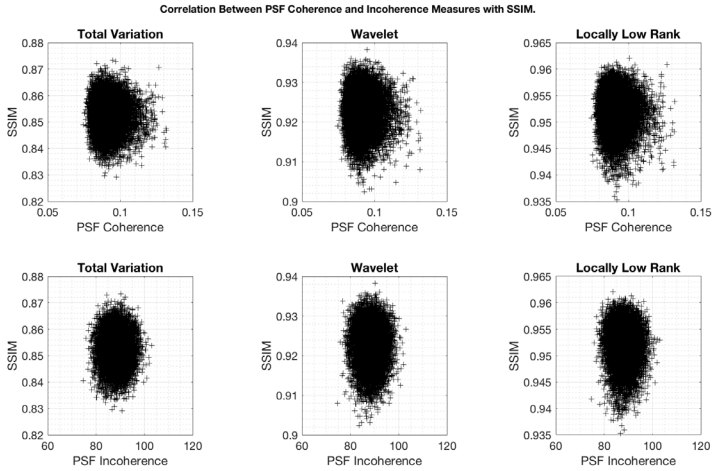

Figure B.11: Correlation between the SSIM scores and PSF coherence and incoherence, for each of the three sparsity regularizations studied.

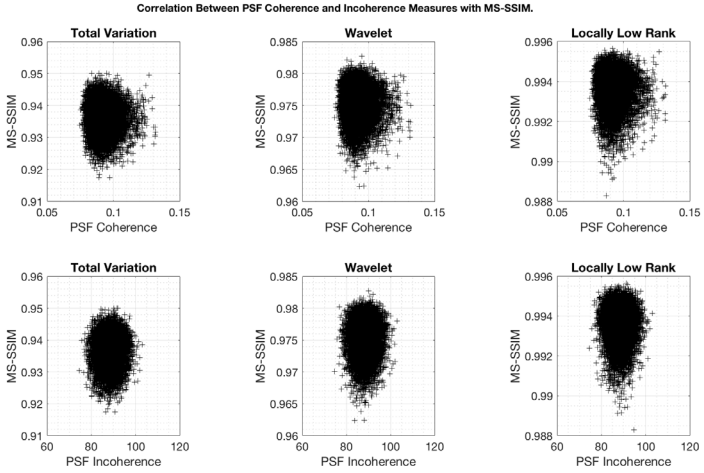

Figure B.12: Correlation between the MS-SSIM scores and PSF coherence and incoherence, for each of the three sparsity regularizations studied.

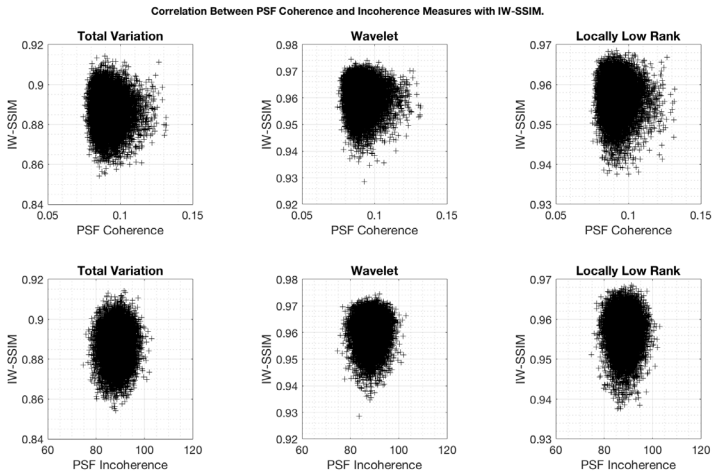

Figure B.13: Correlation between the IW-SSIM scores and PSF coherence and incoherence, for each of the three sparsity regularizations studied.

Appendix C

Image Quality Metric Investigation

C.1 Image Quality Metric Calculation Results

Figure C.1 to Figure C.15 show the objective IQM results for each patient, for each IQM, and for each sparse enforcement technique. Red diamonds have been used to mark the value of λ at each R resulting in the most favourable IQM score. Normalization of the RMSE was done relative to the mean of the reference images.

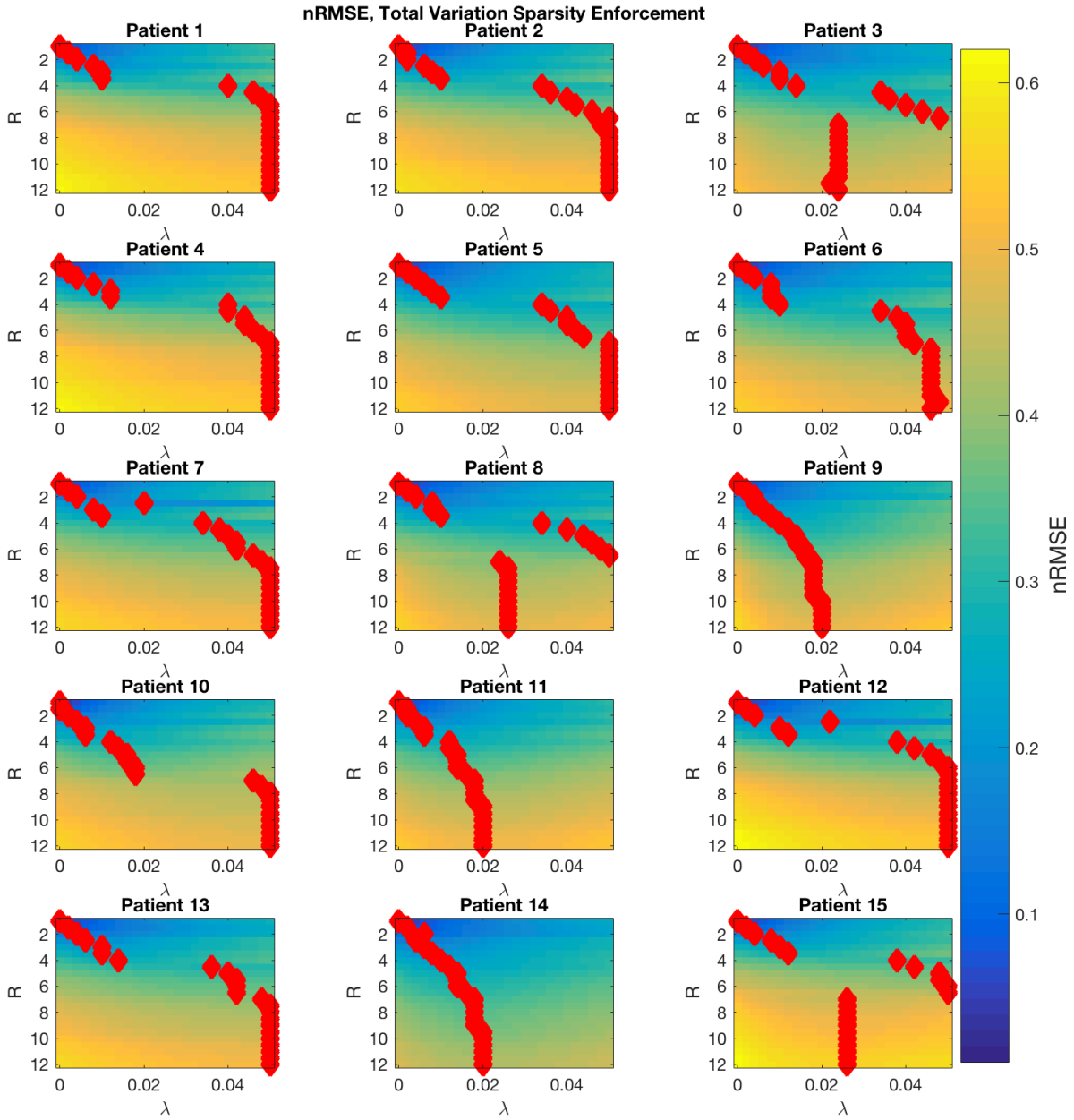

Figure C.1: The normalized RMSE results for each of the 15 patient data sets using a total variation sparse enforcement with 8 coil PI CS reconstruction. The RMSE was normalized to mean signal intensity in each case. Red diamonds mark the value of λ giving the best score at each R.

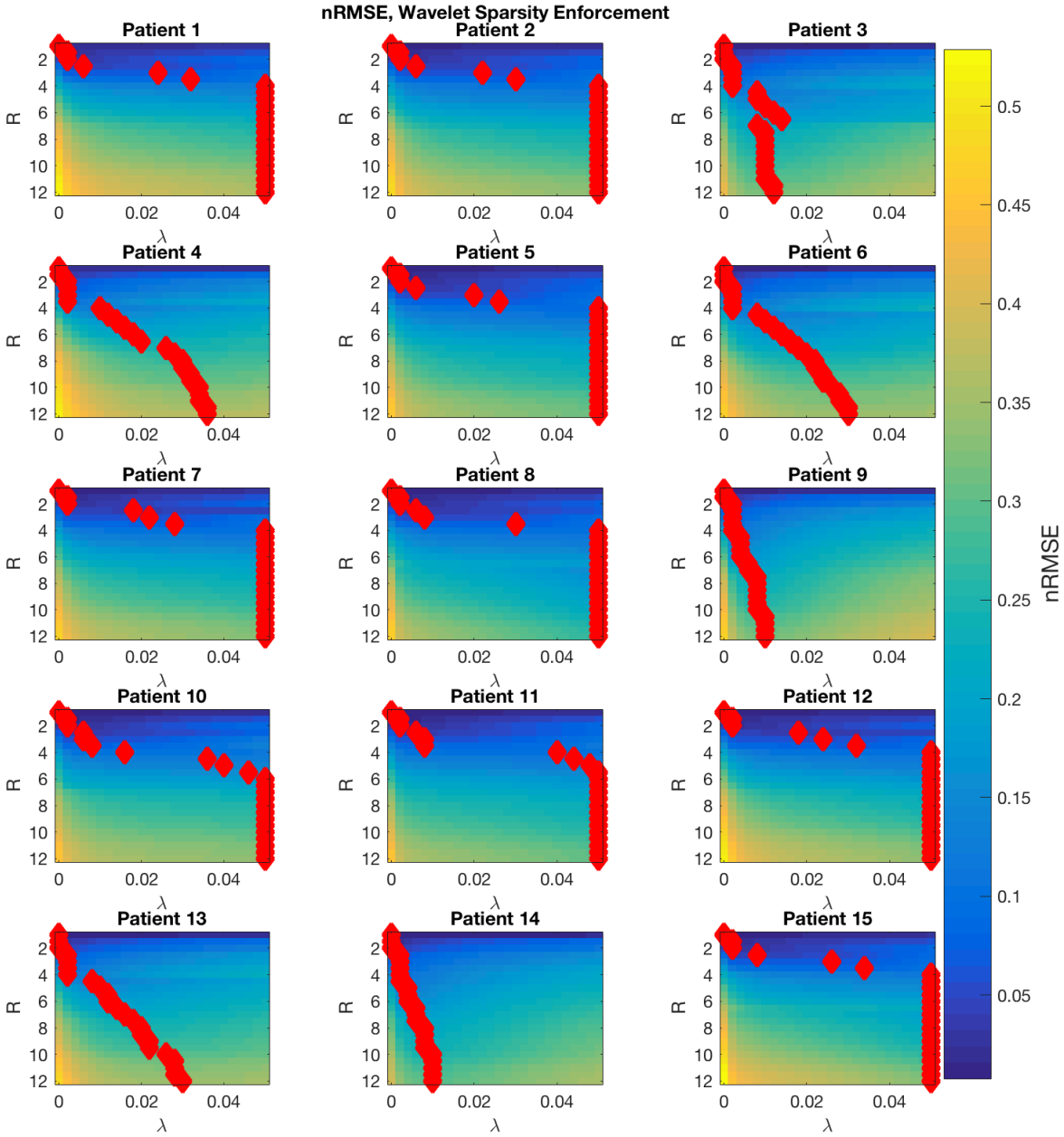

Figure C.2: The normalized RMSE results for each of the 15 patient data sets using a wavelet sparse enforcement with 8 coil PI CS reconstruction. The RMSE was normalized to mean signal intensity in each case. Red diamonds mark the value of λ giving the best score at each R.

