SPARSE MRI

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Michael Lustig
September 2008

© Copyright by Michael Lustig 2009
All Rights Reserved

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(John Pauly) Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(David Donoho)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Dwight Nishimura)

Approved for the University Committee on Graduate Studies.

Abstract

Magnetic Resonance Imaging (MRI) is a non-invasive imaging modality. Unlike Computed Tomography (CT), MRI does not use ionizing radiation. In addition, MRI provides a large number of flexible contrast parameters. These provide excellent soft tissue contrast. Over the years, MRI has improved dramatically in both imaging quality and imaging speed. This revolutionized the field of diagnostic medicine. However, imaging speed, which is essential to many of the MRI applications remains a major challenge.

Imaging speed can be improved by faster collection of data. This can be achieved by using sophisticated non-Cartesian *k*-space trajectories. One of the design challenges is to minimize the gradient waveform duration, subject to both hardware and sequence constraints. Existing methods provide solutions limited to specific trajectories, or solutions which are sub-optimal. Based on optimal control theory, a method for designing gradient waveforms for arbitrary *k*-space trajectories is developed. It is non-iterative, computationally efficient and provides the time-optimal waveforms that traces *k*-space trajectories as fast as possible within the hardware limits.

With current hardware and sequence design methods, a point has nearly been reached in which fundamental physical and physiological effects limits the ability to simply encode data more quickly. This fundamental limit has led many researchers to look for methods to reduce the amount of acquired data without degrading image quality. MR image data are often highly redundant, which can be exploited to reduce the amount of acquired data, and hence the scan time. In this work, a method that exploits the inherent compressibility of MR images is developed. It is based on the recent theory of compressed sensing (CS). In compressed sensing, the data are implicitly compressed in the acquisition process by obtaining fewer, so called, incoherent measurements. This is achieved by various randomized k-space sampling schemes. Images are accurately reconstructed from these measurements using a non-linear recovery processes that enforces the compressibility of the image. As a result, for some applications the scan time can be accelerated up to an order of magnitude.

SPIR-iT is an iTerative Self-consistent Parallel Imaging Reconstruction method. It is auto-calibrating and does not require explicit estimates of the coil sensitivity maps. SPIR-iT formulates the parallel imaging reconstruction through data consistency constraints. It is a general, optimal solution for coil-by-coil parallel imaging from arbitrary k-space trajectories. It is also a general framework for easily incorporating additional image priors, and in particular sparsity/compressibility constraints for combining parallel imaging with compressed sensing.

Acknowledgments

To say the least, I do not particularly like writing. But writing this part of my thesis offers me great pleasure. It is a great opportunity to thank all those people that assisted and supported me through the last few years.

I could not have hoped for a better adviser than Prof. John Pauly. He is the kindest most gentle person I know. John has given me the tools, the means, the guidance, and the freedom to pursue my research directions. His door was always open and no matter how busy he was, he always had the patience to answer all my (sometimes too many) questions. John, when I grow up, I want to be just like you!

I have been fortunate to have had the honor to work closely with Prof. David Donoho. In fact, most of the content of this thesis is based on Dave's own work. Dave has the soul of a mathematician with a vision of an engineer. It amazes me to see the impact his work has had in so many disciplines. Dave has put days and nights providing me with answers and advice, and passionately editing my papers. For that I am for ever in debt.

My first encounter with Prof. Dwight Nishimura was just before I decided to come to Stanford. It was his decision to take me as a TA for his class that made my coming to Stanford possible. I will never forget that. It was from Dwight that I learned most of what I know about MRI. Dwight is a phenomenal teacher with an elegant, systematic and engaging teaching style. He is a man of few words, but each one is worth gold. It was a pleasure to know Dwight better and better over the years. Now that I learned to tune down closer to his level of subtlety I can appreciate him even more.

Steve Conolly has been a great resource for advice and encouragement. He is the

MRSRL psychologist and historian. He always makes sure that everyone are on the path, happy, and know the history of the lab. Steve is an amazingly creative inventor, which inspired me to attempt to be one as well....

Prof. Al Macovski founded and built MRSRL. His fingerprints are present everywhere and I can see a lot of Al in John, Dwight and Steve. I thank Al for providing this amazing environment and for being so positive on just about everything.

I thank Prof. Brian Wandell for all his help and support in my research and career, and for taking the time to serve as the chair of my Ph.D. Oral Committee. Brian is a true scientist, always skeptical, methodical, charismatic, organized, and with amazingly diverse interests. He is a model of how a great scientist should be.

I owe a great deal of gratitude to Marianne Marx. Marianne took a chance on me when I needed the help. She gave me the opportunity to teach and thus the opportunity to come to Stanford. Over the years, our relationship got closer and she became like my grandmother here at Stanford. I thank Marianne for her love and support. I could not have done this without her.

MRSRL has been a unique and exciting place to be at, mostly because of the people in it. Looking back, it still amazes me how much I learned in this period of my life.

From Chuck Cunningham I learned that RF pulses are the coolest thing in the world, dude! His enthusiasm was contagious, and so his keen to flipping roots of polynomials, which I acquired as well.

Juan Santos is THE real-time MRI guy. His knowledge of the internals of the MR system, programming languages and operating systems is second to none. His rtHawk real-time MR platform was an enabler to a large portion of my work. Juan has also been a great friend outside of the lab. We shared many things in common, among them having twins, a nanny and the love for Chile. I thank him and his dear wife Carolla for all the help, love, and support they have given me and my family.

Jin Hyung Lee has been a good friend, office-mate and mentor. As office-mates Jin patiently endured my sometimes childish behavior (sorry about those glue bubble balloons...). As a mentor, she helped me getting started when I joined the group. On the day of the deadline of my first abstract submission, Jin stayed all night to help me with my paper, ... even though I had a "Chutzpa" to leave the submission to her, while I was enjoying a Seal concert.

From Brian Hargreaves I not only learned how to write papers, disclosures, make presentations, write EPIC code and many other things, but I also learned how to do them the right way. Brian is a great educator and the resources he establish are priceless. I thank Brian for his support in my research and the many hours of rigorously editing my papers and abstracts.

Logi Vidarson and I became friend soon after I joined the lab. It is hard not to like this giant Icelander with the slightly twisted sense of humor. Life seems happier when Logi is around.

Peder E.Z. Larson is the only person I know with three initials in his name. Peder and I have grown to be close friend, sharing many great moments in and outside of the office. Playing Quals-Ball, climbing mountains, conference roommates, projects partners, joint papers, the works!

From Phil Beatty I learned all I know about non-Cartesian reconstruction and parallel imaging. Phil and I share an enthusiasm to image reconstruction, and this sparked many intense and engaging discussions from which I have learned so much.

I have always enjoyed spending time with Daeho Lee. If it was a conversation about research, computers, life, or just frequent relaxed procrastination breaks. Daeho has the highest SNR I've seen in a person. My theory is that he is so calm, that he lacks any body noise.

Nikola Stikov is an amazing guy. He has been a good friend to me from the start, even before I joined the group. We went through a lot together. He was my TA, co-TA, my only quals partner, a labmate and an office-mate. I thank him for all he has done for me and for being such a great friend.

I would like to thank the many other former and current MRSRL members that made my time ever more enjoyable and from which I learned so much. When I joined the lab Karla Miller was literally the BOSS. Karla helped me quite a bit to get started with research. Neal Bangerter was always a source of great conversations about life in general and SSFP in particular. From Jongho I learned uncompromising research ethics. Bill Overall was always available for good, almost daily discussions. Ross

Venook had the great quality for filling up those frequently occurring uncomfortable silences with funny stories. I thank Sharon Ungersma for having those wonderful semi-annual parties. Pascal Stang for showing me that you can be the smartest person, but still fail quals....and it's OK. Adam Kerr was an endless source of information. I thank him for great enthusiastic conversation about kids, RF pulses, parallel excitations and wonderful criticism. Greig Scott showed me how to be curious. Krishna Nayak has been my role model for how to start a career. Paul Gurney has been a partner for many late-night discussions. Emine Saritas has been a fun and (very) patient office-mate. Tolga Cukur showed me that paper writing should really take only a week, I greatly enjoyed working with him for a couple of weeks. Uygar Sumbul provided entertainment with the funniest group talks. Hochong Wu showed me that circles are also trajectories. Nate Matter has been a good friend, providing constant supply of humor and great New-Mexico smoked chipotle powder, which I'm addicted to.

The clinicians in the group I worked with, Bob Hu, Garry Gold, and Shreyas Vasanawala, provided me with valuable feedback. They showed me what are the important directions I should focus on, and gave true meaning to my research. They are exceptional people, always open to new ideas and eager to try them in practice.

I would also like to thank my academic collaborators outside of MRSRL. I thank Seung Jean Kim for all the time he spent with me discussing various problems, and the fruitful research that resulted from it. I thank Simon Hu and Dan Vigneron from UCSF for making some of my ideas actually work in practice and made into good use. I thank Yaki Tsaig for the many conversations about compressed sensing. I would also like to thank Marc Alley for helping me with data, implementations and now, as a neighbor across the street, with adjusting to life outside campus. Marc is an amazing human Matlab to C++ cross compiler!

I thank Stephen Boyd for teaching me all I know about linear systems and optimization. Stephen is by far the most engaging and exciting teacher I have had the pleasure of taking classes from. I hope one day, that I will be able to excite my students as much as he excited me.

I would like to thank Jim Pipe for his friendship and support. I thank him for taking me under his wing, promoting my work and giving me the opportunity to present it to others in many occasions.

I would like to thank Nimrod Peleg, Prof. David Malah, Prof. Ron Kimmel and Prof. Freddy Bruckstein from the Technion. They gave me the necessary background, encouragement, support and opportunity to embark on this adventure. I would like to give special thanks to Prof. Michael (Miki) Elad. Miki is not only a good friend, but also my academic "father". Many of my career and research choices has been influenced by Miki.

While at Stanford I have collected many good friends who made my Stanford experience feel more like home, even though my real home is 10,000 miles away. I thank Mattan Erez and Dana Porrat, Eric Setton, Tommer Yahalom, my best friend Daniel Ramot and his wife Kelly, Leith and May-Ling the eternal neighbors, Michal Maimaran and Yuval Salant. I would like to give special thanks to Rafi Cohen and Maya Meidan who have been a second family to us.

My very best friend and partner Yuval Cassuto has been a precious remote companion, going through similar experiences 360 miles to the south. Yuval, his wife Ruth, and their lovely children Assaf and Gili have been family to us all these years.

My very good friend since high-school, Keren Ziv, has been an endless source of moral support. The "optimistic square for hard times" she left on my desk has gotten me through the toughest times. Keren is about to embark on a similar adventure as a postdoc at Stanford. I hope her experience will be as good as mine.

I couldn't have done any of this without the love and support I got from my family, my mom Ahuva, my dad Yosi, and my younger (yet taller) brother Gidi. They have always supported me, encouraged me and provided me with the means to pursue the future I want. Even though they were physically far away, it seemed so they were here with me all the time.

I would like to thank Yonit's family who accepted me as an integral part of their family. I thank Janny, the best mother in-law a man could hope for, Yoav and Yaron the funniest most loving brothers in-law and Yifat, the greatest sister in-law in the world. I thank them for their trust in me and their patience. Yes Yifat, I'm finally done!

My dear, lovely, smart and beautiful wife Yonit is the best thing that ever happened me. I can not imagine my life without her. She is truly the love of my life. Thank you for following me here, thank you for being so patient, thank you for giving me such loving support, thank you for giving me Nogah and Neta-Li, the loveliest most beautiful twin girls in the world, and thank you for making my life so happy.

Finally I would like to thank my late father in-law Yair for trusting me with the hand of his precious daughter, even though I was not tall enough. I thank him for letting me take her so far away at such a difficult time. My pursuit of medical imaging was greatly influenced by his courageous fight against cancer. I dedicate this thesis to him.

Stanford University MICHAEL LUSTIG
Aug. 23, 2008

Contents

Abstract **v**

Acknowledgments **vii**

1 Introduction **1**

1.1 Thesis outline 3

2 Principles of Magnetic Resonance Imaging **5**

2.1 Nuclear Magnetic Resonance Physics 5

2.1.1 Polarization 5

2.1.2 Bloch Equation 6

2.1.3 Resonance 6

2.1.4 Relaxation 7

2.2 Magnetic Resonance Imaging Hardware 7

2.2.1 The Static Magnetic Field B_0 7

2.2.2 Transverse Radio-Frequency Field B_1 7

2.2.3 Receive Coils and Signal Reception 8

2.2.4 Spatial Encoding Gradients 8

2.3 Imaging 9

2.3.1 Excitation and Selective Excitation 9

2.3.2 Spatial Encoding and k-Space 9

2.3.3 Image Acquisition 11

2.3.4 Resolution and Field of View 12

2.3.5 k-space sampling trajectories 12

2.3.6 Rapid Imaging 12

3 Time Optimal Gradient Design 17

3.1 Introduction 17
3.2 Curve Parametrization 18
3.3 The Time-Optimal Gradient Waveform Design Problem 20
3.4 Formulation in the Arc-Length Parametrization 22
3.5 The Time-Optimal Solution in the Phase-Plane 28
3.6 The Time-Optimal Gradient Waveform Solution 30
3.7 Additional Constraints 31
3.8 Examples 31
3.8.1 A Line in k-space 33
3.8.2 Dual Density Spiral Design 33
3.8.3 Time-Optimal Rosette Trajectory Design 34
3.8.4 Randomized Trajectories for Compressed Sensing Application 35
3.9 Discussion 35
3.9.1 Numerical Issues 36
3.9.2 Computational Complexity 36
3.10 Conclusions 36

4 Compressed Sensing MRI 37

4.1 Introduction 37
4.2 Compressed Sensing 39
4.2.1 A simple, intuitive example of compressed sensing 39
4.3 Sparsity 41
4.3.1 Sparsifying Transform 43
4.3.2 The Sparsity of MR Images 44
4.4 Incoherent Sampling 44
4.4.1 PSF and TPSF Analysis 46
4.4.2 Single-slice 2DFT, multi-slice 2DFT and 3DFT Imaging . . . 49
4.4.3 Variable Density Random Undersampling 51
4.4.4 Other Incoherent Sampling Schemes 52

4.4.5 Dynamic Incoherent Sampling: k-t Sparse 55
4.4.6 How Many Samples to Acquire? 55
4.4.7 Monte-Carlo Incoherent Sampling Design 57
4.5 Image Reconstruction 57
4.5.1 Formulation 57
4.5.2 Non-Linear Conjugate-Gradient Solution 59
4.5.3 Low-Order Phase Correction and Phase Constrained Partial *k*-Space 61
4.6 Methods 62
4.6.1 Simulation 62
4.6.2 Undersampled 2D Cartesian Sampling in the Presence of Noise 64
4.6.3 Multi-slice 2DFT Fast Spin-Echo Brain Imaging 64
4.6.4 Contrast-Enhanced 3D Angiography 65
4.6.5 Variable Density Spirals Whole Heart Coronary Imaging . . . 66
4.6.6 k-t Sparse: Application to Dynamic Heart Imaging 68
4.7 Results 68
4.7.1 Simulation 68
4.7.2 2DFT CS Reconstruction in the Presence of Noise 69
4.7.3 Multi-slice Fast Spin-Echo Brain Imaging 72
4.7.4 Contrast Enhanced 3D Angiography 72
4.7.5 Variable Density Spirals, Whole Heart Coronary Imaging . . . 74
4.7.6 k-t Sparse: Application to Dynamic Heart Imaging 74
4.8 Discussion 78
4.8.1 Computational Complexity 78
4.8.2 Reconstruction Artifacts 78
4.8.3 Relation to Other Acceleration Methods 80
4.9 Conclusions 81

5 SPIR-iT **83**
5.1 Introduction 83
5.2 SENSE 85

5.3 Cartesian GRAPPA 86
5.4 SPIR-iT: Self Consistency Formulation 87
    5.4.1 Consistency with the Calibration 89
    5.4.2 Consistency with the Data Acquisition 90
    5.4.3 Constrained Optimization Formulation 90
5.5 Arbitrary Cartesian Sampling 92
5.6 Non-Cartesian Sampling 93
    5.6.1 Calibration 95
    5.6.2 *k*-Space Domain Reconstruction 95
    5.6.3 Image Domain Reconstruction 101
5.7 Off Resonance Correction 103
5.8 Regularization 103
5.9 Results 104
    5.9.1 Noise and Artifacts Measurements 104
    5.9.2 Arbitrary Cartesian Sampling 105
    5.9.3 *k*-Space Based Non-Cartesian Reconstruction 105
    5.9.4 Image Space Based Non-Cartesian Reconstruction 107
    5.9.5 Off Resonance Correction 109
    5.9.6 Total Variation Regularization 109
5.10 Conclusions 113

6 Summary and Recommendation 115

A Derivations for the Time-Optimal Gradient Waveform Design 117
A.1 Arc Length Expression of the Gradient Constraints 117
A.2 Optimality of the Gradient Waveform Design 118

Bibliography 119

List of Tables

3.1 Outline of the time-optimal gradient design algorithm. 23

4.1 Outline of the non-linear CG algorithm 60

5.1 Properties of some parallel imaging reconstruction methods. 85
5.2 A POCS algorithm for SPIR-iT from arbitrary sampling on a Cartesian grid. 93

List of Figures

2.1 The magnetic fields used in MR imaging: The main homogeneous magnetic field B_0 creates a net magnetization that precesses at a resonance frequency $\frac{\gamma}{2\pi}B_0$. The transverse rotating radio-frequency field B_1 is used for exciting the magnetization. The gradient fields G (only G_x is illustrated) are used for spatial encoding. 8

2.2 Fourier spatial encoding. The temporal MRI signal directly samples the spatial frequency domain of the image. Gradient fields cause a linear frequency distribution across the image, which produces a linear phase accrual with time. The received signal samples are spatial frequencies of the image. The corresponding spatial frequencies are proportional to the gradient waveform area. The gradient is limited in amplitude, $G_{\max}$, and slew rate, $S_{\max}$, which are both system specific. 14

2.3 Image resolution is determined by the extent of the k-space that is covered. The supported field of view is determined by the sampling density. Violation of the Nyquist criteria results in aliasing interference in the image domain. The appearance of the artifact depends on the sampling. Coherent folding is produced by equispaced sampling and incoherent interference is produced by irregular sampling 15

2.4 Common sampling trajectories. Top, left to right: Cartesian 2D, Cartesian echo-planar, radial, spiral. Bottom left to right: Cartesian 3D, stack of radial, 3D radial, 3D stack of spirals 15

3.1 Properties of the arc-length parametrization of a curve in $\mathbf{R}^3$. 20

3.2 Example of stage I of the design algorithm: Conversion from arbitrary parametrization to arc-length parametrization. 24

3.3 Example of stage II of the design algorithm: Calculation of the time-optimal "velocity" in the phase-plane. 25

3.4 Example of stage III of the design algorithm: Conversion from the arc-length parametrization to the time-optimal parametrization and the calculation of the gradient waveforms. Note that either the gradient magnitude or the slew-rate are maximized at every time-point, which is a necessary condition for time-optimality. 26

3.5 Examples showing several k-space trajectories with their corresponding time-optimal gradient waveforms and gradient magnitude and slew-rate. (a) A line. (b) Dual density spiral. (c) Rosette trajectory. (d) Randomly perturbed variable density spiral. 32

4.1 Illustration of the domains and operators used in the chapter as well as the requirements of CS: sparsity in the transform domain, incoherence of the undersampling artifacts and the need for nonlinear reconstruction that enforces sparsity. 40

4.2 An intuitive reconstruction of a sparse signal from pseudo-random k-space undersampling. A sparse signal (a) is 8-fold undersampled in k-space (b). Equispaced undersampling results in coherent signal aliasing (c) that can not be recovered. Pseudo-random undersampling results in incoherent aliasing (d). Strong signal components stick above the interference, are detected (e) and recovered (f) by thresholding. The interference of these components is computed (g) and subtracted (h), lowering the total interference level and enabling recovery of weaker components. 42

4.3 Transform sparsity of MR images. Fully sampled images (left) are mapped by a sparsifying transform to a transform domain (middle); the several largest coefficients are preserved while all others are set to zero; the transform is inverted forming a reconstructed image (right). 45

4.4 Empirical statistical evaluation of interference in the PSF for pure random sampling. 1D PSF, 2048 samples, 4-fold undersampling. left: Histogram, showing a Gaussian distribution of the off-center point-spread function. right: Quantile-Quantile plot (QQ plot) of the off-center point spread versus standard normal distribution. In such a QQ plot, if samples come from a normal distribution, the plot is linear. In this case, the plot is indeed close to linear, and therefore the sidelobe values have an approximately normal distribution. 48

4.5 (a) The PSF of random 2D k-space undersampling. (b) The wavelet TPSF of random 2D Fourier undersampling. FDWT and IDWT stand for forward and inverse discrete wavelet transform. Wavelet coefficients are band-pass filters and have limited support both in space and frequency. Random k-space undersampling results in incoherent interference in the wavelet domain. The interference spreads mostly within the wavelet coefficients of the same scale and orientation. 50

4.6 Transform point spread function (TPSF) analysis in the wavelet domain. The k-space sampling patterns and the associated TPSF of coarse-scale and fine-scale wavelet coefficients are shown. (a) Random phase encode undersampling spreads the interference only in 1D and mostly within the same wavelet scale. The result is relatively high peak interference. (b) Sampling differently for each slice, *i.e.*, randomly undersampling the $k_y - z$ plane causes the interference to spread to nearby slices and to other wavelets scales and reduces its peak value. (c) Undersampling the phase encode plane, *i.e.*, $k_y - k_z$ spreads the interference in 2D and results in the lowest peak interference. 51

4.7 (a) For natural images, uniform random undersampling often results in coherent interference. This is attributed to large interference of coarse scale wavelet coefficients. (b) Variable density undersampling, reduces the interference in the coarse scales, effectively achieving better coherence properties. 53

4.8 Point Spread Functions (PSF) of various sampling trajectories. (a) Random lines in 2D (b) Random points in 2D, or cross-section of random lines in 3D (c) Radial (d) Uniform spirals (e) Variable density spirals (f) Variable density perturbed spirals 54

4.9 Top left:Dynamic MRI is a multi-dimensional signal with two or three spatial coordinates and time as an additional dimension. Bottom left: Dynamic images have a sparse representation in an appropriate transform domain. Top right: Traditional $k-t$ sequential sampling. Bottom right: Random ordering is an efficient way to incoherently sample the $k-t$ space. 56

4.10 Low-order phase correction greatly improves the reconstruction. (a) Original phantom magnitude and phase images. (b) Reconstruction without linear-phase correction exhibits residual artifacts. (c) Reconstruction with linear-phase correction. 63

4.11 Simulation: Reconstruction artifacts as a function of acceleration. The LR reconstructions exhibit diffused boundaries and loss of small features. The ZF-w/dc reconstructions exhibit an significant increase of apparent noise due to incoherent aliasing, the apparent noise appears more “white” with variable density sampling. The CS reconstructions exhibit perfect reconstruction at 8 and 12 fold (only var. dens.) accelerations. With increased acceleration there is loss of low-contrast features and not the usual loss of resolution. The reconstructions from variable density random undersampling significantly outperforms the reconstructions from uniform density random undersampling. 70

4.12 2DFT CS reconstruction from noisy data. CS reconstruction can perform denoising of the image as well as interference removal by relaxing the data consistency (a) Reconstruction from complete noisy data. (b) ZF-w/dc: The image suffers from apparent noise due to incoherent aliasing as well as noise. (c) CS reconstruction with TV penalty from noisy undersampled data. Consistency RMS error set to 10^{-5}. (d) CS reconstruction with TV penalty from noisy undersampled data. Consistency RMS error set to 0.1. Note interference removal in both (c) and (d) and the denoising in (d). 71

4.13 Multi-slice 2DFT fast spin echo CS at 2.4 acceleration. (a) The CS-wavelet reconstruction exhibits significant resolution improvement over LR and significant suppression of the aliasing artifacts over ZF-w/dc compared to the full Nyquist sampling. (b) CS wavelet reconstructions from several undersampling schemes. The multi-slice approach outperforms the single-slice approach and variable density undersampling outperforms uniform undersampling. (c) The associated undersampling schemes; variable density (top) and uniform density (bottom), single-slice (left) and multi-slice (right). 73

4.14 Contrast-enhanced 3D angiography reconstruction results as a function of acceleration. Left Column: Acceleration by LR. Note the diffused boundaries with acceleration. Middle Column: ZF-w/dc reconstruction. Note the increase of apparent noise with acceleration. Right Column: CS reconstruction with TV penalty from randomly undersampled *k*-space. 75

4.15 Reconstruction from 5-fold accelerated acquisition of first-pass contrast enhanced abdominal angiography. (a) Reconstruction from a complete data set. (b) LR (c) ZF-w/dc (d) CS reconstruction from random undersampling. The patient has a aorto-bifemoral bypass graft. This carries blood from the aorta to the lower extremities. There is a high-grade stenosis in the native right common illiac artery, which is indicated by the arrows. In figure parts (a) and (d) flow across the stenosis is visible, but it is not on (b) and (c). 76

4.16 Single breath-hold whole-heart coronary artery imaging. Left panel: the imaging sequence timing diagram. Right panel: A slice through the volume of the heart showing the right coronary artery (3). The incoherent artifacts of undersampled variable-density spirals (white arrow) appear as noiselike interference in the linear gridding reconstruction (left). These artifacts are suppressed in the CS reconstruction (right) without compromising image quality. The slice shows: (1) Aorta (2) Chest wall (3) Right coronary artery (4) Right ventricle (5) Left ventricle (6) Liver. 77

4.17 Dynamic imaging of quasi-periodic change. Top: Phantom experiment showing a reconstruction from 4-fold undersampling. Bottom: Dynamic acquisition of the heart motion showing a reconstruction from 7-fold undersampling. 79

5.1 Traditional 2D GRAPPA: Missing k-space data are synthesized from neighboring acquired data. The synthesizing kernel depends on the specific sampling pattern in the neighborhood of the missing point. The reconstruction of a point is independent of the reconstruction of other missing points. 88

5.2 Consistency constrained Cartesian reconstruction: Three consistency equations are illustrated. The reconstruction of each point on the grid is dependent on its entire neighborhood. The reconstruction of missing points depends on the reconstruction of other missing points. The consistency equation is independent of the sampling pattern. 91

5.3 Non-Cartesian consistency reconstruction. A calibration consistency equation describing the relation between neighboring points on a Cartesian grid is shown in red. An acquisition data consistency relation between the Cartesian missing points and the non-Cartesian acquired points is shown in blue. These define a large set of linear equations that is sufficient for reconstruction. 94

5.4 Interpolation kernel in k-space. (a) The specification for the kernel compared to a Kaiser-Bessel kernel. The passband ripple defines the allowed intensity variation in the image. The stopband ripple is associated with the error that leaks into the image due to finite kernel size. The transition band defines the required grid oversampling and image support. The effects of the bandpass ripple are shown in b-j. The original test signal is shown in (b). (c)-(d) show the reconstruction errors with ideal interpolator and imperfect calibration. (e)-(g) show result of using a Kaiser-Bessel interpolator. Large image weighting in (e) results in artifact amplification by compensation in (f) and g. (h)-(j) show the result of using a “flat” pass-band interpolator. The post-intensity correction error in (i) and (h) is significantly reduced. . 98

5.5 k-Space based reconstruction. (a) Illustration of the conjugate-gradient algorithm for non-Cartesian consistency constrained reconstruction in k-space. (b) Illustration of the interpolation operator, D, and its adjoint, D^* (c) Illustration of the calibration consistency operator, G and its adjoint, G^*. The notation $\hat{g}^*_{ji}$ stands for an inverted conjugated version of the filter g_{ji}. 100

5.6 Image-space based reconstruction. (a) Illustration of the conjugate gradient algorithm for non-Cartesian consistency constrained reconstruction in image space. (b) Illustration of the non-uniform Fourier transform operator, D, and its adjoint, D^* (c) Illustration of the calibration consistency operator, G and its adjoint, G^*. 102

5.7 Noise measurements statistics of 100 scans. (a) Fully sampled averaged sum-of-squares image and the individual coil images. (b) sum-of-squares of single 1D GRAPPA reconstructed image (c) sum-of-squares of single 1D SPIR-iT reconstruction. (d) the mean 1D GRAPPA reconstruction error showing residual reconstruction artifacts. (e) the mean 1D SPIR-iT reconstruction error showing reduced reconstruction artifacts. (f) The measured noise amplification (g factor) for the 1D GRAPPA reconstruction. (g) SPIR-iT exhibits reduced g factor. . 106

5.8 SPIR-iT from arbitrary Cartesian sampling. (a) the sampling pattern. (b) sum-of-squares zero-filled reconstruction. (c) consistency constrained reconstruction. 107

5.9 k-Space based non-Cartesian SPIR-iT. (a) Reconstruction from fully sampled data. (b) Properties of the k-space interpolator. (c) Gridding reconstruction at 3-fold acceleration. (d) SPIR-iT at 3-fold acceleration. 108

5.10 Dynamic cardiac imaging with dual density spirals. Two phases of a four chamber view of the heart. (a)-(b) Sum-of-squares of gridding reconstruction exhibits coherent (arrows) and incoherent (noise-like) aliasing artifacts. (c)-(d) Both the coherent and incoherent artifacts are removed by SPIR-iT. (e) One out of the three spiral interleaves. . 110

5.11 Dynamic cardiac imaging with dual density spirals and off-resonance correction. Two phases of a short axis view of the heart. (a)-(b) sum-of-squares of gridding reconstruction exhibits coherent (arrows), incoherent (noise-like) aliasing artifacts and blurring due to off-resonance. (c)-(d) SPIR-iT reduces both the coherent and incoherent artifacts as well as deblurring the image (arrows). 111

5.12 Total Variation regularization (TV): (a) the non-Regularized SPIR-iT exhibits noise amplification due to the g factor. (b) the noise amplification is suppressed by the TV reconstruction, while the edges and features in the image are preserved. 112

Chapter 1

Introduction

Magnetic Resonance Imaging (MRI) is a non-invasive imaging modality. Unlike Computed Tomography (CT), MRI does not use ionizing radiation. In addition, MRI provides a large number of flexible contrast parameters. These provide excellent soft tissue contrast. MRI can also be sensitized to many specific parameters. These include imaging brain oxygen saturation changes due to neuronal activity, measuring blood flow velocities, measuring temperature, and measuring the concentration of metabolites. MRI is also the only way to directly image diffusion of water molecules *in vivo*.

Since its invention more than 30 years ago, MRI has improved dramatically in imaging quality and imaging speed. This has revolutionized diagnostic medicine. Imaging speed is a major part of this revolution and is essential in many of the MRI applications. One effort to improve imaging speed has been focusing on faster data collection. This has been mainly achieved by improvements in MRI hardware, development of fast pulse sequences and efficient scanning trajectories. However, we are currently at the point where fundamental physical and physiological effects limit our ability to simply encode data more quickly.

This fundamental limit has led many researchers to look for methods to reduce the amount of acquired data without degrading image quality. These reduced sampling methods are based on the fact the MRI data are redundant, so the underlying information may be extracted from fewer data than traditionally considered necessary. One of the most significant clinical impacts of reduced sampling methods has been accomplished by parallel imaging with multiple receiver channels [84,97]. Imaging with multiple channels provides more useful data per MRI acquisition, so fewer acquisitions are needed per scan. In recent years, more methods that exploit other sources of redundancy are emerging. Methods such as [26,39,57,70,81,107,110,111] exploit modeled signal properties like spatial and temporal correlations. Methods such as [61,76,106] learn the redundancy by analyzing the data itself.

A property that all images, including MRI, share to some degree is compressibility. It is well established that images can be substantially compressed with almost no visual artifacts. Common compression methods like JPEG and JPEG2000 [103] are based on sparse transform coding. They transform the image content to a sparse representation where the information is concentrated in a few coefficients. Image Compression is mostly conceived as a post-processing operation; First, acquire all the data, then compress it for efficient storage and retrieval. However, recently a new sampling theory called compressed sensing (CS) [11, 21] has emerged. Compressed sensing implicitly compresses the data within the signal acquisition process by obtaining fewer, so called, incoherent measurements. Images can be accurately reconstructed from these measurements using several non-linear recovery processes. The way data are acquired in MRI is compatible with the CS theory. The practical result of CS in the context of MRI is that MR images require much less data for reconstruction.

The contributions of this thesis are three methods for rapid imaging. The time-optimal gradient waveform design method [65] is a fast acquisition approach. The idea behind it is that for any fast imaging technique, the hardware should be maximally exploited to collect data as fast as possible. The other two contributions are reduced sampling approaches. Compressed sensing MRI [64,68] deals with the theoretical and practical issued in applying CS to MRI, whereas the SPIR-iT method [69] introduces a new concept of data consistency for parallel imaging with multiple receiver channels.

Even though it may seem that each of these methods is distinct, they were all motivated by the application of compressed sensing. The time-optimal gradient design method was originally developed as means to design fast randomized sampling trajectories for CS. SPIR-iT was developed as a reconstruction framework to combine CS with parallel imaging.

1.1 Thesis outline

Chapter 2 is a brief introduction to MRI. It provides the necessary basic background about MR imaging for the following more advanced topics. It covers some basic MR physics, hardware description, signal generation, signal reception, spatial encoding and image generation.

Rapid imaging is about rapid collection of data. This is the topic of Chapter 3. There I develop in detail a method to design gradient waveforms that scan k-space trajectories as fast as possible within the system constraints. Many gradient designs have been presented before. Some are sub-optimal or are limited for specific trajectories and specific applications. The proposed method provides a non-iterative, computationally efficient, time-optimal solution for arbitrary k-space trajectories. It provides an efficient way to design waveforms for any application, but in particular designing randomized trajectories for the purpose of applying compressed sensing to MRI.

In Chapter 4 I describe in detail the theory of compressed sensing and its application to rapid MRI. I address the specific difficulties in applying the theory in practice. I present some of the many possible application of CS to reduce scan time and improve the spatial and temporal resolution of MRI.

In Chapter 5 I describe the SPIR-iT method, an iTerative Self-consistent Parallel Imaging Reconstruction method. SPIR-iT formulates the parallel imaging reconstruction through data consistency constraints. It is a general, optimal solution for auto-calibrating parallel imaging from arbitrary k-space trajectories. It is also a general framework for easily incorporating additional image priors, and in particular sparsity/compressibility constraints.

Finally, in Chapter 6 I summarize the contributions of my work to rapid MRI, and provide some insights to possible future research directions.

Chapter 2

Principles of Magnetic Resonance Imaging

2.1 Nuclear Magnetic Resonance Physics

The true description of the Nuclear Magnetic Resonance (NMR) phenomenon is quantum mechanical in nature, but at the macroscopic scale it can be described quite accurately using classical physics. In this thesis I will consider only the classical description [4, 37, 62].

2.1.1 Polarization

The MRI signal is generated by protons in the body, mostly those in water molecules. A strong static field B_0 (See Fig. 2.1) polarizes the protons, yielding a net magnetic moment oriented in the direction of the static field. It is this net magnetic moment, or simply magnetization, which is manipulated and produces the NMR signal. The field direction and its perpendicular plane are often referred to as the longitudinal direction and the transverse plane.

2.1.2 Bloch Equation

The interaction of the magnetization M with an external magnetic field B is governed by the Bloch equation,

$$\frac{dM}{dt} = M \times \gamma B + \frac{M_0 - M_{\mathrm{z}}}{T_1} + \frac{M_{\mathrm{xy}}}{T_2}, \tag{2.1}$$

where M_0, M_{z} and M_{xy} are the equilibrium, longitudinal and transverse magnetization and γ , T_1 and T_2 are constants and are specific to different materials and types of tissues.

2.1.3 Resonance

Applying a radio frequency (RF) excitation field B_1 (See Fig. 2.1) to the net magnetization tips it and produces a magnetization component M_{xy} (or simply m), transverse to the static field. The magnetization precesses at characteristic frequency

$$f_0 = \frac{\gamma}{2\pi} B_0.$$

Here f_0 denotes the precession frequency, B_0 the static field strength, and $\gamma/2\pi$ is a constant (42.57 MHz/T) [37]. A typical 1.5T clinical MR system has a frequency of about 64 MHz. The transverse component of the precessing magnetization produces a signal detectable by a receiver coil. The transverse magnetization at a position r and time t is represented by the complex quantity $m(r,t) = |m(r,t)| \cdot e^{-i\phi(r,t)}$, where $|m(r,t)|$ is the magnitude of the transverse magnetization and $\phi(r,t)$ is its phase. The phase indicates the direction of the magnetization on the transverse plane. The transverse magnetization $m(r)$ can represent many different physical properties of tissue. One very intuitive property is the proton density of the tissue, but other properties, like relaxation, can be emphasized as well. The image of interest in MRI is $m(r)$, the image of the spatial distribution of the transverse magnetization.

2.1.4 Relaxation

While in the transverse plane, the magnetization over a period of time τ experiences relaxation. The longitudinal component experiences exponential recovery $M_z(t+\tau) = M_0(1-e^{-\tau/T_1})+M_z(t)$ with a time constant T_1. The transverse component experiences exponential decay $M_{xy}(t+\tau) = M_{xy}(t)e^{-\tau/T_2}$ with a time constant T_2. The relaxation parameters are one of the most important image contrast mechanisms as different types of tissues have different relaxation parameters.

2.2 Magnetic Resonance Imaging Hardware

The key components of MRI are the interactions of the magnetization with three types of magnetic fields (See Fig. 2.1) and the ability to measure these interactions.

2.2.1 The Static Magnetic Field B_0

This field points in the longitudinal direction. Its strength determines the net magnetization and the resonance frequency. The field homogeneity is very important for imaging. Inhomogeneity often results in image distortion artifacts. In most clinical scanners the field is generated using a superconducting magnet, although some systems use permanent magnets or electromagnets. The field strength for clinical systems ranges from 0.3T to 7T.

2.2.2 Transverse Radio-Frequency Field B_1

This field is a transverse radio-frequency (RF) field produced by coils tuned to the Larmor frequency. This field is used to excite the magnetization from equilibrium by tipping it from the longitudinal direction to the transverse plane. A typical body RF-coil can produce a field strength of $1.6 \cdot 10^{-5}T$.