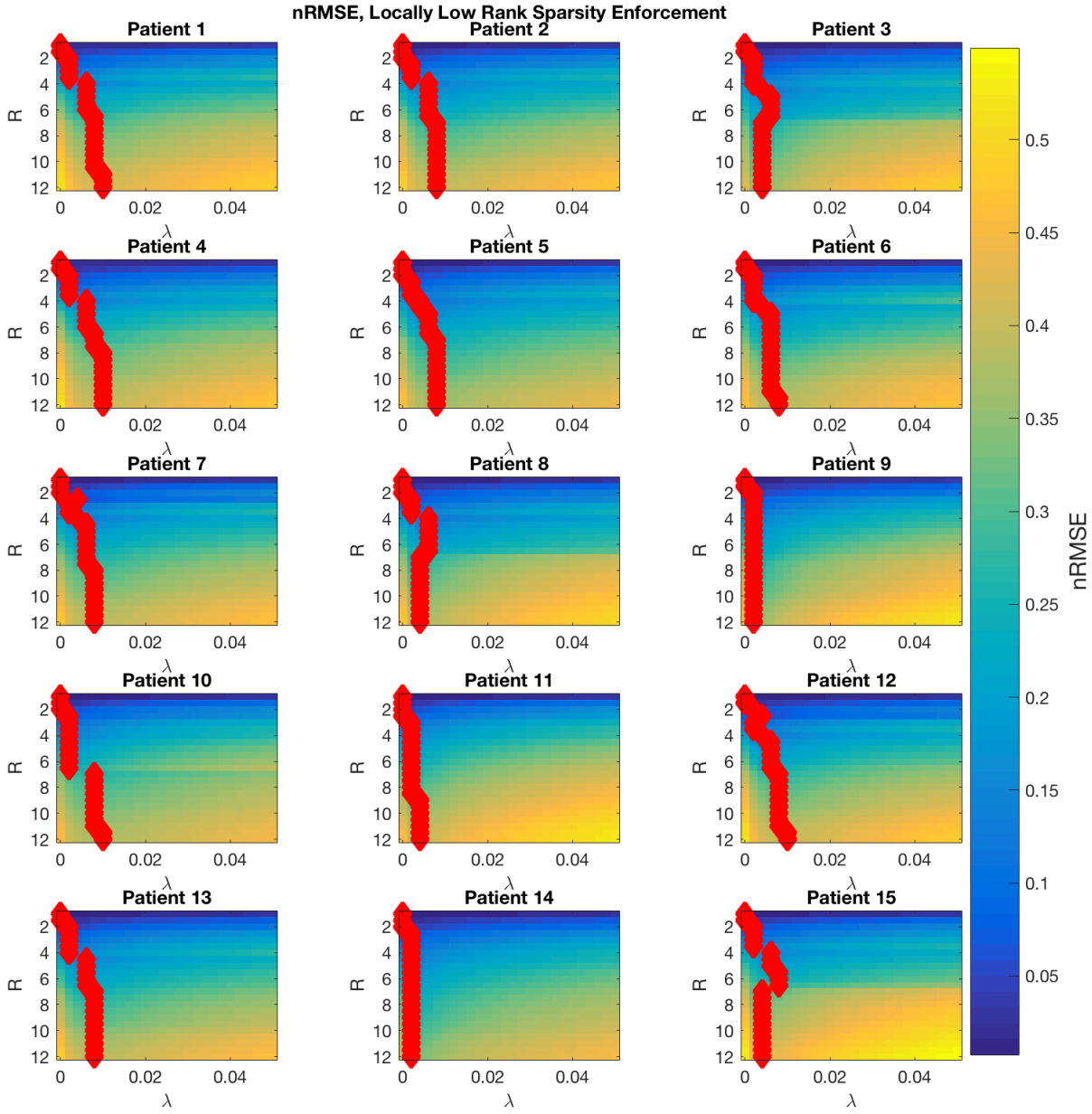

Figure C.3: The normalized RMSE results for each of the 15 patient data sets using a locally low rank sparse enforcement with 8 coil PI CS reconstruction. The RMSE was normalized to mean signal intensity in each case. Red diamonds mark the value of λ giving the best score at each R.

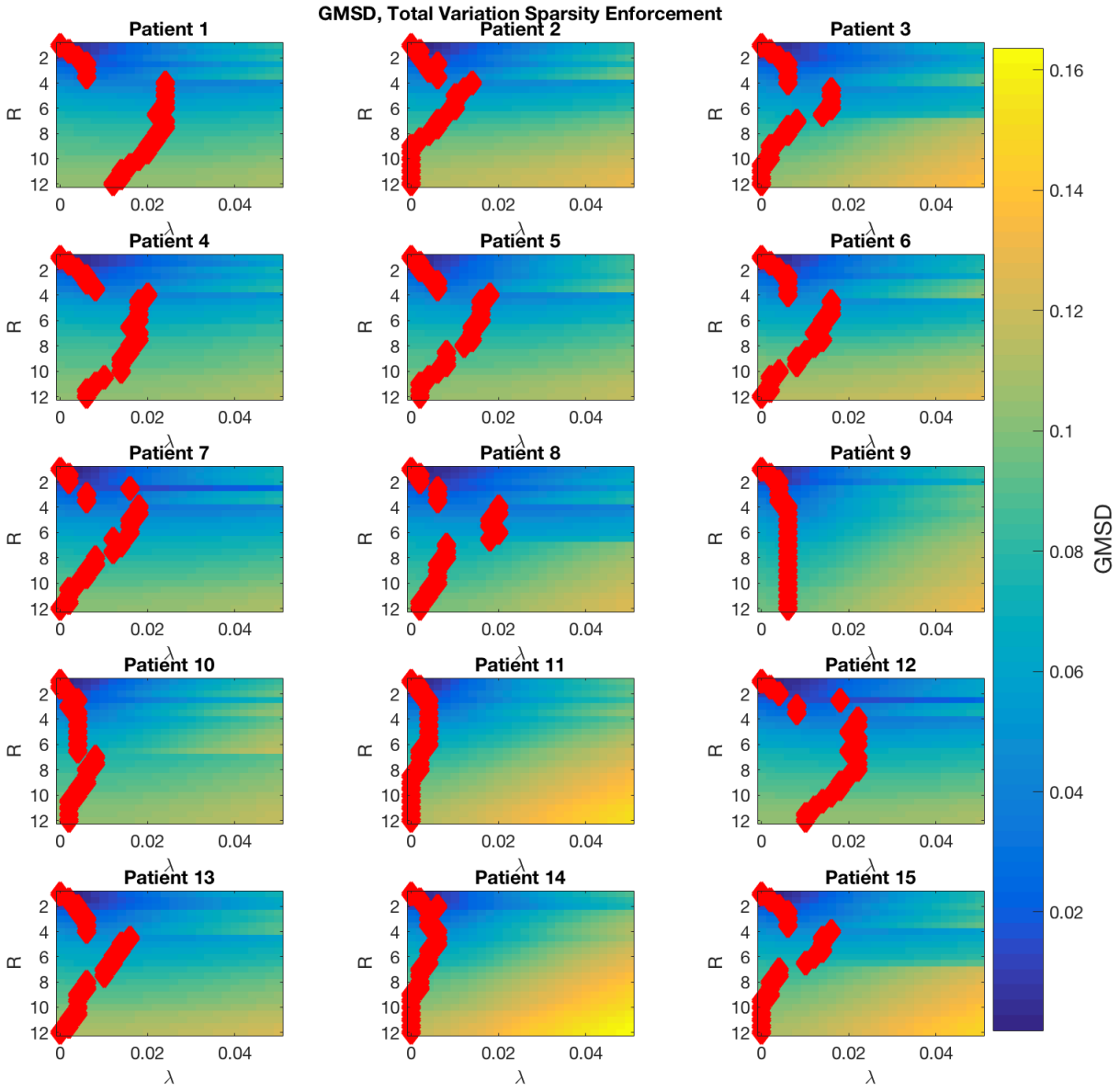

Figure C.4: The GMSD results for each of the 15 patient data sets using a total variation sparsity enforcement with 8 coil parallel imaging compressed sensing reconstruction. Red diamonds mark the value of λ giving the best score at each R.

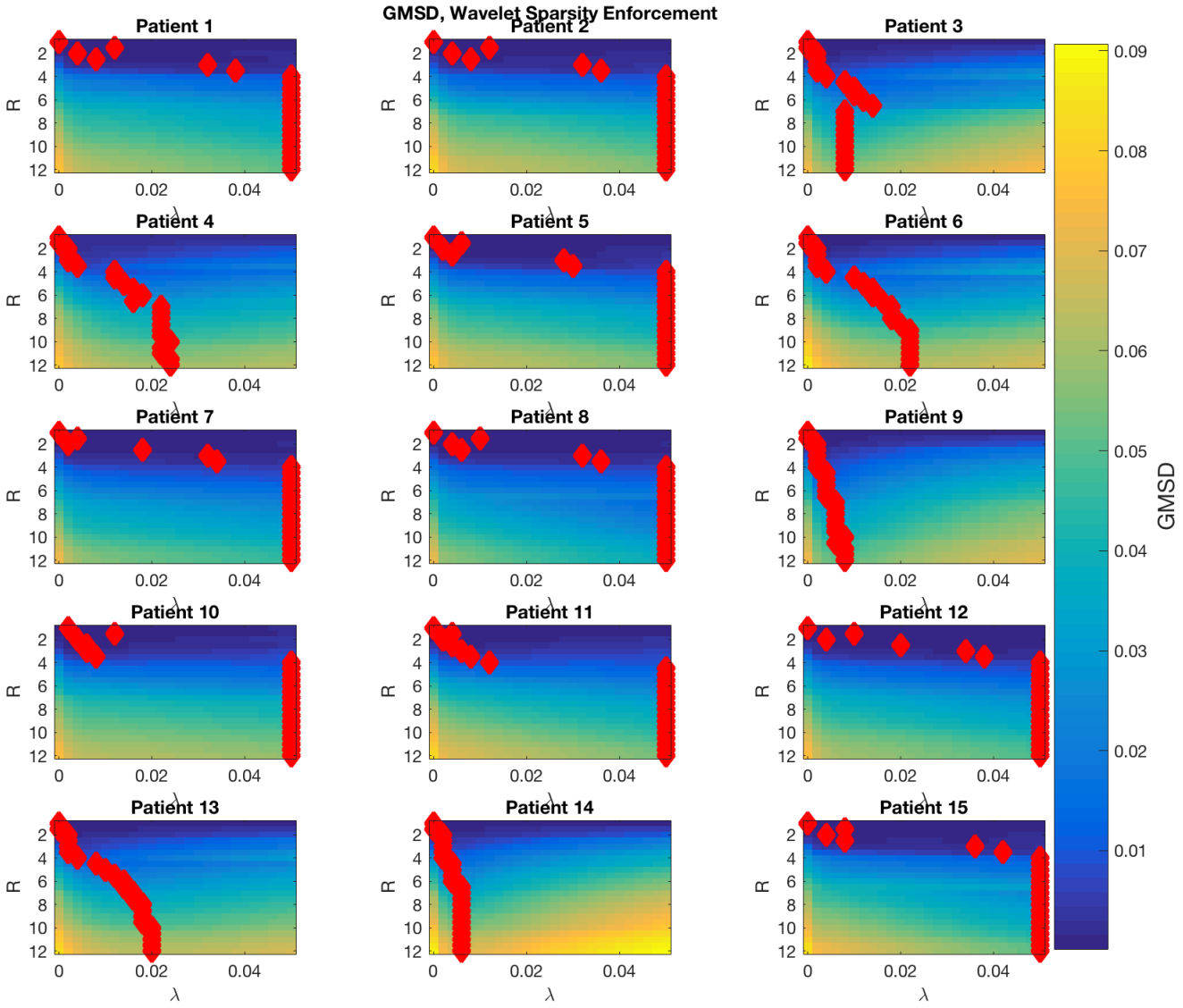

Figure C.5: The GMSD results for each of the 15 patient data sets using a wavelet sparsity enforcement with 8 coil parallel imaging compressed sensing reconstruction. Red diamonds mark the value of λ giving the best score at each R.