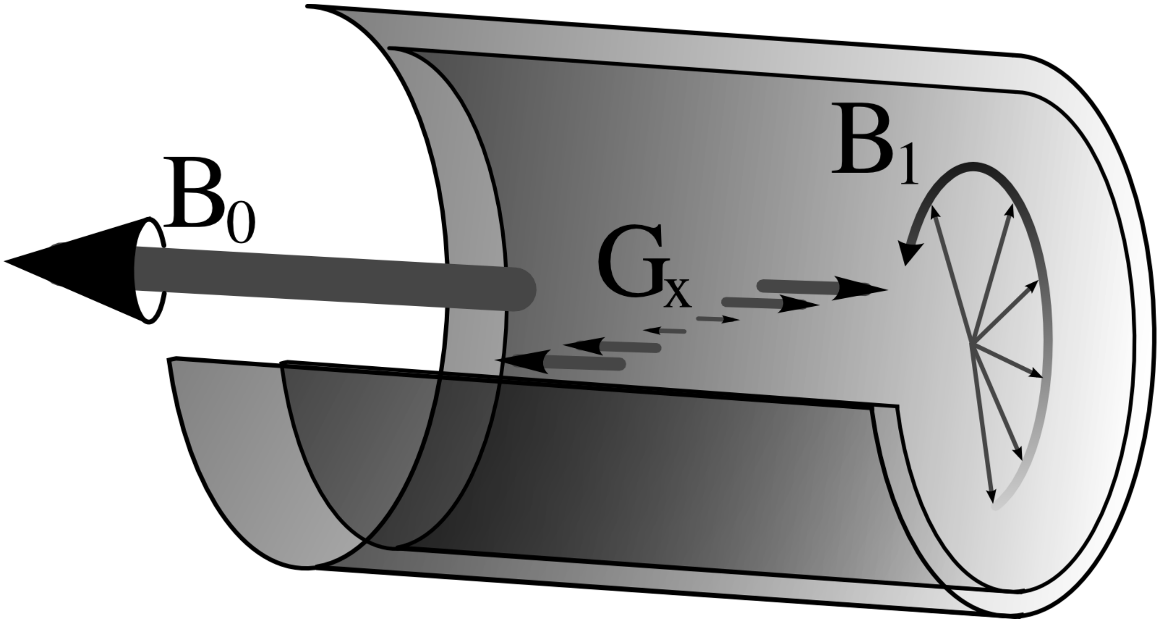

Figure 2.1: The magnetic fields used in MR imaging: The main homogeneous magnetic field B_0 creates a net magnetization that precesses at a resonance frequency $\frac{\gamma}{2\pi}B_0$. The transverse rotating radio-frequency field B_1 is used for exciting the magnetization. The gradient fields G (only G_x is illustrated) are used for spatial encoding.

2.2.3 Receive Coils and Signal Reception

Magnetization that is excited to the transverse plane precesses at the Larmor frequency. The precession creates a changing magnetic flux, which in turn (according to Faraday's law) induces a changing voltage in a receiver coil tuned to the Larmor frequency. This voltage is the MR signal that is used for imaging. The received signal is the cumulative contribution from all the excited magnetization in the volume. With only the homogeneous B_0 field present, the system does not contain any spatial information. The received signal is a complex harmonic with a single frequency peak centered at the Larmor frequency..

2.2.4 Spatial Encoding Gradients

The spatial distribution information comes from three additional fields that vary spatially. Three gradient coils, G_x, G_y and G_z create a linear variation in the longitudinal magnetic field strength as a function of spatial position. For example, when G_x is applied, the magnetic field will vary with position $B(x) = |B_0| + G_x x$. As a result, the resonance frequency of the magnetization will vary in proportion to the gradient field. This variation is used to resolve the spatial distribution as I will show next.

2.3 Imaging

2.3.1 Excitation and Selective Excitation

In general, a B_1 RF field at the resonance frequency excites the whole volume. It is possible through the use of the gradients to selectively excite a smaller portion of it, for example only exciting a slice. The general idea is that only magnetization precessing close to the resonance frequency is affected by the RF field, whereas magnetization at distant frequencies is not affected. When a gradient field is applied, the resonance frequency varies with position. If during that time, a B_1 RF field with a limited bandwidth (for example a *sinc* shaped envelope pulse) is applied, only magnetization at a slice location corresponding to that frequency band is excited. Exciting a slice limits the imaging spatial encoding to two dimension. Exciting a slab or a volume requires three dimensional encoding.

2.3.2 Spatial Encoding and k-Space

As previously mentioned, MR systems can encode spatial information by superimposing the gradient fields on top of the strong static field. Going back to the G_x example, when G_x is applied, the magnetic field will vary with position as $B(x) = |B_0| + G_x x$. This variation causes the precession frequency to vary linearly in space,

$$f(x) = \frac{\gamma}{2\pi}(|B_0| + G_x x).$$

As a result, magnetization at positive x positions will precess at a higher frequency than magnetization at negative x positions.

Spatial encoding using gradients can be understood by a musical instrument analogy; the piano. The pitch of a piano note varies linearly with the position of the key being struck; the sound one hears is the net sum of all notes emitted. A skilled musician listening to the emitted polyharmonic sound can hear which notes are playing and say which keys were pressed (and how hard). The MR signal generated in the presence of a gradient field is likewise a polyphonic mixture. The spatial positions within the patient's body are like piano keys and the emitted RF signal from each position is like a "note", with a frequency linearly varying with position. The polyharmonic MR signal superposes the different "notes"; they encode the spatial position and the magnetization strength at those positions. A signal processing engineer can immediately realize that there is a Fourier relation between the received MR signal and the magnetization distribution and that the magnetization distribution can be decoded by a spectral decomposition.

To see this Fourier relation more concretely consider the following: the gradient-induced variation in precession frequency causes a location dependent phase dispersion to develop. The additional frequency contributed by gradient fields can be written as

$$f(r) = \frac{\gamma}{2\pi} G(t) \cdot r,$$

where $G(t)$ is a vector of the gradient fields' amplitudes. The phase of magnetization is the integral of frequency starting from time zero (immediately following the RF excitation):

$$\begin{aligned} \phi(r,t) &= 2\pi \int_0^t \frac{\gamma}{2\pi} G(s) \cdot r ds \\ &= 2\pi r \cdot k(t), \end{aligned}$$

where

$$k(t) \equiv \frac{\gamma}{2\pi} \int_0^t G(s) ds.$$

The receiver coil integrates over the entire volume, producing a signal

$$s(t) = \int_R m(r) e^{-i2\pi k(t) \cdot r} \, dr.$$

This is the *signal equation for MRI.* In words, the received signal at time t is the Fourier transform of the object $m(r)$ sampled at the spatial frequency $k(t)$. Such information is fundamentally encoded and very different than traditional optical imaging where pixel samples are measured directly.

The design of an MRI acquisition method centers on developing the gradient waveforms $G(t)$ that drive the MR system. These waveforms, along with the associated RF pulses used to produce the magnetization, are called a *pulse sequence*. The integral of the $G(t)$ waveforms traces out a trajectory $k(t)$ in spatial frequency space, or k-space. For illustration, consider the simple example in Fig. 2.2 where, immediately after the RF excitation, a G_x gradient field is applied followed by a G_y gradient. The phases of the magnetization are shown at different time points, along with the k-space trajectory and the MR signal. This encoded sampling and the freedom in choosing the sampling trajectory play a major role in making the compressed sensing idea in chapter 4 ideas naturally applicable to MRI.

2.3.3 Image Acquisition

Constructing a single MR image commonly involves collecting a series of frames of data, called *acquisitions*. In each acquisition, an RF excitation produces new transverse magnetization, which is then sampled along a particular trajectory in k-space.

In principle, a complete MR image can be reconstructed from a single acquisition by using a k-space trajectory that covers a whole region of k-space [5]. This is commonly done in applications such as imaging brain activation. However, for most applications this results in inadequate image resolution and excessive image artifacts.

Magnetization decays exponentially with time. This limits the useful acquisition time window. Also, the gradient system performance and physiological constraints limit the speed at which k-space can be traversed (See Fig. 2.2). These two effects combine to limit the total number of samples per acquisition. As a result, most MRI imaging methods use a sequence of acquisitions; each one samples part of k-space. The data from this sequence of acquisitions is then used to reconstruct an image.

2.3.4 Resolution and Field of View

Traditionally the k-space sampling pattern is designed to meet the Nyquist criterion, which depends on the resolution and field of view (FOV) as shown in Fig. 2.3. Image resolution is determined by the sampled region of k-space: a larger region of sampling gives higher resolution. The supported field of view (FOV) is determined by the sampling density within the sampled region: larger objects require denser sampling to meet the Nyquist criterion. Violation of the Nyquist criterion causes the linear reconstruction to exhibit artifacts. The appearance of such artifacts depends on the details in the sampling pattern, as shown in Fig. 2.3.

2.3.5 k-space sampling trajectories

There is considerable freedom in designing the k-space trajectory for each acquisition. Some 2D and 3D sampling trajectories are illustrated in Fig. 2.4. By far the most popular trajectory uses straight lines from a Cartesian grid. Most pulse sequences used in clinical imaging today are Cartesian. Reconstruction from such acquisitions is wonderfully simple: apply the inverse Fast Fourier Transform (FFT). More importantly, reconstructions from Cartesian sampling are robust to many sources of system imperfections.

While Cartesian trajectories are by far the most popular, many other trajectories are in use, including sampling along radial lines and sampling along spiral trajectories. Radial acquisitions are less susceptible to motion artifacts than Cartesian trajectories [31], and can be significantly undersampled [92], especially for high contrast objects [1,82]. Spirals make efficient use of the gradient system hardware, and are used in real-time and rapid imaging applications [74]. Reconstruction from such non-Cartesian trajectories is more complicated, requiring filtered back-projection algorithms [38] or k-space interpolation schemes (e.g. gridding [48]).

2.3.6 Rapid Imaging

MR acquisition is inherently a process of traversing curves in multi-dimensional k-space. The speed of k-space traversal is limited by physical constraints. In current systems, gradients are limited by maximum amplitude and maximum slew-rate (See Fig. 2.2. In addition, high gradient amplitudes and rapid switching can produce peripheral nerve stimulation in patients [16]. Since this must be avoided, the physiology of the patient provides a fundamental limit to gradient system performance.

Because sampling speed is fundamentally limited, many researchers are striving to reduce the amount of acquired data without degrading image quality. Many such efforts are inspired by the idea that MRI data are, or can be made to be, redundant. Such redundancy can be created by design, for example, using multiple receiver coils [84, 97], which provides more useful data per MR acquisition, so fewer acquisitions are needed per scan. Redundancy can be a known or modeled signal property such as spatial-temporal correlation [26,39,57,70,81,107,110,111] or a redundancy learned and extracted from the data itself [61, 76, 106].

In the following chapters I will present three methods for rapid imaging. They are based on the three concepts just described. The first is a fast scanning approach. It is a time-optimal gradient waveform design method. In fast imaging it is desirable to exploit the hardware capabilities of the system as best as possible. The proposed method provides the gradient waveforms that scan a k-space trajectory as fast as possible within the hardware limits. The next methods, "Sparse MRI" and SPIR-iT are reduced sampling approaches. Sparse MRI is based the theory of compressed sensing as described in [11, 21]. It exploits the inherent compressibility of still and dynamic MR images to vastly undersample the data. SPIR-iT, is a parallel imaging method based on the data redundancy of imaging with multiple receiver coils. It is an extension to the SENSE [84] and GRAPPA [35] reconstruction methods and is based on forcing self consistency of the data (e.g., iTerative, Self-Consistent Parallel imaging Reconstruction). SPIR-iT is also a framework for combining Sparse MRI with parallel imaging.

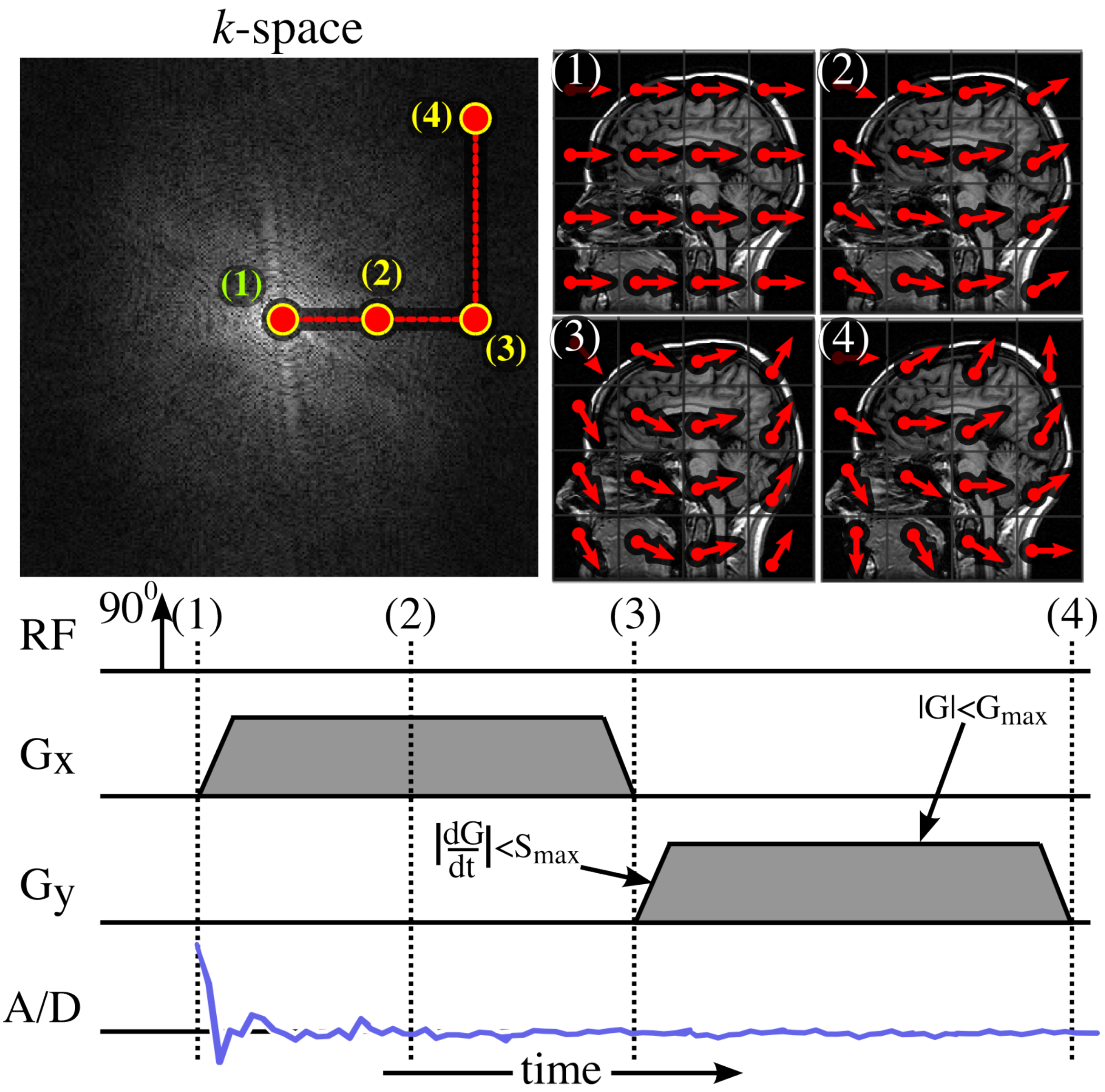

Figure 2.2: Fourier spatial encoding. The temporal MRI signal directly samples the spatial frequency domain of the image. Gradient fields cause a linear frequency distribution across the image, which produces a linear phase accrual with time. The received signal samples are spatial frequencies of the image. The corresponding spatial frequencies are proportional to the gradient waveform area. The gradient is limited in amplitude, $G_{\max}$, and slew rate, $S_{\max}$, which are both system specific.

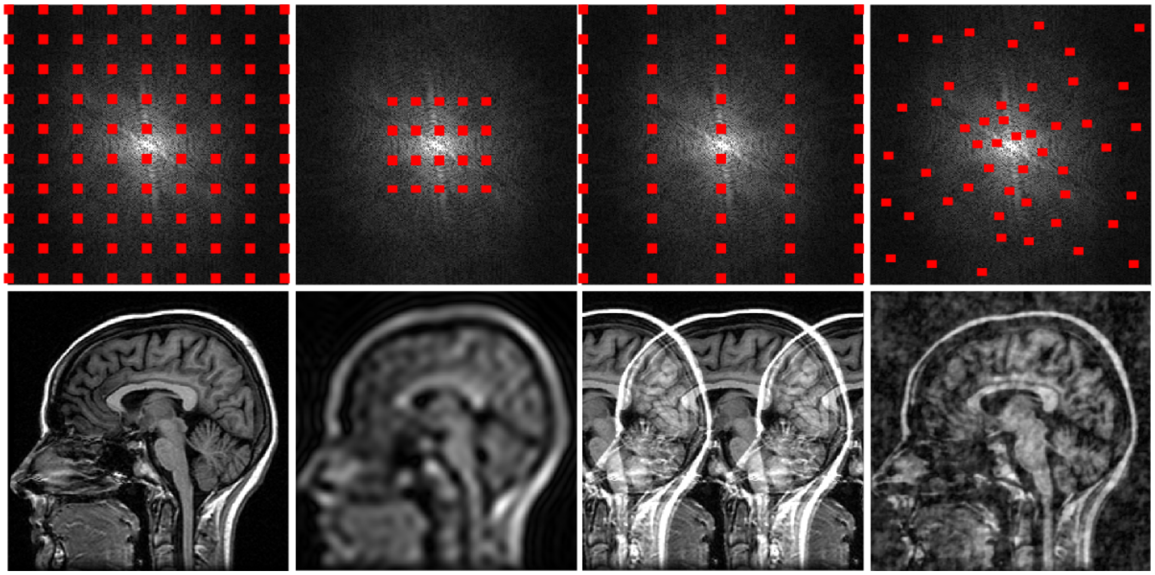

Figure 2.3: Image resolution is determined by the extent of the k-space that is covered. The supported field of view is determined by the sampling density. Violation of the Nyquist criteria results in aliasing interference in the image domain. The appearance of the artifact depends on the sampling. Coherent folding is produced by equispaced sampling and incoherent interference is produced by irregular sampling

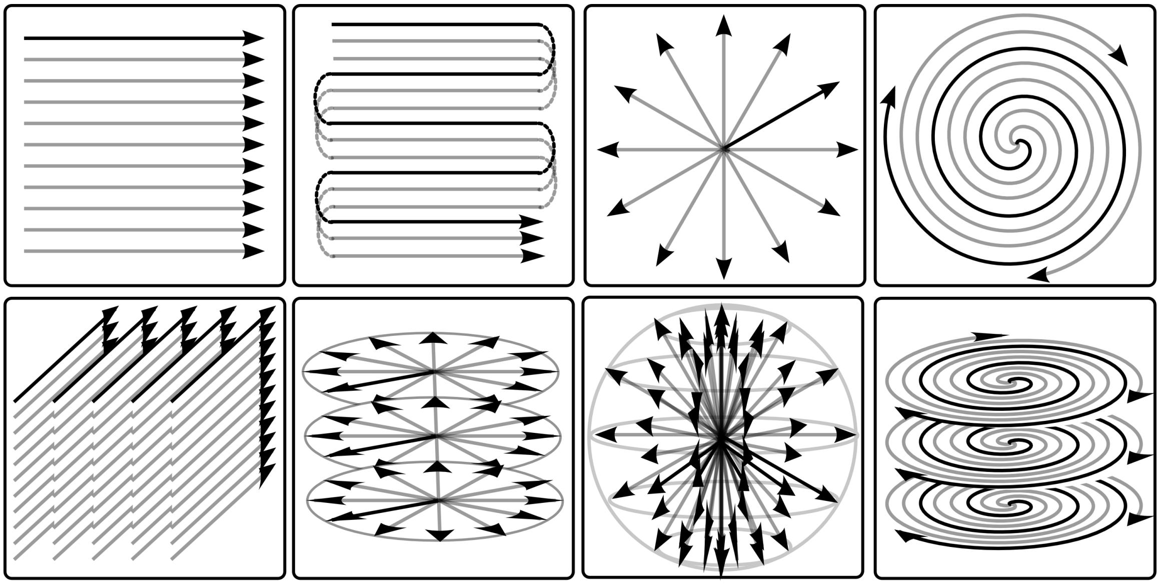

Figure 2.4: Common sampling trajectories. Top, left to right: Cartesian 2D, Cartesian echo-planar, radial, spiral. Bottom left to right: Cartesian 3D, stack of radial, 3D radial, 3D stack of spirals

Chapter 3

Time Optimal Gradient Design

3.1 Introduction

The recent advances in the field of MRI is a result of efforts in many fronts; gradient hardware, high field systems, optimized receiver coil arrays, fast sequences and sophisticated reconstruction methods . These provide the ability to image faster than ever. New acquisition methods are being explored in which k-space is scanned in non-traditional trajectories [36, 66, 74, 79, 83, 91]. In particular, for the application of compressed sensing and even parallel imaging, irregular randomized and incoherent trajectories are desirable.

One of the design challenges in rapid imaging is to minimize the gradient waveform duration, subject to both hardware and sequence constraints. The preferred solution would be to design both the k-space trajectory and the gradient waveforms at the same time. However, because of the complexity and the large number of control variables, often only approximate heuristic methods [18, 75] are used as the optimal solution is generally not tractable.

A simpler, more common approach is to first choose a sampling trajectory and then design the gradient waveforms for it. For example, spirals are often designed this way [30, 52, 55, 74]. In this approach, the problem is to find the gradient waveform that will traverse k-space from one point to another along a specific path and in minimum-time, while meeting the hardware gradient magnitude and slew-rate constraints. For trajectories such as linear, circular [43] or spirals [30, 52, 55, 74] the solution is quite simple; first operate in the slew-rate limited regime till the maximum gradient is reached, and then operate in the maximum gradient regime. For the general arbitrary trajectory, the solution is not simple anymore because there can be numerous switching points between slew-limited acceleration, slew-limited deceleration and gradient-limited regimes. In the current literature, there is no general methodology to design gradient waveforms for arbitrary trajectories that guarantees time-optimality and is computationally inexpensive.

It is important to mention that some of the optimal designs that exist in the literature either solve for 1D waveforms [95, 96], or provide waveforms that traverse k-space from one point to another but **not on a specific path** [17, 40].

In this chapter I develop a fast and simple algorithm based on optimal control theory that provides the complete solution to the time-optimal gradient waveform for arbitrary k-space paths. Using this method a user need only prescribe a parametric curve in k-space (arbitrary parametrization) and the algorithm will output the gradient waveform that traverses the k-space path in minimum time.

3.2 Curve Parametrization

I start by describing some properties of planar and volumetric curves that are essential to the derivation of the time-optimal gradient waveform design.

Suppose we are given a specified path, curve, or trajectory C from C_0 to C_1 in k-space. Suppose the curve C is described as a function of some parameter p:

$$C(p) = (x(p), y(p), z(p)) \in \mathbf{R}^3, \quad p \in [0, p_{\max}]. \tag{3.1}$$

Here, $p = 0$ corresponds to the initial point and $p = p_{\max}$ corresponds to the end point:

$$C(0) = C_0, \qquad C(p_{\max}) = C_1.$$

The first derivative of the curve with respect to its parametrization is the *velocity* or *tangent* vector of the parametrization, which is denoted as

$$T(p) = \frac{dC(p)}{dp} = C'(p). \tag{3.2}$$

The second derivative of the curve with respect to its parametrization is the *acceleration* vector of the parametrization

$$A(p) = \frac{d^2C(p)}{dp^2} = C''(p). \tag{3.3}$$

From here onwards I denote $h'(p)$ as the derivative of the function h with respect to a general parameter p. I make an exception when using the notation $\dot{h}$ to specifically indicate that it is a time derivative.

A very useful parametrization is the Euclidean arc-length s parametrization:

$$C(s) = (x(s), y(s), z(s)), \qquad s \in [0, L], \tag{3.4}$$

where L is the length of the path. This parametrization describes the coordinates as a function of the Euclidean distance along the curve. An important property of this parametrization is that it has unit velocity, *i.e.*,

$$|C'(s)| = 1, \tag{3.5}$$

for all s. Another important property is that the magnitude of the *acceleration* is the curvature $\kappa(s)$ of the curve, *i.e.*

$$|C''(s)| = \kappa(s). \tag{3.6}$$

The curvature of the curve at a given point is the reciprocal of the radius of an osculating circle that has the same first and second derivatives with the curve at that point. Figure 3.1 illustrates the properties of the arc-length parametrization.

When a curve is given in an arbitrary parametrization $C(p) = (x(p), y(p), z(p))$, it

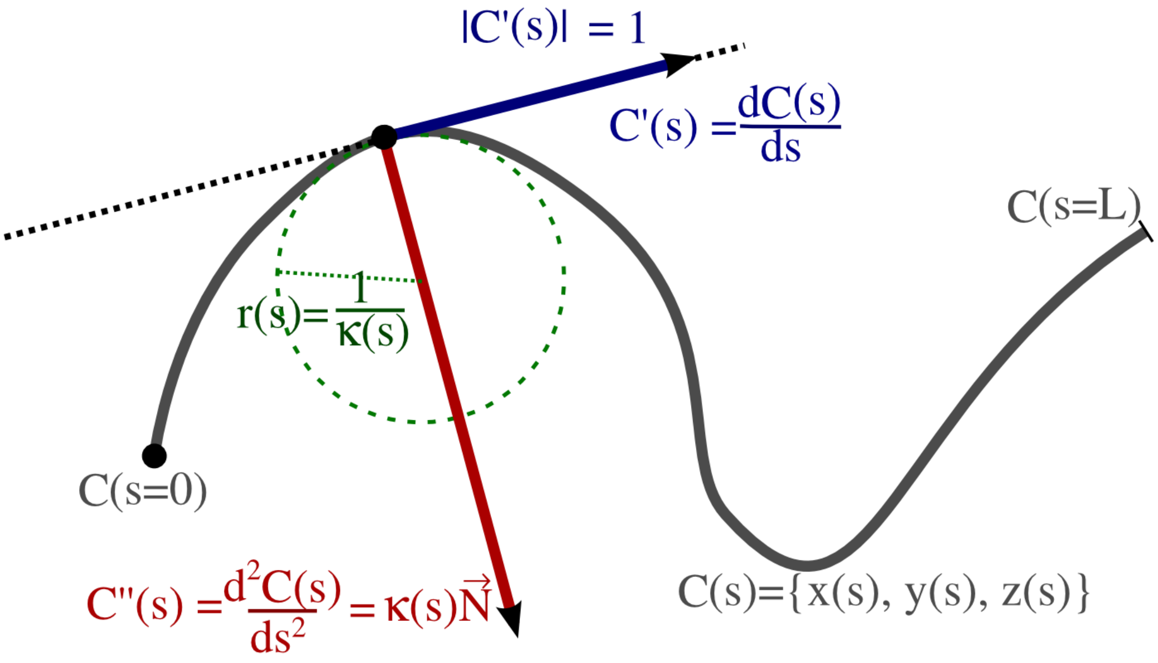

Figure 3.1: Properties of the arc-length parametrization of a curve in $\mathbf{R}^3$.

is always possible to convert into the arc-length parametrization by using the relation

$$s(p) = \int_0^p |C'(q)|dq. \tag{3.7}$$

3.3 The Time-Optimal Gradient Waveform Design Problem

In this chapter I aim to design a gradient waveform as a function of time. This is equivalent to designing a time parametrization of the curve that describes the k-space coordinates as a function of time. Specifically, I aim to design a time function $p = s(t)$ in the arc-length parametrization such that

$$s(0) = 0, \qquad s(T) = L$$

where T is the traversal time. The time trajectory in k-space is given by the composite function $\tilde{C}(t) = C(s(t))$.

First, I derive the relation between the gradient waveform and the curve parametrization. The gradient waveform is proportional to the velocity in the time parametrization and is given by

$$g(t) = \gamma^{-1}\frac{dC(s(t))}{dt} = \gamma^{-1}C'(s(t))\dot{s}(t) \tag{3.8}$$

where γ is the gyro-magnetic ratio. Here, I use $\dot{s}(t)$ to indicate the time derivative of s. The gradient slew-rate is proportional to the *acceleration* in the time parametrization. Using the chain rule I obtain

$$\dot{g}(t) = \gamma^{-1}(C''(s(t))\dot{s}(t)^2 + C'(s(t))\ddot{s}(t)). \tag{3.9}$$

The design variable in the MRI system is the gradient waveform. The gradients are subject to hardware as well as sequence constraints. For the hardware constraints I assume the frequently used slew-limited model as described in [40]. In this model, the gradient amplitude is subject to the maximum amplitude of the system

$$|g(t)| \leq G_{\max}, \qquad t \in [0, T]. \tag{3.10}$$

It is also subject to the maximum slew-rate of the system

$$|\dot{g}(t)| \leq S_{\max}, \qquad t \in [0, T]. \tag{3.11}$$

For the sequence constraints, the gradient waveform is constrained to follow a specific trajectory in k-space such that

$$\tilde{C}(t) = C_0 + \gamma \int_0^t g(\tau)\, d\tau.$$

It is also constrained to have an initial value. For simplicity, I assume an initial value of

$$g(0) = 0.$$

Other optional sequence constraints such as final or intermediate values are possible, but are not assumed here. I discuss some of these optional constraints later in this section.

Now that I have derived the hardware as well as the sequence constraints, I consider the problem of finding the time-optimal gradient waveform that satisfies them. The time-optimal problem can be formulated as

$$\begin{array}{ll} \text{minimize} & T \\ \text{subject to} & |g(t)| \le G_{\max}, \quad t \in [0,T] \\ & |\dot{g}(t)| \le S_{\max}, \quad t \in [0,T] \\ & g(0) = 0 \\ & \tilde{C}(t) = C_0 + \gamma \int_0^t g(\tau)\, d\tau, \quad t \in [0,T] \\ & \tilde{C}(0) = C_0 \\ & \tilde{C}(T) = C_1. \end{array} \tag{3.12}$$

Here the variable is the gradient $g(t)$ defined over the time interval $[0,T]$. The objective is to minimize the traversal time T along the trajectory.

The time-optimal solution is always either gradient or slew-rate limited. Solving for the time-optimal waveform in this formulation requires one to find the optimal switching times between the maximum gradient and maximum slew-rate. This procedure is difficult for complex curves. Instead, I look at an alternative equivalent formulation in which the solution becomes simpler.

3.4 Formulation in the Arc-Length Parametrization

Let s denote the arc length from the initial point. Because the k-space path is given as a constraint, one needs only to design the time function $s(t)$. Note that s is always

INPUTS:

$C(p) = \{x(p), y(p), z(p)\}$ - A k-space curve in an arbitrary parametrization p.

$G_{\max}$ - Maximum gradient amplitude of the system.

$S_{\max}$ - Maximum slew-rate of the system.

OUTPUTS:

$C^\star(t)$ - The k-space curve in a time-optimal parametrization t.

$g^\star(t)$ - The time-optimal gradient waveforms.

I: Convert to arc-length parametrization.

1. Compute $s(p) = \int_0^p \left| \frac{dC(q)}{dq} \right| dq$.
2. Compute $\hat{C}(s) = C(p(s))$ using the inverse of $s(p)$.

II: Find gradient "velocity" $v(s)$ as a function of arc length by integrating the ODE's

1. Compute $\kappa(s) = |\hat{C}''(s)|$.
2. Compute $\alpha(s) \;=\; \min\left\{\gamma G_{\max},\ \sqrt{\frac{\gamma S_{\max}}{\kappa(s)}}\right\}$.
3. Define: $\beta(s, \dot{s}) \equiv \left[\gamma^2 S_{\max}^2 - \kappa^2(s)\dot{s}^4\right]^{1/2}$.
4. Integrate the ODE $\frac{dv_+(s)}{ds} = \begin{cases} \frac{1}{v_+(s)}\beta(s, v_+(s)) & \text{if } v_+(s) < \alpha(s) \\ \frac{d\alpha(s)}{ds} & \text{otherwise,} \end{cases}$ forward with the initial condition $v_+(0) = 0$.
5. Integrate the ODE $\frac{dv_-(s)}{ds} = \begin{cases} -\frac{1}{v_-(s)}\beta(s, v_-(s)) & \text{if } v_-(s) < \alpha(s) \\ -\frac{d\alpha(s)}{ds} & \text{otherwise,} \end{cases}$ backwards with the final condition $v_-(L) = v_+(L)$.
6. Let $v^\star(s) = \min\{v_+(s), v_-(s)\}$.

III: Convert to time parametrization, and compute the gradient waveforms.

1. Compute $s^\star(t)$ using the inverse of $t^\star(s) = \int_0^s \frac{d\sigma}{v^\star(\sigma)}$.
2. Compute $C^\star(t) = \hat{C}(s^\star(t))$.
3. Compute $g^\star(t) = \gamma^{-1}\frac{dC^*(t)}{dt}$.

Table 3.1: Outline of the time-optimal gradient design algorithm.

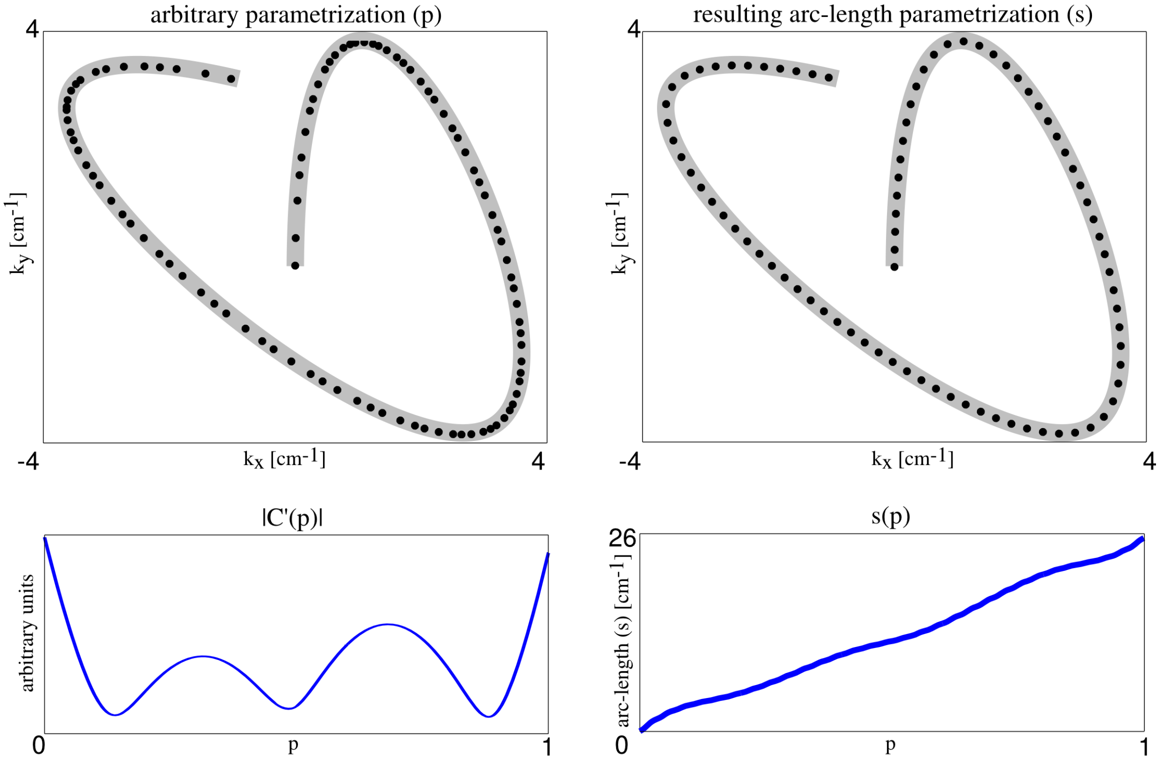

Figure 3.2: Example of stage I of the design algorithm: Conversion from arbitrary parametrization to arc-length parametrization.

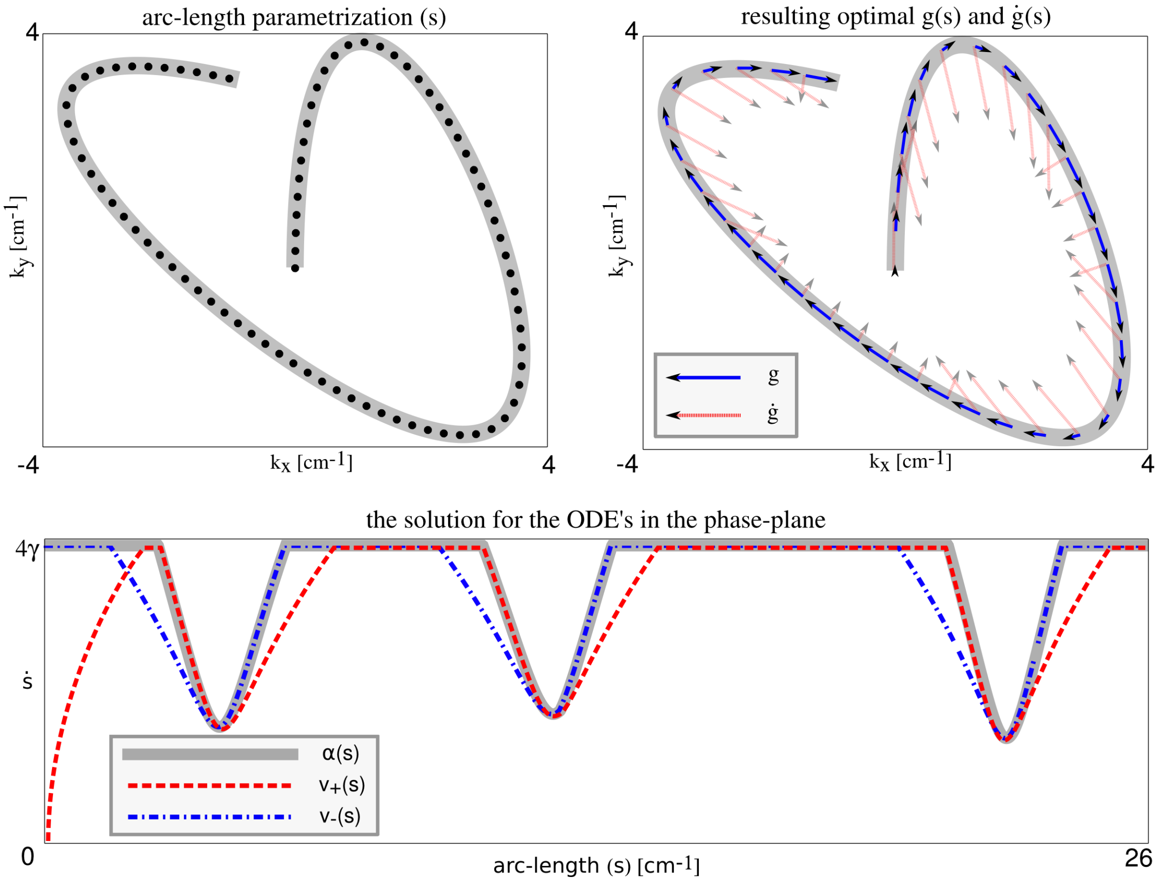

Figure 3.3: Example of stage II of the design algorithm: Calculation of the time-optimal "velocity" in the phase-plane.