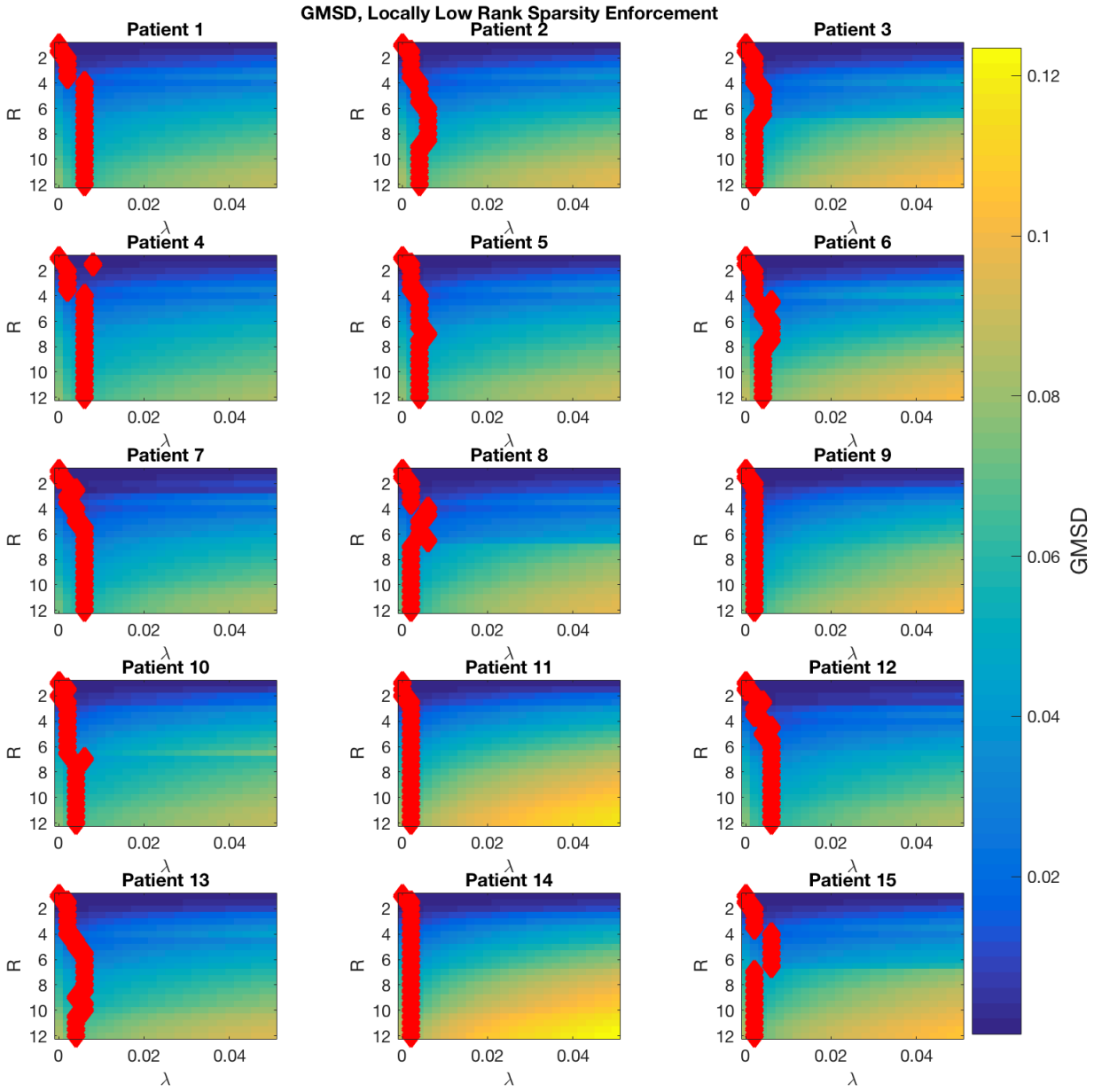

Figure C.6: The GMSD results for each of the 15 patient data sets using a locally low rank sparsity enforcement with 8 coil parallel imaging compressed sensing reconstruction. Red diamonds mark the value of λ giving the best score at each R.

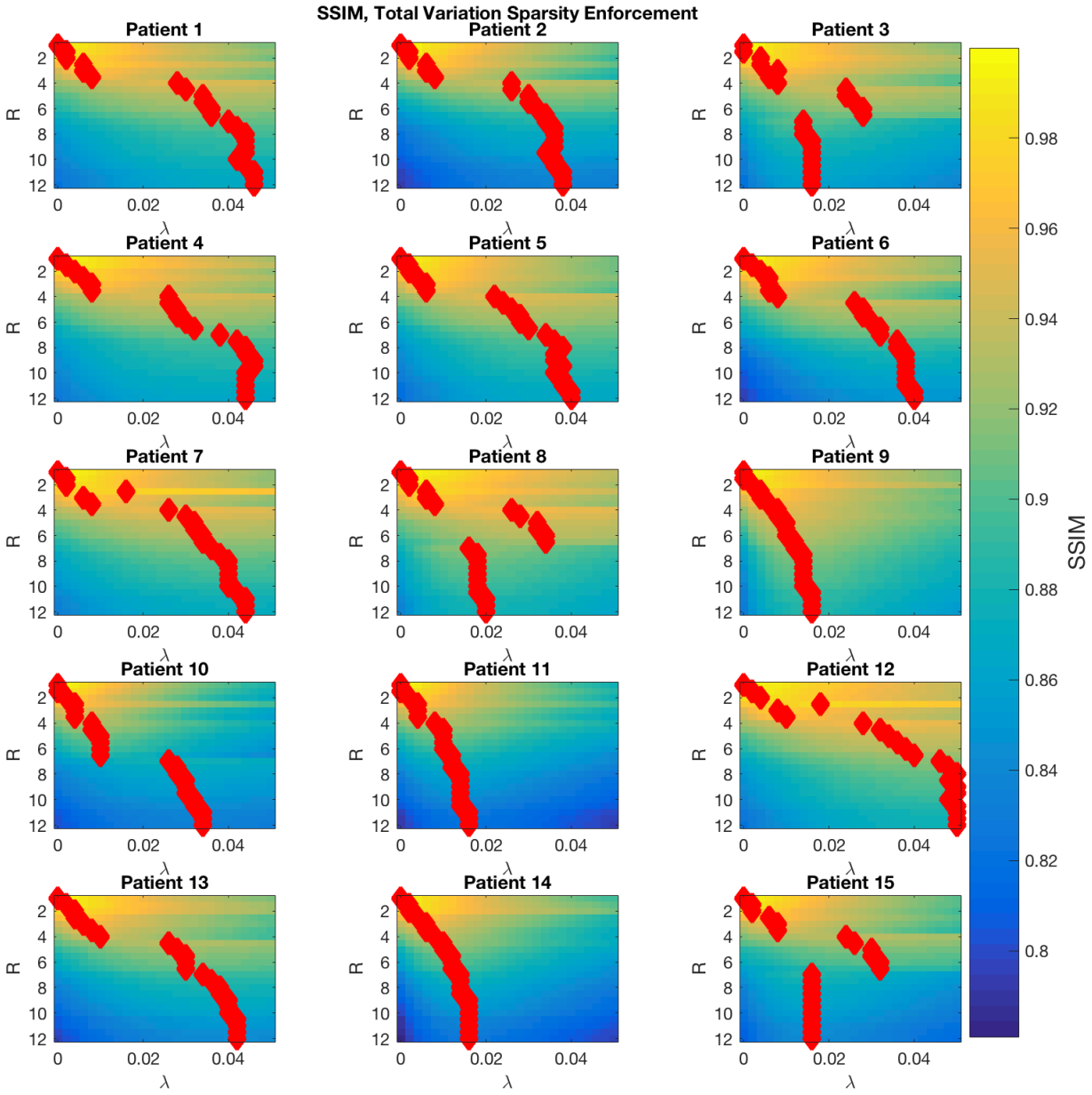

Figure C.7: The SSIM results for each of the 15 patient data sets using a total variation sparsity enforcement with 8 coil parallel imaging compressed sensing reconstruction. Red diamonds mark the value of λ giving the best score at each R.

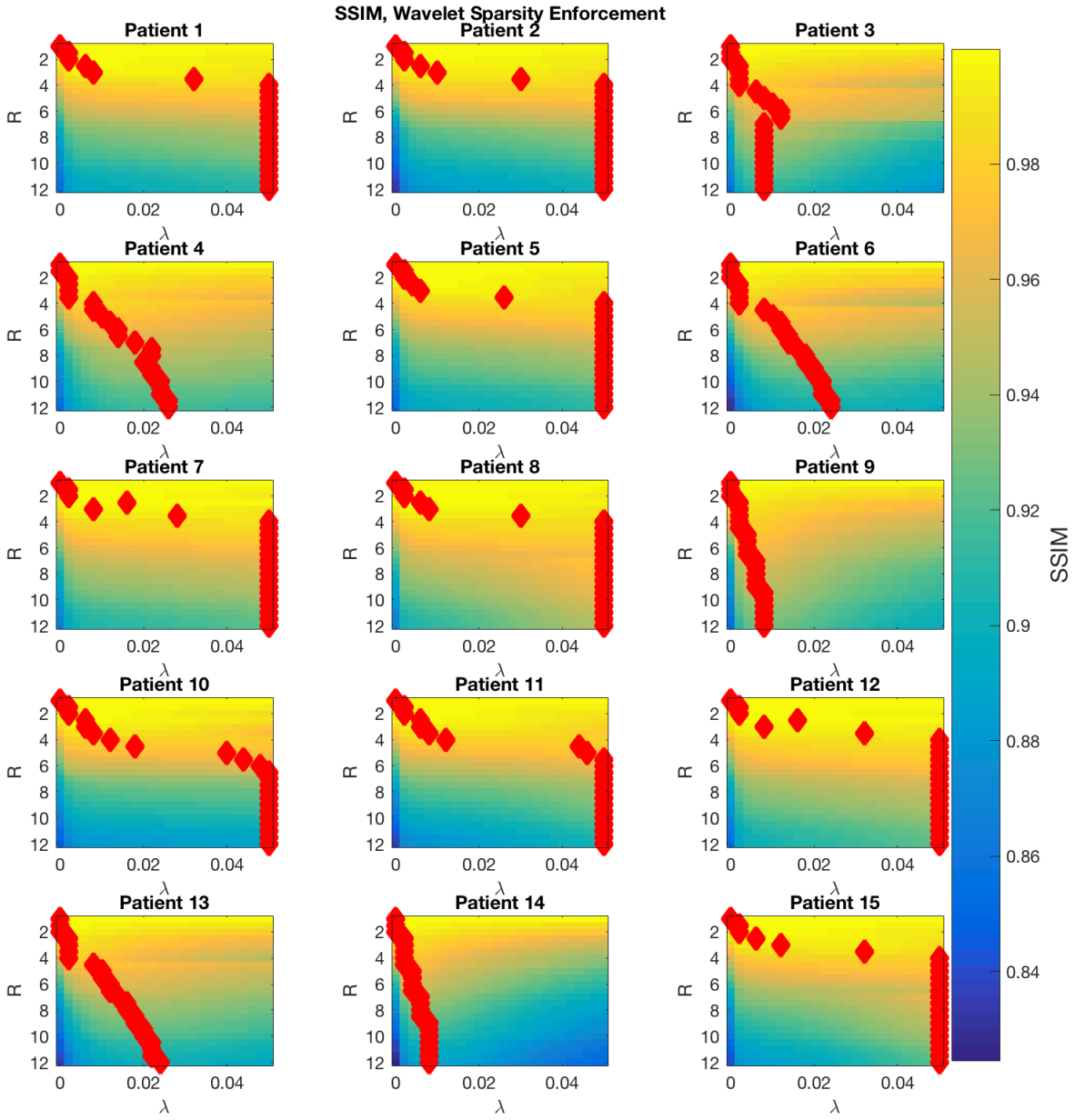

Figure C.8: The SSIM results for each of the 15 patient data sets using a wavelet sparsity enforcement with 8 coil parallel imaging compressed sensing reconstruction. Red diamonds mark the value of λ giving the best score at each R.

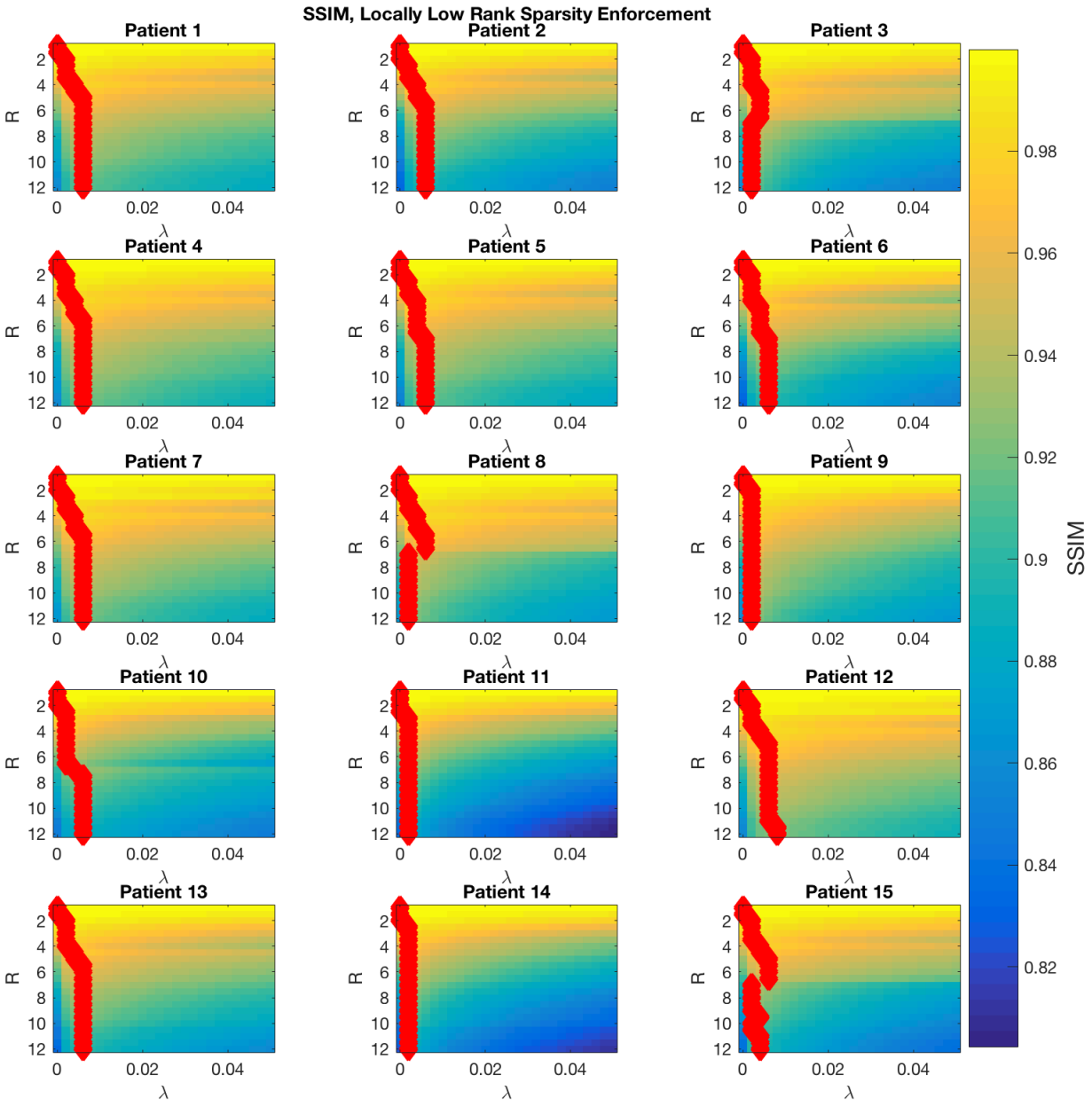

Figure C.9: The SSIM results for each of the 15 patient data sets using a locally low rank sparsity enforcement with 8 coil parallel imaging compressed sensing reconstruction. Red diamonds mark the value of λ giving the best score at each R.

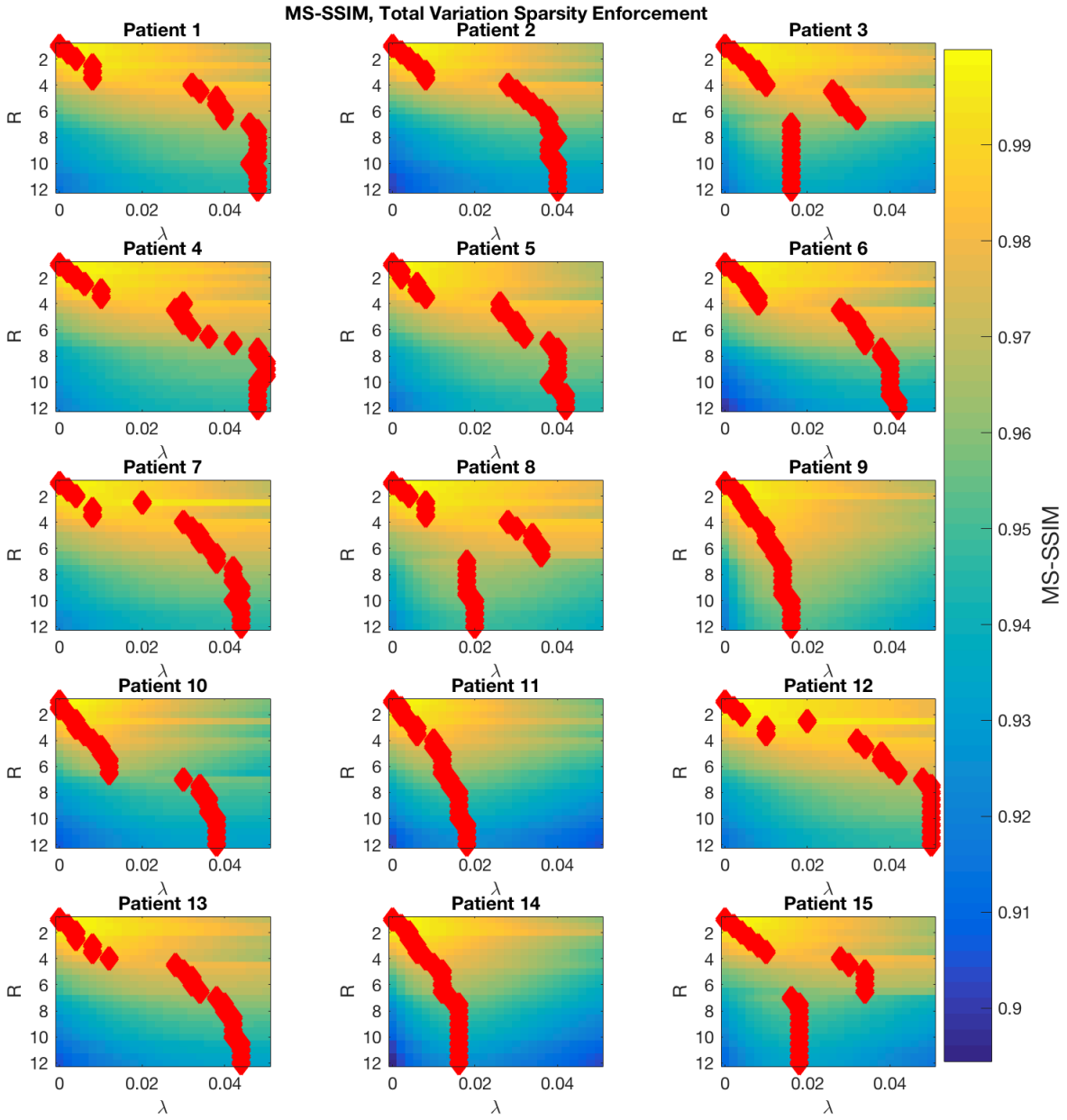

Figure C.10: The MS-SSIM results for each of the 15 patient data sets using a total variation sparsity enforcement with 8 coil parallel imaging compressed sensing reconstruction. Red diamonds mark the value of λ giving the best score at each R.

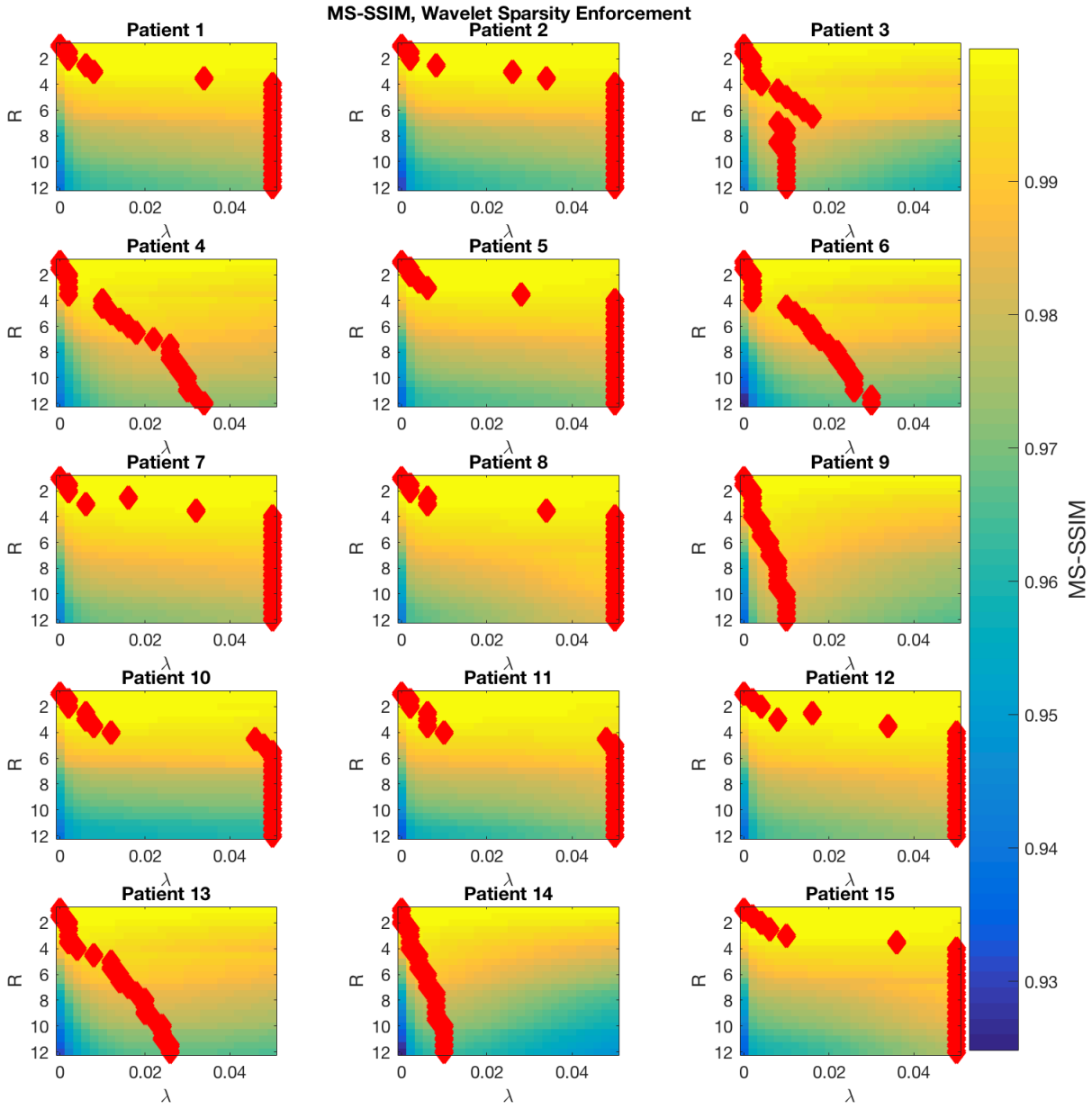

Figure C.11: The MS-SSIM results for each of the 15 patient data sets using a wavelet sparsity enforcement with 8 coil parallel imaging compressed sensing reconstruction. Red diamonds mark the value of λ giving the best score at each R.