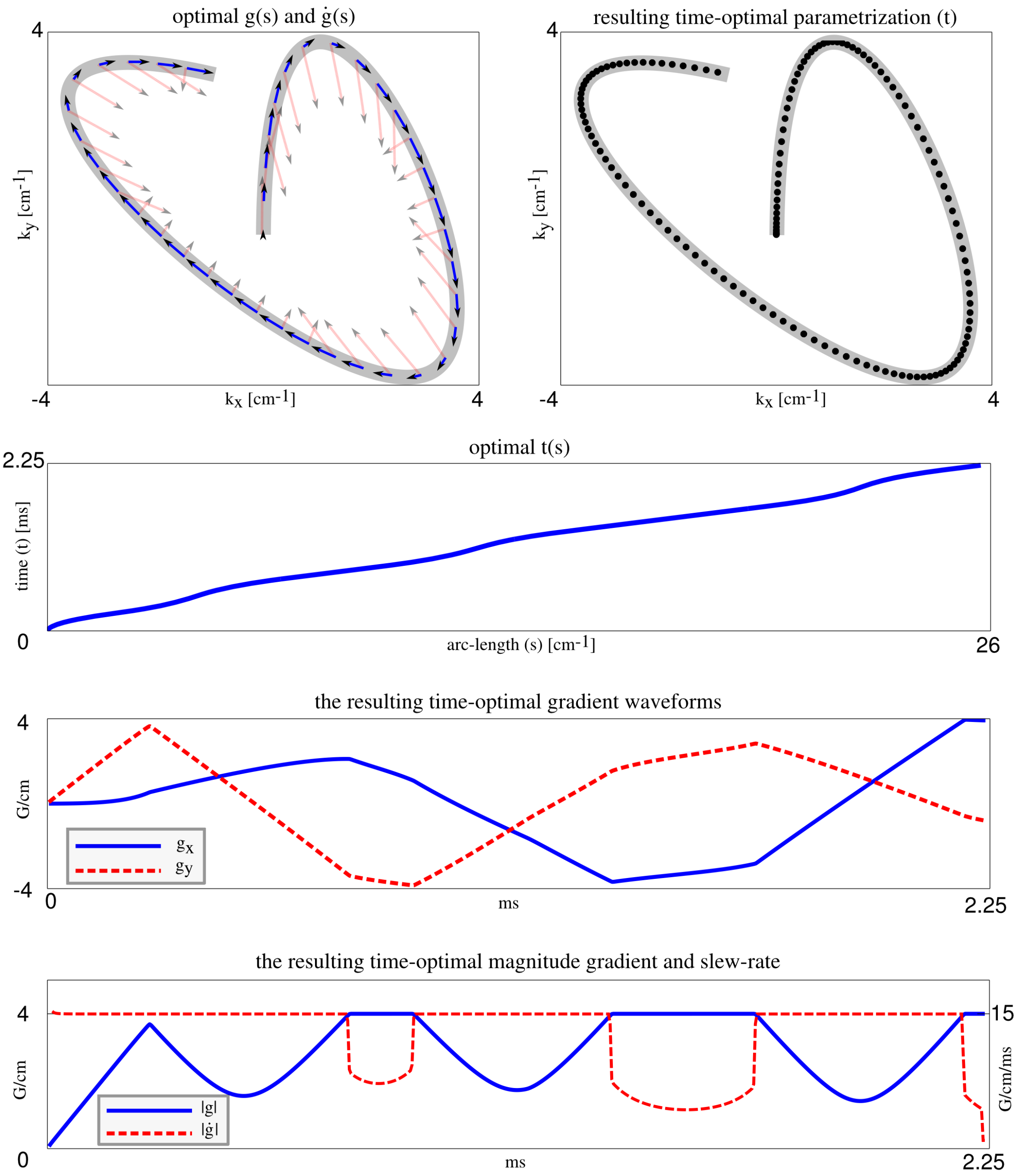

Figure 3.4: Example of stage III of the design algorithm: Conversion from the arc-length parametrization to the time-optimal parametrization and the calculation of the gradient waveforms. Note that either the gradient magnitude or the slew-rate are maximized at every time-point, which is a necessary condition for time-optimality.

increasing, so

$$\dot{s}(t) \geq 0.$$

It follows from (3.5) and (3.8) that

$$|g(t)| = \gamma^{-1}\dot{s}(t). \tag{3.13}$$

This means that it is sufficient to design the gradient magnitude along the path.

I start with the formulation of the hardware constraints in the arc-length parametrization. It follows from (3.8), (3.9), (3.10), and (3.11) that

$$\begin{aligned} \dot{s}(t) &\leq \alpha(s(t)) && (3.14)\\ |\ddot{s}(t)| &\leq \beta(s(t), \dot{s}(t)) && (3.15) \end{aligned}$$

where

$$\begin{aligned} \alpha(s) &= \min\left\{\gamma G_{\max}, \sqrt{\frac{\gamma S_{\max}}{\kappa(s)}}\right\} \\ \beta(s, \dot{s}) &= \left[\gamma^2 S_{\max}^2 - \kappa^2(s)\dot{s}^4\right]^{1/2}. \end{aligned}$$

A complete derivation of (3.14) and (3.15) is deferred to the Appendix. Intuitively, the constraint in (3.14) accounts for the geometry of the trajectory, and is related to the maximum velocity at which a curve can be approached without violating the acceleration constraint, and is independent of past or future velocities. Equation (3.15) is a dynamic constraint: It is a differential inequality that describes the allowed change in the velocity at a specific point on the path given the velocities in its proximity.

Finally, problem (3.12) can be equivalently formulated in the arc-length parametrization as

$$\begin{array}{ll} \text{minimize} & T \\ \text{subject to} & \dot{s}(t) \leq \alpha(s(t)), \quad t \in [0, T] \\ & |\ddot{s}(t)| \leq \beta(s(t), \dot{s}(t)), \quad t \in [0, T] \\ & s(0) = 0, \quad \dot{s}(0) = 0 \\ & s(T) = L. \end{array} \tag{3.16}$$

Here the variable is the time function $s(t)$. Once the optimal solution $s^\star(t)$ of problem (3.16) is known, one can find the solution to the original problem (3.12) using

$$g^\star(t) = \gamma^{-1} C'(s^\star(t))\dot{s}^\star(t).$$

3.5 The Time-Optimal Solution in the Phase-Plane

Up until now, I have only provided the formulation of the problem. Here, I provide a complete solution to (3.16). The solution is obtained in the velocity vs. arc-length plane ($\dot{s}$ vs. s). This plane is often referred to as the *phase-plane* in optimal-control theory literature [47, 53, 94]. The outline is the following: I first find the optimal velocity as a function of arc length ($v^\star(s)$) and then find the optimal time function $s^\star(t)$ which can be used to derive the optimal gradient waveform. (This reformulation method has been used in solving time-optimal path planning problems for robotic manipulators along specified paths [47, 53, 94].)

In the phase-plane, I represent the velocity as a function of arc length:

$$\dot{s}(t) = v(s(t)). \tag{3.17}$$

Note that the traversal time T is a function of the velocity given by

$$T = \int dt = \int \frac{dt}{ds}\, ds = \int_0^L \frac{1}{\dot{s}}\, ds = \int_0^L \frac{1}{v(s)}\, ds.$$

Also, note that the acceleration is also a function of the velocity given by

$$\ddot{s} = \frac{d\dot{s}}{ds}\frac{ds}{dt} = \frac{d\dot{s}}{ds}\dot{s} = v'(s)v(s).$$

Then, the time-optimal control problem (3.16) amounts to solving the following optimization problem in the phase-plane:

$$\begin{array}{ll} \text{minimize} & \displaystyle\int_0^L \frac{1}{v(s)}\, ds. \\ \text{subject to} & v(s) \leq \alpha(s), \quad s \in [0, L] \\ & |v'(s)| \leq \dfrac{1}{v(s)} \beta(s, v(s)), \quad s \in [0, L] \\ & v(0) = 0 \end{array} \tag{3.18}$$

where the optimization variable is the function $v(s)$ defined over $[0, L]$.

The optimal solution $v^\star$ to this problem describes the relation between the optimal time function $s^\star$ and its derivative in the phase-plane:

$$\dot{s} = v^\star(s).$$

Using this relation, I can readily recover $s^\star$ from $v^\star$.

I now describe a complete solution to the optimization problem (3.18). In order to find the optimal velocity, one needs to integrate two ordinary differential equations (ODEs). The first ODE is given by

$$\frac{dv_+(s)}{ds} = \begin{cases} \dfrac{1}{v_+(s)} \beta(s, v_+(s)) & \text{if } v_+(s) < \alpha(s) \\ \frac{d\alpha(s)}{ds} & \text{otherwise,} \end{cases} \tag{3.19}$$

which I integrate forward with the initial condition $v_+(0) = 0$. The second ODE is

$$\frac{dv_-(s)}{ds} = \begin{cases} -\dfrac{1}{v_-(s)} \beta(s, v_-(s)) & \text{if } v_-(s) < \alpha(s) \\ -\frac{d\alpha(s)}{ds} & \text{otherwise,} \end{cases} \tag{3.20}$$

which I integrate backwards with the final condition $v_-(L) = v_+(L)$. (Recall that $\alpha(s)$ and $\beta(s, \dot{s})$ are given in (3.15).)

The optimal velocity $v^\star(s)$ is simply given by

$$v^\star(s) = \min\{v_+(s), v_-(s)\}. \tag{3.21}$$

A proof of optimality is given in the Appendix.

3.6 The Time-Optimal Gradient Waveform Solution

Here I describe the procedure to recover the time-optimal gradient waveform $g^\star(t)$ from the optimal velocity $v^\star(s)$. $s^\star(t)$ can be obtained by computing the inverse function of $t^\star(s)$, using the relation

$$t^\star(s) = \int_0^s \frac{d\sigma}{v^\star(\sigma)}.$$

In particular, the traversal time, which is the optimal value of (3.16), is given by

$$T^\star = \int_0^L \frac{ds}{v^\star(s)}.$$

It follows from (3.8) that the time-optimal gradient waveform is

$$g^\star(t) = \gamma^{-1}\frac{dC(s^\star(t))}{dt}. \tag{3.22}$$

A summary of the design algorithm is given in table 3.1. Figs. 3.2 through 3.4 show a simplified example of the stages of the algorithm applied to a simple trajectory. Figure 3.2 shows the transition from the arbitrary parametrization to the arc-length parametrization. Figure 3.3 shows the calculation of the gradient “velocity” as a function of arc length in the phase-plane. Figure 3.4 shows the transition from the arc-length parametrization to the time parametrization and the calculation of the gradient waveforms. Although the trajectory in Figs. 3.2-3.4 has little practical value, it was chosen for its educational value as it shows clearly and simply the stages of the design.

3.7 Additional Constraints

To handle nonzero initial gradient $g(0)$, it suffices to integrate the ODE (3.19) forward with the initial condition $v_+(0) = \gamma|g(0)|$ instead of $v_+(0) = 0$. However, if the value of $g(0)$ is infeasible, the outcome of the design will be the maximum feasible one.

In the same way, to handle final gradient value $g(L)$, it suffices to integrate the ODE (3.20) backward with the final condition $v_-(0) = \gamma|g(L)|$ instead of $v_-(L) = v_+(L)$. Again, if the value is infeasible, the outcome will be the maximum feasible one.

In general, intermediate gradient and slew-rate magnitude constraint can be applied by replacing the constraints in Eq. (3.10) and Eq. (3.11) to be a function of the arc length, s:

$$|g(t)| \leq G_{\max}(s), \qquad t \in [0, T],\ s \in [0, L] \tag{3.23}$$

and

$$|\dot{g}(t)| \leq S_{\max}(s), \qquad t \in [0, T],\ s \in [0, L]. \tag{3.24}$$

3.8 Examples

In this section I present a few examples demonstrating some of the applications of the method to gradient waveform design.

The design algorithm was implemented in Matlab (The MathWorks, Inc., Natick, MA, USA) and in the C programming language. All simulations were performed on a Mandriva Linux workstation with an AMD Athlon 3800+ 64bit processor and 2.5GB memory. Derivative operations were approximated by finite differences. Numerical integrations were approximated by the trapezoid method. The ODEs were solved using a 4^{th} order Runge-Kutte method [7]. I used the cubic-spline interpolation method

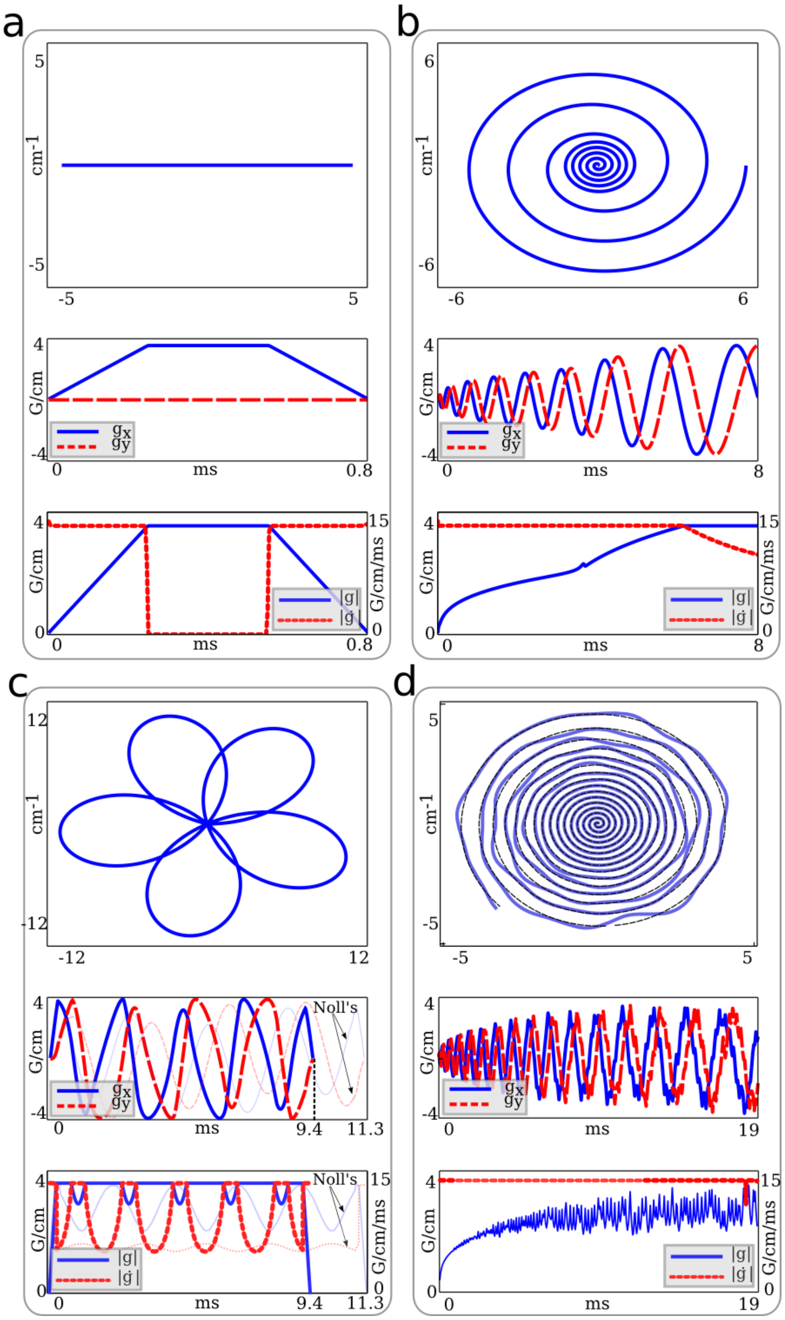

Figure 3.5: Examples showing several k-space trajectories with their corresponding time-optimal gradient waveforms and gradient magnitude and slew-rate. (a) A line. (b) Dual density spiral. (c) Rosette trajectory. (d) Randomly perturbed variable density spiral.

for interpolating the curve when needed. In all the designs, I assumed gradients capable of 40 mT/m, slew-rate of 150 mT/m/ms and a sampling rate of 4 μs.

3.8.1 A Line in k-space

As a sanity test I have applied the algorithm to a horizontal line in k-space. Constraining zero initial and final gradient amplitude I expected the resulting gradient waveform to be a slew and gradient limited trapezoid. The k-space line was chosen to be 10 cm^{-1} long. The result of the experiment are shown in Fig. 3.5a. As expected, the resulting gradient waveform is a trapezoid. Note, that the trapezoid is either slew limited or gradient limited, which guarantees time optimality in this simple case.

3.8.2 Dual Density Spiral Design

Spiral trajectories are used in real-time and rapid imaging applications. In some applications, for example MR fluoroscopy [98] and autocalibrated parallel imaging [42], spirals with varying densities are desired. Here I demonstrate an alternative design by a simple curve parametrization followed by the time-optimal gradient design.

The spiral is parametrized using the following equation

$$\begin{aligned} k &= r \cdot e^{i\theta(r)} \\ \theta(r) &= \frac{2\pi}{N_{\text{itlv}}} \int_0^r FOV(\rho) d\rho,, \end{aligned} \tag{3.25}$$

where N_{itlv} is the number of spiral interleaves and $FOV(r)$ is a function that describes the supported field of view (FOV) as a function of radius in k-space.

To test the method, I designed a dual density spiral with 16 interleaves, 0.83mm resolution (maximum k-space radius of 6cm^{-1}), a FOV of 55cm for $r \in [0, 1.2]$cm^{-1} and 10cm for $r \in [1.8, 6]$cm^{-1}. The transition region of $r \in [1.2, 1.8]$cm^{-1} was linearly interpolated. The results of the design for a single interleaf are presented in Fig. 3.5b. As expected the gradient waveforms are gradient and slew limited. The transition between the two density regions is clearly seen in the magnitude gradient plot.

3.8.3 Time-Optimal Rosette Trajectory Design

The rosette trajectory was first introduced in [79] as a spectrally selective imaging method. Because of the multiple crossings of the trajectory, off-resonance often causes destructive interference and therefore off-resonance voxels do not show up in the image. Recently, rosettes have been reconsidered as a rapid sampling trajectory for CS-MRI with application to hyperpolarized ^{13}C imaging [9], where scan time is crucial.

In [79] Noll describes rosettes as the parametric curve in time

$$k(t) = k_{\max} \sin(\omega_1 t) e^{i\omega_2 t}. \tag{3.26}$$

The gradient waveforms according to Noll are

$$\begin{aligned} G(t) &= \frac{1}{\gamma}\frac{d}{dt}k(t) \\ &= \frac{k_{\max}}{2\gamma}((\omega_1 + \omega_2)e^{i(\omega_1+\omega_2)t} \\ &\quad +(\omega_1 - \omega_2)e^{-i(\omega_1-\omega_2)t}. \end{aligned} \tag{3.27}$$

The advantage of this design is that there is an analytic expression for the gradient waveform, and no further calculations are needed. However, it has a couple of limitations. The first is that the waveforms are neither gradient limited nor slew limited and are not time-optimal. The scan efficiency is especially reduced when designing high resolution trajectories, where the waveforms constraints are set mostly by the maximum gradient amplitude. The second is that at $t = 0$, Eq. (3.27) does not have a zero value, therefore a correction is needed to ramp the gradients from zero initial value to meet the slew-rate constraints. Instead, I use the parametrization of Eq. (3.26) and apply the design method to get the desired time-optimal gradient waveforms. To test the design I used $\omega_1 = 1.419$, $\omega_2 = 0.8233$, $k_{\max} = 12$ as the parameters in Eq. (3.26). The results of applying the proposed design are presented in Fig. 3.5c. The faint lines in the figure show the result for Noll's design. The figure shows that as expected, the waveforms of the proposed method are gradient and slew limited. It also shows that for this specific trajectory Noll's design requires a traversal time of 11.32 ms. This is a 20% increase over the time-optimal design which requires only 9.46 ms.

3.8.4 Randomized Trajectories for Compressed Sensing Application

It was shown in [66] that under-sampled randomly perturbed spiral trajectories can provide faster imaging when used with a special non-linear reconstruction. When under-sampling to save scan time, it is essential that the gradient waveforms be time-optimal. Therefore, I applied the design algorithm to a randomly perturbed variable density spiral where the optimal waveform requires numerous switching between acceleration-deceleration slew-rate limited regions and gradient magnitude limited regions. I designed a 4 interleave variable density spiral, chosen to have a resolution of 1 mm and a FOV of 20 cm at the k-space origin, that linearly decreases to 5 cm on the periphery. The spiral was perturbed in the radial direction by a randomly generated smooth waveform Figure 3.5d shows the result of the randomly perturbed spiral design. Again, as expected the waveforms are gradient and slew limited.

3.9 Discussion

The results in the previous section show the main advantage of the proposed design method; it can be applied to *any* type of k-space trajectory: from simple rectilinear to the complicated randomized trajectories. The trajectory and gradient design is simplified to designing a parametric curve, and then designing the gradients for it. The latter, is also a disadvantage of the method, since a poor choice of trajectory that has sharp curvatures may lead to a poor design. Nevertheless, as shown in the examples, the simplicity and generality of the method makes it a powerful design tool.

3.9.1 Numerical Issues

The differential equations Eq. (3.19) and Eq. (3.20) have the gradient magnitude in the denominator. When the gradient magnitude is very small, the right hand side of the equation can become very large and cause inaccuracies in the integration. This can be mitigated by choosing a smaller step size, which increases the computation time and memory consumption. However, since slowing gradients are associated with large curvature locations in the k-space trajectory that are known in advance, one can use a variable step size that is adapted to the curvature.

3.9.2 Computational Complexity

The proposed method is non-iterative and provides a direct time-optimal solution. The computational complexity is linear with respect to the length of the trajectory curve, and requires a solution to an ODE propagated forward and backwards in time. The implementation used a fixed step 4^{th} order Runge-Kutte method, which required a relatively small step size to maintain accuracy. Nevertheless, the C programming language implementation required about half a second to design the waveforms in the examples.

3.10 Conclusions

I have provided a fast, simple and non-iterative method for designing the time-optimal gradient waveforms for any k-space trajectory. It is the complete solution for the gradient waveform design with a k-space path constraint. I have demonstrated some simple time-optimal waveforms by first designing a parametric curve and using the design method to produce the gradient waveforms.

Chapter 4

Compressed Sensing MRI

4.1 Introduction

Imaging speed is important in many MRI applications. However, the speed at which data can be collected in MRI is fundamentally limited by physical (gradient amplitude and slew-rate) and physiological (nerve stimulation) constraints. Therefore many researchers are seeking methods to reduce the amount of acquired data without degrading the image quality.

When k-space is undersampled, the Nyquist criterion is violated, and Fourier reconstructions exhibit aliasing artifacts. Many previous proposals for reduced data imaging try to mitigate undersampling artifacts. They fall in three groups: (a) Methods generating artifacts that are incoherent or less visually apparent, at the expense of reduced apparent SNR [32,72,82,92,104]; (b) Methods exploiting redundancy in k-space, such as partial-Fourier, parallel imaging etc. [73,84,97]; (c) Methods exploiting either spatial or temporal redundancy or both [58, 70, 77, 106, 108].

This chapter's aim is to exploit the *sparsity* which is implicit in MR images, and develop an approach combining elements of approaches a and c. Implicit sparsity means *transform sparsity*, *i.e.*, the underlying object of interest happens to have a sparse representation in a known and fixed mathematical transform domain. To begin with, consider the identity transform, so that the transform domain is simply the image domain itself. Here sparsity means that there are relatively few significant pixels with nonzero values. For example, angiograms are extremely sparse in the pixel representation. More complex medical images may not be sparse in the pixel representation, but they do exhibit transform sparsity, since they have a sparse representation in terms of spatial finite differences, in terms of their wavelet coefficients, or in terms of other transforms.

Sparsity is a powerful constraint, generalizing the notion of finite object support. It is well understood why support constraints in image space (*i.e.*, small FOV or bandpass sampling) enable sparser sampling of k-space. Sparsity constraints are more general because nonzero coefficients do not have to be bunched together in a specified region. Transform sparsity is even more general because the sparsity needs only to be evident in some transform domain, rather than in the original image (pixel) domain. Sparsity constraints, under the right circumstances, can enable sparser sampling of k-space as well [11, 21].

The possibility of exploiting transform sparsity is motivated by the widespread success of data compression in imaging. Natural images have a well-documented susceptibility to compression with little or no visual loss of information. Medical images are also compressible [56]. Underlying the most well-known image compression tools such as JPEG, and JPEG-2000 [103] are the Discrete Cosine transform (DCT) and wavelet transform. These transforms are useful for image compression because they transform image content into a vector of sparse coefficients; a standard compression strategy is to encode the few significant coefficients and store them, for later decoding and reconstruction of the image.

The widespread success of compression algorithms with real images raises the following questions: Since the images intend to be acquired will be compressible, with most transform coefficients negligible or unimportant, is it really necessary to acquire all that data in the first place? Can we not simply measure the compressed information directly from a small number of measurements, and still reconstruct the same image which would arise from the fully sampled set? Furthermore, since MRI measures Fourier coefficients, and not pixels, wavelet or DCT coefficients, the question is whether it is possible to do the above by measuring only a subset of k-space.

A substantial body of mathematical theory has recently been published establishing the possibility to do exactly this. The formal results can be found by searching for the phrases *compressed sensing* (CS) or *compressive sampling* [11, 21]. According to these mathematical results, if the underlying image exhibits transform sparsity, and if k-space undersampling results in incoherent artifacts in that transform domain, then the image can be recovered from randomly undersampled frequency domain data, provided an appropriate nonlinear recovery scheme is used.

4.2 Compressed Sensing

CS was first proposed in the literature of Information Theory and Approximation Theory in an abstract general setting. One measures a small number of random linear combinations of the signal values – much smaller than the number of signal samples nominally defining it. The signal is reconstructed with good accuracy from these measurements by a non-linear procedure. Sampling in MRI is a special case of CS, where the sampled linear combinations are simply individual Fourier coefficients (k-space samples). In that setting, CS is claimed to be able to make accurate reconstructions from a small subset of k-space, rather than an entire k-space grid.

The CS approach requires that: (a) the desired image have a sparse representation in a known transform domain (*i.e.*, is compressible), (b) the aliasing artifacts due to k-space undersampling be incoherent (noise like) in that transform domain. (c) a non-linear reconstruction be used to enforce both sparsity of the image representation and consistency with the acquired data. To help keep in mind these ingredients, consider Fig. 4.1, which depicts relationships among some of these main concepts. It shows the image, the k-space and the transform domains, the operators connecting these domains and the requirements for CS.

4.2.1 A simple, intuitive example of compressed sensing

To get intuition for the importance of incoherence and the feasibility of CS in MRI, consider the example in Fig. 4.2. A sparse 1D signal (Fig. 4.2a), 256 samples long, is

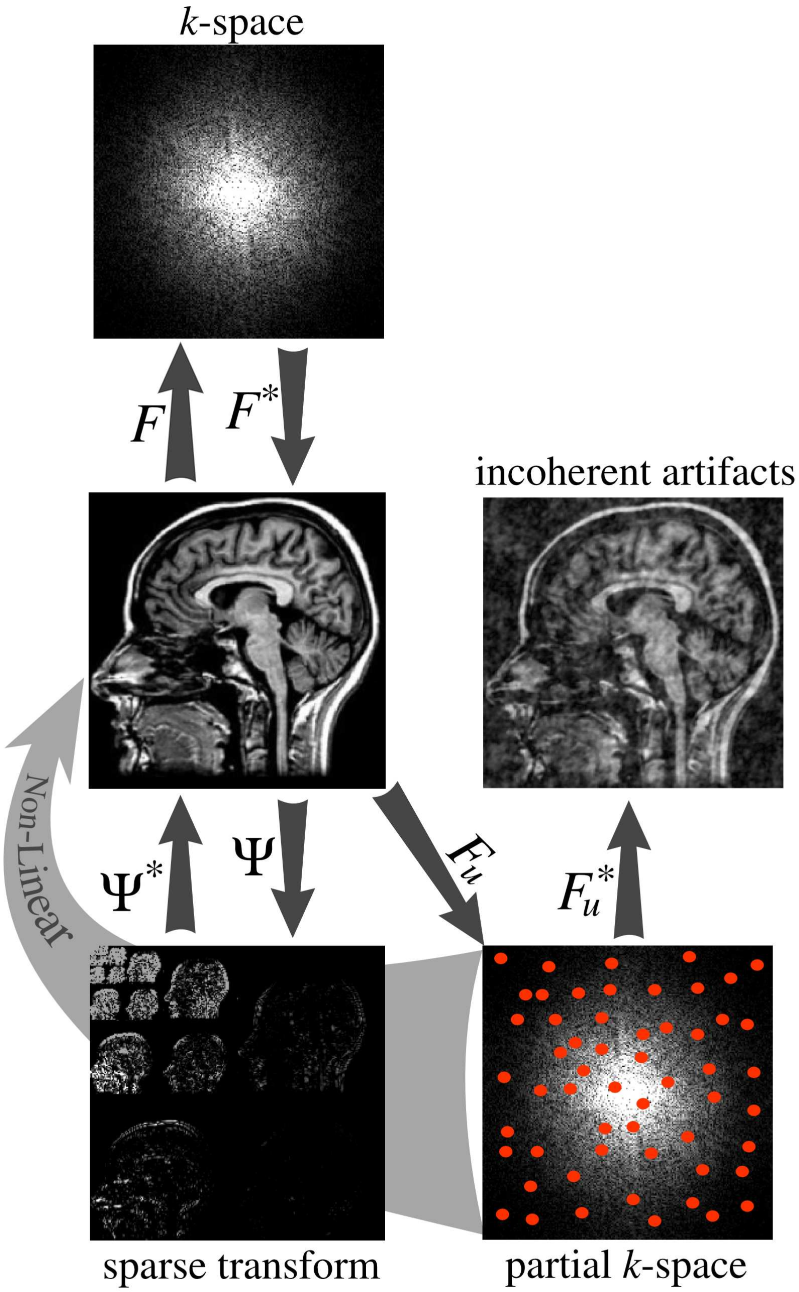
k-space
F
F^*
incoherent artifacts
Non-Linear
Ψ^*
Ψ
F_u
F_u^*
sparse transform
partial k-space

Figure 4.1: Illustration of the domains and operators used in the chapter as well as the requirements of CS: sparsity in the transform domain, incoherence of the undersampling artifacts and the need for nonlinear reconstruction that enforces sparsity.

undersampled in k-space (Fig. 4.2b) by a factor of eight. Here, the sparse transform is simply the identity. Later, I will consider the case where the transform is nontrivial.

Equispaced k-space undersampling and reconstruction by zero-filling results in coherent aliasing, a superposition of shifted replicas of the signal as illustrated in Fig. 4.2c. In this case, there is an inherent ambiguity; it is not possible to distinguish between the original signal and its replicas, as they are all equally likely.

Random undersampling results in a very different situation. The zero-filling Fourier reconstruction exhibits incoherent artifacts that actually behave much like additive random noise (Fig. 4.2d). Despite appearances, the artifacts are not noise; rather, undersampling causes leakage of energy away from each individual nonzero coefficient of the original signal. This energy appears in other reconstructed signal coefficients, including those which had been zero in the original signal.

It is possible, if all the underlying original signal coefficients are known, to calculate this leakage analytically. This observation enables the signal in Fig. 4.2d to be accurately recovered although it was 8-fold undersampled. An intuitive plausible recovery procedure is illustrated in Fig. 4.2e-h. It is based on thresholding, recovering the strong components, calculating the interference caused by them and subtracting it. Subtracting the interference of the strong components reduces the total interference level and enables recovery of weaker, previously submerged components. By iteratively repeating this procedure, one can recover the rest of the signal components. A recovery procedure along these lines was proposed by Donoho et. al [24] as a fast approximate algorithm for CS reconstruction.

4.3 Sparsity

There are two ways to express the mathematical notion of sparsity of a vector of coefficients. In *strong sparsity* most of the coefficients are required to be exact zeros; this can be quantified by the fraction of nonzero entries. In *weak sparsity* most of the coefficients are very small, but need not be exactly zero; this can be quantified by the rate at which the sorted nonzero amplitudes decay. For real data, it is rare for transform coefficients to be exactly zero; hence, in MR imaging weak sparsity is

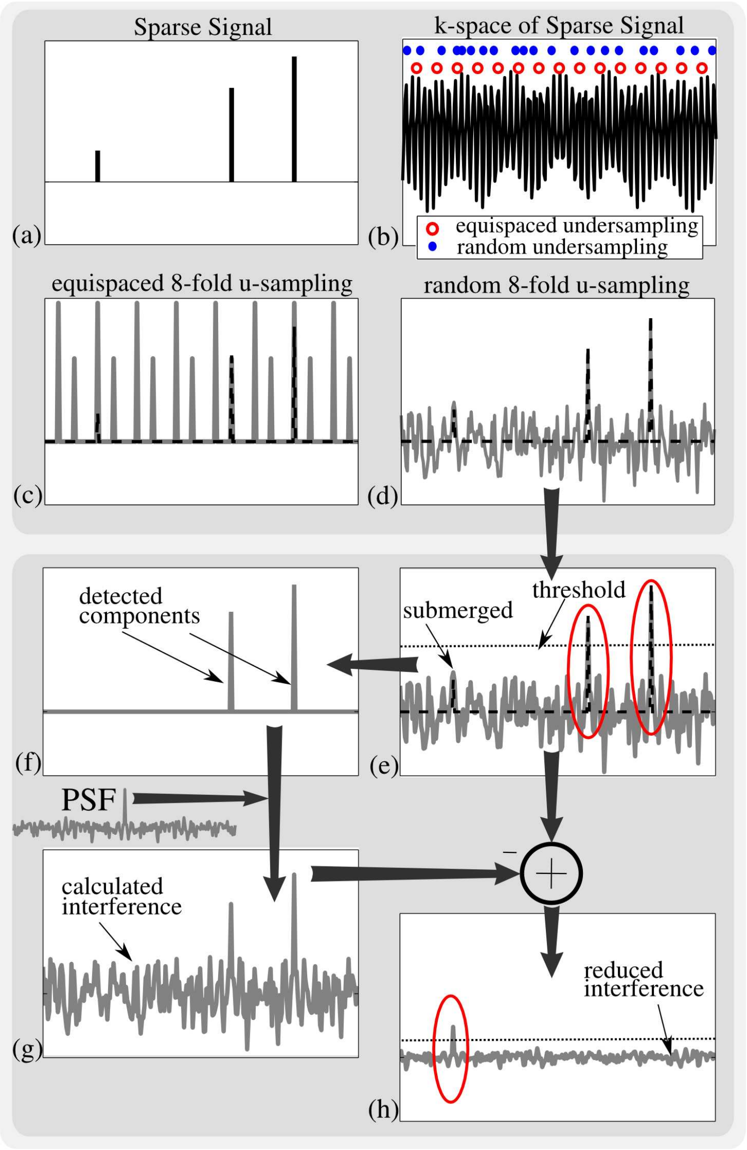
Sparse Signal
k-space of Sparse Signal
(a)
(b)
equispaced undersampling
random undersampling
equispaced 8-fold u-sampling
random 8-fold u-sampling
(c)
(d)
detected components
submerged
threshold
(f)
(e)
PSF
−
calculated interference
reduced interference
(g)
(h)

Figure 4.2: An intuitive reconstruction of a sparse signal from pseudo-random k-space undersampling. A sparse signal (a) is 8-fold undersampled in k-space (b). Equispaced undersampling results in coherent signal aliasing (c) that can not be recovered. Pseudo-random undersampling results in incoherent aliasing (d). Strong signal components stick above the interference, are detected (e) and recovered (f) by thresholding. The interference of these components is computed (g) and subtracted (h), lowering the total interference level and enabling recovery of weaker components.

the better model. For mathematical analysis in compressed sensing, the conclusions are frequently stronger and the arguments more transparent when assuming strong sparsity, so it is helpful to keep both models in mind.

4.3.1 Sparsifying Transform

A sparsifying transform is an operator mapping a vector of image data to a sparse vector. In recent years, there has been extensive research in sparse image representation. As a result we currently possess a library of diverse transformations that can sparsify many different type of images [100].

For example, piecewise constant images can be sparsely represented by spatial finite-differences (*i.e.*, computing the differences between neighboring pixels) ; indeed, away from boundaries, the differences vanish. Real-life MR images are of course not piecewise smooth. But in some problems, where boundaries are the most important information (angiograms for example) computing finite-differences results in a sparse representation.

Natural, real-life images are known to be sparse in the discrete cosine transform (DCT) and wavelet transform domains [10, 29, 80, 103]. The DCT is central to the JPEG image compression standard and MPEG video compression, and is used billions of times daily to represent images and videos. The wavelet transform is used in the JPEG-2000 image compression standard [103]. The wavelet transform is a multi-scale representation of the image. Coarse-scale wavelet coefficients represent the low resolution image components and fine-scale wavelet coefficients represent high resolution components. Each wavelet coefficient carries both spatial position and spatial frequency information at the same time (see Fig. 4.5b (top) for a spatial position and spatial frequency illustrations of a mid-scale wavelet coefficient).

Since computing finite-differences of images is a high-pass filtering operation, the finite-differences transform can also be considered as computing some sort of fine-scale wavelet transform (without computing coarser scales).

Sparse representation is not limited to still imagery. Many still images can be compressed 5 to 10-fold without perceptible loss of visual information, but often videos can safely be compressed much more heavily. This is demonstrated by the success of MPEG, which uses the fact that some parts of a movie are either constant or else undergo motion that is similar between neighboring pixels. Inter-frame temporal differences of video content are often sparse, so movies are sparsified by temporal finite differences. Dynamic MR images are highly compressible as well. For example, heart images are quasi-periodic. Therefore, their temporal Fourier transform is sparse. The hemodynamic response of brain activation in some functional MRI experiments can also be sparsified by temporal Fourier transform.

4.3.2 The Sparsity of MR Images

The transform sparsity of MR images can be demonstrated by applying a sparsifying transform to a fully sampled image and reconstructing an approximation to the image from a subset of the largest transform coefficients. The sparsity of the image is the percentage of transform coefficients sufficient for diagnostic-quality reconstruction. Of course the term ‘diagnostic quality’ is subjective. Nevertheless for specific applications it is possible to get an empirical sparsity estimate by performing a clinical trial and evaluating reconstructions of many images quantitatively or qualitatively.

To illustrate this, consider Fig. 4.3 in which a typical brain image was compressed with wavelets, a slice of an angiogram was compressed with finite-differences, and the time series of a cross section of a dynamic heart sequence was compressed by temporal Fourier transform. The important information is captured by 10%, 5% and 5% of the largest transform coefficients, respectively.