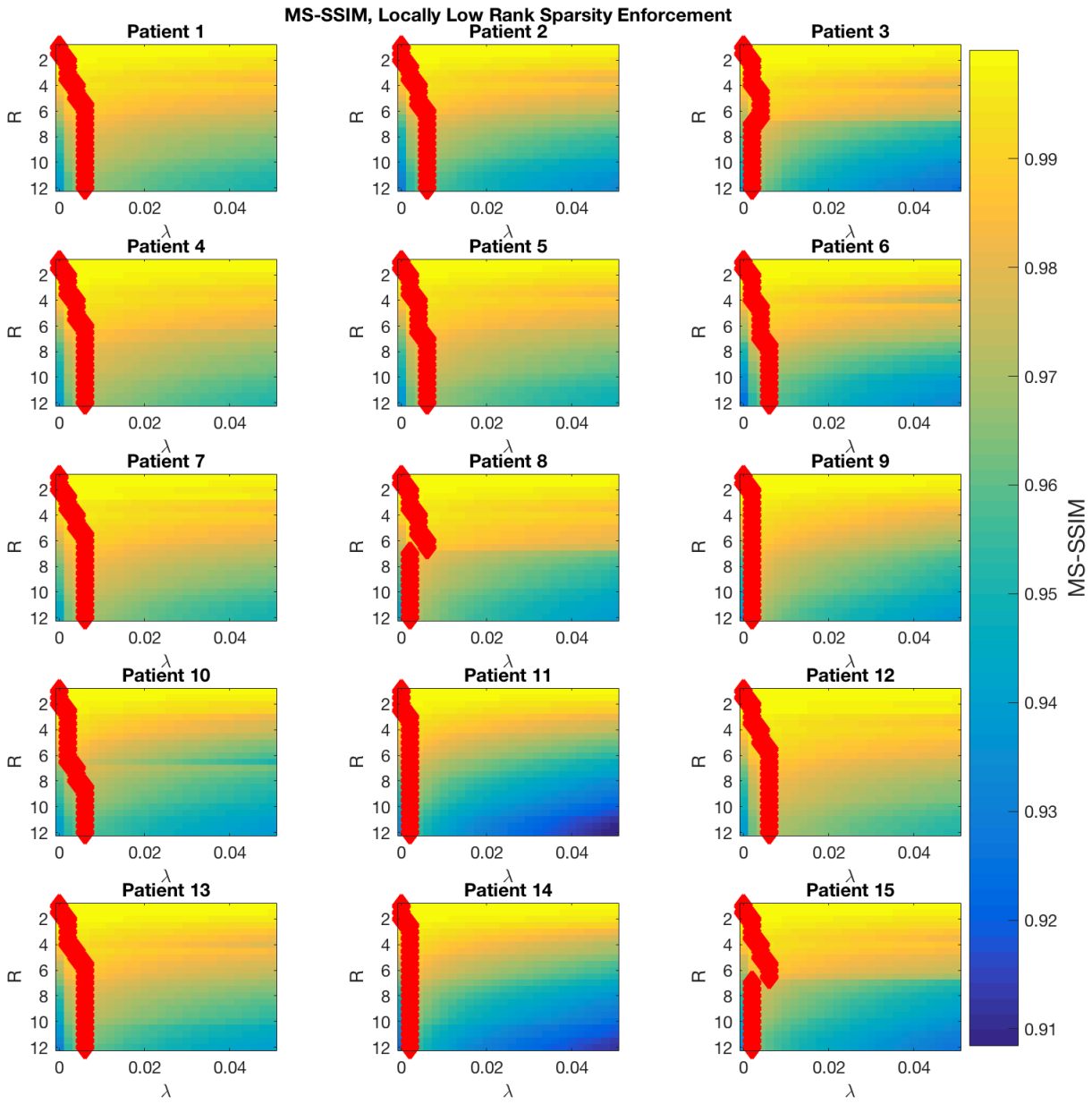

Figure C.12: The MS-SSIM results for each of the 15 patient data sets using a locally low rank sparsity enforcement with 8 coil parallel imaging compressed sensing reconstruction. Red diamonds mark the value of λ giving the best score at each R.

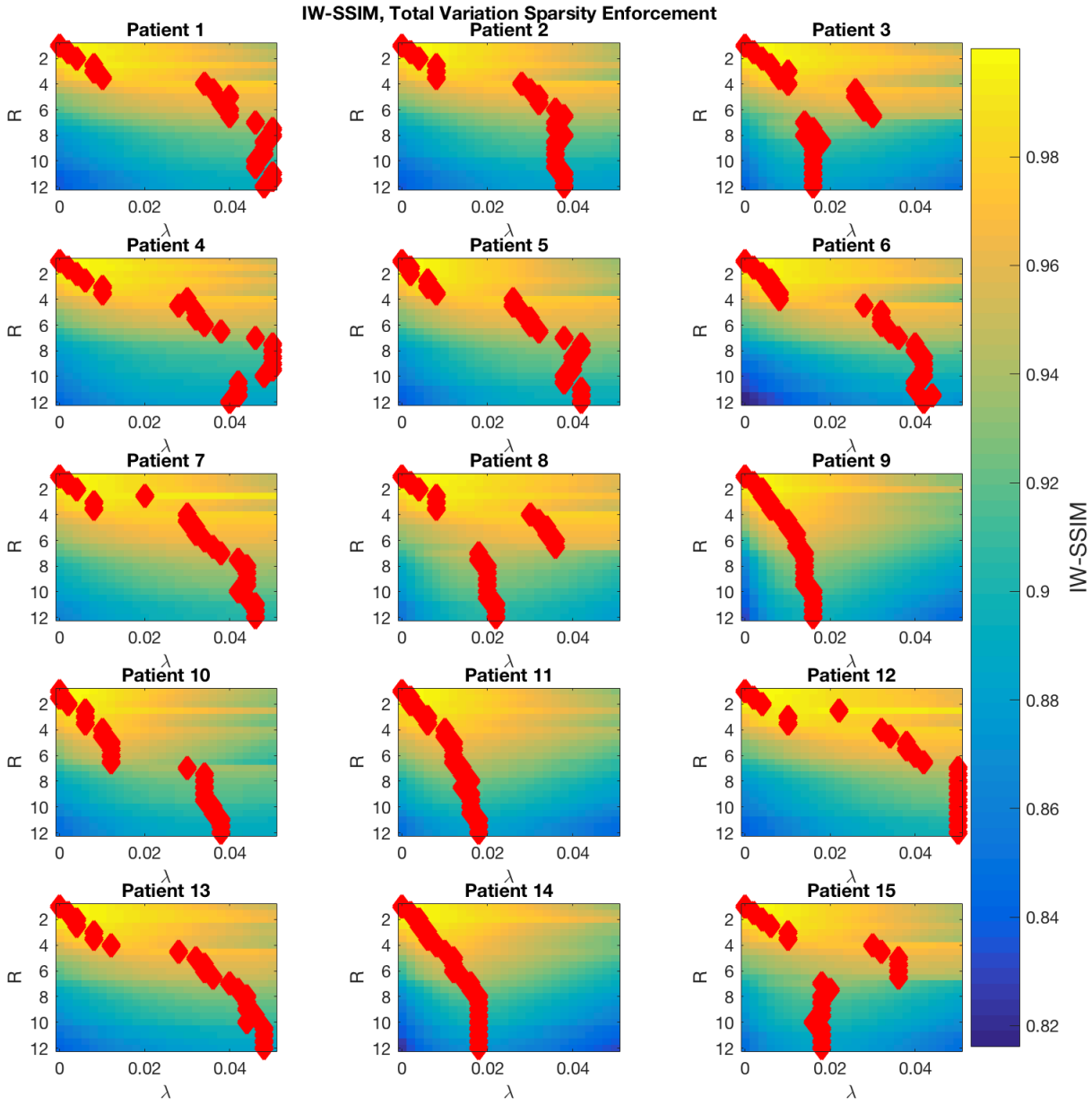

Figure C.13: The IW-SSIM results for each of the 15 patient data sets using a total variation sparsity enforcement with 8 coil parallel imaging compressed sensing reconstruction. Red diamonds mark the value of λ giving the best score at each R.

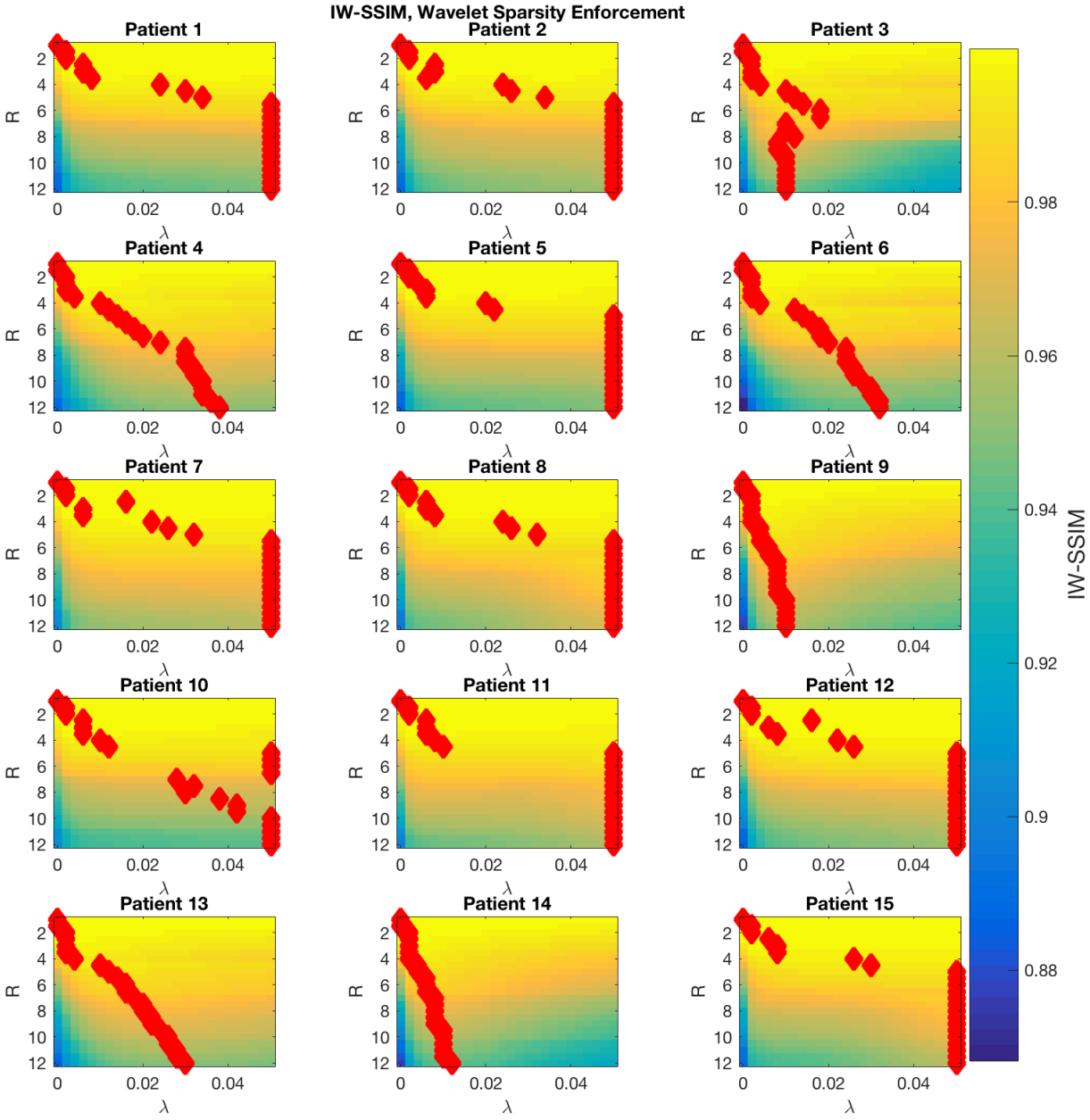

Figure C.14: The IW-SSIM results for each of the 15 patient data sets using a wavelet sparsity enforcement with 8 coil parallel imaging compressed sensing reconstruction. Red diamonds mark the value of λ giving the best score at each R.