4.4 Incoherent Sampling

“Randomness is too important to be left to chance”

Incoherent aliasing interference in the sparse transform domain is an essential ingredient for CS. This can be well understood from the previous simple 1D example. In the original CS papers [11, 21], sampling a completely random subset of k-space was chosen to simplify the mathematical proofs and in particular to guarantee a very high

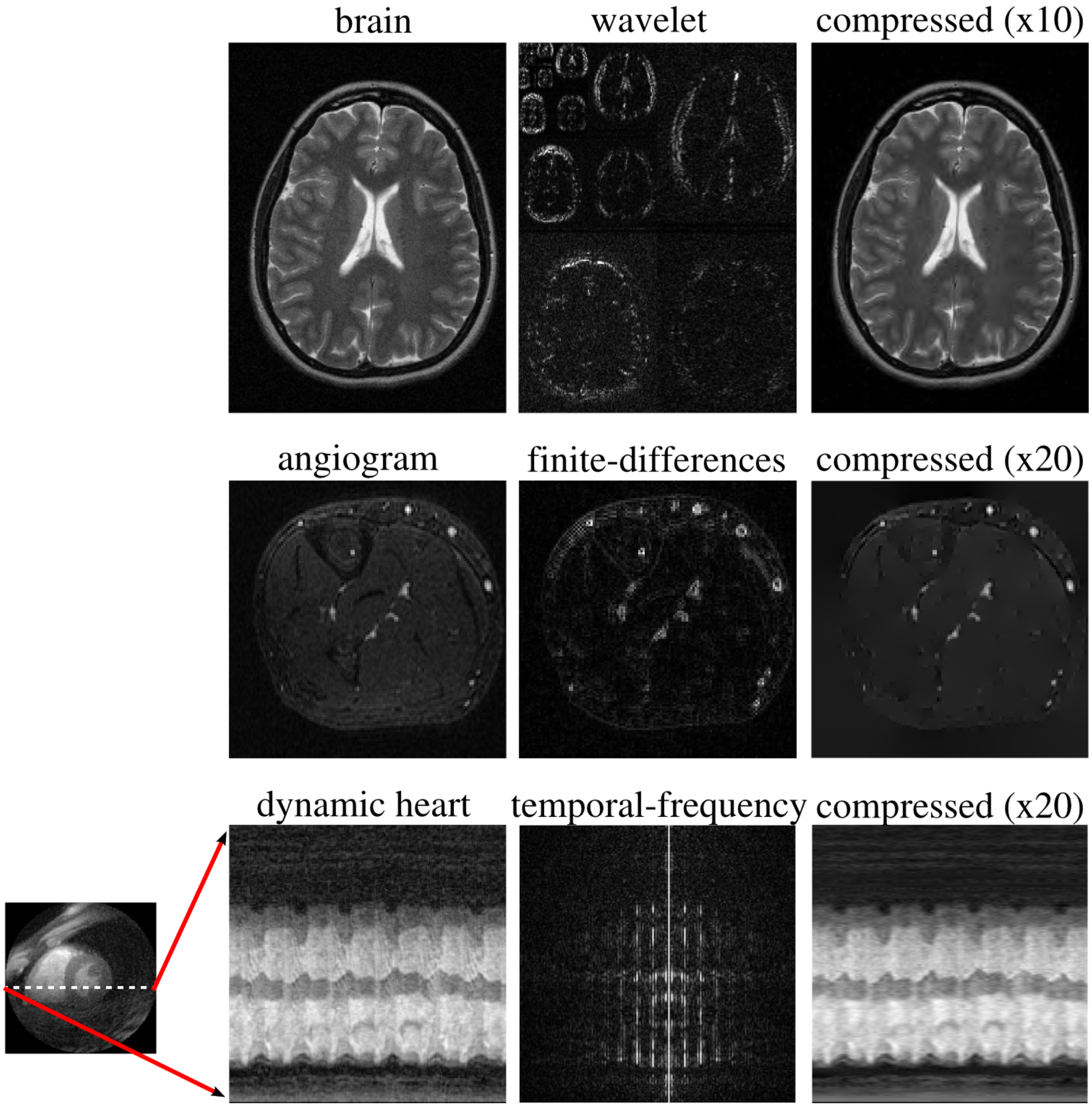
brain
wavelet
compressed (x10)
angiogram
finite-differences
compressed (x20)
dynamic heart
temporal-frequency
compressed (x20)

Figure 4.3: Transform sparsity of MR images. Fully sampled images (left) are mapped by a sparsifying transform to a transform domain (middle); the several largest coefficients are preserved while all others are set to zero; the transform is inverted forming a reconstructed image (right).

degree of incoherence.

Random point k-space sampling in all dimensions is generally impractical as the k-space trajectories have to be relatively smooth due to hardware and physiological considerations. Instead, the aim is to design a practical *incoherent sampling scheme* that mimics the interference properties of pure random undersampling as closely as possible yet allows rapid collection of data.

There are numerous ways to design incoherent sampling trajectories. In order to focus and simplify the discussion, I consider mostly the case of Cartesian grid sampling where the sampling is restricted to undersampling the phase-encodes and fully sampled readouts. In addition, I briefly discuss alternative non-Cartesian sampling trajectories.

I focus on Cartesian sampling because it is by far the most widely used in practice. It is simple and also highly robust to numerous sources of imperfection. Non-uniform undersampling of phase encodes in Cartesian imaging has been proposed in the past as an acceleration method because it produces incoherent artifacts [32, 72, 104]– exactly what is necessary. Undersampling phase-encode lines offers pure randomness in the phase-encode dimensions, and a scan time reduction that is exactly proportional to the undersampling. Finally, implementation of such an undersampling scheme is simple and requires only minor modifications to existing pulse sequences.

4.4.1 PSF and TPSF Analysis

When designing an incoherent sampling trajectory, how does one guarantee that the choice of sampling is indeed incoherent? Is choosing indices at random "incoherent enough"? Is one random choice better than other choices? In order to answer these questions, one needs a metric to evaluate the incoherence of the aliasing interference.

The point spread function (PSF) is a natural tool to measure incoherence. Let $\mathcal{F}_u$ be the undersampled Fourier operator and let e_i be the i^{th} vector of the natural basis (*i.e.*, having '1' at the i^{th} location and zeroes elsewhere). Then $PSF(i;j) = e_j^* \mathcal{F}_u^* \mathcal{F}_u e_i$ measures the contribution of a unit-intensity pixel at the i^{th} position to a pixel at the j^{th} position. Under Nyquist sampling there is no interference between pixels and $PSF(i;j)|_{i \neq j} = 0$. Undersampling causes pixels to interfere and $PSF(i;j)|_{i \neq j}$ to assume nonzero values. A simple measure to evaluate the incoherence is the maximum of the sidelobe-to-peak ratio (SPR), $\max_{i \neq j} \left| \frac{PSF(i,j)}{PSF(i,i)} \right|$.

The PSF of pure 2D random sampling, where samples are chosen at random from a Cartesian grid, offers a standard for comparison. In this case $PSF(i;j)|_{i \neq j}$ looks random as illustrated in Fig. 4.5a. Empirically, the real and the imaginary parts separately behave much like zero-mean random white Gaussian noise as illustrated in Fig. 4.4. The standard deviation of the observed SPR depends on the number, N, of samples taken and the number, D, of grid points defining the underlying image. For a constant sampling reduction factor $p = \frac{D}{N}$ the standard deviation obeys the formula:

$$\sigma_{\text{SPR}} = \sqrt{\frac{p-1}{D}}. \tag{4.1}$$

Eq. 4.1 is easily derived. The total energy in the PSF is $\frac{N}{D}$ and the energy of the main lobe is $\left(\frac{N}{D}\right)^2$. The off-center energy is therefore $\frac{N}{D} - \left(\frac{N}{D}\right)^2$. Normalizing by the number of off-center pixels and also by the main lobe's energy and setting $p = \frac{D}{N}$ we get Eq. 4.1.

The MR images of interest are typically sparse in a transform domain rather than the usual image domain. In such a setting, incoherence is analyzed by generalizing the notion of PSF to *Transform Point Spread Function* (TPSF) which measures how a single transform coefficient of the underlying object ends up influencing other transform coefficients of the measured undersampled object.

Let Ψ be an orthogonal sparsifying transform (non-orthogonal TPSF analysis is beyond the scope and is not discussed here). The $TPSF(i;j)$ is given by the following equation,

$$TPSF(i;j) = e_j^* \Psi \mathcal{F}_u^* \mathcal{F}_u \Psi^* e_i \ . \tag{4.2}$$

In words, a single point in the transform space at the i^{th} location is transformed to the image space and then to the Fourier space. The Fourier space is subjected to undersampling, then transformed back to the image space. Finally, a return is made to the transform domain and the j^{th} location of the result is selected. An example using an orthogonal wavelet transform is illustrated by Fig. 4.5b. The size

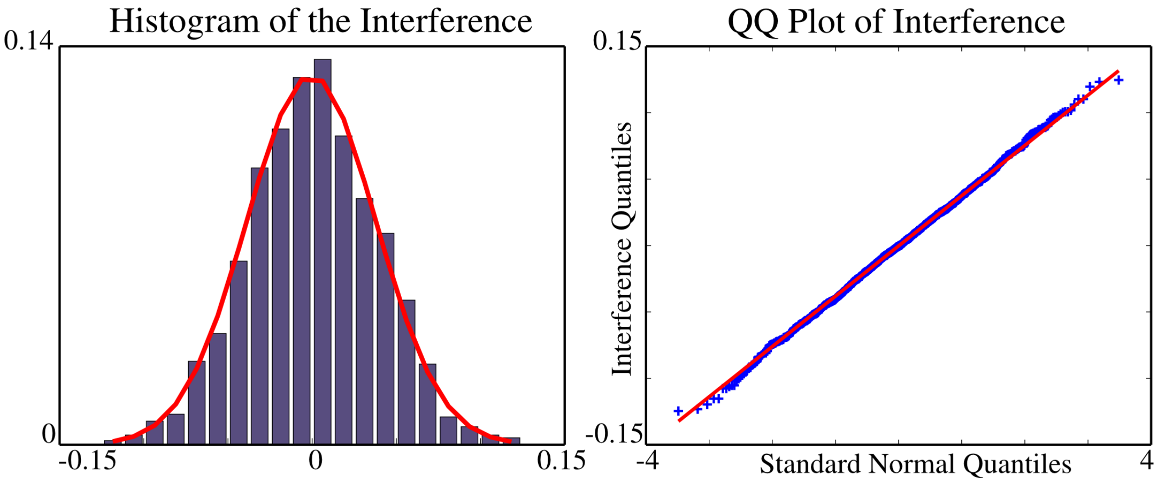

Figure 4.4: Empirical statistical evaluation of interference in the PSF for pure random sampling. 1D PSF, 2048 samples, 4-fold undersampling. left: Histogram, showing a Gaussian distribution of the off-center point-spread function. right: Quantile-Quantile plot (QQ plot) of the off-center point spread versus standard normal distribution. In such a QQ plot, if samples come from a normal distribution, the plot is linear. In this case, the plot is indeed close to linear, and therefore the sidelobe values have an approximately normal distribution.

of the sidelobes in $TPSF(i;j)|_{i\neq j}$ are used to measure the incoherence of a sampling trajectory. A small $TPSF(i;j)|_{i\neq j}$ with random noise-like statistics is desirable.

4.4.2 Single-slice 2DFT, multi-slice 2DFT and 3DFT Imaging

Equipped with the PSF and $TPSF$ analysis tools, I consider three cases of Cartesian sampling: 2DFT, multi-slice 2DFT and 3DFT. In single-slice 2DFT, only the phase encodes are undersampled and the interference spreads only along a single dimension. The interference standard deviation as calculated in Eq. 4.1 is $D^{1/4}$ times larger than the theoretical pure random 2D case for the same acceleration – (16 times for a 256×256 image). Therefore in 2DFT one can expect relatively modest accelerations because mostly 1D sparsity is exploited.

In multi-slice 2DFT sampling is done in a hybrid k-space vs. image space ($k_y - z$ space). Undersampling differently the phase-encodes of each slice randomly undersamples the $k_y - z$ space. This can reduce the peak sidelobe in the TPSF of some appropriate transforms, such as wavelets, as long as the transform is also applied in the slice dimension. Hence, it is possible to exploit some of the sparsity in the slice dimension as well. Figure 4.6a-b shows that undersampling each slice differently has reduced peak sidelobes in the TPSF compared to undersampling the slices the same way. However, it is important to mention that for wavelets, randomly undersampling in the hybrid $k_y - z$ space is not as effective, in terms of reducing the peak sidelobes, as randomly undersampling in a pure 2D k-space (Fig. 4.6c). The method of multi-slice 2DFT will work particularly well when the slices are thin and finely spaced. When the slices are thick and with gaps, there is little spatial redundancy in the slice direction and the performance of the reconstruction would be reduced to the single-slice 2DFT case. Undersampling with CS can be used to bridge gaps or acquire more thinner slices without compromising the scan time.

Randomly undersampling the 3DFT trajectory is the preferred method. Here, it is possible to randomly undersample the 2D phase encode plane ($k_y - k_z$) and achieve the theoretical high degree of 2D incoherence. Additionally, 2D sparsity is fully exploited,

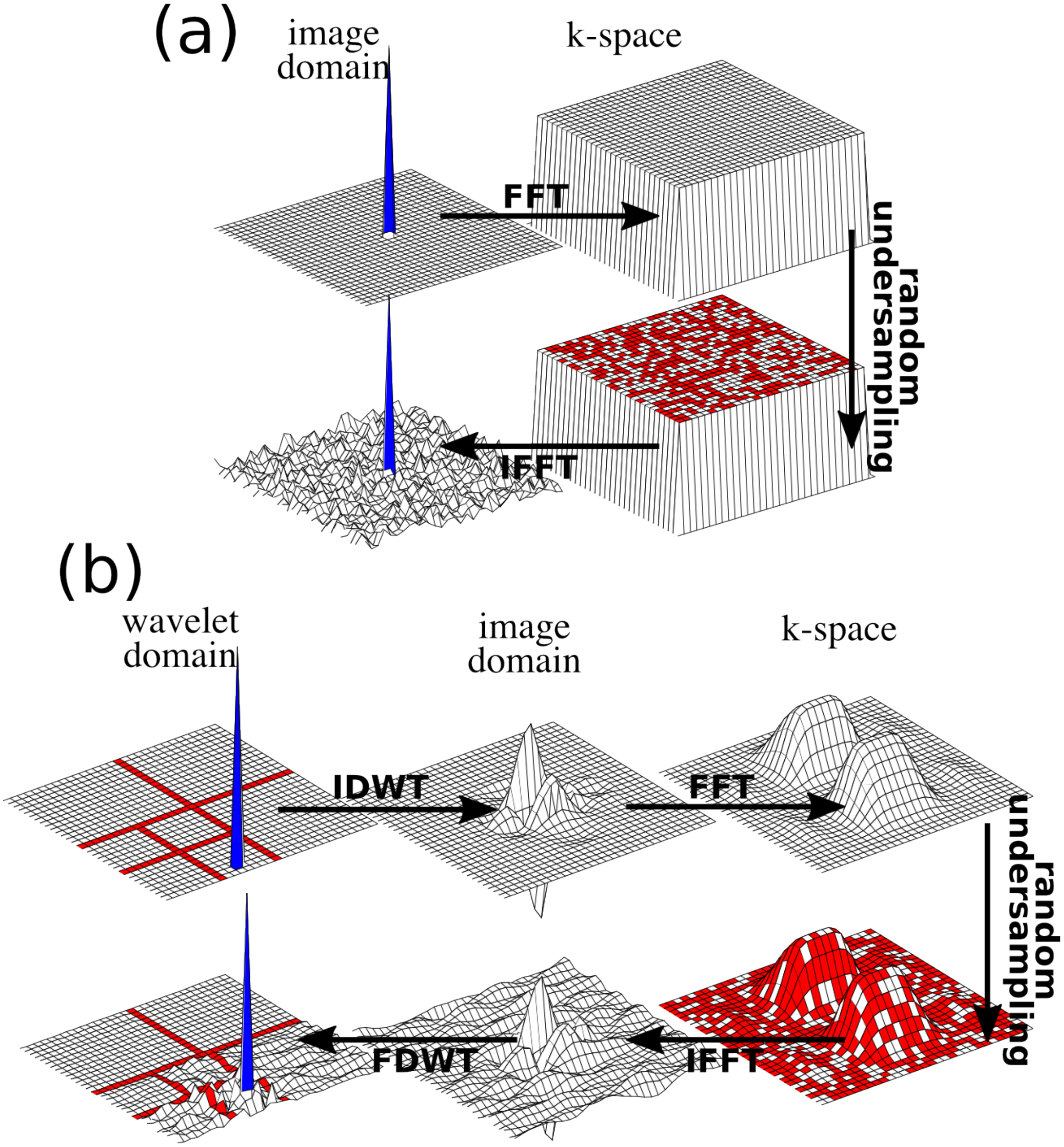
(a)
image domain
k-space
FFT
random undersampling
IFFT
(b)
wavelet domain
image domain
k-space
IDWT
FFT
random undersampling
FDWT
IFFT

Figure 4.5: (a) The PSF of random 2D k-space undersampling. (b) The wavelet TPSF of random 2D Fourier undersampling. FDWT and IDWT stand for forward and inverse discrete wavelet transform. Wavelet coefficients are band-pass filters and have limited support both in space and frequency. Random k-space undersampling results in incoherent interference in the wavelet domain. The interference spreads mostly within the wavelet coefficients of the same scale and orientation.

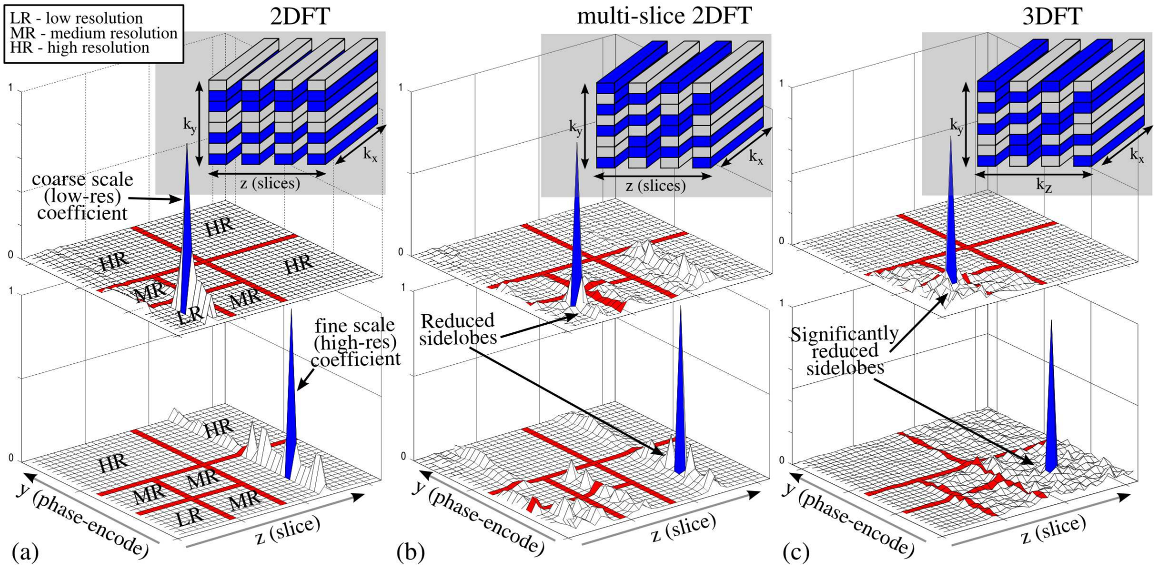

Figure 4.6: Transform point spread function (TPSF) analysis in the wavelet domain. The k-space sampling patterns and the associated TPSF of coarse-scale and fine-scale wavelet coefficients are shown. (a) Random phase encode undersampling spreads the interference only in 1D and mostly within the same wavelet scale. The result is relatively high peak interference. (b) Sampling differently for each slice, *i.e.*, randomly undersampling the $k_y - z$ plane causes the interference to spread to nearby slices and to other wavelets scales and reduces its peak value. (c) Undersampling the phase encode plane, *i.e.*, $k_y - k_z$ spreads the interference in 2D and results in the lowest peak interference.

and images have a sparser representation in 2D. 3D imaging is particularly attractive because it is often time consuming and scan time reduction is a higher priority than 2D imaging. Figure 4.6c illustrates the proposed undersampled 3DFT trajectory and its wavelet TPSF. The peak interference of the wavelet coefficients is significantly reduced compared to multi-slice and plain 2DFT undersampling.

4.4.3 Variable Density Random Undersampling

The incoherence analysis so far assumes the few non-zeros are scattered at random among the entries of the transform domain representation. Representations of natural images exhibit a variety of significant non-random structures. First, most of the energy of images is concentrated close to the k-space origin. Furthermore, using wavelet analysis one can observe that coarse-scale image components tend to be less sparse than fine-scale components. Figure 4.7a. shows that uniformly undersampling at random affects coarse scale coefficients more than fine scale. The coarse scale wavelet coefficients are mostly submerged in the interference, resulting in coherent low-resolution interference in the image.

These observations show that, for a better performance with 'real images', one should be undersampling less near the k-space origin and more in the periphery of k-space. For example, one may choose samples randomly with sampling density scaling according to a power of distance from the origin. Empirically, using density powers of 1 to 6 greatly reduces the total interference and, as a result, iterative algorithms converge faster with better reconstruction. The optimal sampling density is beyond the scope of this thesis, and should be investigated in future research. Figure 4.7b shows that using a variable density scheme significantly reduces the interference in the coarse scale, achieving an overall better coherence property.

4.4.4 Other Incoherent Sampling Schemes

Getting completely away from a Cartesian grid allows far greater flexibility in designing sampling trajectories with low coherence. Popular non-Cartesian schemes include sampling along radial lines or spirals. Traditionally, undersampled radial trajectories have been used [1, 82, 92] to accelerate acquisitions, because the artifacts from linear reconstruction seem benign and incoherent – much like adding noise to the image. Variable-density spirals [60, 104] and also variable density Cartesian acquisitions [32, 72, 78] have been proposed for similar reasons. From ones perspective, it is recognized that such artifacts are benign because *the corresponding PSFs are incoherent*. Figure 4.8c-f shows the PSF of several such trajectories: radial, uniform spirals, variable density spirals and variable density perturbed spirals. These trajectories are strong candidates for CS: with appropriate nonlinear reconstruction, the seeming noise-like artifacts can be suppressed without degrading image quality.

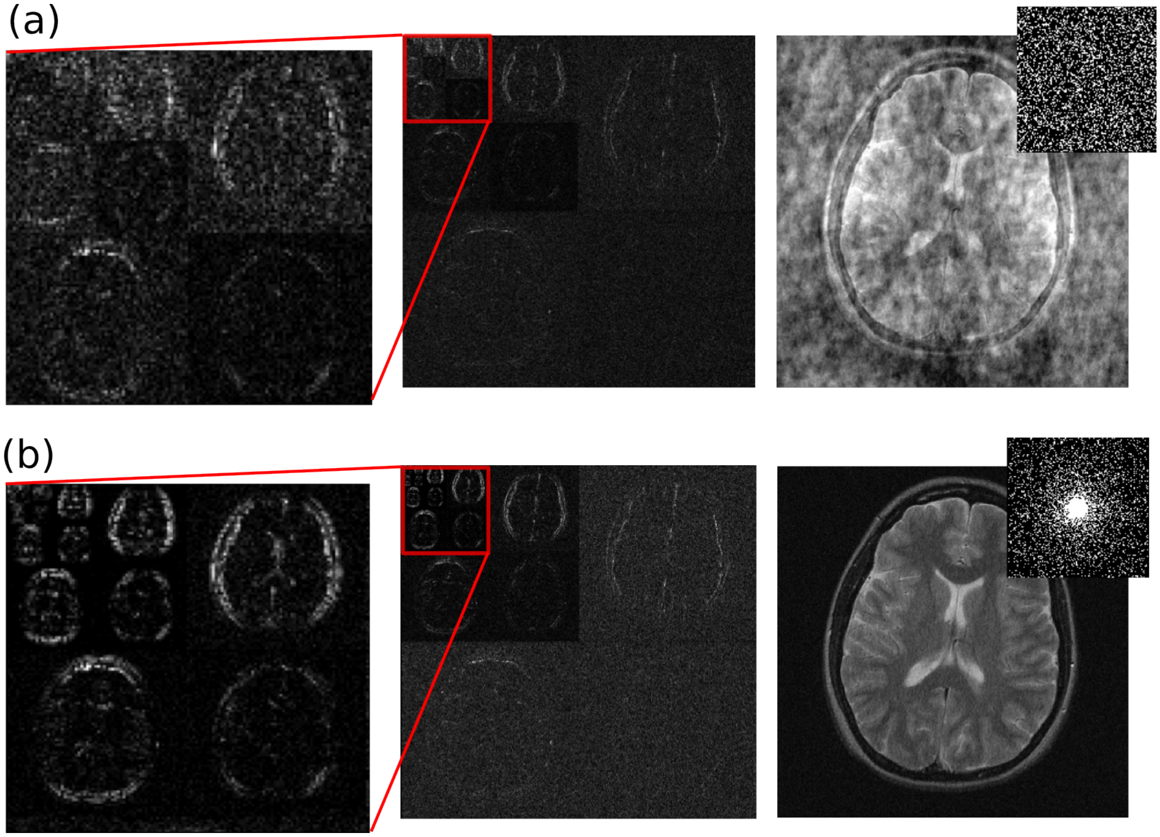

Figure 4.7: (a) For natural images, uniform random undersampling often results in coherent interference. This is attributed to large interference of coarse scale wavelet coefficients. (b) Variable density undersampling, reduces the interference in the coarse scales, effectively achieving better coherence properties.

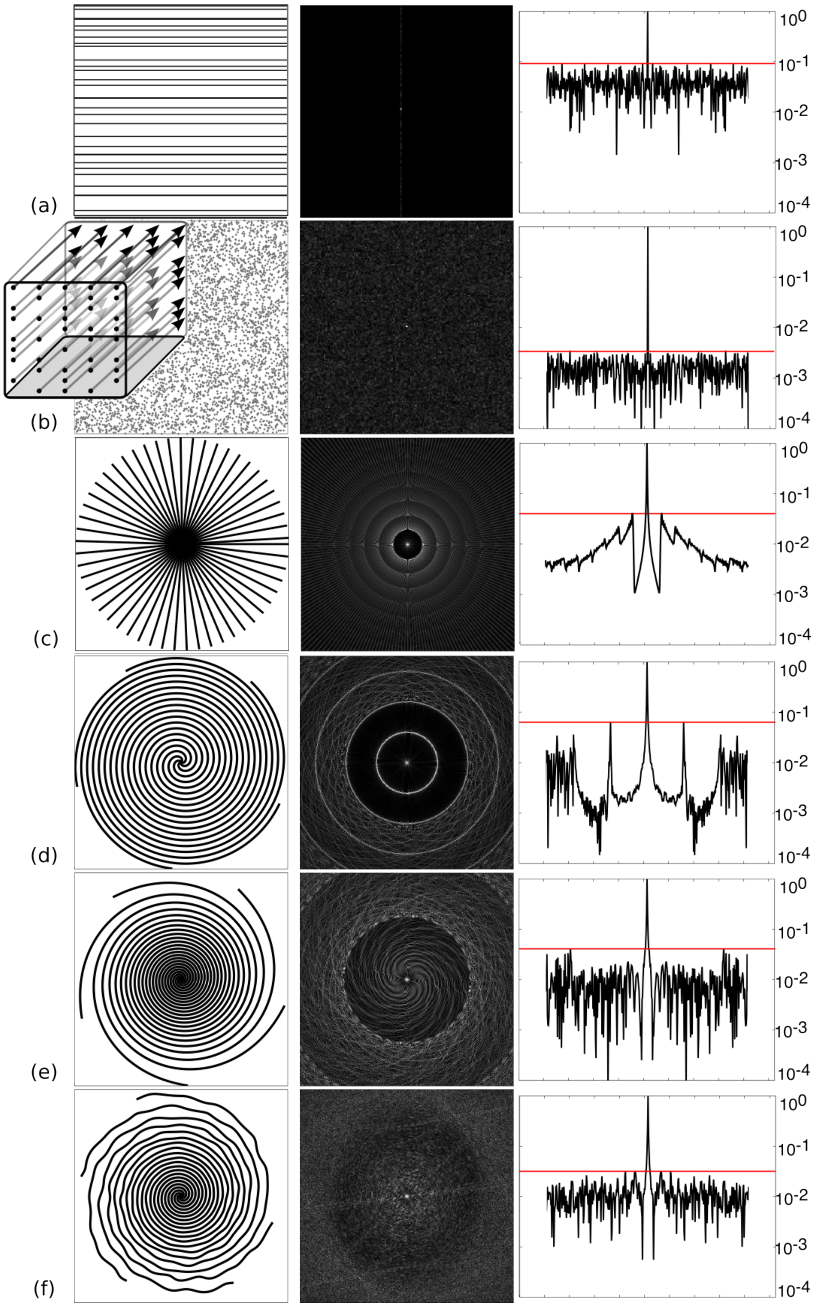

Figure 4.8: Point Spread Functions (PSF) of various sampling trajectories. (a) Random lines in 2D (b) Random points in 2D, or cross-section of random lines in 3D (c) Radial (d) Uniform spirals (e) Variable density spirals (f) Variable density perturbed spirals

4.4.5 Dynamic Incoherent Sampling: k-t Sparse

A dynamic sequence of images is a multi-dimensional signal with two or three spatial coordinates and time as an additional dimension (See Fig. 4.9 top-left panel). Dynamic MRI data are acquired in the spatial frequency vs time ($k-t$) domain. Traditional collection schemes sample the $k-t$ domain on a regular set of congruent lines (Fig. 4.9 top-right). Often, it is impossible to meet the spatio-temporal Nyquist-rate this way. Then, sub-Nyquist sampling, followed by linear reconstruction, causes coherent aliasing of the signal in the spatial-temporal frequency ($x-f$) domain. As an alternative, randomly ordering a subset of k-space lines (Fig. 4.9 bottom-right) is incoherent with respect to the $x-f$ domain and produces benign artifacts in linear reconstructions [81]. So random ordering of lines is an effective and inexpensive way to incoherently sample dynamic data. Of course, the same ideas of random ordering apply to non-Cartesian sampling such as radial lines and spirals, improving incoherence and better exploiting the hardware.

Dynamic imaging CS has major advantages over static imaging: sequences of images, like videos, are highly compressible – much more than static images as illustrated in both Figs. 4.3 and 4.9 bottom-left panel. At the same time, dynamic imaging requires several orders of magnitude more samples than static imaging and it is often impossible to meet the Nyquist rate. CS compensates for this by incoherent sampling and by exploiting the inherent sparsity of the dynamic sequence.

4.4.6 How Many Samples to Acquire?

A theoretical bound on the number of Fourier sample points that need be collected with respect to the number of sparse coefficients is derived in [11, 21]. However, I as well as other researchers have observed that in practice, for a good reconstruction, the number of k-space samples should be roughly two to five times the number of sparse coefficients (The number of sparse coefficients can be calculated in the same way as in the *The Sparsity of MR Images* section). The results, presented in this chapter, support this claim. Similar observations were reported by Candès et al. [13] and by [105].

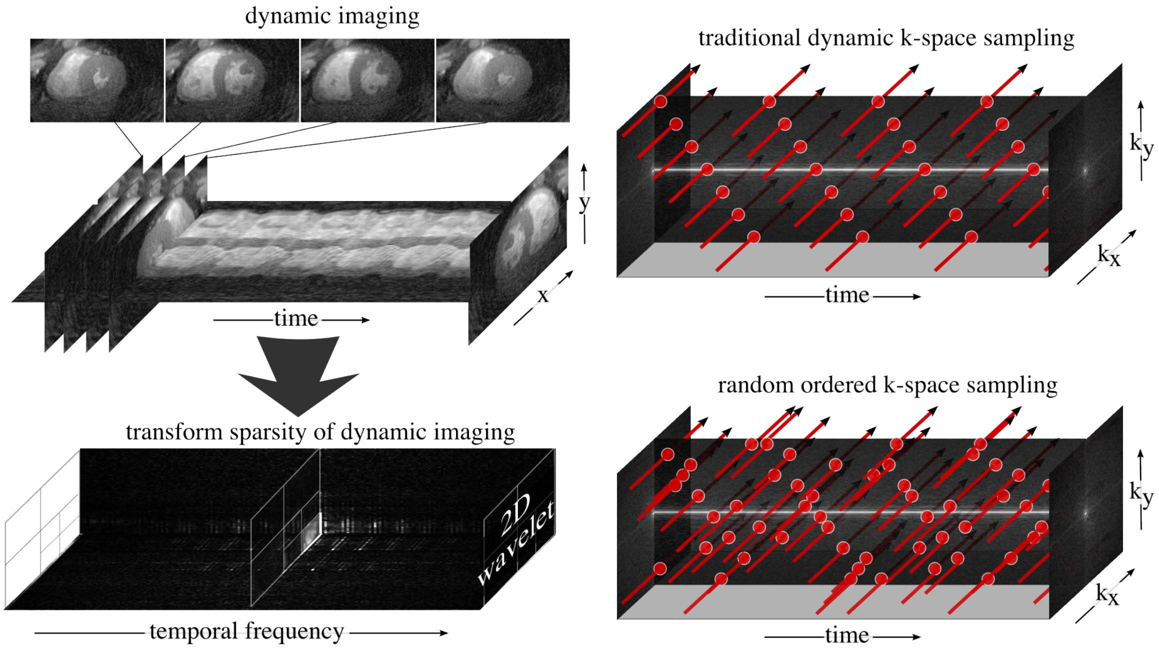
dynamic imaging
y
x
time
traditional dynamic k-space sampling
ky
kx
time
transform sparsity of dynamic imaging
2D wavelet
temporal frequency
random ordered k-space sampling
ky
kx
time

Figure 4.9: Top left:Dynamic MRI is a multi-dimensional signal with two or three spatial coordinates and time as an additional dimension. Bottom left: Dynamic images have a sparse representation in an appropriate transform domain. Top right: Traditional $k-t$ sequential sampling. Bottom right: Random ordering is an efficient way to incoherently sample the $k-t$ space.

4.4.7 Monte-Carlo Incoherent Sampling Design

Finding an optimal sampling scheme that maximizes the incoherence for a given number of samples is a combinatorial optimization problem and might be considered intractable. However, choosing samples at random often results in a good, incoherent, near-optimal solution. Therefore I propose the following Monte-Carlo design procedure: Choose a grid size based on the desired resolution and FOV of the object. Undersample the grid by constructing a probability density function (*pdf*) and randomly draw indices from that density. Variable density sampling of k-space is controlled by the *pdf* construction. A plausible choice is diminishing density according to a power of distance from the origin as previously discussed. Because the procedure is random, one might accidentally choose a sampling pattern with a “bad” $TPSF$. To prevent such situation, repeat the procedure many times, each time measure the peak interference in the TPSF of the resulting sampling pattern. Finally, choose the pattern with the lowest peak interference. Once a sampling pattern is determined it can be used again for future scans.

4.5 Image Reconstruction

4.5.1 Formulation

I now describe in more detail the processes of nonlinear image reconstruction appropriate to the CS setting. Suppose the image of interest is a vector m, let Ψ denote the linear operator that transforms from pixel representation into a sparse representation, and let $\mathcal{F}_u$ be the undersampled Fourier transform, corresponding to one of the k-space undersampling schemes discussed earlier. The reconstruction is obtained by solving the following constrained optimization problem:

$$\begin{aligned} \textit{minimize} \quad & ||\Psi m||_1 \\ s.t. \quad & ||\mathcal{F}_u m - y||_2 < \epsilon \end{aligned} \tag{4.3}$$

Here m is the reconstructed image, where y is the measured k-space data from the scanner and ϵ controls the fidelity of the reconstruction to the measured data. The threshold parameter ϵ is usually set below the expected noise level.

The objective function in Eq. 4.3 is the ℓ_1 norm, which is defined as $||x||_1 = \sum_i |x_i|$. Minimizing $||\Psi m||_1$ promotes sparsity [15]. The constraint $||\mathcal{F}_u m - y||_2 < \epsilon$ enforces data consistency. In words, among all solutions which are consistent with the acquired data, Eq. 4.3 finds a solution which is compressible by the transform Ψ.

When finite-differences operator is used as a sparsifying transform, the objective in Eq. 4.3 is often referred to as Total-Variation (TV) [87], since it is the sum of the absolute variations in the image. The objective then is usually written as $TV(m)$. Even when using other sparsifying transforms in the objective, it is often useful to include a TV penalty as well [105]. This can be considered as requiring the image to be sparse by both the specific transform and finite-differences at the same time. In this case Eq. 4.3 is written as

$$\begin{aligned} minimize \quad & ||\Psi m||_1 + \alpha TV(m) \\ s.t. \quad & ||\mathcal{F}_u m - y||_2 < \epsilon, \end{aligned}$$

where α trades Ψ sparsity with finite-differences sparsity.

The ℓ_1 norm in the objective is a crucial feature of the whole approach. Minimizing the ℓ_1 norm of an objective often results in a sparse solution. On the other hand, minimizing the ℓ_2 norm, which is defined as $||x||_2 = \left(\sum_i |x_i|^2\right)^{1/2}$ and commonly used for regularization because of its simplicity, does not result in a sparse solution and hence is not suitable for use as objective function in Eq. 4.3. Intuitively, the ℓ_2 norm penalizes large coefficients heavily, therefore solutions tend to have many smaller coefficients – hence not be sparse. In the ℓ_1 norm, many small coefficients tend to carry a larger penalty than a few large coefficients, therefore small coefficients are suppressed and solutions are often sparse.

Special purpose methods for solving Eq. 4.3 have been a focus of research interest since CS was first introduced. Proposed methods include: interior point methods [15, 54], projections onto convex sets [13], homotopy [23], iterative soft thresholding [19, 27, 99], and iteratively reweighted least squares [20, 50]. Next I describe the approach which is similar to [6, 8, 14], using non-linear conjugate gradients and backtracking line-search.

It is important to mention that some of the above iterative algorithms for solving the optimization in Eq. 4.3 in effect perform thresholding and interference cancellation at each iteration. Therefore there is a close connection between the previous simple intuitive example of interference cancellation and the more formal approaches that are described above.

4.5.2 Non-Linear Conjugate-Gradient Solution

Eq. 4.3 poses a constrained convex optimization problem. Consider the unconstrained problem in so-called Lagrangian form:

$$\underset{\boldsymbol{m}}{\operatorname{argmin}} \quad ||\mathcal{F}_u m - y||_2{}^2 + \lambda ||\Psi m||_1, \tag{4.4}$$

where λ is a regularization parameter that determines the trade-off between the data consistency and the sparsity. As is well-known, the parameter λ can be selected appropriately such that the solution of Eq. 4.4 is exactly as Eq. 4.3. The value of λ can be determined by solving Eq. 4.4 for different values, and then choosing λ so that $||F_u m - y||_2 \approx \epsilon$.

I propose solving Eq. 4.4 using a non-linear conjugate gradient descent algorithm with backtracking line search where $f(m)$ is the cost-function as defined in Eq. 4.4. The outline of the algorithm is given in table 4.1.

The conjugate gradient requires the computation of $\nabla f(m)$ which is,

$$\nabla f(m) = 2F_u^*(F_u m - y) + \lambda \nabla ||\Psi m||_1. \tag{4.5}$$

The ℓ^1 norm is the sum of absolute values. The absolute value function however, is not a smooth function and as a result Eq. 4.5 is not well defined for all values of m. Instead, I approximate the absolute value with a smooth function by using the relation $|x| \approx \sqrt{x^*x + \mu}$, where μ is a positive smoothing parameter. With this

INPUTS:

y - k-space measurements
$\mathcal{F}_u$ - undersampled Fourier operator associated with the measurements
Ψ - sparsifying transform operator
λ - a data consistency tuning constant

OPTIONAL PARAMETERS:

TolGrad - stopping criteria by gradient magnitude (default 10^{-4})
MaxIter - stopping criteria by number of iterations (default 100)
α, β - line search parameters (defaults $\alpha = 0.05$, $\beta = 0.6$)

OUTPUTS:

m - the numerical approximation to Eq. 4.4

```
% Initialization
$k = 0;\ m = 0;\ g_0 = \nabla f(m_0);\ \Delta m_0 = -g_0$
% Iterations
$\textit{while}\ (||g_k||_2 > \text{TolGrad and } k < \text{maxIter})\ \{$
    % Backtracking line-search
    $t = 1;\ \textit{while}\ (f(m_k + t\Delta m_k)\ >\ f(m_k)\ +\ \alpha t \cdot \textit{Real}\,(g_k^* \Delta m_k))\ \{\ t = \beta t\}$
    $m_{k+1}\ =\ m_k\ +\ t\Delta m_k$
    $g_{k+1}\ =\ \nabla f(m_{k+1})$
    $\gamma\ =\ \frac{||g_{k+1}||_2^2}{||g_k||_2^2}$
    $\Delta m_{k+1}\ =\ -\,g_{k+1}\ +\ \gamma \Delta m_k$
    $k = k + 1\ \}$
```

Table 4.1: Outline of the non-linear CG algorithm

approximation, $\frac{d|x|}{dx} \approx \frac{x}{\sqrt{x^*x+\mu}}$.