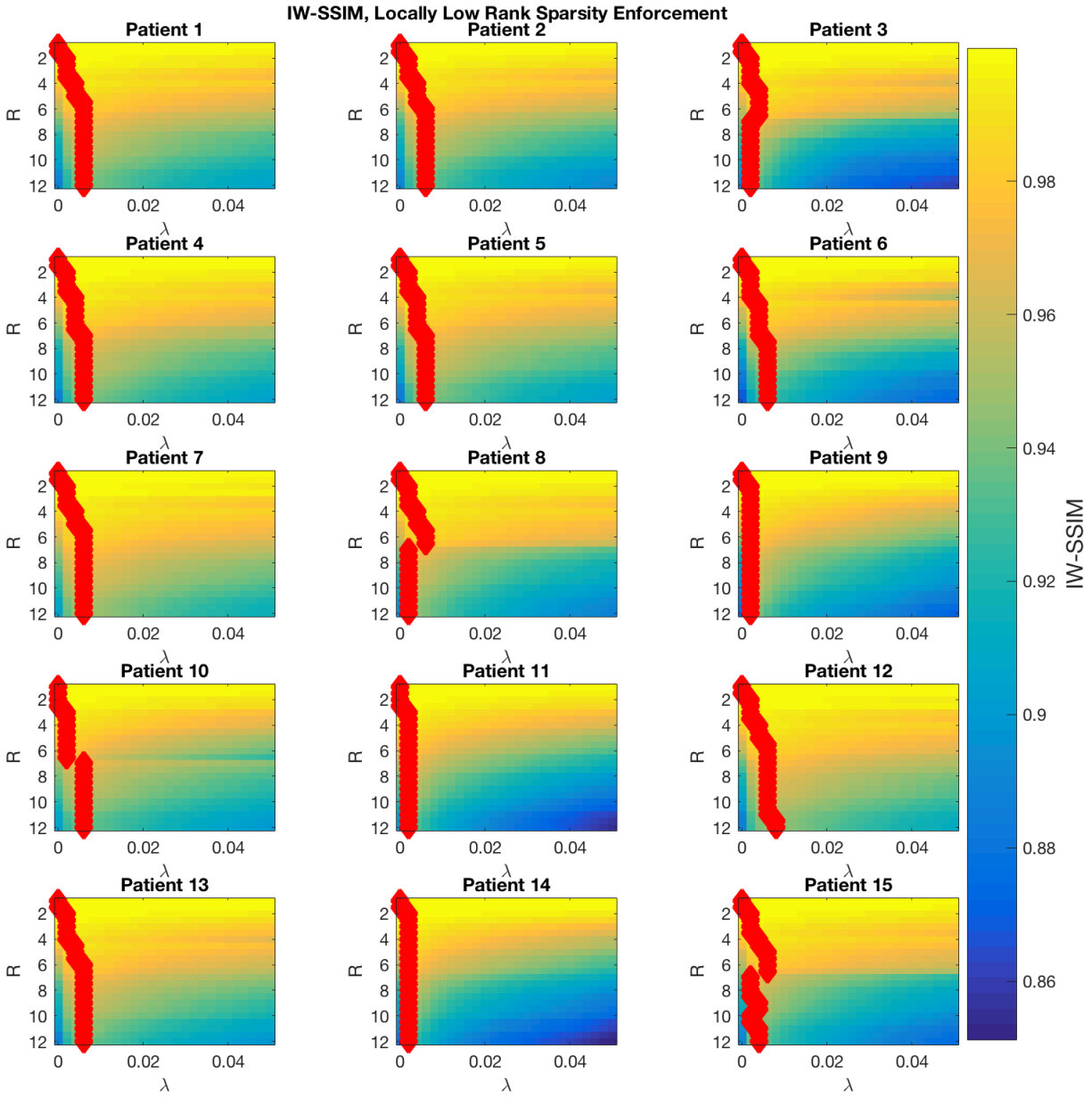

Figure C.15: The IW-SSIM results for each of the 15 patient data sets using a locally low rank sparsity enforcement with 8 coil parallel imaging compressed sensing reconstruction. Red diamonds mark the value of λ giving the best score at each R.

Appendix D

Simulations of Temporally Dynamic Synthetic Phantom

D.1 Simulator Framework Organization

The sections of the simulation framework are briefly described below. The full simulation framework, complete with extensive comments, can be found at `https://github.com/biotic-imaging/dynamic_sim`.

The simulation is implemented in MATLAB, using parallel computations and a sparse data class to aid in speed and RAM usage.

D.1.1 User Defined Parameters

The user may specify the in-plane dimensions of the synthetic phantom, with support for minimum dimensions of 175×175 and no upper limit (though it is ideally intended for dimensions on the order of 200×200). The number of slices is fixed at 32.

The user may specify the addition of noise, where noise is added to the image domain during k-space sampling (see section D.1.5) and is modelled with a Rician distribution[49].

The length of the simulation is controlled by choosing the number of CIRCUS pairs to acquire, where a CIRCUS pair consists of a calibration region followed by an "effective" quanta (see section D.1.2). The user may also select the number of effective CIRCUS quanta to pre-generate for the k-space sampling procedure, defining the properties of the effective quanta (e.g. b and c, number of individual quanta placed into a single effective quanta for sampling, and so on). Note that the number of pre-generated effective quanta is not necessarily equal to the number of CIRCUS pairs specified for acquisition. For example the user may specify that 150 CIRCUS pairs (and thus 150 effective CIRCUS quanta) are to be acquired using 75 pre-generated effective CIRCUS quanta, meaning that the pre-generated effective CIRCUS quanta will be recycled and sampled twice.

The user may select snapshot sampling of the CIRCUS patterns by choosing a repetition time (TR) of zero seconds, or choose a realistic sampling of k-space by selecting a nonzero TR (see section D.1.5). For snapshot sampling, a temporal lag between samples to allow for phantom evolution is set by the parameter *dTimePerfect*.

The user may specify the number of simulated coils to use during k-space sampling. One coil with uniform sensitivity across the phantom volume may be selected, or any even number of coils with sinusoidal sensitivity profile (see section 3.4) may be chosen.

The user may select the k-space data interleaving parameters. These include the number of calibration regions and effective CIRCUS quanta to include in each of the interleaved volumes, how many calibration regions to jump ahead during the interleaving process, and whether the effective time of the interleaved k-space volume should result from equal contributions of calibration region and effective CIRCUS quanta sampling times or just from calibration region sampling times.

Finally, the CS regularization weight and sparse regularization to be utilized by BART during CS image reconstruction are set.

D.1.2 Generate of CIRCUS Patterns

CIRCUS was implemented by alternately acquiring effective CIRCUS quanta and calibration regions. A calibration region is a 6×6 fully sampled region at the center of k-space. This provides k-space data where low frequencies are sampled, as well as permitting the use of advanced techniques that estimate coil sensitivity for PI. An effective quanta, for the purposes of sampling k-space, is defined as a grouping of several individual CIRCUS quanta. For example, an effective quanta may consist of a group of 10 combined CIRCUS quanta.

All calibration regions are identical, but the effective CIRCUS quanta are unique. The phase encoding coordinates for the calibration region and each effective CIRCUS quanta are pregenerated. The information for each are stored in seperate structure arrays.

D.1.3 Determine Sampling Times

For convenience during parallel computations, a vector containing the times at which sampling occurs is pregenerated. For snapshot sampling, the number of CIRCUS

pairs acquired and the temporal lag between them controls the sampling times. For realistic sampling, the number of phase encoded lines across all CIRCUS pairs that will be sampled and the TR used in sampling control the sampling times.

D.1.4 Generate Numerical Phantom

The synthetic phantom is composed of a static base, stored in a 3D array, in which vacancies exist for the placement of the phantom features. A structure array contains phantom feature shapes, intensity evolutions, and intensity evolution parameters. The phantom features are represented by 3D arrays, such that they may be conveniently added to the phantom base There are five planes of phantom features, which are shown in Figure 6.1.

- Slices 3-5 contain cylindrical embedded features. These features decay exponentially in intensity, all beginning at the same initial magnitude but decaying at different rates.
- Slices 9-11 contain cylindrical embedded features. These features decay exponentially in intensity, all beginning at the different initial magnitudes but decaying at the same rate.
- Slices 15-17 contain cylindrical embedded features. These features oscillate sinusoidally in intensity, all with the same amplitudes but with different periods.
- Slices 21-23 contain static "pin grids".
- Slices 27-29 contain pin grids that increase linearly in intensity.

The intensity evolution for exponentially decaying features are generated from a model of the form of equation D.1, the intensity evolution for sinusoidally oscillating features are generated from a model of the form of equation D.2, and the intensity evolution for linearly evolving features are generated from a model of the form of equation D.3.

$$I(t) = Ae^{-t/B} + C \tag{D.1}$$

$$I(t) = A\sin\left(\frac{2\pi}{B}\left[t - C\right]\right) + D \tag{D.2}$$

$$I(t) = \left(\frac{t}{t_{max}}\right) A + B \tag{D.3}$$

Where $C(t)$ represents the feature intensity at sampling time t, and t_{max} is the maximum sampling time in the simulation. The vector of sampling times described in section D.1.3 is used to generate vectors containing the temporal evolution information for each feature.

D.1.5 Sample K-Space of Temporally Evolving Phantom

At each sampling time, specified by the vector generated in section D.1.3, the synthetic phantom is assembled with the appropriate feature intensity evolutions. If specified by the user, Rician noise is added to the phantom image. For each simulated coil used in k-space sampling, the phantom image is weighted by the appropriate coil sensitivity map (see section 3.4). The Fourier transform of the resulting phantom image is then taken, and the resulting k-space sampled with the appropriate CIRCUS pattern for the current iteration. The sampling acquires a calibration region and then an effective CIRCUS quanta (pregenerated as described in section D.1.2), and repeats this pattern until the number of pairs specified by the user for sampling is met. If the number of CIRCUS pairs to be acquired exceeded the number of pre-generated effective CIRCUS quanta, the algorithm simply recycles the quanta patterns. Sampling of k-space is specified by the user to be either "snapshot" sampling or "realistic" sampling.

1. **Snapshot sampling** of CIRCUS patterns involves setting TR to zero seconds, such that each of the CIRCUS patterns (i.e. the individual calibration regions and quanta) are sampled instantaneously. A temporal lag, controlled by a parameter *dTimePerfect*, is set to allow temporal evolution between the snapshot samples.

2. **Realistic sampling** of CIRCUS patterns utilizes a nonzero TR. One line of k-space is sampled every TR seconds, where the phase encode coordinates are selected from the center of k-space outwards and frequency encoding is assumed to be instantaneous.

The output of the sampling process is a 5D array, containing 3D sampled k-space data for each coil and for each of the CIRCUS pairs. The intial, mean, and final

sampling times for each CIRCUS pattern are also returned as vectors[i], allowing the user to weight the effective sampling times for each CIRCUS pattern as they desire.

D.1.6 Interleave Sampled K-Space Data

The individually sampled k-space volumes from section D.1.5, time stamped with an effective sampling time (either time of first sampling, time of end sampling, or mean sampling time) are passed to the interleaving algorithm. The number of calibration regions N_{cal} and the number of effective quanta N_{quan} specified by the user for interleaving are combined by summing the k-space volumes, averaging any overlapping k-space points according to the number of overlaps that occurred at each location.

The interleaving algorithm is summarized in Figure D.5, taking $N_{cal} = 2$, $N_{quan} = 3$ and a jump of 1 calibration region between "anchor" calibration regions as an example. The interleaving algorithm "anchors" on calibration regions (marked with coloured circles) and begins by seeking out the nearest N_{cal} calibration regions relative to the current anchor calibration region. Following this, the nearest N_{quan} effective CIRCUS quanta are found. During both of these searches, the algorithm looks symmetrically on either side of the anchor calibration region, wrapping at the boundaries of acquired k-space volumes as needed (e.g. blue interleaved volume in Figure D.5). The $N_{cal} + N_{quan}$ sampled volumes are then interleaved into a new k-space volume. The effective time $\bar{t}_i$ of each interleaved k-space volume is calculated to be either the average of the effective sampling times t_i of all the constituent k-space volumes (i.e. equal weighting between calibration regions and effective quanta) or the average of the effective sampling times of only the constituent calibration regions. For example, the effective sampling time options of I_1 in Figure D.5 would give:

$$\begin{aligned} (\bar{t}_1)_{equal} &= \frac{1}{5}(t_1 + t_2 + t_3 + t_4 + t_6) \\ (\bar{t}_1)_{calib} &= \frac{1}{5}(t_1 + t_3) \end{aligned} \tag{D.4}$$

The output of the interleaving algorithm is a 5D array containing 3D k-space volumes for each coil and each of the interleaved volumes, as well as a vector of the effective sampling time for each of the interleaved volumes.

[i]Note that for snapshot sampling, these will all be equivalent.

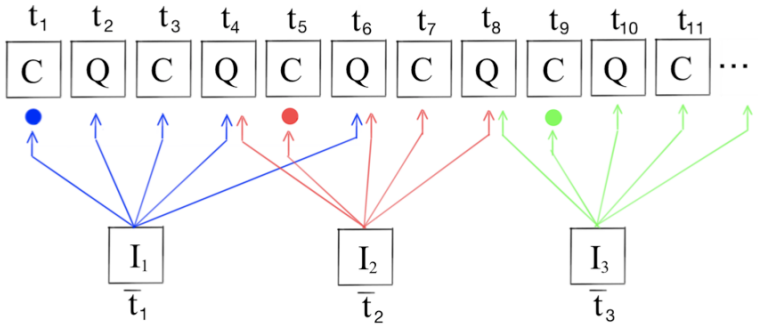

Figure D.1: The individual k-space volumes are interleaved prior to CS reconstruction. The algorithm anchors on calibration regions (shown with coloured circles), jumping the anchor point by a user specified amount at each iteration (jump of 2 shown in the figure). The algorithm then searches for the nearest N_{cal} calibration regions and N_{quan} effective quanta to interleave into the final k-space volume I_i. The effective time $\bar{t}_i$ of each interleaved volume is calculated from either an equal contribution of constituent t_i, or from t_i taken only from constituent calibration regions.