Now, let W be a diagonal matrix with the diagonal elements $w_i = \sqrt{(\Psi m)_i^*(\Psi m)_i + \mu}$. Equation 4.5 can be approximated by,

$$\nabla f(m) \approx 2F_u^*(F_u m - y) + \lambda \Psi^* W^{-1} \Psi m. \tag{4.6}$$

In practice, Eq. 4.6 is used with a smoothing factor $\mu \in [10^{-15}, 10^{-6}]$. The number of CG iterations varies with different objects, problem size , accuracy and undersampling. Examples in this chapter required between 80 and 200 CG iterations.

4.5.3 Low-Order Phase Correction and Phase Constrained Partial k-Space

In MRI, instrumental sources of phase errors can cause low-order phase variation. These carry no physical information, but create artificial variation in the image which makes it more difficult to sparsify, especially by finite differences. By estimating the phase variation, the reconstruction can be significantly improved. This phase estimate may be obtained using very low-resolution fully sampled k-space information. Alternatively, the phase is obtained by solving Eq. 4.3 to estimate the low-order phase, and repeating the reconstruction while correcting for the phase estimate.

The phase information is incorporated by a slight modification of Eq. 4.3,

$$\begin{aligned} minimize \quad & ||\Psi m||_1 \\ s.t. \quad & ||\mathcal{F}_u P m - y||_2 < \epsilon \end{aligned} \tag{4.7}$$

where P is a diagonal matrix whose entries give the estimated phase of each pixel.

Figure 4.10 shows a simulation of a phantom reconstructed with and without a low-order phase correction. The reconstruction using the phase estimate is significantly better.

Phase-constrained partial k-space has been used clinically for many years to accelerate the scan time by a factor of up to two [73]. The basic idea exploits the conjugate symmetry of the Fourier transform of real-valued images. A phase estimate can be used not only for phase correction, but also to include phase-constrained partial k-space with CS reconstruction. The conjugate Fourier symmetry is obtained by constraining the reconstructed image to be real after the phase correction. This is implemented by a slight modification of Eq. 4.7

$$\begin{aligned} \textit{minimize} \quad & ||\Psi m||_1 \\ s.t. \quad & ||\mathcal{F}_u P m - y||_2 < \epsilon \\ & m \in \mathbb{R}. \end{aligned} \tag{4.8}$$

4.6 Methods

Unless stated otherwise, most experiments were performed on a 1.5T Signa Excite scanner. All CS reconstructions were implemented in Matlab (The MathWorks, Inc., Natick, MA, USA) using the non-linear conjugate gradient method as described in Appendix I. Two linear schemes were used for comparison, zero-filling with density compensation (ZF-w/dc) and low-resolution (LR). ZF-w/dc consists of a reconstruction by zero-filling the missing k-space data and k-space density compensation. The k-space density compensation is computed from the probability density function from which the random samples were drawn. LR consists of reconstruction from a Nyquist sampled low-resolution acquisition. The low-resolution acquisition contained centric-ordered data with the same number of data points as the undersampled sets.

4.6.1 Simulation

To test the CS reconstruction performance and reconstruction artifacts with increased undersampling, I designed a simulation. For the simulation I constructed a phantom by placing 18 features with 6 different sizes (3 to 75 pixel area) and 3 different intensities (0.33, 0.66 and 1). The features were distributed randomly in the phantom to simulate an angiogram. The phantom had 100×100 pixels out of which 575 are

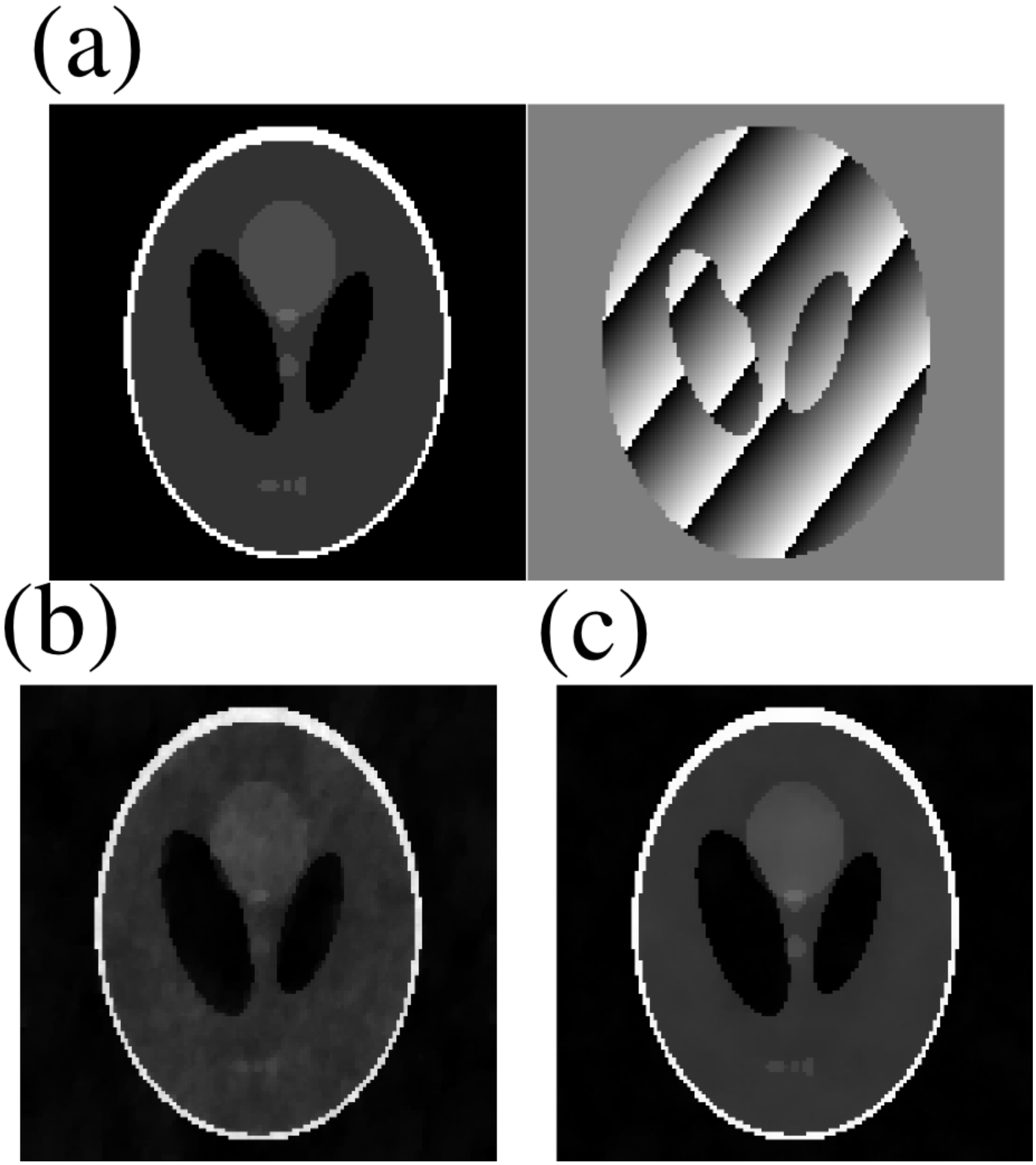

Figure 4.10: Low-order phase correction greatly improves the reconstruction. (a) Original phantom magnitude and phase images. (b) Reconstruction without linear-phase correction exhibits residual artifacts. (c) Reconstruction with linear-phase correction.

non-zero (5.75%). The finite-differences of the phantom consisted of 425 non-zeros (4.25%).

The first aim of the simulation was to examine the performance of the CS reconstruction and its associated artifacts with increased undersampling compared to the LR and ZF-w/dc methods. The second aim was to demonstrate the advantage of variable density random undersampling over uniform density random undersampling.

From the full k-space I constructed sets of randomly undersampled data with uniform density as well as variable density (density power of 12) with corresponding accelerations factors of 8, 12 and 20 (1250, 834 and 500 k-space samples). Since the phantom is sparse both in image space and by finite differences, the data were CS reconstructed by using an ℓ^1 penalty on the image as well as a TV penalty (finite differences as the sparsifying transform) in Eq. 4.3. The result was compared to the ZF-w/dc and LR linear reconstructions.

4.6.2 Undersampled 2D Cartesian Sampling in the Presence of Noise

CS reconstruction is known to be stable in the presence of noise [12, 41], and can also be used to further perform non-linear edge preserving denoising [22, 87] of the image. To document the performance of CS in the presence of noise, I scanned a phantom using a 2D Cartesian spin-echo sequence with scan parameters yielding measured $SNR = 6.17$. The k-space was undersampled by a factor of 2.5 by randomly choosing phase-encodes lines with a quadratic variable density. A CS reconstruction using a TV penalty in Eq. 4.3 was obtained, with two different consistency RMS errors of $\epsilon = 10^{-5}$ and $\epsilon = 0.1$. The result was compared to the ZF-w/dc reconstruction, and the reconstruction based on complete Nyquist sampling. Finally, the image quality as well as the resulting SNR of the reconstructions were compared.

4.6.3 Multi-slice 2DFT Fast Spin-Echo Brain Imaging

In the theory section it was shown that brain images exhibit transform sparsity in the wavelet domain. Brain scans are a significant portion of MRI scans in the clinic, and most of these are multi-slice acquisitions. CS has the potential to reduce the acquisition time, or improve the resolution of current imagery.

In this experiment I acquired a T2-weighted multi-slice k-space data of a brain of a healthy volunteer using a FSE sequence ($256\times192\times32$, $res = 0.82mm$, $slice = 3mm$, $echo - train = 15$, $TR/TE = 4200/85ms$). For each slice I acquired different sets of 80 phase-encodes chosen randomly with quadratic variable density from 192 possible phase encodes, for an acceleration factor of 2.4. The image was CS reconstructed by using a wavelet transform (Daubechies 4) as sparsifying transform together with a TV penalty in Eq. 4.3. To reduce computation time and memory load, I separated the 3D problem into many 2D CS reconstructions, *i.e.*, iterating between solving for the $y-z$ plane slices, and solving for the $x-y$ plane slices. To demonstrate the reduction in scan time, as well as improved resolution, the multi-slice reconstruction was then compared to the ZF-w/dc and LR linear reconstructions and to the reconstruction based on complete Nyquist sampling.

The TPSF analysis shows that the multi-slice approach has considerable advantage over the 2DFT in recovering coarse scale image components. To demonstrate this, the multi-slice CS reconstruction was compared to a reconstruction from data in which each slice was undersampled in the same way. To further enhance the effect, I repeated the reconstructions for data that was randomly undersampled with uniform density where the coarse scale image components are severely undersampled.

4.6.4 Contrast-Enhanced 3D Angiography

Angiography is a very promising application for CS. First, the problem matches the assumptions of CS. Angiograms appear to be sparse already to the naked eye. The blood vessels are bright with a very low background signal. Angiograms are sparsified very well by both the wavelet transform and by finite-differences. This is illustrated in Fig. 4.3 ; blood vessel information is preserved in reconstructions using only 5% of the transform coefficients. Second, the benefits of CS are of real interest in this application. In angiography there is often a need to cover a very large FOV with relatively high resolution, and the scan time is often crucial.

To test the behavior of CS for various degrees of undersampling in a controlled way, I simulated k-space data by computing the Fourier transform of a magnitude post-contrast 3DFT angiogram of the peripheral legs. The scan was RF-spoiled gradient echo (SPGR) sequence with the following parameters: $TR = 6\ ms$, $TE = 1.5\ ms$, $Flip = 30°$. The acquisition matrix was set to $480 \times 480 \times 92$ with corresponding resolution of $1 \times 0.8 \times 1\ mm$. The imaging plane was coronal with a superior-inferior readout direction.

From the full k-space set, five undersampled data sets with corresponding acceleration factors of $5,\ 6.7,\ 8,\ 10,\ 20$ were constructed by randomly choosing phase encode lines with the quadratic variable k-space density. To reduce complexity, prior to reconstruction, a 1D Fourier transform was applied in the fully sampled readout direction. This effectively creates 480 separable purely random undersampled 2D reconstructions. Finally, the images were CS reconstructed by using a TV penalty in Eq. 4.3. The result was compared to the ZF-w/dc and LR linear reconstructions.

I further tested the method, now with true k-space data on a first-pass abdominal contrast enhanced angiogram with the following scan parameters: $TR/TE = 3.7/0.96\ ms$, $FOV = 44\ cm$, $matrix = 320 \times 192 \times 32$ (with 0.625 fractional echo), $BW = 125\ kHz$.

The fully sampled data were undersampled 5-fold in retrospect with a quadratic k-space density effectively reducing the scan time from $22\ s$ to $4.4\ s$. The images were CS reconstructed from the undersampled data using a TV penalty in Eq. 4.3 and the result was again compared to the ZF-w/dc and LR linear reconstructions. To compensate for the fractional echo, a Homodyne partial-Fourier reconstruction [73] was performed in the readout direction.

4.6.5 Variable Density Spirals Whole Heart Coronary Imaging

X-ray coronary angiography is the gold standard for evaluating coronary artery disease, but it is invasive. Multi-slice x-ray CT is a non-invasive alternative, but generates high doses of ionizing radiation. MRI is emerging as a non-invasive, non-ionizing alternative [25, 74, 88, 101, 109].

Coronary arteries are constantly subject to heart and respiratory motion; high-resolution imaging is therefore a challenging task. Heart motion can be handled by synchronizing acquisitions to the cardiac cycle (cardiac gating). Respiratory motion can be mitigated by long scans with navigated breathing compensation [109], or simply by short breath-held acquisitions [25, 74]. However, breath-held cardiac-triggered collection schemes face strict timing constraints and very short imaging windows. The number of acquisitions is limited to the number of cardiac cycles in the breath-hold period. The number of heart-beats per period is itself limited – patients in need of coronary diagnosis cannot be expected to hold their breath for long! Also, each acquisition must be very short to avoid motion blurring. On top of this, many slices need to be collected to cover the whole volume of the heart. Because of these constraints, traditionally breath-held cardiac triggered acquisitions have limited spatial resolution and only partial coverage of the heart [25, 109]. Compressed sensing can accelerate data acquisition, allowing the entire heart to be imaged in a single breath-hold [89].

Figure 4.16 shows a diagram of the multi-slice acquisition. To meet the strict timing requirements, the hardware efficient spiral k-space trajectory is used. For each cardiac trigger, a single spiral in k-space is acquired for each slice. The heart does move considerably during the imaging period, but because each acquisition is very short, each slice is relatively immune to motion and inter-slice motion is manifested as geometric distortion across the slices rather than blurring. Geometric distortion has little effect on the clinical diagnostic value of the image. Even though spirals are very efficient, the strict timing limitations make it necessary to undersample k-space twofold. To do so, undersampled variable density spirals [104] are used. Such spirals have an incoherent PSF (see Fig. 4.8e). When used with linear gridding reconstruction [48] undersampling artifacts are incoherent and appear simply as added noise. Coronary images are generally piece-wise smooth and are sparsified well by finite-differences. CS reconstruction can suppress undersampling-induced interference without degrading the image quality.

4.6.6 k-t Sparse: Application to Dynamic Heart Imaging

Dynamic imaging of time-varying objects is challenging because of the spatial and temporal sampling requirements of the Nyquist criterion. Often temporal resolution is traded off against spatial resolution (or *vice versa*). Artifacts appear in the traditional linear reconstruction when the Nyquist criterion is violated.

Now consider a special case: dynamic imaging of time-varying objects undergoing quasi-periodic changes. Special cases include heart imaging, which I focus on here. Heart motion is quasi-periodic: the time series of intensity in a single voxel is sparse in the temporal frequency domain (See Fig. 4.3). At the same time, a single frame of the heart 'movie' is sparse in the wavelet domain. A simple transform can exploit both effects: apply a spatial wavelet transform followed by a temporal Fourier transform (see Fig. 4.9 bottom-left panel).

Can the natural sparsity of dynamic sequences be exploited to reconstruct a time-varying object sampled at significantly sub-Nyquist rates? Consider the Cartesian sampling scheme that acquires for each time slice a single line in k-space, following an orderly progression through the space of lines as time progresses (see Fig. 4.9 top-right panel). For the desired FOV and resolution it is impossible, using this scheme, to meet the spatial-temporal Nyquist rate. In fact, this scheme is particularly inefficient for dynamic imaging with traditional acquisitions and reconstruction methods. Instead, one change is made: make the k-space line ordering *random* instead of orderly [67,81]. The random ordering comes much closer to randomly sampling $k - t$ space (See Fig. 4.9 bottom-right panel) and the sampling operator becomes much less coherent with the sparsifying transform.

4.7 Results

4.7.1 Simulation

Figure 4.11 presents the simulation results. The LR reconstruction, as expected, shows a decrease in resolution with acceleration characterized by loss of small structures and diffused boundaries. The ZF-w/dc reconstructions exhibit a decrease in apparent SNR due to the incoherent interference, which completely obscures small and dim features. The uniform density undersampling interference is significantly larger and more structured than the variable density. In both ZF-w/dc reconstructions the features that are brighter than the interference appear to have well-defined boundaries. In the CS reconstructions, at 8-fold acceleration (approximately 3 times more Fourier samples than sparse coefficients) I get exact recovery from both uniform density and variable density undersampling! At 12-fold acceleration (approximately 2 times more Fourier samples than sparse coefficients) I still get exact recovery from the variable density undersampling, but lose some of the low-contrast features in the uniform density undersampling. At 20-fold acceleration (similar number of Fourier samples as sparse coefficients) I get loss of image features in both reconstructions. The reconstruction errors are severe from the uniform density undersampling. However, in reconstruction from the variable density undersampling, only the weak intensity objects have reconstruction errors; the bright, high contrast features are well reconstructed.

4.7.2 2DFT CS Reconstruction in the Presence of Noise

Figure 4.12 presents the reconstruction results. Figure 4.12a shows the reconstruction of a fully sampled phantom scan. The measured SNR is 6.17. The ZF-w/dc reconstruction result in Fig. 4.12b exhibits significant apparent noise in the image with measured SNR of 3.79. The apparent noise is mostly incoherent aliasing artifacts due to the undersampling as well as noise increase from the density compensation (which is essential to preserve the resolution). Some coherent aliasing artifacts are also visible (arrows). In Fig. 4.12c the artifacts are suppressed by the CS reconstruction, recovering the noisy image with an SNR of 9.84. The SNR is slightly better because the CS reconstruction is inherently a denoising procedure. By increasing the RMS consistency parameter to $\epsilon = 0.1$ (less consistency) the CS reconstruction recovers and denoises the phantom image. Measured SNR increases dramatically to 26.9 without damaging the image quality. The denoising is non-linear edge-preserving TV denoising and is shown in Fig. 4.12d.

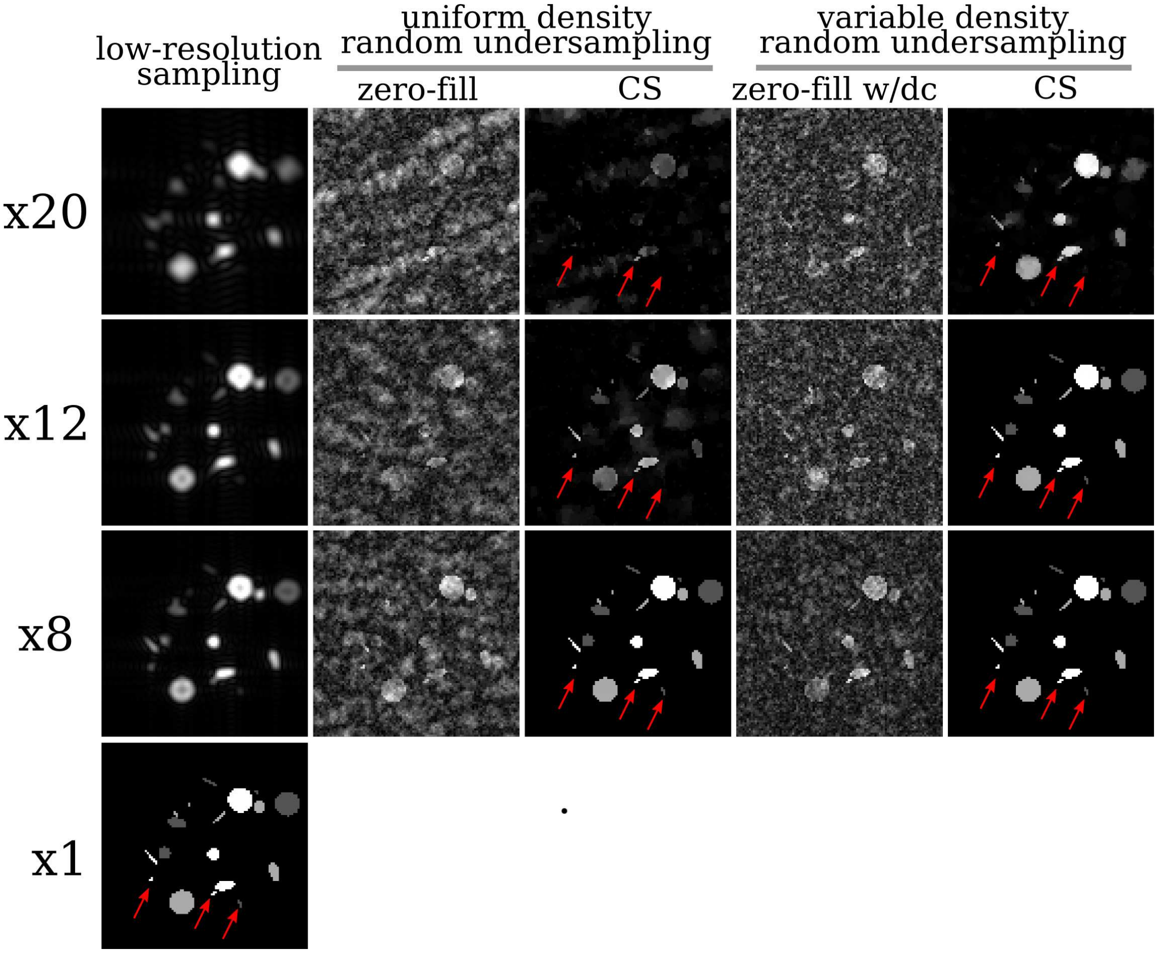
low-resolution sampling
uniform density random undersampling
zero-fill
CS
variable density random undersampling
zero-fill w/dc
CS
x20
x12
x8
x1

Figure 4.11: Simulation: Reconstruction artifacts as a function of acceleration. The LR reconstructions exhibit diffused boundaries and loss of small features. The ZF-w/dc reconstructions exhibit an significant increase of apparent noise due to incoherent aliasing, the apparent noise appears more "white" with variable density sampling. The CS reconstructions exhibit perfect reconstruction at 8 and 12 fold (only var. dens.) accelerations. With increased acceleration there is loss of low-contrast features and not the usual loss of resolution. The reconstructions from variable density random undersampling significantly outperforms the reconstructions from uniform density random undersampling.

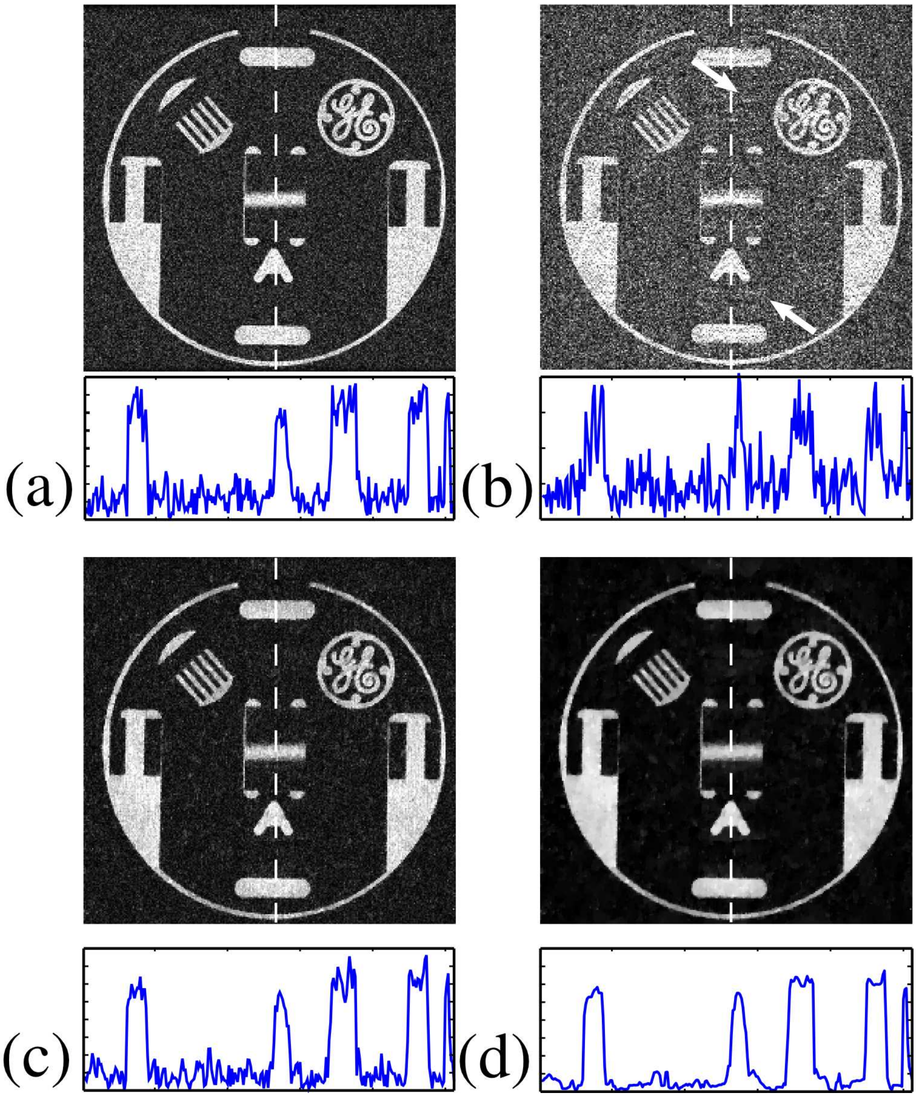

Figure 4.12: 2DFT CS reconstruction from noisy data. CS reconstruction can perform denoising of the image as well as interference removal by relaxing the data consistency (a) Reconstruction from complete noisy data. (b) ZF-w/dc: The image suffers from apparent noise due to incoherent aliasing as well as noise. (c) CS reconstruction with TV penalty from noisy undersampled data. Consistency RMS error set to 10^{-5}. (d) CS reconstruction with TV penalty from noisy undersampled data. Consistency RMS error set to 0.1. Note interference removal in both (c) and (d) and the denoising in (d).

4.7.3 Multi-slice Fast Spin-Echo Brain Imaging

Figure 4.13 shows the experimental results. In Fig. 4.13a coronal and axial slices of the multi-slice CS reconstruction are compared to the full Nyquist sampling, ZF-w/dc and LR reconstructions. CS exhibits significant resolution improvement over LR and significant suppression of the aliasing artifacts over ZF-w/dc compared to the full Nyquist sampling.

Figure 4.13b shows CS reconstructions from several undersampling schemes. The corresponding undersampling schemes are given in Fig. 4.13c. Low-resolution aliasing artifacts are observed in the reconstructions in which the data was undersampled the same way for all slices. The artifacts are more pronounced for uniform undersampling. The reason is that some of the coarse-scale wavelet components in these reconstructions were not recovered correctly because of the large peak interference of coarse-scale components that was documented in the $TPSF$ theoretical analysis (see Fig. 4.6a). These artifacts are significantly reduced when each slice is undersampled differently. This is because the theoretical $TPSF$ peak interference in such sampling scheme is significantly smaller (see Fig. 4.6b), which enables better recovery of these components. The results in Fig. 4.13b show again that a variable density undersampling scheme performs significantly better than uniform undersampling.

4.7.4 Contrast Enhanced 3D Angiography

Figure 4.14 shows a region of interest in the maximum intensity projection (MIP) of the reconstruction results as well as a slice reconstruction from 10-fold acceleration. The LR reconstruction (left column), as expected, shows a decrease in resolution with acceleration characterized by loss of small structures and diffused blood vessel boundaries. The ZF-w/dc reconstruction (middle column), exhibits a decrease in apparent SNR due to the incoherent interference, which obscures small and dim vessels. Interestingly, the boundaries of the very bright vessels remain sharp and are diagnostically more useful than the LR. The CS reconstruction (right column), on the other hand, exhibits good reconstruction of the blood vessels even at very high accelerations. The resolution as well as the contrast are preserved with almost no

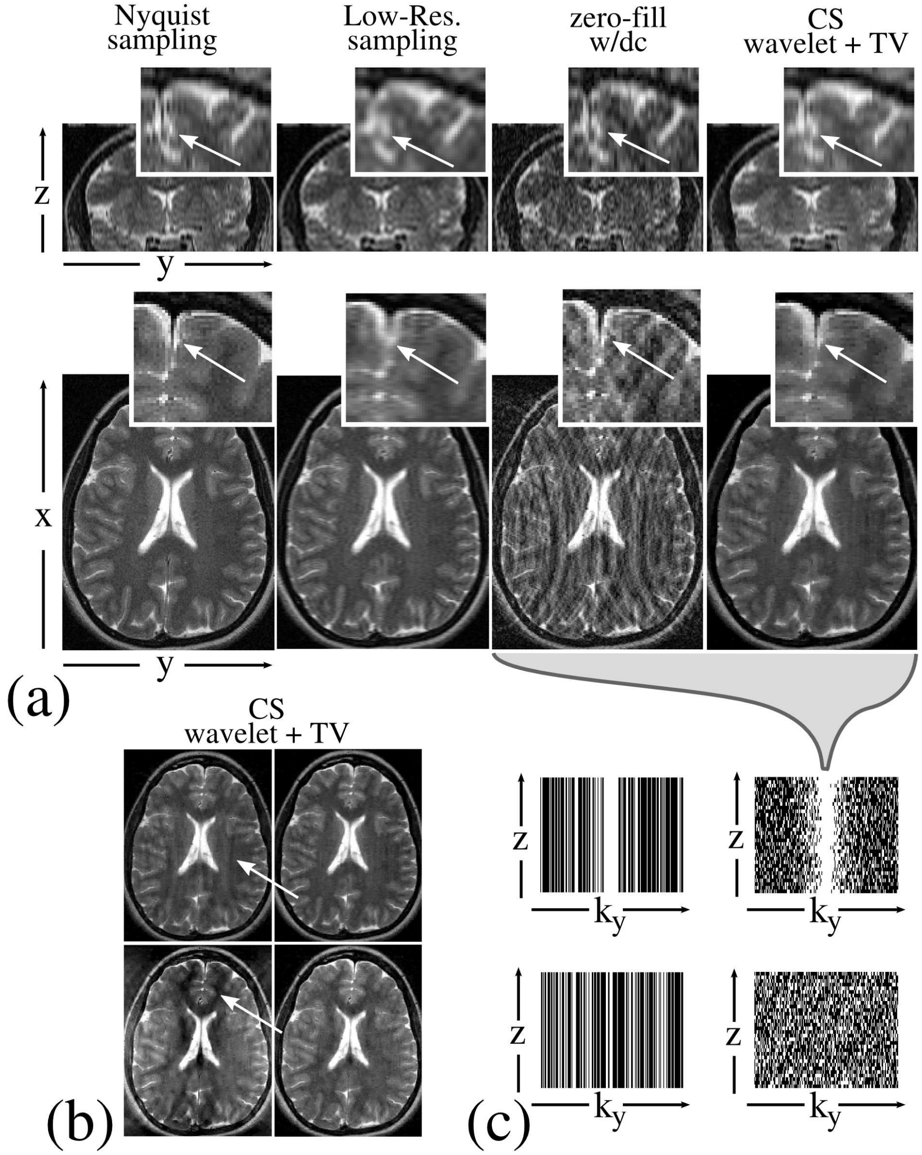
Nyquist sampling
Low-Res. sampling
zero-fill w/dc
CS wavelet + TV
z
y
x
y
(a)
CS wavelet + TV
(b)
z
k_y
z
k_y
z
k_y
z
k_y
(c)

Figure 4.13: Multi-slice 2DFT fast spin echo CS at 2.4 acceleration. (a) The CS-wavelet reconstruction exhibits significant resolution improvement over LR and significant suppression of the aliasing artifacts over ZF-w/dc compared to the full Nyquist sampling. (b) CS wavelet reconstructions from several undersampling schemes. The multi-slice approach outperforms the single-slice approach and variable density undersampling outperforms uniform undersampling. (c) The associated undersampling schemes; variable density (top) and uniform density (bottom), single-slice (left) and multi-slice (right).

loss of information at up to 10-fold acceleration. Even at acceleration of 20-fold the bright blood vessel information is well preserved. These results conform with the thresholding experiment in Fig. 4.3 as well as the simulation results in Fig 4.11.

Figure 4.15 shows the reconstruction result from the first-pass contrast experiment. The imaged patient has an aorto-bifemoral bypass graft. This carries blood from the aorta to the lower extremities, and is seen on the left side of the aorta (right in the image). There is a high-grade stenosis in the native right common illiac artery, which is indicated by the arrows. Again, at 5-fold acceleration the LR acquisition exhibits diffused boundaries and the ZF-w/dc exhibits considerable decrease in apparent SNR. The CS reconstruction exhibits a good reconstruction of the blood vessels. In particular, in Fig. 4.15d flow across the stenosis is visible, but it is not visible in Figs. 4.15b-c.

4.7.5 Variable Density Spirals, Whole Heart Coronary Imaging

Figure 4.16 shows a comparison of the linear direct gridding reconstruction and CS, on the right coronary artery reformatted from a single breath-hold whole-heart acquisition. The linear gridding reconstruction suffers from apparent noise artifacts actually caused by undersampling. Indeed, the CS reconstruction suppresses those artifacts, without impairing the image quality.

4.7.6 k-t Sparse: Application to Dynamic Heart Imaging

Fig. 4.17 shows results from two experiments. The first result used synthetic data: a motion phantom, periodically changing in a cartoon of heart motion. The figure depicts an image sequence reconstructed from a sampling rate 4 times slower than the Nyquist rate, using randomly-ordered acquisition and nonlinear reconstruction. The second result involved dynamic real-time acquisition of heart motion. The given FOV (16cm), resolution (2.5mm) and repetition time (4.4ms) allows a Nyquist rate of 3.6 frames per second (FPS). This leads to temporal blurring and artifacts in the traditionally-reconstructed image. By instead using random ordering and CS

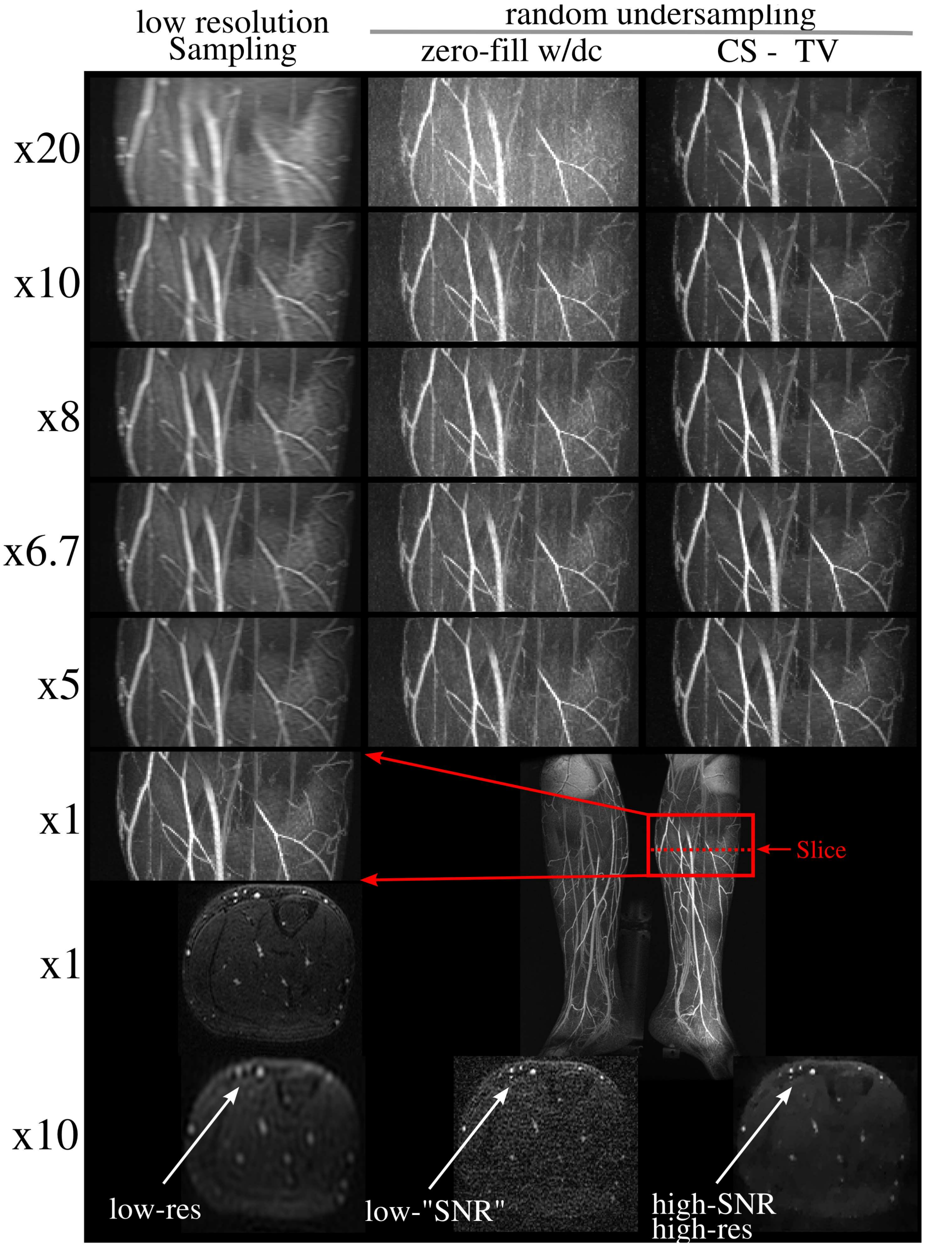
low resolution
Sampling
random undersampling
zero-fill w/dc
CS - TV
x20
x10
x8
x6.7
x5
x1
x1
x10
Slice
low-res
low-"SNR"
high-SNR
high-res

Figure 4.14: Contrast-enhanced 3D angiography reconstruction results as a function of acceleration. Left Column: Acceleration by LR. Note the diffused boundaries with acceleration. Middle Column: ZF-w/dc reconstruction. Note the increase of apparent noise with acceleration. Right Column: CS reconstruction with TV penalty from randomly undersampled k-space.