D.1.7 Perform CS Image Reconstructions

Coil sensitivity profiles for the appropriate number of coils (through knowledge of the size of the interleaved k-space array) are generated. The k-space data and coil sensitivity profiles are then passed to BART for CS reconstruction using the regularization method and weight specified by the user. The output is a 4D array containing reconstructed 3D images for each of the corresponding effective sampling times.

D.1.8 Calculate Objective IQM Scores

For each reconstructed image, the algorithm seeks the sampled time t (stored in a vector as described in section D.1.3) that most closely matches the effective sampling time of the reconstructed image. The synthetic phantom at time t is then generated, and serves as the reference image. The RMSE, GMSD, SSIM, MS-SSIM, and IW-SSIM are then calculated between the reconstructed image and the reference image, where each IQM is implemented as described in section **??**.

D.1.9 Recover Quantitative Parameters

For each voxel, the reconstructed time series is extracted. The appropriate temporal model for the feature (equations D.1 to D.3) is then fit to the data using a least-squares fitting routine (specifically the *lsqcurvefit* routine from the MATLAB optimization toolbox).

D.2 Quantitative Parameter Recovery Errors

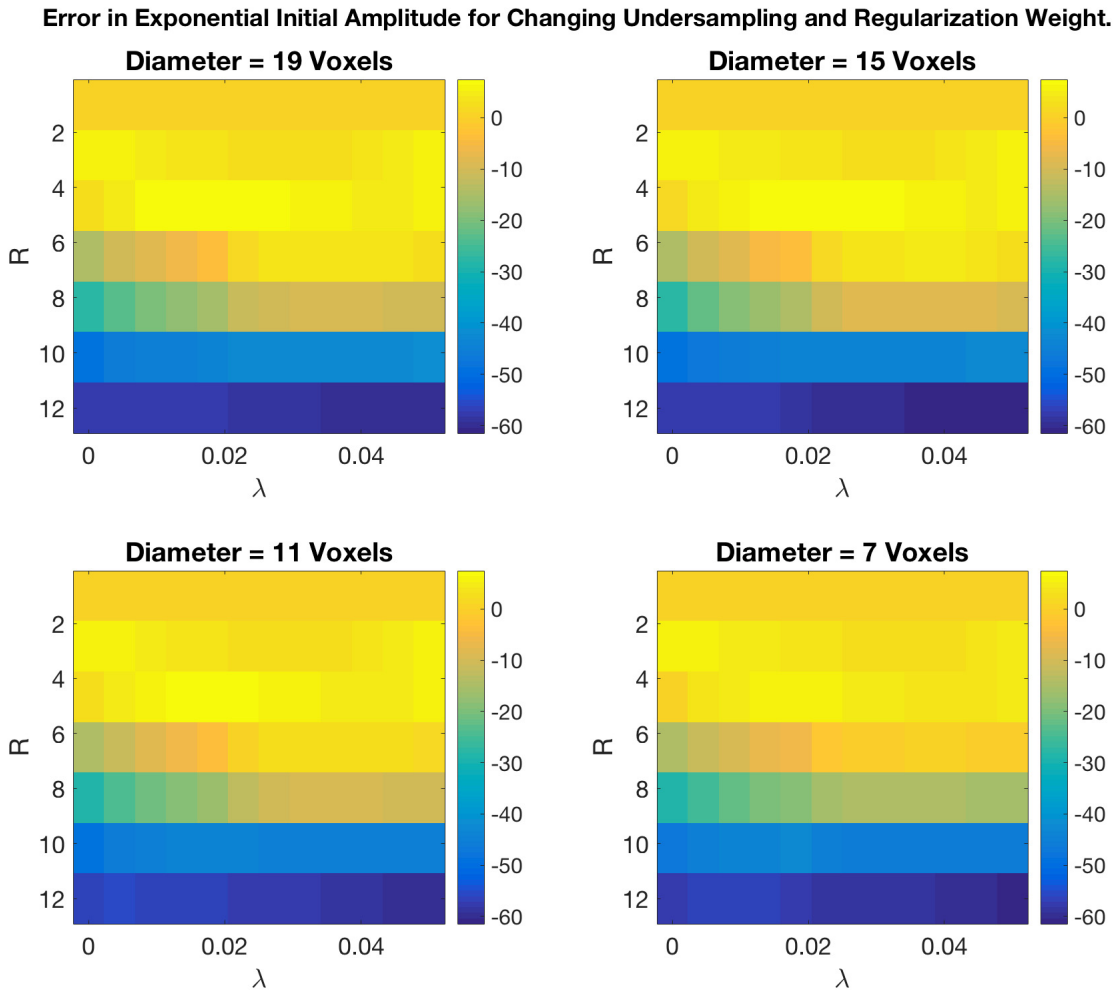

Figure D.2: The change in mean error for the recovered initial magnitude parameter in the exponentially decaying features across each embedded cylindrical feature, as undersampling factor and regularization weight increase.

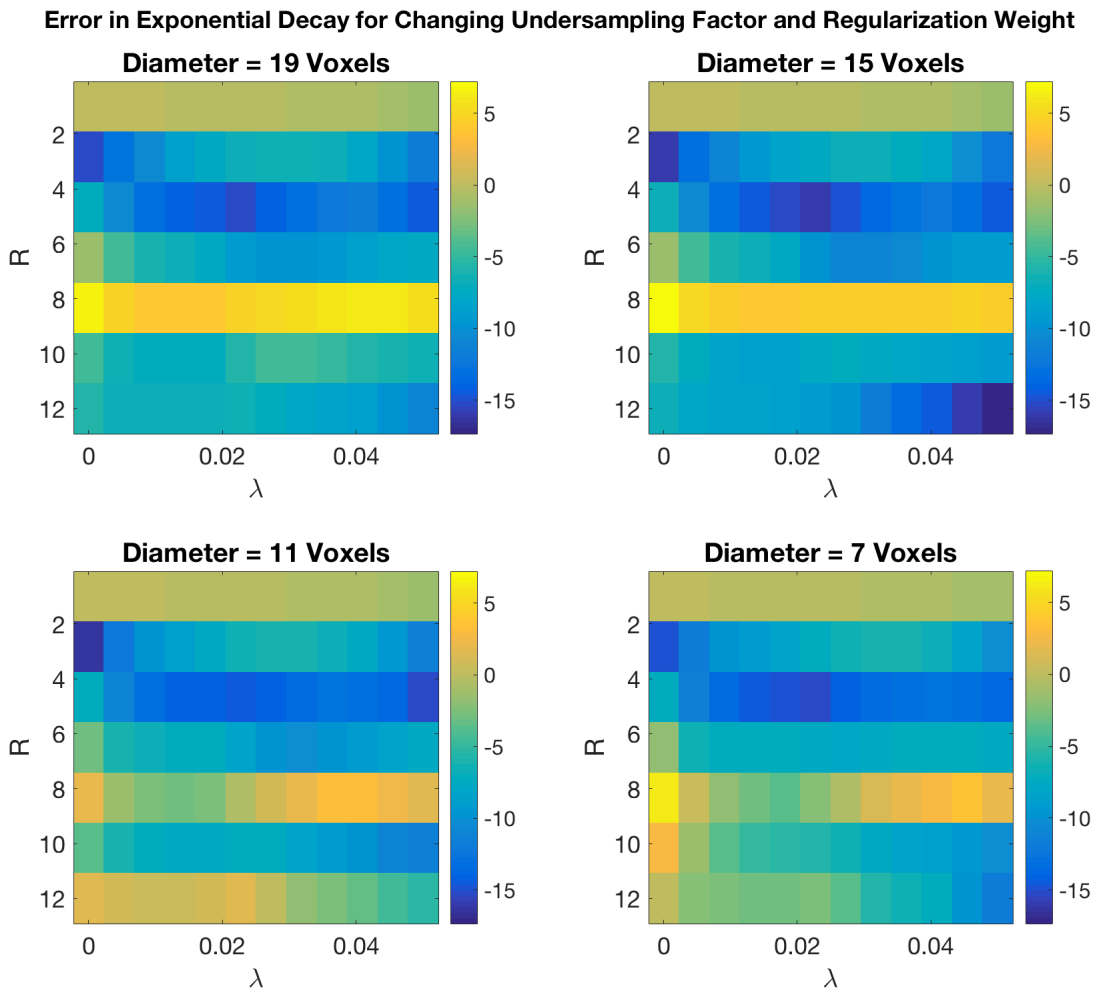

Figure D.3: The change in mean error for the recovered decay parameter in the exponentially decaying features across each embedded cylindrical feature, as undersampling factor and regularization weight increase.

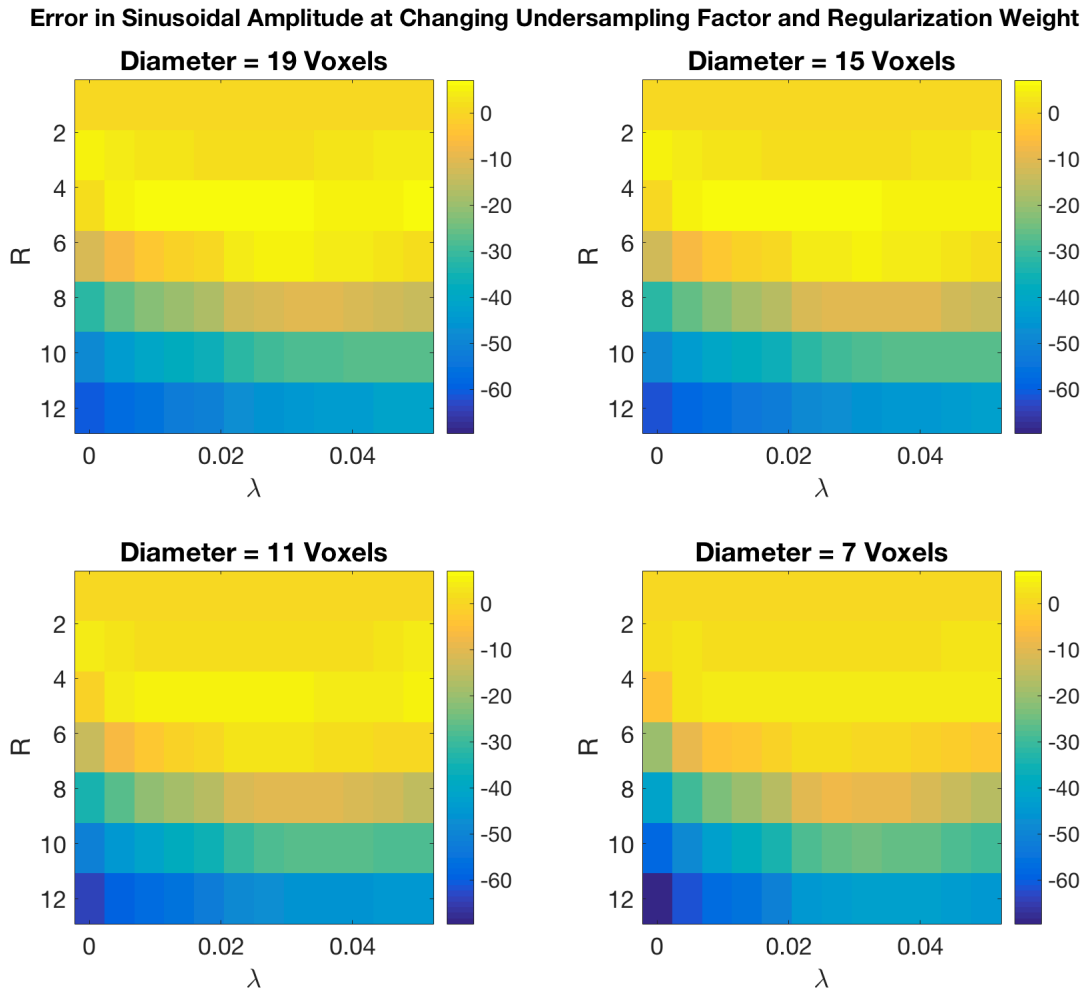

Figure D.4: The change in mean error for the recovered amplitude parameter in the sinusoidally evolving features across each embedded cylindrical feature, as undersampling factor and regularization weight increase.