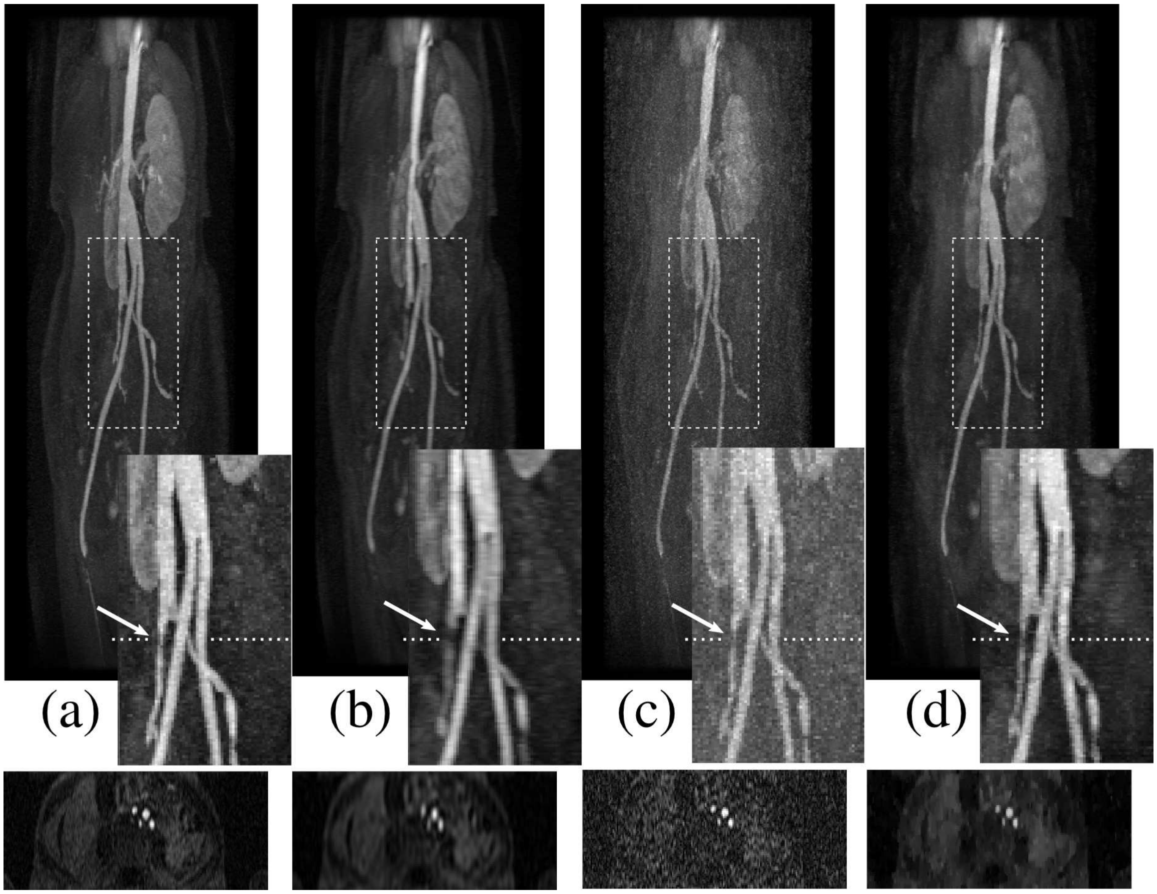

Figure 4.15: Reconstruction from 5-fold accelerated acquisition of first-pass contrast enhanced abdominal angiography. (a) Reconstruction from a complete data set. (b) LR (c) ZF-w/dc (d) CS reconstruction from random undersampling. The patient has a aorto-bifemoral bypass graft. This carries blood from the aorta to the lower extremities. There is a high-grade stenosis in the native right common illiac artery, which is indicated by the arrows. In figure parts (a) and (d) flow across the stenosis is visible, but it is not on (b) and (c).

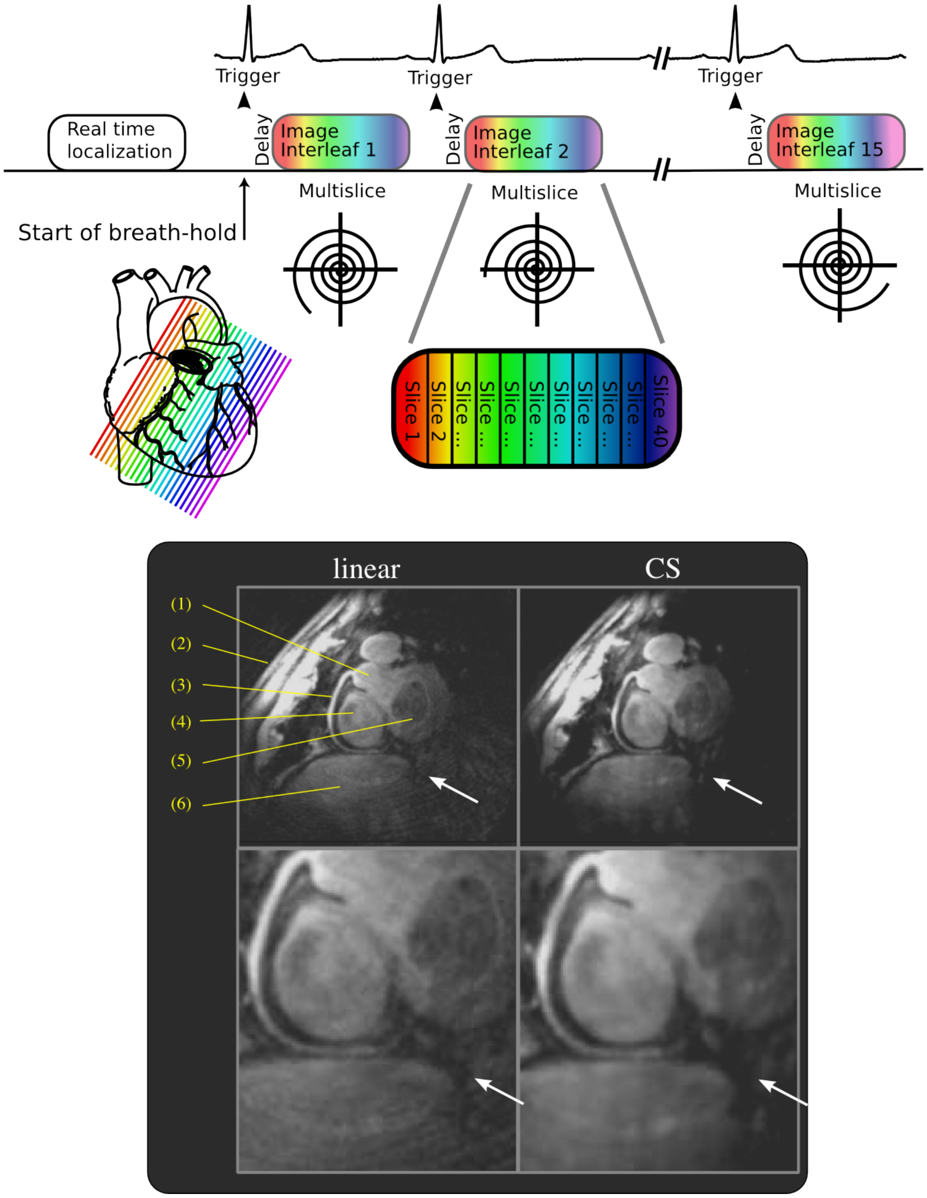
Trigger
Trigger
Trigger
Real time localization
Delay
Image Interleaf 1
Delay
Image Interleaf 2
Delay
Image Interleaf 15
Multislice
Multislice
Multislice
Start of breath-hold
Slice 1
Slice 2
Slice ...
Slice ...
Slice ...
Slice ...
Slice ...
Slice ...
Slice ...
Slice ...
Slice 40
linear
CS
(1)
(2)
(3)
(4)
(5)
(6)

Figure 4.16: Single breath-hold whole-heart coronary artery imaging. Left panel: the imaging sequence timing diagram. Right panel: A slice through the volume of the heart showing the right coronary artery (3). The incoherent artifacts of undersampled variable-density spirals (white arrow) appear as noiselike interference in the linear gridding reconstruction (left). These artifacts are suppressed in the CS reconstruction (right) without compromising image quality. The slice shows: (1) Aorta (2) Chest wall (3) Right coronary artery (4) Right ventricle (5) Left ventricle (6) Liver.

reconstruction it was possible to recover the dynamic sequence at the much higher rate of 25FPS with significantly reduced image artifacts.

4.8 Discussion

4.8.1 Computational Complexity

Development of fast algorithms for solving Eq. 4.3 accurately or approximately is an increasingly popular research topic. Many of these methods have been mentioned in the theory section. Overall, the reconstruction is iterative and more computationally intensive than linear reconstruction methods. However, some of the methods proposed show great potential to significantly reduce the overall complexity.

The examples in this chapter were reconstructed using a non-linear conjugate gradient method with backtracking line-search. In a Matlab (The MathWorks, Inc., Natick, MA, USA) implementation, it takes about 150 CG iterations (approximately 30 seconds) to reconstruct a 480×92 angiogram using a TV-penalty at 5-fold acceleration. I expect a significant reduction in the reconstruction time by code optimization.

4.8.2 Reconstruction Artifacts

The ℓ^1 reconstruction tends to slightly shrink the magnitude of the reconstructed sparse coefficients. The resulting reconstructed coefficients are often slightly smaller than in the original signal. This coefficient shrinkage decreases when the reconstruction consistency parameter ϵ in Eq. 4.3 is small.

In some wavelet-based CS reconstructions, small high-frequency oscillatory artifacts may appear in the reconstruction. This is due to false detection of fine-scale wavelet components. To mitigate these artifacts it is recommended to add a small TV penalty on top of the wavelet penalty. This can be considered as requiring the image to be sparse in both wavelet and finite-differences transforms.

In CS, the contrast in the image plays a major part in the ability to vastly undersample and reconstruct images. High contrast often results in large distinct sparse coefficients. These can be recovered even at very high accelerations. For example, a

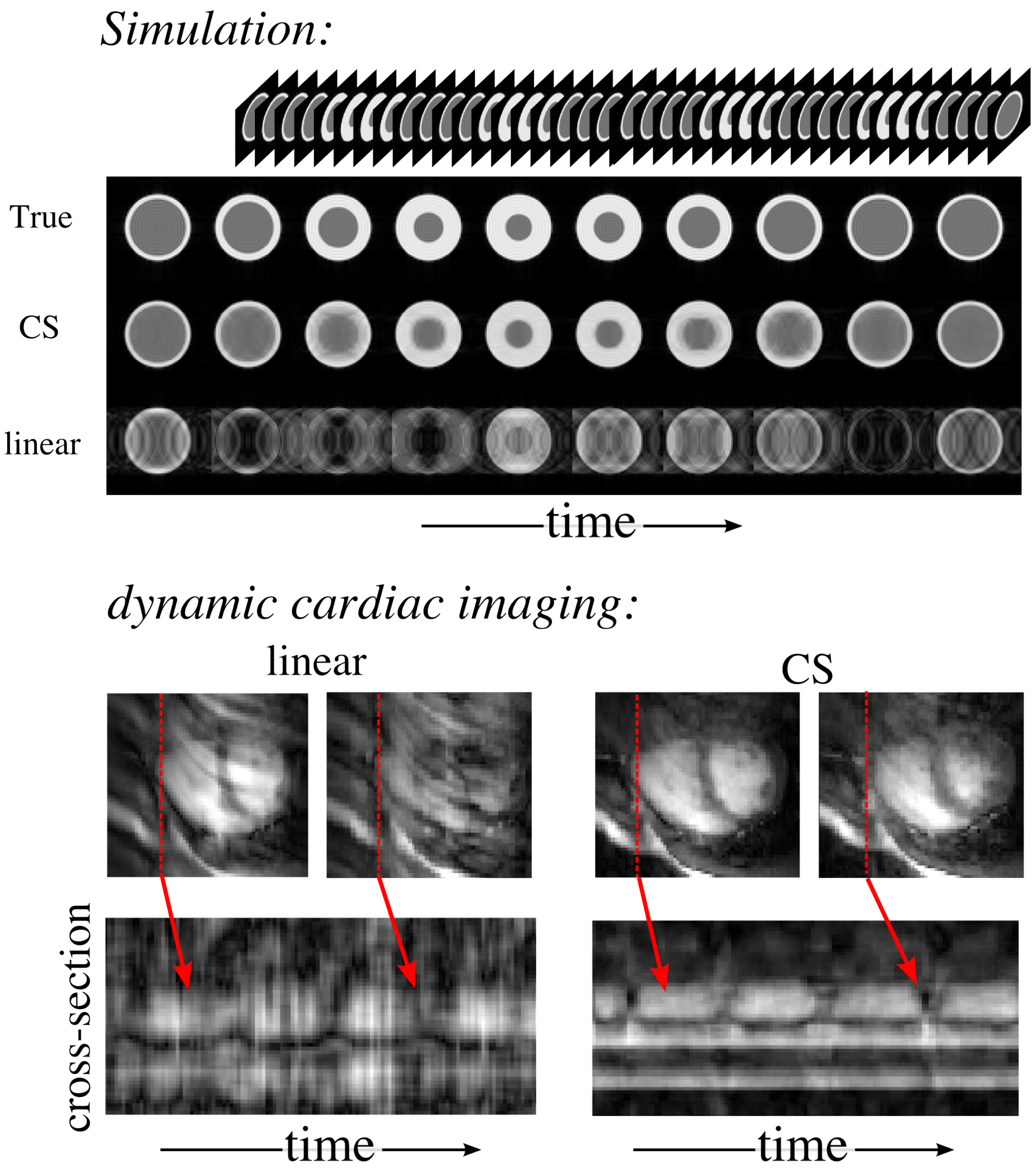
Simulation:
True
CS
linear
time
dynamic cardiac imaging:
linear
CS
cross-section
time
time

Figure 4.17: Dynamic imaging of quasi-periodic change. Top: Phantom experiment showing a reconstruction from 4-fold undersampling. Bottom: Dynamic acquisition of the heart motion showing a reconstruction from 7-fold undersampling.

single bright pixel will most likely appear in the reconstruction even with vast undersampling (See Figs. 4.11 and 4.14 for an example). However, features with lower contrast at the same accelerations will be so deeply submerged by the interference that they would not be recoverable. As such, with increased acceleration the most distinct artifacts in CS are not the usual loss of resolution or increase in aliasing interference, but loss of low-contrast features in the image. Therefore, CS is particularly attractive in applications that exhibit high resolution high contrast image features, and rapid imaging is required.

4.8.3 Relation to Other Acceleration Methods

Vastly undersampled 3D radial trajectories – VIPR [1] have demonstrated high acceleration for angiography. The VIPR trajectory is a 3D incoherent sampling scheme in which the interference spreads in all 3 dimensions. As such, reconstruction from VIPR acquisitions can be further improved by using the CS approach.

Wajer's PhD thesis [108] suggested undersampling k-space and employing a Bayesian reconstruction to randomized trajectories. This approach, although different, is related to finite difference sparsity.

Non uniform sampling with maximum entropy reconstruction has been used successfully to accelerate multi-dimensional NMR acquisitions [86]. Maximum entropy reconstruction is also related to sparsity of finite differences.

CS reconstruction exploits sparsity and compressibility of MR images. It can be combined with other acceleration methods that exploit different redundancies. For example, constraining the image to be real in Eq. 4.7 effectively combines phase constrained partial k-space with the CS reconstruction. In a similar way, CS can be combined with SENSE reconstruction by including the coil sensitivity information in Eq. 4.3. In general, any other prior on the image that can be expressed as a convex constraint can be incorporated in the reconstruction.

4.9 Conclusions

I have presented the theory of CS and the details of its implementation for rapid MR imaging. I demonstrated experimental verification of several implementations for 2D, 2D multi-slice, 3D and dynamic imaging. I showed that the sparsity of MR images can be exploited to significantly reduce scan time, or alternatively, improve the resolution of MR imagery. I demonstrated high acceleration in *in-vivo* experiments. CS can play a major part in many applications that are limited by the scan time, when the images exhibit transform sparsity.

Chapter 5

SPIR-iT

5.1 Introduction

Multiple receiver coils have been used since the beginning of MRI, mostly for the benefit of increased signal to noise ratio (SNR). In the late 80's, Kelton, Magin and Wright proposed in an abstract [51] to use multiple receivers for scan acceleration. However, it was not until the late 90's when Sodickson *et al.* presented their method SMASH [97] and later Pruessmann *et al.* presented SENSE [84], that accelerated scans using multiple receivers became a practical and viable option.

Multiple receiver scans can be accelerated because the data obtained for each coil are acquired in parallel and each coil image is weighted differently by the spatial sensitivity of its coil. This sensitivity information in conjunction with gradient encoding reduces the required number of data samples that is needed for reconstruction. This concept of reduced data acquisition by combined sensitivity and gradient encoding is called parallel imaging.

Over the years, a variety of methods for parallel imaging reconstruction have been developed. These methods differ by the way the sensitivity information is used. Methods like SMASH [97], SENSE [84], SPACE-RIP [59], PARS [112] and kSPA [63] explicitly require the coil sensitivities to be known. In practice, it is very difficult to measure the coil sensitivities with high accuracy. Errors in the sensitivity are often amplified and even small errors can result in visible artifacts in the image.

On the other hand, autocalibrating methods like AUTO-SMASH [44,49], PILS [33], GRAPPA [35] and APPEAR [2] implicitly use the sensitivity information for reconstruction and avoid some of the difficulties associated with explicit estimated of the sensitivities. Another major difference is in the reconstruction target. SMASH, SENSE, SPACE-RIP, kSPA and AUTO-SMASH attempt to directly reconstruct a single combined image. Coil-by-coil methods, PILS, PARS and GRAPPA directly reconstruct the individual coil images leaving the choice of combination to the user. In practice, coil-by-coil methods tend to be more robust to inaccuracies in the sensitivity estimation and often exhibit fewer visible artifacts [2,34]. Table 5.1 summarizes some of the existing methods and their properties.

SENSE is an explicit sensitivity, single image reconstruction method. Among all methods, the SENSE approach is the most general. It provides a framework for arbitrary k-space sampling and to easily incorporate additional image priors. When the sensitivities are known, SENSE is the optimal solution [2,34]. To the best of the author's knowledge, none of the coil-by-coil autocalibrating methods are as flexible and optimal as SENSE. Some proposals [42,45,46,93] adapt GRAPPA to reconstruct some non-Cartesian trajectories , but these are approximations and lose some of the ability to remove all the aliasing artifacts. The APPEAR [2] method can provide a more accurate reconstruction, but is not flexible in incorporating additional priors and regularizations.

Following SENSE, while using a coil-by-coil autocalibrating approach similar to GRAPPA, I propose here a general and optimal autocalibrating coil-by-coil reconstruction method called SPIR-iT. It is based on self-consistency with the calibration and acquisition data. The SPIR-iT method is flexible and can reconstruct data from arbitrary k-space sampling patterns and easily incorporates additional image priors. SPIR-iT stands for iTerative Self-consistent Parallel Imaging Reconstruction.

In this chapter I first review the foundations upon which SPIR-iT is built, the SENSE and GRAPPA methods. I then define the SPIR-iT consistency constraints and show that the reconstruction is a solution to an optimization problem. I then extend the method to arbitrary sampling patterns with k-space-based and image-based approaches. Finally I show that the method can easily introduce additional

method	**sensitivity**	**image**	**comments**
SMASH	explicit	single	restricted sensitivities
SENSE	explicit	single	optimal, iterative non-Cartesian, inverse-problem
AUTO-SMASH	implicit	single	Cartesian, direct
SPACE-RIP	explicit	single	direct 1D non-Cartesian
PILS	implicit	coil-by-coil	localized coil sensitivities
GRAPPA	implicit	coil-by-coil	Cartesian w/ approximate direct non-Cartesian variants
PARS	explicit	coil-by-coil	direct non-Cartesian
APPEAR	implicit	coil-by-coil	direct non-Cartesian
kSPA	explicit	single	direct non-Cartesian, approximate
SPIRiT	implicit	coil-by-coil	optimal, iterative non-Cartesian, inverse problem

Table 5.1: Properties of some parallel imaging reconstruction methods.

priors such as off-resonance correction and regularization.

5.2 SENSE

SENSE poses parallel imaging reconstruction as an inverse problem in image space. Provided that the coil sensitivities are known or can be measured with sufficient accuracy, the problem can be formulated as a set of linear equations. Let m be the underlying magnetization image. Let s_i be the sensitivity of the i^{th} coil and let D be a partial Fourier operator corresponding to the k-space undersampling scheme. The received signal for the i^{th} coil can be written as,

$$y_i = Ds_im. \tag{5.1}$$

It is often more convenient to write the entire set of equations in matrix form as

$$y = Em, \tag{5.2}$$

where the matrix E is an encoding matrix that incorporates the coil sensitivities and the partial Fourier operators. The system in Eq. 5.2 is often solved by least-squares either directly [84] or iteratively [85].

The advantage of the SENSE approach is that it is very general. It can be used with arbitrary sampling trajectories, and other priors on the image can be easily incorporated. It is the optimal solution when the coil sensitivities are exact. However, it is often very difficult to accurately and robustly measure the sensitivities, and even small errors can lead to visible artifacts in the image.

5.3 Cartesian GRAPPA

The GRAPPA reconstruction poses the parallel imaging reconstruction as a synthesis problem in k-space. It is a self-calibrating coil-by-coil reconstruction. Unlike SENSE, which attempts to reconstruct a single combined image, GRAPPA attempts to reconstruct the individual coil images directly.

In the traditional GRAPPA [35] algorithm a non-acquired k-space value in the i^{th} coil, at position r, $x_i(r)$, is synthesized by a linear combination of acquired neighboring k-space data from all coils. Let ξ_{ir} be a vector of all points on a Cartesian grid in the neighborhood of $x_i(r)$ in the i^{th}coil, and let ξ_{ir}^p be a subset of ξ_{ir}, choosing only those points that are acquired. The recovery of $x_i(r)$ is given by:

$$x_i(r) = \sum_j \langle g_{ji}^p, \xi_{jr}^p \rangle, \tag{5.3}$$

where g_{ji}^p is the vector set of weights obtained by calibration. The full k-space grid is reconstructed by solving Eq. 5.3 for each missing point in all coils and all positions.

The linear combination weights, or calibration kernel, used in Eq. 5.3 are obtained by calibration from a fully acquired k-space region. The calibration finds the set of weights that is the most consistent with the calibration data in the least-squares sense. In other words, the calibration stage looks for a set of weights such that if one tries to synthesize each of the calibration points from their neighborhood the result should be as close as possible to the true calibration points. More formally, the calibration is described by the following equation:

$$arg\min_{\boldsymbol{g}_{ji}^{p}} \sum_{r\in\text{Calib}} \left|\left| \sum_{j} \langle g_{ji}^{p}, \xi_{jr}^{p} \rangle - x_i(r) \right|\right|^2 . \tag{5.4}$$

The assumption is that if the calibration consistency holds within the calibration area, it also holds in other parts of k-space. Therefore Eq. 5.3 can be used for the reconstruction. In the original GRAPPA paper, the undersampling pattern was such that the k-space sampling in the neighborhood of the missing points is the same for all points. Therefore a single set of weights is enough for reconstruction and the calibration can be performed only once. In 2D acceleration however, the sampling pattern around each missing point can be quite different. Different sets of weights must be obtained for each sampling pattern. As an illustration, consider the 2D GRAPPA reconstruction in Fig. 5.1. The figure portrays two equations to solve two missing data points. Each of the equations uses a different set of calibration weights. The neighborhood size is a square of three k-space pixels.

5.4 SPIR-iT: Self Consistency Formulation

Inspired by the generality of SENSE and the autocalibrating coil-by-coil GRAPPA method, I take a slightly different approach to GRAPPA. My aim is to describe the reconstruction as an inverse problem governed by data consistency constraints. The key in the approach is to separate the consistency constraints into 1) consistency with the calibration, and 2) consistency with the data acquisition. I formulate these constraints as sets of linear equations. The desired reconstruction is the solution that satisfies the sets of equations best according to a suitable error measure criteria.

Even though the acquired k-space data may, or may not be Cartesian, ultimately, the desired reconstruction is a complete Cartesian k-space grid for each of the coils. Therefore, I define the entire Cartesian k-space grid for all the coils as the unknown variables in the equations. This step makes the formulation very general especially when considering noisy data, regularization and non-Cartesian sampling.

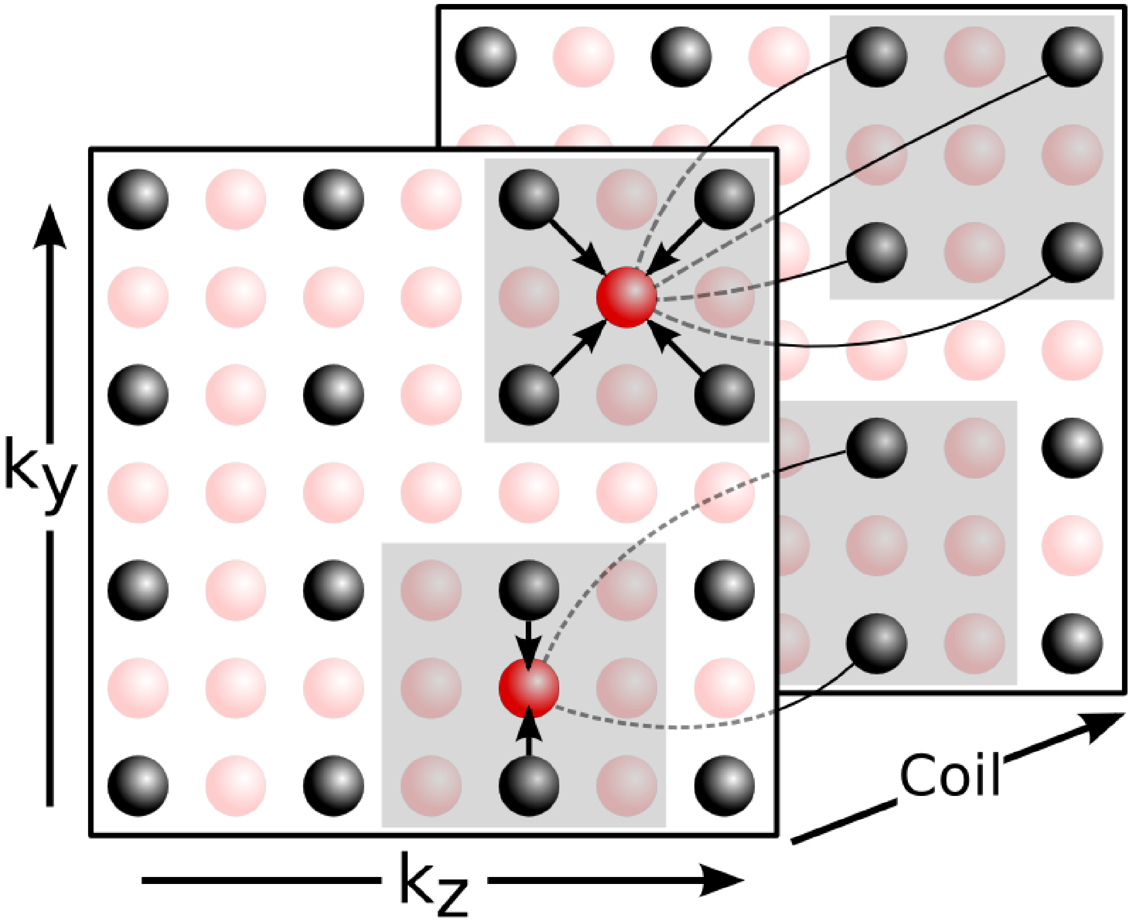

Figure 5.1: Traditional 2D GRAPPA: Missing k-space data are synthesized from neighboring acquired data. The synthesizing kernel depends on the specific sampling pattern in the neighborhood of the missing point. The reconstruction of a point is independent of the reconstruction of other missing points.

5.4.1 Consistency with the Calibration

Traditional GRAPPA enforces calibration consistency only between synthesized points and the acquired points in their associated neighborhoods. The proposed approach expands the notion of consistency by enforcing consistency between every point on the grid, $x_i(r)$, and its entire neighborhood across all coils, ξ_{jr}. It is important to emphasize that the "entire neighborhood" includes all the k-space points near $x_i(r)$ in all coils whether they were acquired or not. The consistency equation for *any* k-space data point is given by,

$$x_i(r) = \sum_j \langle g_{ji}, \xi_{jr} \rangle. \tag{5.5}$$

The vector g_{ji} the vector set of weights obtained by calibration, similar to Eq. 5.3. The difference from the traditional GRAPPA weights, g^p_{ji}, is that g_{ji} is a full kernel independent of the actual k-space sampling pattern. Equation 5.3 defines a large set of *decoupled* linear equations that can be solved separately. On the other hand, Eq. 5.5 defines a large set of *coupled* linear equations. As an illustration, consider Fig. 5.2 which has a similar sampling setup as the 2D GRAPPA problem in Fig. 5.1. In the figure, three equations are portrayed. It shows that the synthesis of a missing central point depends on both acquired and missing points in its neighborhood and that the equations are coupled.

In the same way as before, it is convenient to write the entire coupled system of equations in matrix form. Let x be the entire k-space grid for all coils, and let G be a matrix containing the g_{ji}'s in the appropriate locations. The system of equations can simply be written as,

$$x = Gx. \tag{5.6}$$

The matrix G is in fact a series of convolution operators that convolve the entire k-space with the appropriate calibration kernels,

$$x_i = \sum_j g_{ij} * x_j. \tag{5.7}$$

Equations 5.6 and 5.7 can be explained intuitively. Applying the operation G on

x is the same as attempting to synthesize every point from its neighborhood. If x is indeed the right solution then synthesizing every point from its neighborhood should yield the exact same k-space data!

5.4.2 Consistency with the Data Acquisition

Of course, any plausible reconstruction has to be consistent with the real data that were acquired by the scanner. This constraint can also be expressed as a set of linear equations in matrix form. Let y be the vector of acquired data from all coils (concatenated). Let the operator D be a linear operator that relates the reconstructed Cartesian k-space, x, to the acquired data. The data acquisition consistency is given by

$$y = Dx. \tag{5.8}$$

This formulation is very general in the sense that x is always Cartesian k-space data, whereas y can data acquired with arbitrary k-space sampling patterns. In Cartesian acquisitions, the operator D selects only acquired k-space locations. The selection can be arbitrary, such as uniform, variable density or pseudo random patterns. In non-Cartesian sampling, the operator D represents an interpolation matrix. It interpolates data from a Cartesian k-space grid onto non-Cartesian k-space locations in which the data were acquired.

5.4.3 Constrained Optimization Formulation

Equations 5.6 and 5.8 describe the calibration and data acquisition consistency constraints as sets of linear equations that the reconstruction must satisfy. However, due to noise and calibration errors these equations can only be solved approximately. Therefore I propose as reconstruction the solution to an optimization problem given by,

$$\begin{array}{ll} \text{minimize} & ||(G-I)x||^2 \\ \text{s.t.} & ||Dx-y||^2 \leq \epsilon. \end{array} \tag{5.9}$$

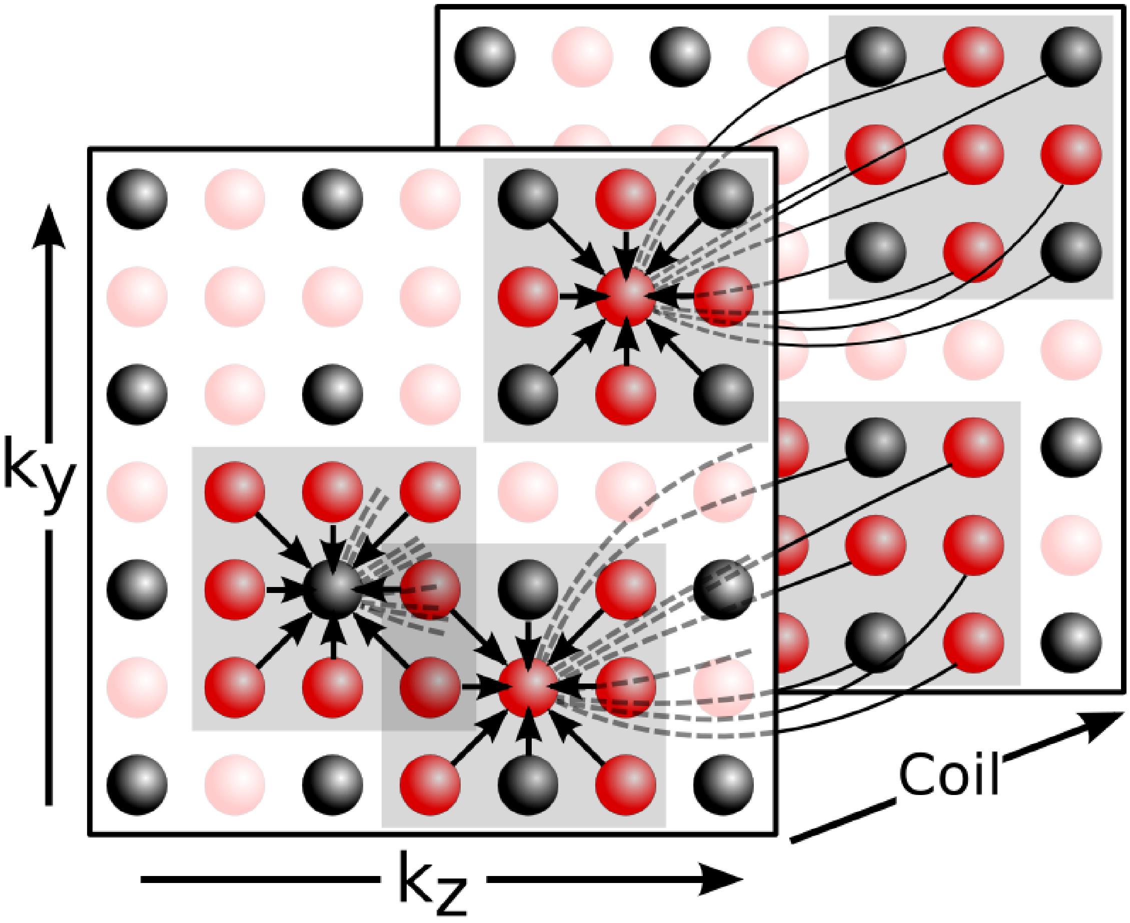

Figure 5.2: Consistency constrained Cartesian reconstruction: Three consistency equations are illustrated. The reconstruction of each point on the grid is dependent on its entire neighborhood. The reconstruction of missing points depends on the reconstruction of other missing points. The consistency equation is independent of the sampling pattern.

The parameter ϵ is introduced as a way to control the consistency. It trades off data acquisition consistency with calibration consistency. The beauty in this formulation is that the calibration consistency is always applied to a Cartesian k-space, even though the acquired data may be non-Cartesian. The treatment of non-Cartesian sampling appears only in the data acquisition constraint.

A useful reformulation of Eq. 5.9 is the unconstrained Lagrangian form,

$$arg\min_{\boldsymbol{x}} \quad ||Dx - y||^2 \ + \ \lambda(\epsilon)||(G - I)x||^2. \tag{5.10}$$

As is well known, the parameter λ can be chosen appropriately such that the solution of Eq. 5.10 is exactly as Eq. 5.9. It is worth pointing out that in most cases both the G and D operation and their adjoints can be calculated very quickly. In general the optimization in Eq. 5.9 and variations of it can often be solved very efficiently by iterative methods such as the conjugate gradient algorithm and projection over convex sets (POCS) iterations. In the next sections I will provide the reconstruction details for more specific examples, including the calculation of the G and D operators and their adjoints.

5.5 Arbitrary Cartesian Sampling

In some cases it may be useful to enforce a data acquisition equality constraint (*i.e.*, $\epsilon = 0$ in Eq. 5.9). One way to implement this is to iteratively reduce λ, and repeatedly solving Eq. 5.10. A more simple and efficient approach is to incorporate the equality constraint in the objective, such that the problem is formulated as simple least-squares. Let y be the vector of acquired points, let x_{nacq} be a vector representing only the missing points, and let G_{acq} and G_{nacq} be the part of the operator G that operates on the acquired and non-acquired points respectively. Rewrite Eq. 5.9 as

$$\arg\min_{\boldsymbol{x}_{\textbf{nacq}}} ||(G_{\text{nacq}} - I)x_{\text{nacq}} + (G_{\text{acq}} - I)y||^2 , \tag{5.11}$$

```
POCS SPIR-iT CARTESIAN RECONSTRUCTION
INPUTS:
        $y$ - $k$ - space measurements from all coils
        $n_{\text{acq}}$ - indeces of acquired $k$-space
        $G$ - operator matrix obtained by calibration
        errToll - stopping tolerance
OUTPUTS:
        $x_k$ - reconstructed $k$-space for all coils
ALGORITHM:
        $x_k = 0; x_{k-1} = 0;$
        do {
            $x_k = Gx_{k-1}$ % Calibration consistency projection
            $x_k[n_{\text{acq}}] = y$ % Data acquisition consistency projection
            $e = ||x_k - x_{k-1}||$ % Error stopping criteria
            $x_{k-1} = x_k$
        }while $e > errToll$
```

Table 5.2: A POCS algorithm for SPIR-iT from arbitrary sampling on a Cartesian grid.

which has the usual format of a least-squares problem (*i.e.*, $\text{arg}min_{\boldsymbol{x}} ||Ax - b||$) and can be solved directly using sparse matrices or more preferably iteratively using standard conjugate gradient algorithms. A simpler alternative is to implement the reconstruction as a POCS algorithm. The POCS approach is described in table 5.2 and involves repeated application of the G operator.

5.6 Non-Cartesian Sampling

In theory, Eq. 5.9 is the solution for the non-Cartesian case. However, the practical success of the reconstruction depends on how well the operators G and D approximate the true data, and how fast they can be computed in practice. The main difficulty facing the reconstruction is an accurate and efficient interpolation scheme in k-space. Next, I present two approaches. The first operates entirely in k-space. The other also applies calculations in the image domain. The former is fast but approximate, whereas the latter is of higher accuracy but with increased computation.

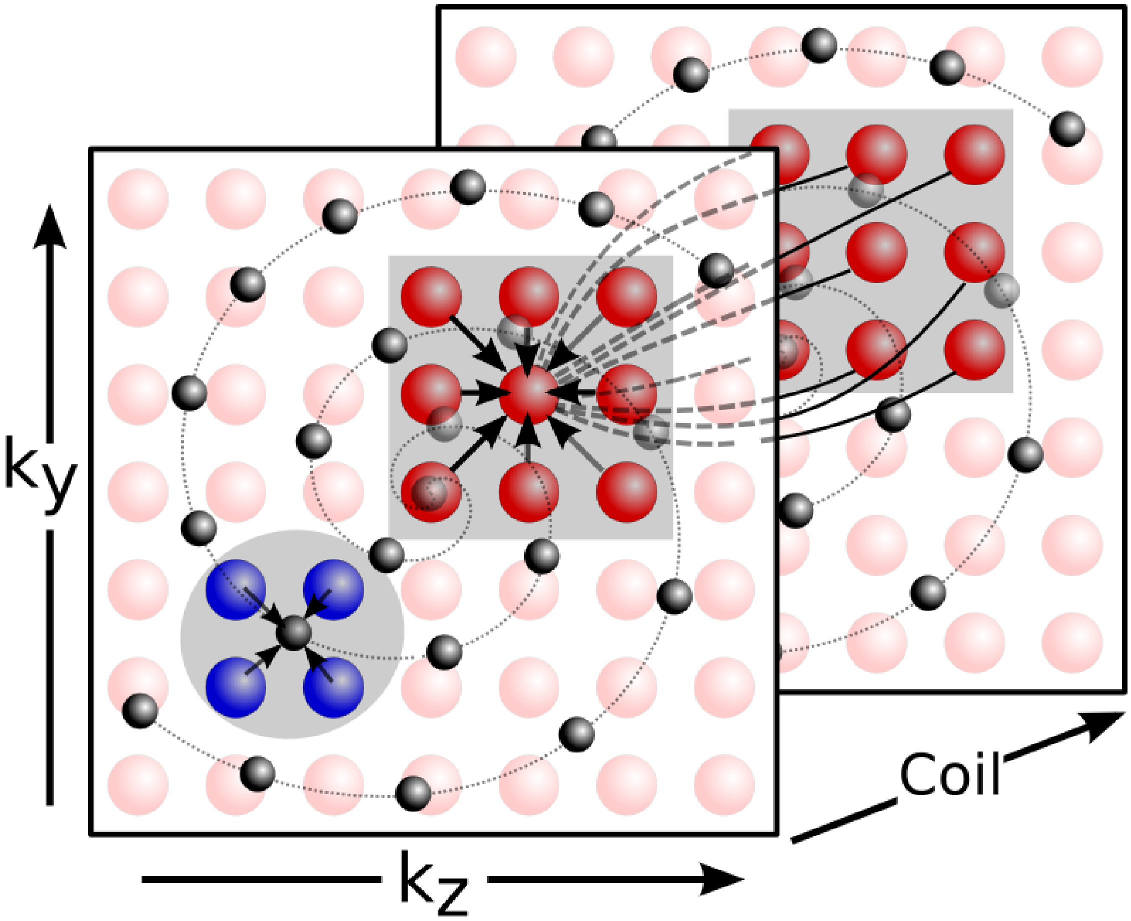

Figure 5.3: Non-Cartesian consistency reconstruction. A calibration consistency equation describing the relation between neighboring points on a Cartesian grid is shown in red. An acquisition data consistency relation between the Cartesian missing points and the non-Cartesian acquired points is shown in blue. These define a large set of linear equations that is sufficient for reconstruction.