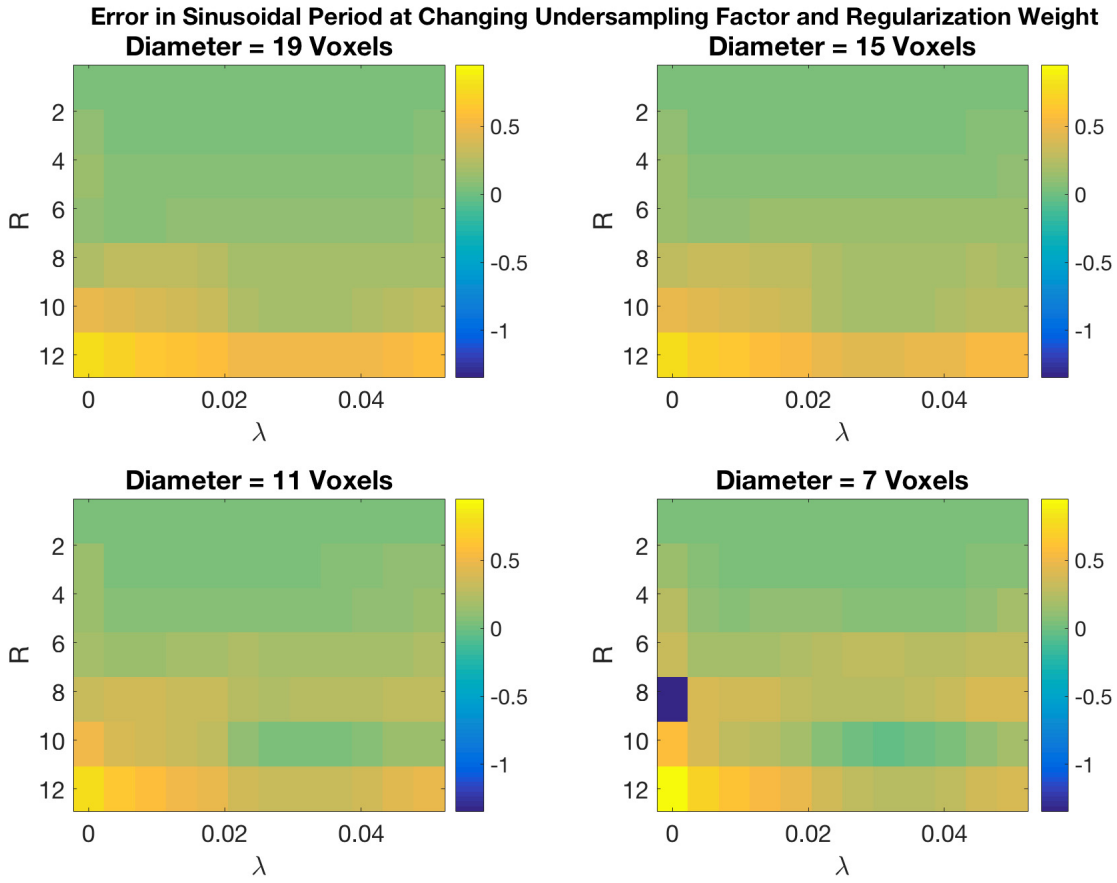

Figure D.5: The change in mean error for the recovered period parameter in the sinusoidally evolving features across each embedded cylindrical feature, as undersampling factor and regularization weight increase.

Bibliography

[1] T. Ai, J. N. Morelli, X. Hu, D. Hao, F. L. Goerner, B. Ager, and V. M. Runge. A historical overview of magnetic resonance imaging, focusing on technological innovations. *Investigative radiology*, 47(12):725–741, Dec 2012.

[2] P. S. Tofts. Modeling tracer kinetics in dynamic gd-dtpa mr imaging. *Journal of magnetic resonance imaging : JMRI*, 7(1):91–101, Jan-Feb 1997.

[3] S. P. Sourbron and D. L. Buckley. On the scope and interpretation of the tofts models for DCE-MRI. *Magnetic resonance in medicine*, 66(3):735–745, Sep 2011.

[4] S. Verma, B. Turkbey, N. Muradyan, A. Rajesh, F. Cornud, M. A. Haider, P. L. Choyke, and M. Harisinghani. Overview of dynamic contrast-enhanced MRI in prostate cancer diagnosis and management. *AJR.American journal of roentgenology*, 198(6):1277–1288, Jun 2012.

[5] C. M. Hoeks, J. O. Barentsz, T. Hambrock, D. Yakar, D. M. Somford, S. W. Heijmink, T. W. Scheenen, P. C. Vos, H. Huisman, I. M. van Oort, J. A. Witjes, A. Heerschap, and J. J. Futterer. Prostate cancer: multiparametric mr imaging for detection, localization, and staging. *Radiology*, 261(1):46–66, Oct 2011.

[6] C. H. Lee, O. Akin-Olugbade, and A. Kirschenbaum. Overview of prostate anatomy, histology, and pathology. *Endocrinology and metabolism clinics of North America*, 40(3):565–75, viii–ix, Sep 2011.

[7] Kai H. Hammerich, Gustavo E. Ayala, and Thomas M.Wheeler. *Anatomy of the prostate gland and surgical pathology of prostate cancer*. Prostate Cancer. Cambridge University Press, 2008.

[8] M. Lustig, D. Donoho, and J. M. Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic resonance in medicine*, 58(6):1182–1195, Dec 2007.

[9] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly. Compressed sensing mri. *IEEE Signal Processing Magazine*, 25(2):72–82, 2008.

[10] E. J. Candes and M. B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2):21–30, 2008.

[11] E. J. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006.

[12] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.

[13] S. Winkelmann, T. Schaeffter, T. Koehler, H. Eggers, and O. Doessel. An optimal radial profile order based on the golden ratio for time-resolved MRI. *IEEE Transactions on Medical Imaging*, 26(1):68–76, Jan 2007.

[14] J. Liu and D. Saloner. Accelerated MRI with CIRcular cartesian undersampling (CIRCUS): a variable density cartesian sampling strategy for compressed sensing and parallel imaging. *Quantitative Imaging in Medicine and Surgery*, 4(1):57–67, Feb 2014.

[15] Ieee standard for floating-point arithmetic, 2008.

[16] H. Wang, Z. Su, H. Ye, X. Xu, Z. Sun, L. Li, F. Duan, Y. Song, T. Lambrou, and L. Ma. Reproducibility of dynamic contrast-enhanced MRI in renal cell carcinoma: A prospective analysis on intra- and interobserver and scan-rescan performance of pharmacokinetic parameters. *Medicine*, 94(37), Sep 2015.

[17] B. Kumar, G. R. Sinha, and K. Thakur. Quality assessment of compressed mr medical images using general regression neural network. *International Journal of Pure and Applied Sciences and Technology*, pages 158–169, 2011.

[18] R. Kumar and M. Rattan. Analysis of various quality metrics for medical image processing. *International Journal of Advanced Research in Computer Science and Software Engineering*, 2(11):137–144, 2012.

[19] Jos V. Manjn, Pierrick Coup, Antonio Buades, D. Louis Collins, and Montserrat Robles. New methods for mri denoising based on sparseness and self-similarity, January 2012 2012.

[20] Y. Suzuki, A. Kunimatsu, K. Kamiya, M. Katsura, H. Mori, K. Maruyama, T. Feiweier, K. Ino, Y. Watanabe, J. Sato, K. Yano, and K. Ohtomo. Introducing a structural similarity index and map for quality control in tractography performed using multi-band epi. *Proceedings of the International Society for Magnetic Resonance in Medicine*, 24:2074, 2016.

[21] Jun Miao, Donglai Huo, and David L. Wilson. Quantitative image quality evaluation of mr images using perceptual difference models. *Medical physics*, 35(6Part1):2541–2553, 2008.

[22] T. Akasaka, K. Fujimoto, T. Yamamoto, T. Okada, Y. Fushumi, A. Yamamoto, T. Tanaka, and K. Togashi. Optimization of regularization parameters in compressed sensing of magnetic resonance angiography: Can statistical image metrics mimic radiologists' perception? *PloS one*, 11(1):e0146548, Jan 8 2016.

[23] Li Sze Chow and Raveendran Paramesran. Review of medical image quality assessment. *Biomedical Signal Processing and Control*, 27:145–154, 5 2016.

[24] Haacke E.M., Brown R.W., Thompson M.R., and Venkatesan R. *Magnetic Resonance Imaging Physical Principles and Sequence Design*. John Wiley & Sons Inc., 1999.

[25] Katherine L. Wright, Jesse I. Hamilton, Mark A. Griswold, Vikas Gulani, and Nicole Seiberlich. Non-cartesian parallel imaging reconstruction. *Journal of Magnetic Resonance Imaging*, 40(5):1022–1040, 2014.

[26] L. Feng, R. Grimm, K. T. Block, H. Chandarana, S. Kim, J. Xu, L. Axel, D. K. Sodickson, and R. Otazo. Golden-angle radial sparse parallel MRI: Combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 72(3):707–717, Sep 2014.

[27] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger. Sense: sensitivity encoding for fast mri. *Magnetic resonance in medicine*, 42(5):952–962, Nov 1999.

[28] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase. Generalized autocalibrating partially parallel acquisitions (grappa). *Magnetic resonance in medicine*, 47(6):1202–1210, Jun 2002.

[29] L. Feng. *Rapid and Continuous Magnetic Resonance Imaging Using Compressed Sensing*. PhD thesis, 2015.

[30] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok. *Introduction to Compressed Sensing*, chapter 1, pages 1–65. Compressed Sensing: Theory and Applications. Cambridge University Press, 2012.

[31] 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.

[32] Ingrid Daubechies. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1992.

[33] F. Ong and M. Lustig. Beyond low rank + sparse: Multiscale low rank matrix decomposition. *IEEE Journal of Selected Topics in Signal Processing*, 10(4):672–687, 2016.

[34] W. Xue, L. Zhang, X. Mou, and A. C. Bovik. Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. *IEEE Transactions on Image Processing*, 23(2):684–695, 2014.

[35] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

[36] H. Wang, E. P. Simoncelli, and A. C. Bovik. Multi-scale structural similarity for image quality assessment. In *Proc. IEEE Asilomar Conf. Signals, Syst., Comput.*, pages 1398–1402, 2003.

[37] Z. Wang and Q. Li. Information content weighting for perceptual image quality assessment. *IEEE Transactions on Image Processing*, 20(5):1185–1198, 2011.

[38] Zhou Wang and Alan C. Bovik. *Modern Image Quality Assessment*. 2006.

[39] Martin J. Wainwright and Eero P. Simoncelli. Scale mixtures of gaussians and the statistics of natural images. In *Proceedings of the 12th International Conference on Neural Information Processing Systems*, NIPS'99, pages 855–861, Cambridge, MA, USA, 1999. MIT Press.

[40] P. Burt and E. Adelson. The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31(4):532–540, 1983.

[41] H. R. Sheikh and A. C. Bovik. Image information and visual quality. In *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pages iii–709–12 vol.3, 2004.

[42] Manojkumar Saranathan, Dan W. Rettmann, Brian A. Hargreaves, Sharon E. Clarke, and Shreyas S. Vasanawala. Differential subsampling with cartesian ordering (disco): A high spatio-temporal resolution dixon imaging sequence for multiphasic contrast enhanced abdominal imaging. *Journal of Magnetic Resonance Imaging*, 35(6):1484–1492, 2012.

[43] W. T. Dixon. Simple proton spectroscopic imaging. *Radiology*, 153(1):189–194, 10/01; 2017/06 1984.

[44] Martin Uecker, Frank Ong, Jonathan Tamir, Dara Bahri, Patrick Virtue, Joseph Cheng, Tao Zhang, and Michael Lustig. Berkeley advanced reconstruction toolbox. *Proceedings of the International Society for Magnetic Resonance in Medicine*, page 2486, 2015.

[45] G. h. Chen, C. l. Yang, and S. l. Xie. Gradient-based structural similarity for image quality assessment. In *2006 International Conference on Image Processing*, pages 2929–2932, 2006.

[46] A. K. Moorthy and A. C. Bovik. Blind image quality assessment: From natural scene statistics to perceptual quality. *IEEE Transactions on Image Processing*, 20(12):3350–3364, 2011.

[47] M. A. Saad, A. C. Bovik, and C. Charrier. Blind image quality assessment: A natural scene statistics approach in the dct domain. *IEEE Transactions on Image Processing*, 21(8):3339–3352, 2012.

[48] Jeffrey P. Woodard and Monica Carley-Spencer. No-reference image quality metrics for structural mri. *Neuroinformatics*, 4(3):243–262, 2006.

[49] H. Gudbjartsson and S. Patz. The rician distribution of noisy mri data. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 34(6):910–914, Dec 1995.