5.6.1 Calibration

The SPIR-iT approach is autocalibrating. I assume that there is always a subset of k-space that is sufficiently sampled for the purpose of calibration. For example, radial trajectories and dual density spiral [42] trajectories have a region around the k-space origin that is sufficiently sampled. From such data, Cartesian calibration data are obtained by interpolation. Alternatively, such a Cartesian calibration region can be obtained by a separate scan. The full kernel weights are calculated from calibration data similar to Eq. 5.4.

5.6.2 k-Space Domain Reconstruction

The most natural way to approximate the data consistency constraint $y = Dx$ is to use convolution interpolation to interpolate from Cartesian k-space onto the non-Cartesian grid and vice versa. This is very similar to the gridding algorithm. However, several details need to be taken into consideration.

Consider a convolution interpolation from a Cartesian grid to non-Cartesian of the i^{th} coil k-space data. Let c be an interpolation convolution kernel, k be the non-Cartesian k-space coordinates, and $\delta(x)$ the usual impulse function. One can write the operation D applied to the i^{th} coil more specifically as,

$$y_i(n) = \int_r \delta(k(n) - r) \left\{ c * x_i \right\} (r) dr \tag{5.12}$$

Ideally, the interpolation kernel, c, should be a sinc function. However, this has a prohibitively large kernel size. Traditionally, in the gridding algorithm [48], a very small kernel is used to reduce the computation. The errors associated with a small kernel are mitigated by oversampling the grid, and the associated image weighting is mitigated by a deapodization of the resulting image to correct the intensity weighting. The gridding algorithm in effect trades off complexity with tolerable accuracy errors. In the proposed reconstruction, the data acquisition consistency in Eq. 5.8 is enforced. Therefore, the consistent solution when using a non-sinc kernel can not be exactly x_i, but a function of it, $\tilde{x}_i$. It is given by,

$$\tilde{x}_i = \tilde{c} * x_i, \tag{5.13}$$

for which $c * \tilde{x}_i = c * \tilde{c} * x_i \approx x_i$.

In the image domain this function is manifested as a multiplication by a deapodizing function

$$\begin{aligned} \tilde{m}_i(r) &= \tilde{C}(r) m_i(r) \\ &= \frac{1}{C(r)} m_i(r), \end{aligned} \tag{5.14}$$

where m_i, $\tilde{m}_i$, $\tilde{C}$, and C denote the inverse Fourier transform of of x_i, $\tilde{x}_i$, $\tilde{c}$, and c respectively. In the proposed reconstruction, after $\tilde{x}$ and $\tilde{m}$ are recovered by solving the optimization in Eq. 5.10, the image intensity weighting can be corrected by multiplication with the apodizing function,

$$\begin{aligned} \hat{m}_i(r) &= C(r) \tilde{m}_i(r) \\ &= \frac{C(r)}{C(r)} m_i(r) \\ &= m_i(r). \end{aligned} \tag{5.15}$$

Choosing an Interpolation Kernel

Some published papers [3,28] provide ways of choosing appropriate interpolation kernels and grid oversampling factors such that the computational complexity is small and the reconstruction error is minimized. Most of these interpolation kernels (including the popular Kaiser-Bessel) produce significant weighting across the image field of view. This weighting can sometimes be as high as a factor of 10. In an ideal world, where there is no noise, and the calibration weights are perfect, the image weighting can be completely corrected with no apparent artifacts [3]. However, in real life, calibration errors and noise are amplified when the image weighting is compensated.

The error in a reconstructed voxel, $\hat{m}_i(r)$, due to calibration errors is a complicated weighted sum of the rest of the voxels in all coils. The weighted sum depends on the specific calibration weight's errors and on the point spread function (PSF) of the sampling pattern. Let $w_i(r)$ be a set of coefficients representing the error contributed by voxels, then the error in $\hat{m}_i(r)$ can be written as,

$$\begin{aligned} \hat{m}_i(r) - m_i(r) &= C(r)\left(\frac{m_i(r)}{C(r)} + \sum_{i,\rho} w_i(\rho - r)\frac{m_i(\rho)}{C(\rho)}\right) - m_i(r) \qquad (5.16) \\ &= \sum_{i,\rho} \frac{C(r)}{C(\rho)} w_i(\rho - r) m_i(\rho). \end{aligned}$$

Of course the error depends on the image, the calibration and the sampling. But, it is obvious that a large ratio of $C(r)/C(\rho)$ can cause significant error amplification.

In order to mitigate this error amplification, one needs to design an interpolation kernel which causes an acceptable variation in the image. For example, a windowed-sinc kernel would introduce much less variation than the Kaiser-Bessel kernel would. An alternative is to modify the min-max SOCP interpolator by Beatty [3] to include the image intensity variation constraint. In fact, the interpolation kernel design is much like a filter design problem. The bandpass ripple of the interpolation filter corresponds to the intensity variation in the FOV. The stopband ripple corresponds to the interpolation errors associated with a finite interpolation kernel and the transition-width corresponds to the required grid oversampling. Figure 5.4 illustrates these concepts. It shows the requirements for the interpolation kernel design. It also shows the reconstruction error amplification that is associated with a large image intensity variation and how it can be corrected.

Revisiting the Calibration

After going through the reconstruction details, I now revisit the details of the calibration. The grid oversampling in k-space requires that the calibration kernel support be increased by the same amount. In addition, the grid oversampling corresponds to a larger FOV support in the reconstructed image. Therefore, the sampling trajectory of the calibration region must support this larger FOV, or else the fidelity and integrity of the calibration may be compromised. In addition, the interpolation kernel that

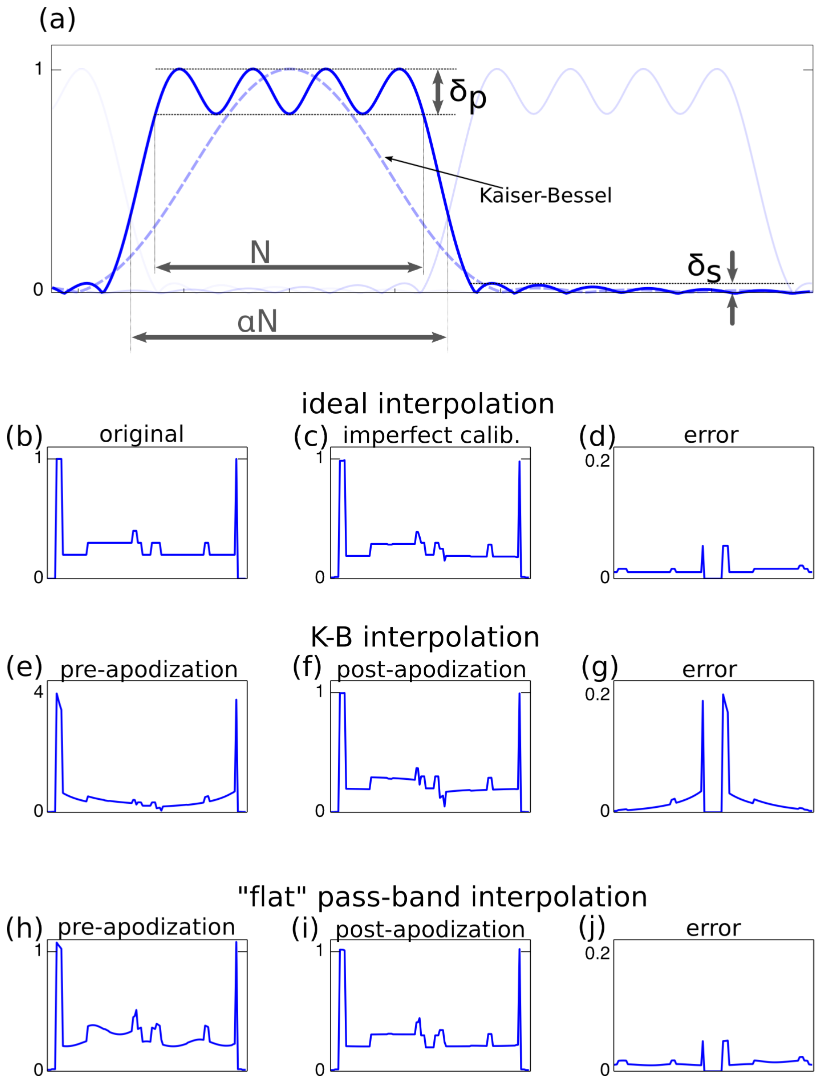

Figure 5.4: Interpolation kernel in k-space. (a) The specification for the kernel compared to a Kaiser-Bessel kernel. The passband ripple defines the allowed intensity variation in the image. The stopband ripple is associated with the error that leaks into the image due to finite kernel size. The transition band defines the required grid oversampling and image support. The effects of the bandpass ripple are shown in b-j. The original test signal is shown in (b). (c)-(d) show the reconstruction errors with ideal interpolator and imperfect calibration. (e)-(g) show result of using a Kaiser-Bessel interpolator. Large image weighting in (e) results in artifact amplification by compensation in (f) and g. (h)-(j) show the result of using a “flat” pass-band interpolator. The post-intensity correction error in (i) and (h) is significantly reduced.

grids the calibration region should have as little passband ripple as possible, again to ensure the fidelity of the calibration process.

Reconstruction

The steps for non-Cartesian k-space based reconstruction are as follows:

- Design an appropriate interpolation kernel based on a specific maximum aliasing amplitude, image weighting, kernel width and grid oversampling
- Perform a calibration that supports the oversampled grid size to obtain the calibration weights
- Solve Eq. 5.10 using the conjugate gradients algorithm to obtain reconstructed Cartesian k-space data for all coils
- Reconstruct the images by computing the inverse Fourier transform
- Crop to the desired grid size
- Apodize the images by multiplying with the inverse Fourier transform of the interpolation kernel.

A diagram for the conjugate gradient algorithm implementation is described in detail in Fig 5.5. The computation of the operators G and D is also illustrated in that figure.

The k-space implementation has the advantage of not requiring a Fourier transform during the iterations. However, it has some disadvantages. To comply with the bandpass ripple requirement the interpolation kernel is inevitably larger than what is commonly used for gridding (for the same interpolation error and grid oversampling). In addition, the calibration consistency convolution has to be performed on an oversampled grid size, with a larger calibration kernel. These increase the complexity of the operations in each iteration of the reconstruction. At some point, depending on the size of the interpolation kernel, calibration kernel, grid size and the specific machine that is used for reconstruction, the operations in k-space may be more costly

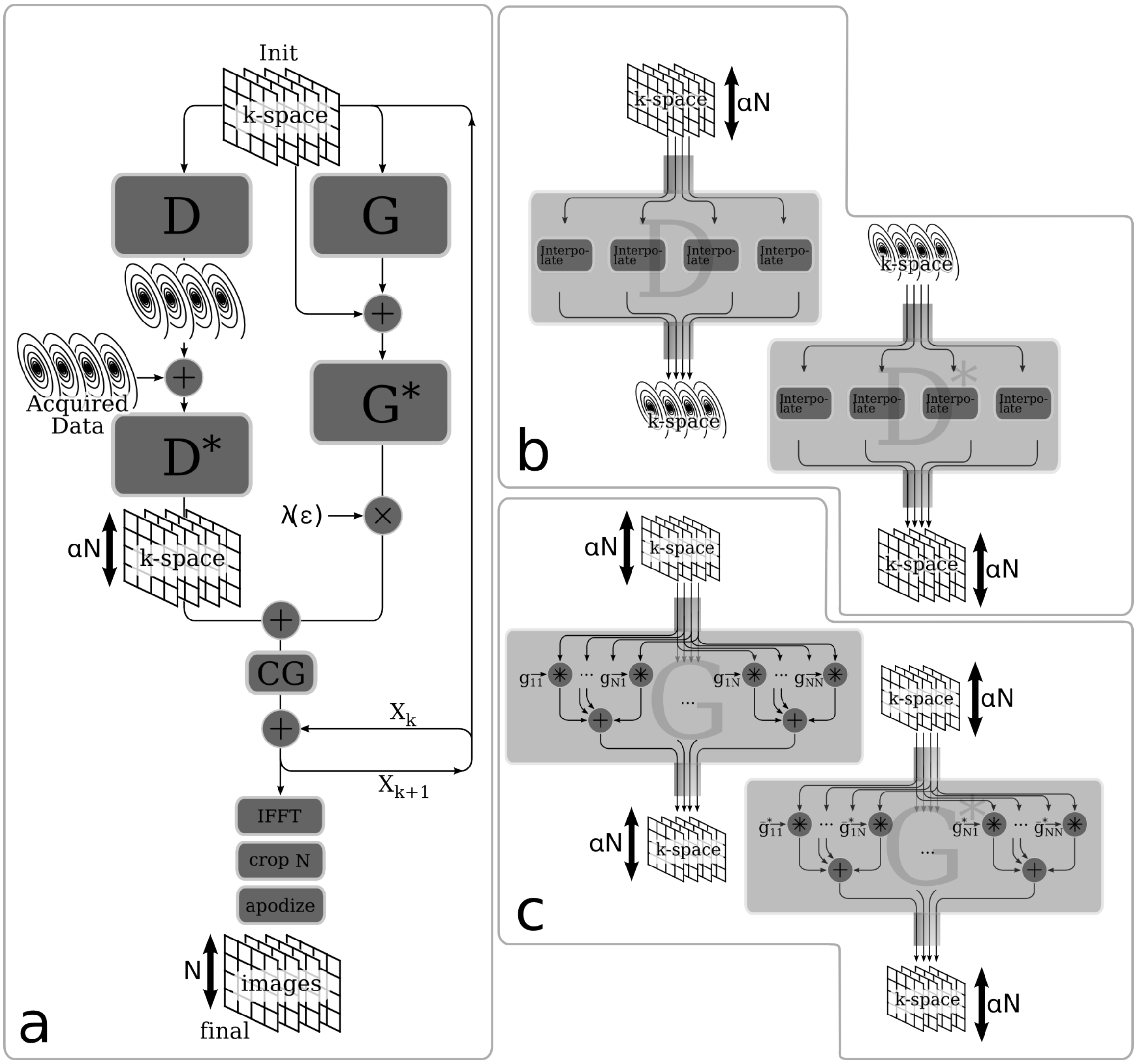
Init
k-space
D
G
G*
Acquired
Data
D*
λ(ε)
αN
k-space
CG
X_k
X_{k+1}
IFFT
crop N
apodize
N
images
final
a
k-space
αN
Interpo-
late
D
k-space
b
D*
k-space
αN
αN
k-space
g11
gN1
g1N
gNN
G
...
αN
k-space
G*
c
αN
k-space
k-space
αN

Figure 5.5: *k*-Space based reconstruction. (a) Illustration of the conjugate-gradient algorithm for non-Cartesian consistency constrained reconstruction in *k*-space. (b) Illustration of the interpolation operator, D, and its adjoint, D^* (c) Illustration of the calibration consistency operator, G and its adjoint, G^*. The notation $\hat{g}_{ji}^*$ stands for an inverted conjugated version of the filter g_{ji}.

than performing similar operations in image space. Therefore, I turn to describe the reconstruction in image space as well.

5.6.3 Image Domain Reconstruction

The reconstruction can be described in the image domain by adjusting the operators D and G appropriately and solving for the full Cartesian images, m, instead of the full Cartesian k-space, x. The new optimization is now,

$$arg\min_{\boldsymbol{m}} \quad ||Dm - y||^2 \; + \; \lambda(\epsilon)||(G - I)m||^2 \qquad | \quad m = \mathrm{IFFT_N(x)}. \tag{5.17}$$

which is, not surprisingly, similar to Eq. 5.10.

When represented as acting in image space, the operator D becomes non-uniform Fourier transform (the nuFFT is historically referred to as inverse gridding) operations. Each nuFFT operates on an individual coil image and computes the k-space values at given k-space locations. It can be written more formally as,

$$y_i(n) = \int \delta(k(n) - r) \left\{ c * \mathrm{FFT}_{\alpha \mathrm{N}} \left(\frac{m_i}{C} \right) \right\} (r) dr \tag{5.18}$$

The advantage here is that the convolution interpolation kernel can be very small, since it is possible to pre-compensate for the weighting in image space prior to taking the Fourier transform.

The modification to the G operator involves conversion of the convolution operations to multiplication with the inverse Fourier transform of the convolution kernels,

$$m_i(r) = \sum_j G_{ji}(r) m_j(r) \qquad | \quad G_{ji} = \mathrm{IFFT_N(g_{ji})}. \tag{5.19}$$

The advantage here is that large calibration kernels do not incur further increase in computational complexity since the convolution is implemented as a multiplication in image space. Figure 5.6 illustrates the conjugate gradient reconstruction of Eq. 5.17, and the implementation of the D and G operators.

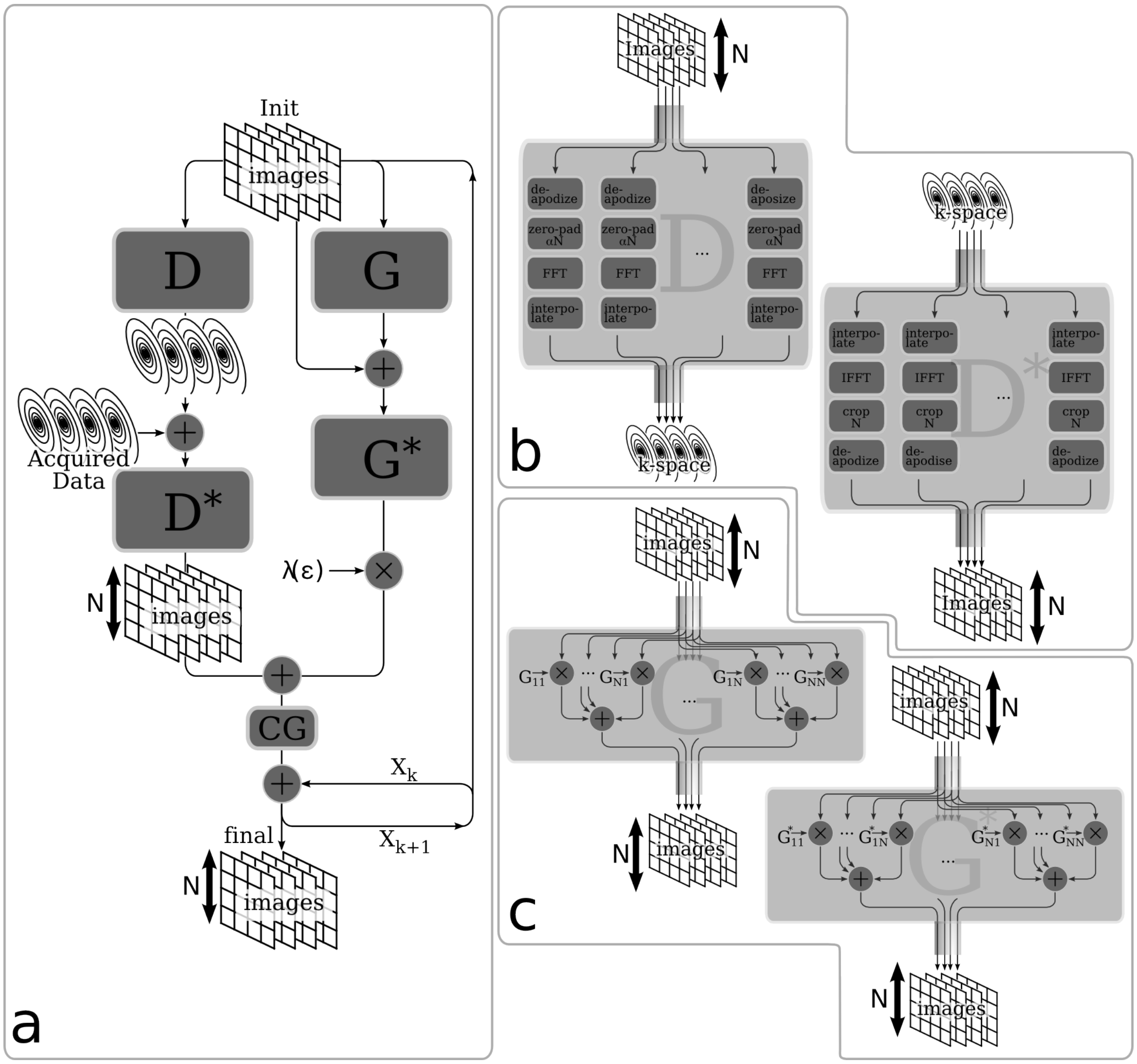
Init
images
D
G
Acquired Data
D*
G*
λ(ε)
N
images
CG
X_k
X_{k+1}
final
N
images
a
Images
N
de-apodize
zero-pad αN
FFT
interpo-late
de-aposize
D
...
k-space
b
k-space
interpo-late
IFFT
crop N
de-apodize
de-apodise
D*
Images
N
images
N
G_{11}
G_{N1}
G_{1N}
G_{NN}
G
images
N
G^*_{11}
G^*_{1N}
G^*_{N1}
G^*_{NN}
G*
c
N
images
N
images

Figure 5.6: Image-space based reconstruction. (a) Illustration of the conjugate gradient algorithm for non-Cartesian consistency constrained reconstruction in image space. (b) Illustration of the non-uniform Fourier transform operator, D, and its adjoint, D^* (c) Illustration of the calibration consistency operator, G and its adjoint, G^*.

5.7 Off Resonance Correction

One of the advantages of representing the reconstruction as a solution to linear equations is that it is possible to incorporate off-resonance variation in the reconstruction. This is particularly important for non-Cartesian trajectories, where off-resonance frequencies lead to image blurring. The off-resonance correction appears in the data consistency constraint in Eq. 5.17 by modifying the nuFFT operator D to include off-resonance information. Denote $\phi_n(r)$ to be a complex vector with unit magnitude elements. The phase of $\phi_n(r)$ represents the phase of the image, contributed by off-resonance at the time the sample $k(n)$ is acquired. The part of the operator D that operates on the i^{th} coil image is,

$$y_i(n) = \int \delta(k(n) - r) \left\{ c * \mathrm{FFT}_{\alpha \mathrm{N}} \left(\frac{m_i \phi_n}{C} \right) \right\} (r) dr. \tag{5.20}$$

Efficient ways to implement this operator is to approximate it using a multi-frequency reconstruction approach as described by [71, 102].

5.8 Regularization

Since Eq. 5.21 is described as an optimization, one can include additional terms in the objective function, as well as constraints that express prior knowledge in the reconstruction. Consider the optimization problem

$$arg\min_{\boldsymbol{x}} \quad ||Dx - y||^2 + \lambda_1 ||(G - I)x||^2 + \lambda_2 R(x) \tag{5.21}$$

where the function $R(x)$ is a penalty function that incorporates the prior knowledge. This formulation is very flexible because the penalty can be applied on the image as well as k-space data. Denoting W as a data weighting function, $\nabla\{\}$ as a finite-difference operator, and $\Psi\{\}$ as a wavelet operator, here are some examples of potential penalties:

$$\begin{aligned} R(x) &= ||x||_2, \quad \text{Tychonov regularization} \\ R(x) &= ||Wx||_2, \quad \text{Weighted Tychonov regularization} \\ R(x) &= ||\nabla\{IFFT(x)\}||_1, \quad \text{Total Variation (TV)} \\ R(x) &= ||\Psi\{IFFT(x)\}||_1, \quad \ell_1 \text{ wavelet.} \end{aligned}$$

The former two are ℓ_1 penalties and are increasingly popular due to the theory of compressed sensing [68] that was introduced in Chapter 4.

5.9 Results

5.9.1 Noise and Artifacts Measurements

To evaluate the noise and reconstruction properties of SPR-iT in comparison to the original GRAPPA method, I constructed the following experiment: I scanned a watermelon 100 times using a balanced-steady-state-free-precessession(b-SSFP) sequence with scan parameters $TE = 2.5\ ms$, $TR = 6.4\ ms$, $flip = 60^\circ$, $BW = 62.5\ Khz$. The slice thickness was 5 mm, field of view of $16.5 \times 16.5\ cm$ and a matrix size of 256×256 corresponding to 0.65×0.65 in-plane resolution. The scan was performed on a GE Signa-Excite 1.5T scanner using an 8-channel receive coil. The data were undersampled in retrospect by a factor of 3 in the phase-encode direction. Then, each image was reconstructed with traditional GRAPPA and with SPIR-iT. Both reconstructions used 30 calibration lines and $9x9$ 2D calibration kernel. SPIR-iT used an equality data consistency constraint and was solved iteratively using the conjugate gradient algorithm with 20 iterations. The resulting reconstructed coil images were combined with square root of sum-of-squares. Following the reconstructions, the mean difference error with the fully sampled data set was calculated. In addition the standard deviation for each pixel was computed and normalized by the standard deviation of the pixels in the full set to obtain the noise amplification (g factor) estimate.

Figure 5.7 shows the result of the experiment. It shows that SPIR-iT has less errors compared to the GRAPPA reconstruction. This is because the data consistency is enforced. In addition, it exhibits a reduced overall g factor noise amplification because the equations are coupled and because early termination of the conjugate gradient algorithm is a form of regularization.

5.9.2 Arbitrary Cartesian Sampling

To demonstrate the generality of SPIR-iT I tested the reconstruction on an irregularly undersampled data set, by retrospectively undersampling (2-fold) the phase-encodes of a T_1 weighted, Cartesian 3D spoiled gradient echo sequence of a brain (see Fig. 5.8a). The scan parameters were $TE = 8\ ms$, $TR = 17.6\ ms$, $flip = 20°$, $BW = 6.94\ Khz$. The field of view was $16 \times 16 \times 22\ cm$ and a matrix size of $192 \times 160 \times 220$ corresponding to a phase-encode plane resolution of $1 \times 1\ mm$. The scan was performed on a GE Signa-Excite 1.5T scanner using an 8-channel receive coil. The data were reconstructed using the conjugate gradient algorithm with 22×22 calibration lines and a 7×7 calibration kernel size. The number of iterations was 10. Figure 5.8 shows the SPIR-iT reconstruction compared to a zero-filled reconstruction. The reconstruction is able to remove all the aliasing artifacts.

5.9.3 k-Space Based Non-Cartesian Reconstruction

To demonstrate the k-space based non-Cartesian reconstruction, I scanned a phantom using a spiral gradient echo sequence. The spiral trajectory was designed with 20 interleaves, 30 cm field of view, and 0.75 mm in-plane resolution. The readout time was kept short, only 5 ms to avoid off-resonance effects. The scan was performed on a GE Signa-Excite 1.5T scanner using a 4-channel cardiac coil.

First, an interpolation kernel was designed (Fig. 5.9b) with time-bandwidth product of 9, 0.1 passband ripple, maximum aliasing error (stop-band) of 0.005 and a transition width corresponding to grid oversampling of 1.25. Then, the complete data set was gridded with density compensation (360×360 matrix size). A central Cartesian k-space of 50×50 pixels was extracted and used for calibrating a 9×9 kernel. The calibration was used in SPIR-iT reconstruction from 6 out of the 20 interleaves. The

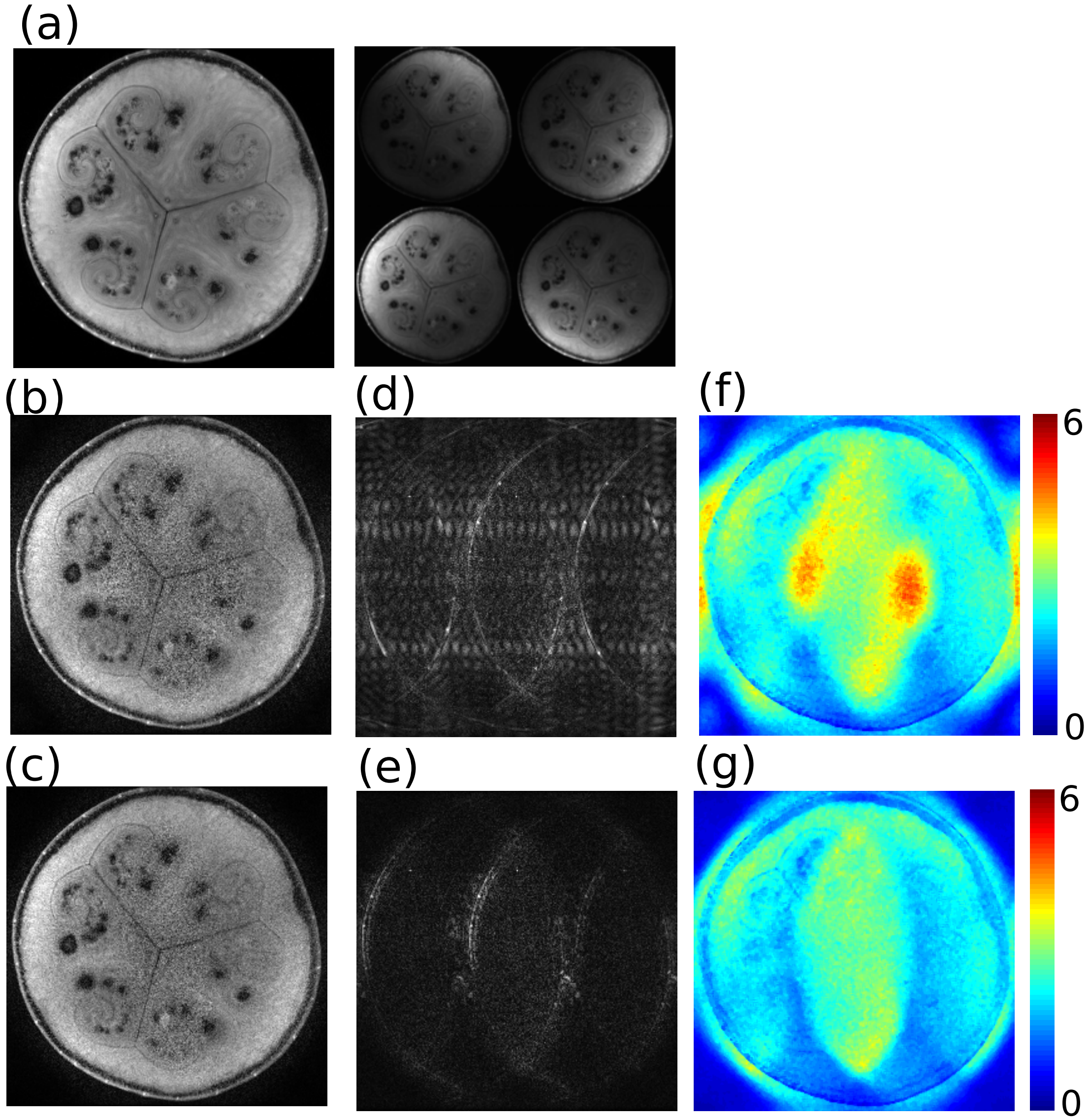

Figure 5.7: Noise measurements statistics of 100 scans. (a) Fully sampled averaged sum-of-squares image and the individual coil images. (b) sum-of-squares of single 1D GRAPPA reconstructed image (c) sum-of-squares of single 1D SPIR-iT reconstruction. (d) the mean 1D GRAPPA reconstruction error showing residual reconstruction artifacts. (e) the mean 1D SPIR-iT reconstruction error showing reduced reconstruction artifacts. (f) The measured noise amplification (g factor) for the 1D GRAPPA reconstruction. (g) SPIR-iT exhibits reduced g factor.

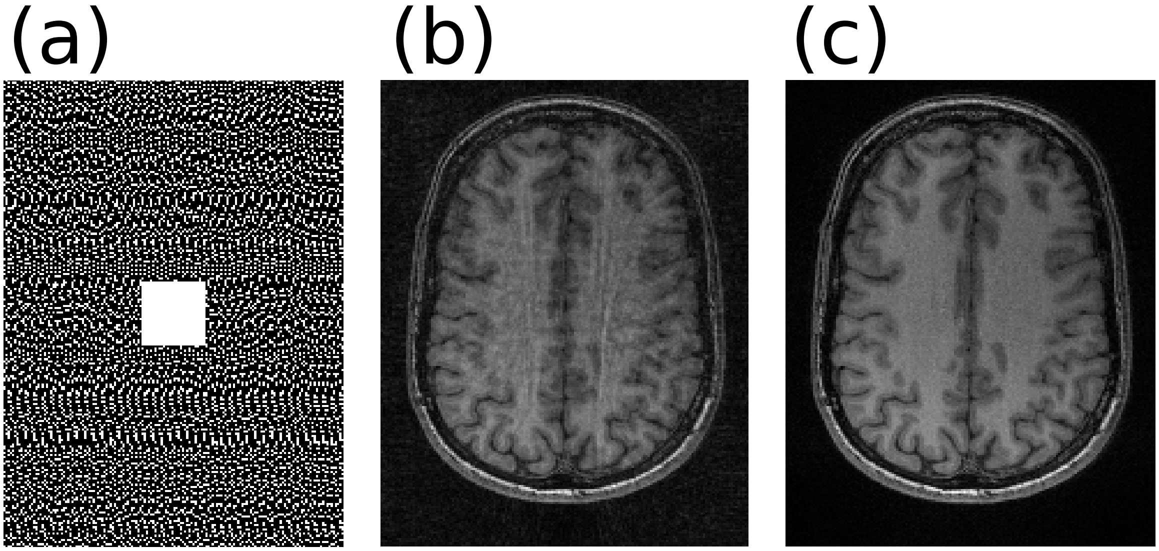

Figure 5.8: SPIR-iT from arbitrary Cartesian sampling. (a) the sampling pattern. (b) sum-of-squares zero-filled reconstruction. (c) consistency constrained reconstruction.

number of iterations was 30.

Figure 5.9 shows the result of the experiment. It shows the SPIR-iT result compared to gridding and the full acquisition. Most of the aliasing artifacts that appear in the gridding reconstruction are removed in SPIR-iT. However, the reconstruction is noisier than the fully sampled reconstruction due to the shorter acquisition time and the g factor.

5.9.4 Image Space Based Non-Cartesian Reconstruction

To demonstrate the image-based non-Cartesian reconstruction I scanned a dynamic heart using a 3 interleave dual density spiral gradient-echo sequence. The spiral trajectory was designed to support field of view ranging from 35 cm in the center to 10 cm (3-fold acceleration) in the periphery. The in-plane resolution was 1.5 mm with readout time of 16 ms, slice thickness was 5 mm. The TR was set to 25 ms achieving a sliding window temporal resolution of 40 FPS. The trajectory is illustrated in Fig. 5.10e. The scan was performed using the rtHawk real-time MRI system [90] installed on a GE Signa-Excite 1.5T scanner with a 4-channel cardiac coil. The data for each

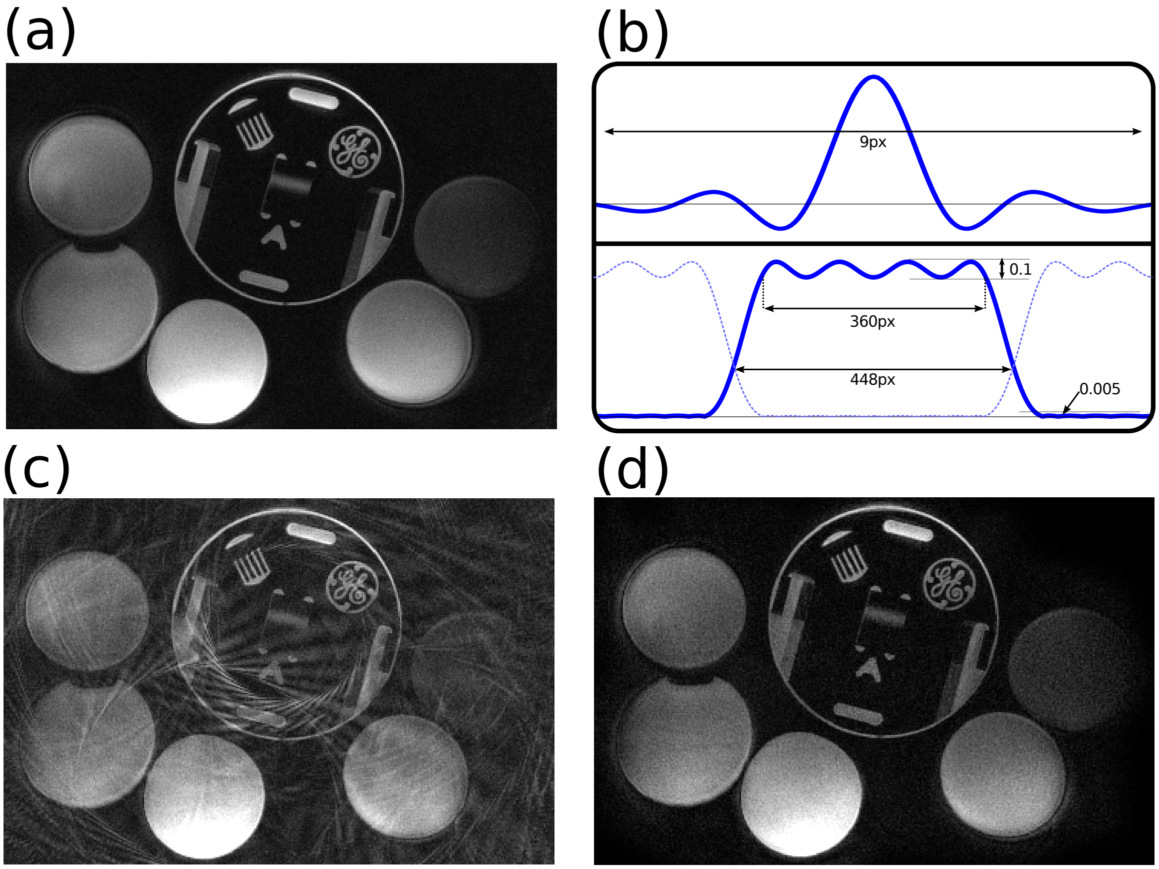
(a)
(b)
9px
0.1
360px
448px
0.005
(c)
(d)

Figure 5.9: k-Space based non-Cartesian SPIR-iT. (a) Reconstruction from fully sampled data. (b) Properties of the k-space interpolator. (c) Gridding reconstruction at 3-fold acceleration. (d) SPIR-iT at 3-fold acceleration.

sliding window image was reconstructed using the image based SPIR-iT algorithm. To speed up the reconstruction the previous image was used as an initial image for the reconstruction of the next image frame. This way, 7 conjugate gradient iterations were sufficient for good image quality.

Figure 5.10 shows two frames in the cardiac cycle. The gridding reconstruction images (a) and (b) suffer from coherent (arrows) and incoherent artifacts due to the undersampling. These artifacts are removed by SPIR-iT.

5.9.5 Off Resonance Correction

To demonstrate the iterative multi-frequency off-resonance correction capabilities of the reconstruction I applied the same scan parameters as in the previous section to a short axis dynamic view of the heart. Prior to the acquisition a field map measurement was taken. As a data consistency operator I used the approach described by Sutton [102].

Figure 5.11 shows the result of off-resonance corrected SPIR-iT compared to a gridding reconstruction with no off-resonance correction. SPIR-iT is able to suppress the coherent and incoherent aliasing, as well as deblurring the image.

5.9.6 Total Variation Regularization

To demonstrate the regularization capabilities of SPIR-iT, I retrospectively undersampled the 3D brain data (described previously) by a factor of four. I reconstructed the data using SPIR-iT with and without total variation regularization.

Figure 5.12 illustrates the result of the reconstructions. It shows that the non-regularized SPIR-iT exhibits increased noise, especially in the center of the image, due to the g factor. The noise is suppressed by the total variation regularization, while the edges and features in the image are preserved.

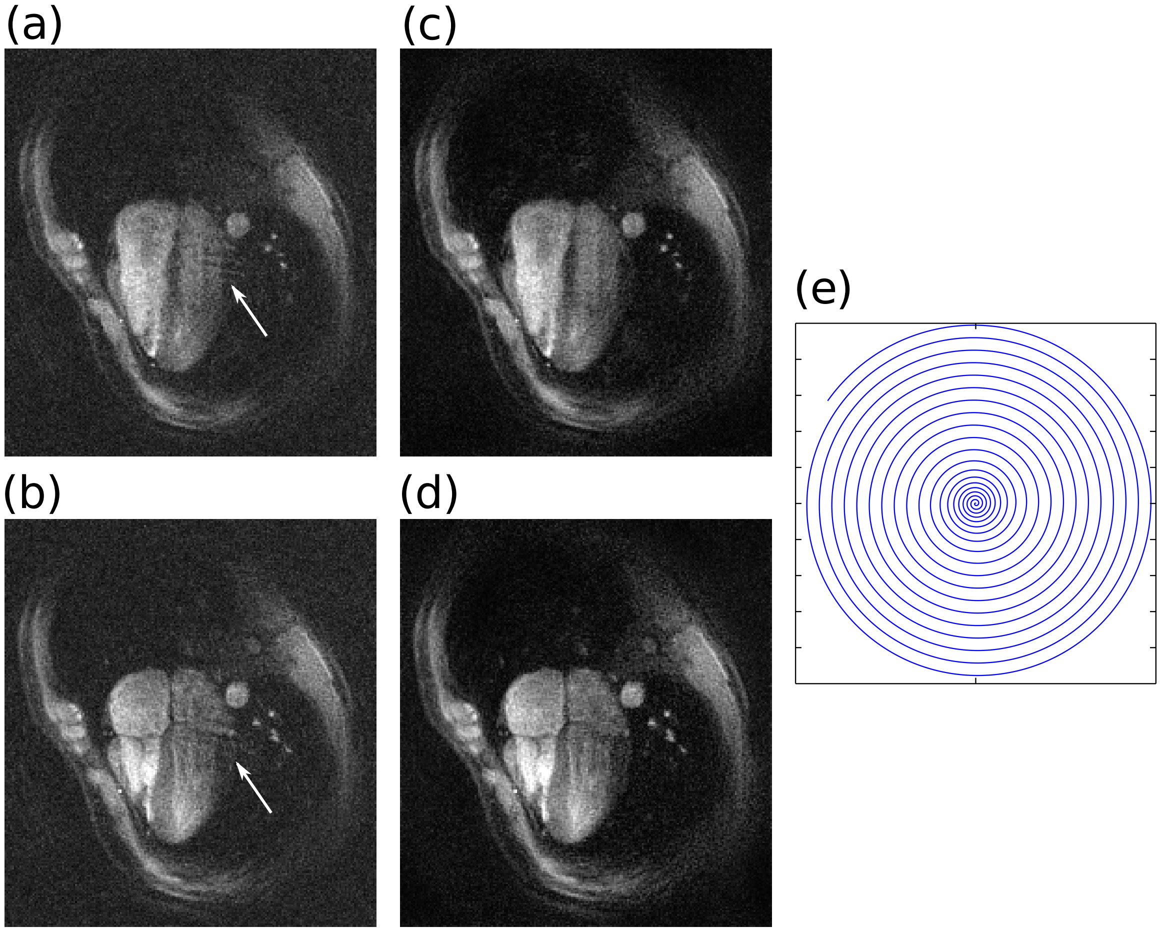

Figure 5.10: Dynamic cardiac imaging with dual density spirals. Two phases of a four chamber view of the heart. (a)-(b) Sum-of-squares of gridding reconstruction exhibits coherent (arrows) and incoherent (noise-like) aliasing artifacts. (c)-(d) Both the coherent and incoherent artifacts are removed by SPIR-iT. (e) One out of the three spiral interleaves.

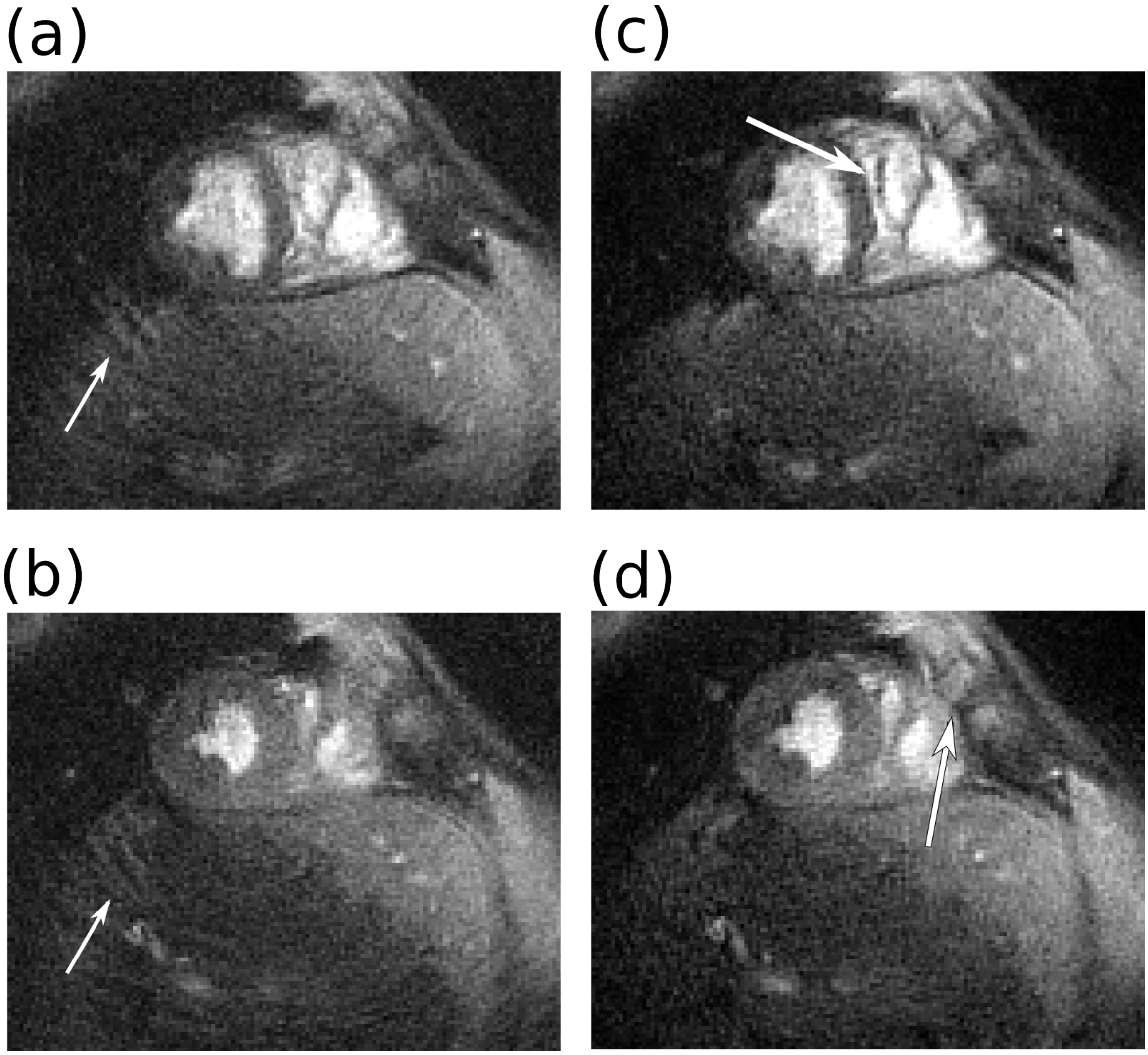

Figure 5.11: Dynamic cardiac imaging with dual density spirals and off-resonance correction. Two phases of a short axis view of the heart. (a)-(b) sum-of-squares of gridding reconstruction exhibits coherent (arrows), incoherent (noise-like) aliasing artifacts and blurring due to off-resonance. (c)-(d) SPIR-iT reduces both the coherent and incoherent artifacts as well as deblurring the image (arrows).

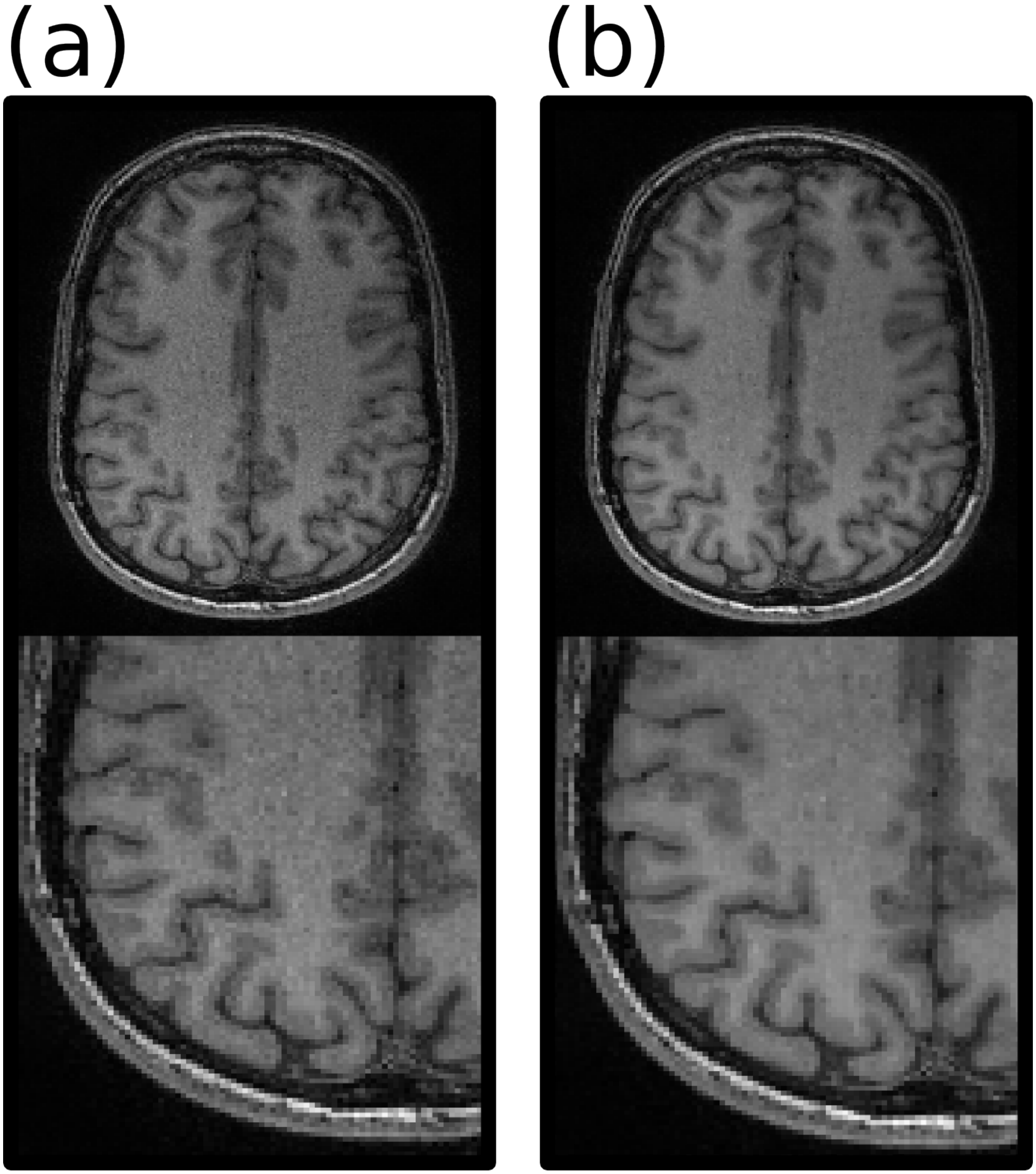

Figure 5.12: Total Variation regularization (TV): (a) the non-Regularized SPIR-iT exhibits noise amplification due to the g factor. (b) the noise amplification is suppressed by the TV reconstruction, while the edges and features in the image are preserved.

5.10 Conclusions

In conclusion, I have presented SPIR-iT, a self consistency formulation for coil-by-coil parallel imaging reconstruction. The reconstruction can be solved efficiently using iterative methods for data acquired on arbitrary k-space trajectories. I presented two variants of the reconstruction; image-space based, and k-space based. In the k-space based reconstruction I introduced new requirements for k-space interpolation. In addition, I demonstrated that the formulation is easily extendible to include other techniques such as regularization and off-resonance correction. I showed that enforcing consistency constraints results in better conditioning, more accurate reconstruction, and better noise behavior than the traditional GRAPPA approach.

Chapter 6

Summary and Recommendation

The title “Sparse MRI” summarized well the theme of this thesis. On the one hand, sparse sampling of k-space is used to accelerate the acquisitions and reduce the overall scan time. On the other hand, sparsity of representation allows the sparsely sampled data to be reconstructed into accurate full resolution, full field of view images. Continuing with the theme, I present a somewhat sparse summary. The key contributions of the thesis are:

- A new method to design time-optimal gradient waveforms. It provides a general, computationally efficient solution for arbitrary k-space trajectories.
- The application of the compressed sensing theory to rapid MRI.
- A new iTerative Self-consistent Parallel Imaging Reconstruction (SPIR-iT) method. It is a general, optimal solution for auto-calibrating parallel imaging from arbitrary k-space trajectories. It is also a general framework for easily incorporating additional image priors, and in particular sparsity/compressibility constraints.

There are several directions for future work that I find interesting:

- Extension of the time-optimal gradient waveform design to non-rotationally invariant hardware constraints.
- Extension of the time-optimal gradient waveform design to include the L-C model for gradient hardware constraints.

- Finding the optimal sampling density for compressed sensing.
- Improving the reconstruction time for compressed sensing.
- Extending compressed sensing to non-periodic dynamic imaging.
- Exploring the combination of compressed sensing, parallel imaging and non-Cartesian trajectories.
- Training sparsifying transformations tailored for specific applications. Instead of wavelets use brainlets, angiolets, cardiaclets, diffusionlets, etc.

Appendix A

Derivations for the Time-Optimal Gradient Waveform Design

A.1 Arc Length Expression of the Gradient Constraints

I start by observing that

$$\begin{aligned} |\dot{g}(t)|^2 &= \left|\gamma^{-1}\left(C''(s(t))\dot{s}(t)^2 + C'(s(t))\ddot{s}(t)\right)\right|^2 \\ &= \gamma^{-2}\left(|C''(s(t))|^2\dot{s}(t)^4 + |C'(s(t))|^2\ddot{s}(t)^2\right. \\ &\quad \left. + C''(s(t))^T C'(s(t))\dot{s}(t)^2\ddot{s}(t)\right). \end{aligned}$$

The tangent and the normal vectors are always orthogonal, that is, $C''(s(t))^T C'(s(t)) = 0$. Also, recall that $|C'(s(t))| = 1$ and that $|C''(s(t))| = \kappa(s(t))$. The slew-rate constraint can therefore be expressed as

$$\gamma^{-2}\left(k(s(t))^2\dot{s}(t)^4 + \ddot{s}(t)^2\right) \leq S_{\max}^2,$$

which in turn implies that

$$\ddot{s}(t)^2 \leq \left[\gamma^2 S_{\max}^2 - \kappa(s(t))^2\dot{s}(t)^4\right].$$

The right-hand side is non-negative when

$$\dot{s}(t) \leq \sqrt{\frac{\gamma S_{\max}}{\kappa(s(t))}}.$$

Finally, the gradient magnitude is limited $|g(t)| \leq G_{\max}$ resulting in the desired second constraint,

$$\dot{s}(t) \leq \min\left\{\gamma G_{\max}, \sqrt{\frac{\gamma S_{\max}}{\kappa(s(t))}}\right\}.$$

A.2 Optimality of the Gradient Waveform Design

Let v denote any velocity profile along the trajectory that satisfies the constraints of (3.18). Recall that the traversal time T_v of v is given by

$$T_v = \int_0^L \frac{ds}{v(s)}.$$

It follows from Lemma 3 in [53] that

$$v(s) \leq v^\star(s) = \min\{v_+(s), v_-(s)\}, \quad \forall\, s \in [0, L].$$

Therefore, $T_v \geq T_{v^\star}$. It can be shown through some arguments similar to those used to prove Theorem 3 in [53] that $v^\star$ satisfies the constraints of (3.18). Therefore, $v^\star$ is optimal for (3.18).

Bibliography

[1] A. V. Barger, W. F. Bloch, Y. Toropov, T. M. Grist, and C. A. Mistretta. Time-resolved contrast-enhanced imaging with isotropic resolution and broad coverage using an undersampled 3D projection trajectory. *Magn Reson Med*, 48(2):297–305, 2002.

[2] P. J. Beatty. *Reconstruction Methods for Fast Magnetic Resonance Imaging.* PhD thesis, Stanford University, 2006.

[3] P.J. Beatty, D.G. Nishimura, and J.M. Pauly. Rapid gridding reconstruction with a minimal oversampling ratio. *IEEE Trans Med Imaging*, 24:799–808, Jun 2005.

[4] M. A. Bernstein, F. K. King, and X. J. Zhou. *Handbook of MRI pulse sequences.* Elsevier Academic Press, Burlington, MA, 2004.

[5] M. A. Bernstein, F. K. King, and X. J. Zhou. *Handbook of MRI pulse sequences*, chapter 16. Elsevier Academic Press, Burlington, MA, 2004.

[6] K. T. Block, M. U., and J. Frahm. Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint. *Magn Reson Med*, 57(6):1086–1098, Jun 2007.

[7] W. E. Boyce and R. C. DiPrima. *Elementary Differential Equations and Boundary Value Problems.* John Wiley&Sons, Inc., 6 edition, 1997.

[8] M. M. Bronstein, A. M. Bronstein, M. Zibulevsky, and H. Azhari. Reconstruction in diffraction ultrasound tomography using nonuniform FFT. *IEEE Trans Med Imaging*, 21(11):1395–1401, Nov 2002. Evaluation Studies.

[9] E. K. Buc, S. Kohler, and I. Hancu. 13C multi-spectral 2D rosette imaging. In *Proceedings of the 15th Annual Meeting of ISMRM*, page 1379, Berlin, 2007.

[10] R. W. Buccigrossi and E. P. Simoncelli. Image compression via joint statistical characterization in the wavelet domain. *IEEE Trans. Image Processing*, 8:1688–1701, 1999.

[11] Emmanuel Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52:489–509, 2006.

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

[13] E. Candés and J. K. Romberg. Signal recovery from random projections. In *Proceedings of SPIE Computational Imaging III*, page 5674, San Jose, 2005.

[14] T. C. Chang, L. He, and T. Fang. MR image reconstruction from sparse radial samples using bregman iteration. In *Proceedings of the 14th Annual Meeting of ISMRM*, page 696, Seattle, 2006.

[15] S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM J. Sci Comp*, 20(1):33–61, 1999.

[16] B. A. Chronik, B. K. Rutt, B. A. Chronik, and B. K. Rutt. A comparison between human magnetostimulation thresholds in whole-body and head/neck gradient coils simple linear formulation for magnetostimulation specific to MRI gradient coils. *Magn Reson Med*, 45(5):916–919, 2001.

[17] B. M. Dale and J. L. Duerk. Time-optimal control of gradients. In *Proceedings of the 10th Annual Meeting of ISMRM*, page 2361, Honolulu, 2002.

[18] B. M. Dale, J. S. Lewin, and J. L. Duerk. Optimal design of k-space trajectories using a multi-objective genetic algorithm. *Magn Reson Med*, 52(4):831–841, Oct 2004.

[19] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Applied Mathematics*, 57:1413 – 1457, 2004.

[20] D. L. Donoho, M. Elad, and V. Temlyakov. Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE Trans. Info. Thry*, 52:6–18, 2006.

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

[22] D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81:425–455, 1994.

[23] D. L. Donoho and Y. Tsaig. Fast solution of ℓ_1-norm minimization problems when the solution may be sparse. *IEEE Transactions on Information Theory*, 2008. Preprint.

[24] D. L. Donoho, Y. Tsaig, I. Drori, and J-L. Starck. Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit. *Technical Report, Department of Statistics, Stanford University*, 2006. Preprint.

[25] M. Doyle, M. B. Scheidegger, R. G. Degraaf, J. Vermeulen, and G. M. Pohost. Coronary artery imaging in multiple 1-sec breath holds. *Magnetic Resonance Imaging*, 11(1):3 – 6, 1993.

[26] M. Doyle, E. G. Walsh, G. G. Blackwell, and G. M. Pohost. Block regional interpolation scheme for k-space (BRISK): a rapid cardiac imaging technique. *Magn Reson Med*, 33(2):163–170, Feb 1995. Comparative Study.

[27] M. Elad, B. Matalon, and M. Zibulevsky. Coordinate and subspace optimization methods for linear least squares with non-quadratic regularization. *Journal on Applied and Computational Harmonic Analysis*, 2006. In Press.

[28] J. A. Fessler and B. P. Sutton. Nonuniform fast Fourier transforms using min-max interpolation. *IEEE Transactions on Signal Processing*, 51(2):560 – 74, February 2003.

[29] D. J. Field. Relations between the statistics of natural images and the response profiles of cortical cells. *Journal of the Optical Society of America*, A(4):2379–2394, 1987.

[30] G. H. Glover. Simple analytic spiral *k*-space algorithm. *Magn Reson Med*, 42(2):412–415, 1999.

[31] G. H. Glover and J. M. Pauly. Projection reconstruction technique for reduction of motion effects in MRI. *Magn Reson Med*, 28:275–289, 1992.

[32] A. Greiser and M. von Kienlin. Efficient k-space sampling by density-weighted phase-encoding. *Magn Reson Med*, 50(6):1266–1275, Dec 2003.

[33] M. A. Griswold, P. M. Jakob, M. Nittka, J. W. Goldfarb, and A. Haase. Partially parallel imaging with localized sensitivities (PILS). *Magn Reson Med*, 44:602–609, Oct 2000.

[34] M. A. Griswold, S. Kannengiesser, R. M. Heidemann, J. Wang, and P. M. Jakob. Field-of-view limitations in parallel imaging. *Magn Reson Med*, 52:1118–1126, Nov 2004.

[35] 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). *Magn Reson Med*, 47(6):1202–10, 2002.

[36] P. T. Gurney, B. A. Hargreaves, and D. G. Nishimura. Design and analysis of a practical 3D cones trajectory. *Magn Reson Med*, 55(3):575–582, Mar 2006.

[37] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. Wiley-Liss, New York, 1st edition, 1999.

[38] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*, chapter 14. Wiley-Liss, New York, 1st edition, 1999.

[39] J. M. Hanson, Z. P. Liang, R. L. Magin, J. L. Duerk, and P. C. Lauterbur. A comparison of RIGR and SVD dynamic imaging methods. *Magn Reson Med*, 38(1):161–167, Jul 1997. Comparative Study.

[40] B. A. Hargreaves, D. G. Nishimura, and S. M. Conolly. Time-optimal multidimensional gradient waveform design for rapid imaging. *Magn Reson Med*, 51(1):81–92, 2004.

[41] J. Haupt and R. Nowak. Signal reconstruction from noisy random projections. *IEEE Transactions in Information Theory*, 52(9):4036–4048, 2006.

[42] K. Heberlein and X. Hu. Auto-calibrated parallel spiral imaging. *Magnetic resonance in medicine*, 55(3):619–625, 2006.

[43] O. Heid. The fastest circular k-space trajectories. In *Proceedings of the 10th Annual Meeting of ISMRM*, page 2364, Honolulu, 2002.

[44] R. M. Heidemann, M. A. Griswold, A. Haase, and P. M. Jakob. VD-AUTO-SMASH imaging. *Magn Reson Med*, 45:1066–1074, Jun 2001.

[45] R. M. Heidemann, M. A. Griswold, N. Seiberlich, G. Krger, S. A. Kannengiesser, B. Kiefer, G. Wiggins, L. L. Wald, and P. M. Jakob. Direct parallel image reconstructions for spiral trajectories using GRAPPA. *Magn Reson Med*, 56:317–326, Aug 2006.

[46] P. Hu and C. H. Meyer. Bosco: Parallel image reconstruction based on successive convolution operations. In *Proceedings of the 14th Annual Meeting of ISMRM*, page 10, Seattle, 2006.

[47] S. Dubosky J. Bobrow and J. Gibson. Time-optimal control of robotic manipulators along specified paths. *The International Journal of Robotics Research*, 4(3):3–17, 1985.

[48] J. I. Jackson, C. H. Meyer, D. G. Nishimura, and A. Macovski. Selection of a convolution function for Fourier inversion using gridding. *IEEE Trans Med Imaging*, 10(3):473–478, 1991.

[49] P. M. Jakob, M. A. Griswold, R. R. Edelman, and D. K. Sodickson. AUTO-SMASH: a self-calibrating technique for SMASH imaging. SiMultaneous Acquisition of Spatial Harmonics. *MAGMA*, 7:42–54, Nov 1998.

[50] H. Jung, J. C. Ye, and E. Y. Kim. Improved k-t BLAST and k-t SENSE using FOCUSS. *Phys Med Biol*, 52(11):3201–3226, Jun 2007.

[51] J. R. Kelton, R. L. Magin, and S. M. Wright. An algorithm for rapid image acquisition using multiple receiver coils. In *Proc., SMRM, 8th Annual Meeting*, page 1172, Amsterdam, 1989.

[52] D-H. Kim, E. Adalsteinsson, and D. M. Spielman. Simple analytic variable density spiral design. *Magn Reson Med*, 50(1):214–219, 2003.

[53] S-J. Kim, D. Choi, and I. Ha. A comparison principle for state-constrained differential inequalities and its application to time-optimal control. *IEEE Transactions on Automatic Control*, 50(7):967–983, 2005.

[54] S-J. Kim, K. Koh, M. Lustig, and S. Boyd. An efficient method for compressed sensing. In *Proceedings of IEEE International Conference on Image Processing (ICIP)*, San Antonio, 2007. in press.

[55] K. F. King, T. K. Foo, and C. R. Crawford. Optimized gradient waveforms for spiral scanning. *Magn Reson Med*, 34(2):156–160, 1995.

[56] D. A. Koff and H. Shulman. An Overview of Digital Compression of Medical Images: Can We Use Lossy Compression in Radiology? *CARJ*, 4:211–217, 2006.

[57] F. R. Korosec, R. Frayne, T. M. Grist, and C. A. Mistretta. Time-resolved contrast-enhanced 3D MR angiography. *Magn Reson Med*, 36(3):345–351, Sep 1996.

[58] F. R. Korosec, R. Frayne, T. M. Grist, and C. A. Mistretta. Time-resolved contrast-enhanced 3D MR angiography. *Magn Reson Med*, 36(3):345–351, Sep 1996.

[59] W. E. Kyriakos, L. P. Panych, D. F. Kacher, C. F. Westin, S. M. Bao, R. V. Mulkern, and F. A. Jolesz. Sensitivity profiles from an array of coils for encoding and reconstruction in parallel (SPACE RIP). *Magn Reson Med*, 44:301–308, Aug 2000.

[60] J. H. Lee, B. A. Hargreaves, B. S. Hu, and D. G. Nishimura. Fast 3D imaging using variable-density spiral trajectories with applications to limb perfusion. *Magn Reson Med*, 50(6):1276–1285, 2003.

[61] Z. P. Liang and P. C. Lauterbur. *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*, chapter 10. IEEE Press, Picataway, NJ, 1999.

[62] Z. P. Liang and P. C. Lauterbur. *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective.* IEEE Press, Picataway, NJ, 1999.

[63] C. Liu, R. Bammer, and M. E. Moseley. Parallel imaging reconstruction for arbitrary trajectories using k-space sparse matrices (kSPA). *Magn Reson Med*, 58:1171–1181, Dec 2007.

[64] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly. Compressed Sensing MRI. *IEEE Signal Processing Magazine*, 27:72–82, March 2008.

[65] M. Lustig, S. J. Kim, and J. M. Pauly. A fast method for designing time-optimal gradient waveforms for arbitrary k-space trajectories. *IEEE Trans Med Imaging*, 27:866–873, Jun 2008.

[66] M. Lustig, J. H. Lee, D. L. Donoho, and J. M. Pauly. Faster imaging with randomly perturbed, under-sampled spirals and ℓ^1 reconstruction. In *Proceedings of the 13th Annual Meeting of ISMRM*, page 685, Miami Beach, 2005.

[67] M. Lustig, J. M. Santos, D. L. Donoho, and J. M. Pauly. k-t SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity. In *Proceedings of the 14th Annual Meeting of ISMRM*, page 2420, Seattle, 2006.

[68] M. Lustig, D. L. Donoho, and J. M. Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med*, 58:1182–1195, 2007.

[69] M. Lustig and J. M. Pauly. Iterative GRAPPA: A general solution for the GRAPPA reconstruction from arbitrary *k*-space trajectories. In *Proceedings of the 15th Annual Meeting of ISMRM*, page 333, Berlin, 2007.

[70] B. Madore, G. H. Glover, and N. J. Pelc. Unaliasing by fourier-encoding the overlaps using the temporal dimension (UNFOLD), applied to cardiac imaging and fMRI. *Magn Reson Med*, 42(5):813–828, 1999.

[71] L. C. Man, J. M. Pauly, and A. Macovski. Multifrequency interpolation for fast off-resonance correction. *Magn Reson Med*, 37:785–792, May 1997.

[72] G. J. Marseille, R. de Beer, M. Fuderer, A. F. Mehlkopf, and D. van Ormondt. Nonuniform phase-encode distributions for MRI scan time reduction. *J Magn Reson*, 111(1):70–75, 1996.

[73] G. McGibney, M. R. Smith, S. T. Nichols, and A. Crawley. Quantitative evaluation of several partial Fourier reconstruction algorithms used in MRI. *Magn Reson Med*, 30(1):51–59, 1993.

[74] C. H. Meyer, B. S. Hu, D. G. Nishimura, and A. Macovski. Fast spiral coronary artery imaging. *Magn Reson Med*, 28(2):202–213, 1992.

[75] R. Mir, A. Guesalaga, J. Spiniak, M. Guarini, and P. Irarrazaval. Fast three-dimensional k-space trajectory design using missile guidance ideas. *Magn Reson Med*, 52(2):329–336, Aug 2004.

[76] C. A. Mistretta, O. Wieben, J. Velikina, W. Block, J. Perry, Y. Wu, K. Johnson, and Y. Wu. Highly constrained backprojection for time-resolved MRI. *Magn Reson Med*, 55(1):30–40, Jan 2006.

[77] C. A. Mistretta, O. Wieben, J. Velikina, W. Block, J. Perry, Y. Wu, K. Johnson, and Y. Wu. Highly constrained backprojection for time-resolved MRI. *Magn Reson Med*, 55(1):30–40, Jan 2006.

[78] K. S. Nayak and D. G. Nishimura. Randomized trajectories for reduced aliasing artifacts. In *Proceedings of the 6th Annual Meeting of ISMRM*, page 670, Sydney, 1998.

[79] D. C. Noll. Multishot rosette trajectories for spectrally selective MR imaging. *IEEE Trans Med Imaging*, 16(4):372–377, Aug 1997.

[80] B. A. Olshausen and D. J. Field. Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14:481–487, 2004.

[81] T. Parrish and X. Hu. Continuous update with random encoding (CURE): a new strategy for dynamic imaging. *Magn Reson Med*, 33(3):326–336, Mar 1995.

[82] D. C. Peters, F. R. Korosec, T. M. Grist, W. F. Block, J. E. Holden, K. K. Vigen, and C. A. Mistretta. Undersampled projection reconstruction applied to MR angiography. *Magn Reson Med*, 43(1):91–101, 2000.

[83] J. G. Pipe. An optimized center-out k-space trajectory for multishot MRI: Comparison with spiral and projection reconstruction. *Magn Reson Med*, 42(4):714–720, 1999.

[84] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger. SENSE: Sensitivity encoding for fast MRI. *Magn Reson Med*, 42(5):952–962, 1999.

[85] K. P. Pruessmann, M. Weiger, P. Brnert, and P. Boesiger. Advances in sensitivity encoding with arbitrary k-space trajectories. *Magn Reson Med*, 46:638–651, Oct 2001.

[86] D. Rovnyak, D.P. Frueh, M. Sastry, Z. J. Sun, A. S. Stern, J. C. Hoch, and G. Wagner. Accelerated acquisition of high resolution triple-resonance spectra using non-uniform sampling and maximum entropy reconstruction. *Journal of Magnetic Resonance*, 170(1):15–21, 2004.

[87] L. Rudin, S. Osher, and E. Fatemi. Non-linear total variation noise removal algorithm. *Phys. D*, 60:259–268, 1992.

[88] T. S. Sachs, C. H. Meyer, P. Irarrazabal, B. S. Hu, D. G. Nishimura, and A. Macovski. The diminishing variance algorithm for real-time reduction of motion artifacts in MRI. *Magn Reson Med*, 34(3):412–422, 1995.

[89] J. M. Santos, C. H. Cunningham, M. Lustig, B. A. Hargreaves, B. S. Hu, D. G. Nishimura, and J. M. Pauly. Single breath-hold whole-heart MRA using variable-density spirals at 3T. *Magnetic Resonance in Medicine*, 55(2):371–379, 2006.

[90] J.M. Santos, G.A. Wright, and J.M. Pauly. Flexible real-time magnetic resonance imaging framework. *Conf Proc IEEE Eng Med Biol Soc*, 2:1048–1051, 2004.

[91] K. Scheffler and J. Hennig. Frequency resolved single-shot MR imaging using stochastic k-space trajectories. *Magn Reson Med*, 35(4):569–576, Apr 1996.

[92] K. Scheffler and J. Hennig. Reduced circular field-of-view imaging. *Magn Reson Med*, 40(3):474–480, 1998.

[93] N. Seiberlich, F. Breuer, R. Heidemann, M. Blaimer, M. Griswold, and P. Jakob. Reconstruction of undersampled non-Cartesian data sets using pseudo-Cartesian GRAPPA in conjunction with GROG. *Magn Reson Med*, 59:1127–1137, May 2008.

[94] K. Shin and N. Mckay. Minimum-time control of robotic manipulators with geometric path constraints. *IEEE Transactions on Automatic Control*, 30(6):531–541, 1985.

[95] O. P. Simonetti, J. L. Duerk, and V. Chankong. MRI gradient waveform design by numerical optimization. *Magn Reson Med*, 29(4):498–504, 1993.

[96] O. P. Simonetti, J. L. Duerk, and V. Chankong. An optimal design method for magnetic resonance imaging gradient waveforms. *IEEE Trans Med Imaging*, 12(2):350–360, 1993.

[97] D. K. Sodickson and W. J. Manning. Simultaneous acquisition of spatial harmonics (SMASH): Fast imaging with radiofrequency coil arrays. *Magn Reson Med*, 38(4):591–603, 1997.

[98] D. M. Spielman, J. M. Pauly, and C. H. Meyer. Magnetic resonance fluoroscopy using spirals with variable sampling densities. *Magnetic resonance in medicine*, 34(3):388–394, 1995.

[99] J-L. Starck, M. Elad, and D. L. Donoho. Image decomposition via the combination of sparse representation and a variational approach. *IEEE Trans. Image Proc.*, 14(10):1570–1582, 2005.

[100] J-L. Starck, M. Elad, and D. L. Donoho. Image decomposition via the combination of sparse representations and a variational approach. *IEEE Trans. On Image Processing*, 14(10):1570–1582, 2005.

[101] M. Stuber, R. M. Botnar, P. G. Danias, D. K. Sodickson, K. V. Kissinger, M. Van Cauteren, J. De Becker, and W. J. Manning. Double-oblique free-breathing high resolution three-dimensional coronary magnetic resonance angiography. *Journal of the American College of Cardiology*, 34(2):524 – 531, Aug., 1999.

[102] B. P. Sutton, D. C. Noll, and J. A. Fessler. Fast, iterative image reconstruction for MRI in the presence of field inhomogeneities. *IEEE Trans Med Imaging*, 22:178–188, Feb 2003.

[103] David S. Taubman and Michael W. Marcellin. *JPEG 2000: Image Compression Fundamentals, Standards and Practice.* Kluwer International Series in Engineering and Computer Science., 2002.

[104] C-M. Tsai and D. G. Nishimura. Reduced aliasing artifacts using variable-density *k*-space sampling trajectories. *Magn Reson Med*, 43(3):452–458, 2000.

[105] Y. Tsaig and D. L. Donoho. Extensions of compressed sensing. *Signal Processing*, 86:533–548, 2006.

[106] J. Tsao, P. Boesiger, and K. P. Pruessmann. *k-t* BLAST and *k-t* SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations. *Magn Reson Med*, 50(5):1031–1042, 2003.

[107] J. J. van Valls, M. E Brumrner, W. T. Dixon, H. H. Tuithof, H. Engels, R. C Nelson, B. M. Gerety, J. L. Chezmar, and J. A. Boer. Keyhole method for accelerating imaging of contrast agent uptake. *J Magn Reson Imag*, 3:671–675, 1993.

[108] F. Wajer. *Non-Cartesian MRI Scan Time Reduction through Sparse Sampling.* PhD thesis, Delft University of Technology, 2001.

[109] O. M. Weber, A. J. Martin, and C. B. Higgins. Whole-heart steady-state free precession coronary artery magnetic resonance angiography. *Magnetic Resonance in Medicine*, 50(6):1223 – 8, December 2003.

[110] N. P. Willis and Y. Bresler. Optimal scan for time-varying tomography. I. Theoretical analysis and fundamental limitations. *IEEE Trans Image Process*, 4:642–653, 1995.

[111] N. P. Willis and Y. Bresler. Optimal scan for time-varying tomography. II. Efficient design and experimental validation. *IEEE Trans Image Process*, 4:654–666, 1995.

[112] E. N. Yeh, C. A. McKenzie, M. A. Ohliger, and D. K. Sodickson. Parallel magnetic resonance imaging with adaptive radius in k-space (PARS): constrained

image reconstruction using k-space locality in radiofrequency coil encoded data. *Magn Reson Med*, 53:1383–1392, Jun 2005.