Coherence Analysis of Compressive Sensing-Based Magnetic Resonance Imaging Reconstruction

Kai Zhu

Faculty of Science, Engineering and Technology
Swinburne University of Technology

This dissertation is submitted for the degree of
Doctor of Philosophy

Melbourne, Australia 2017
Abstract

This thesis focuses on the coherence analysis of compressive sensing (CS)-based magnetic resonance imaging (MRI) reconstruction. MRI is widely used in the field of medical imaging due to its excellent resolution in tissue depiction, flexible orientation, and freedom from harmful radiation. However, one major problem of MRI is its relatively long data acquisition time, which may lead to blur and artifacts in magnetic resonance (MR) images, and restrict its applications in certain areas.

Intensive research activities have been carried out in order to shorten the data acquisition time. In traditional MRI, images of interest are encoded with Fourier encoding, so that the acquired data are stored in the Fourier domain, or k-space. Normally the k-space data are acquired row by row following a Cartesian trajectory. Each row of data is collected in one phase encoding step. A common practice to shorten the data acquisition time is to sub-sample the k-space data by deliberately skipping some phase encoding steps. Therefore, specially designed reconstruction methods are required to reconstruct images from sub-sampled k-space data. Reconstruction based on CS is a new and promising approach, which exploits the sparsity of images. In CS, a sparse signal can be recovered from many fewer randomly-sampled measurements compared to the length of the original signal.

For data encoded with Fourier encoding, most of the energy is stored in the low frequency part of k-space, which controls the main structure and brightness of images. In the CS-based MRI reconstruction method, the low frequency part needs to be densely sampled in order to preserve the main structure and brightness. Accordingly, the high frequency part should be sparsely sampled to achieve the intended reduction factor ($R$). This kind of variable density sub-sampling violates the randomness of sensing matrix to some extent, leading to deteriorated performance of CS-based MRI reconstruction.

Recently, noiselet encoding has been proposed in order to spread the energy of k-space data evenly across all phase encoding steps. Therefore, random sub-sampling can be applied to the data acquisition process. Two-dimensional (2D) image reconstruction results have shown that by adopting noiselet encoding and random sub-sampling, improved image quality and decreased relative error can be obtained.
Unfortunately, the question why this encoding method improves reconstruction results has not been well analyzed and understood.

In this thesis, the first coherence analysis of CS-based MRI reconstruction from sub-sampled data encoded by Fourier encoding and noiselet encoding is conducted. Coherence serves as an important measure of the quality of the sensing matrix, or from another perspective, the quality of the sparsifying dictionary in CS. Combinations of different encoding methods and sub-sampling techniques result in different coherence values. Generally speaking, a smaller coherence value leads to better reconstruction results. Simulation results show that the coherence value of noiselet encoding with random sub-sampling is smaller than that of Fourier encoding with variable density sub-sampling.

The relationship between coherence value and reconstruction quality is not linear. Phase transition analysis is therefore conducted in order to compare the quantitative performance of CS-based MRI reconstruction from sub-sampled data encoded with two encoding techniques. Phase transition results also show that the performance of CS-based MRI reconstruction from sub-sampled data encoded with noiselet encoding is better than that encoded with Fourier encoding.

Image reconstruction simulation is also conducted in order to compare the quality of CS-based MRI reconstruction from sub-sampled data encoded with two different encoding techniques. Data are assumed to be chosen from a slice of three-dimensional (3D) volume imaging. Both directions of data are phase encoding directions, so that sub-sampling can be applied in two directions. Phantom and MR images are used in the simulation. Reconstruction results show that images recovered from sub-sampled data encoded with noiselet encoding is of better quality.

Overall, the original contribution of this thesis is the first coherence analysis of CS-based MRI reconstruction from sub-sampled data encoded with Fourier encoding and noiselet encoding. This analysis provides understanding of the Fourier and non-Fourier encoding in MRI when using the CS technique to reconstruct images. Phase transition analysis shows the quantitative difference in reconstruction. Image reconstruction performance comparison are conducted when Fourier encoding and noiselet encoding are used in 3D volume imaging. The results of this thesis provide guidance on pulse sequence design for 3D volume imaging using novel encoding techniques.
I dedicate this thesis to my loving parents.
Acknowledgements

PhD study has been a long and very important part of my life. During this period, I have received help and encouragement from many people and institutions. I would like to use this part of my thesis to acknowledge them.

First of all, I would like to express my sincere thanks to Prof. Cishen Zhang. Without his guidance, I could not apply for and finish my PhD study. During my research, he has always been patient, supportive, and resourceful. His constant kindness and encouragement helped me a lot throughout this period.

Secondly, I would like to thank Prof. Jingxin Zhang and my co-supervisor Dr. Jinchuan Zheng. We had intensive discussions on research during my study. I thank them for their guidance and caring towards me.

I would also like to thank Swinburne University of Technology and National ICT Australia (NICTA) for providing scholarships and research facilities to me, so that I could enjoy a peaceful life and concentrate my time on research. Staff from these two institutions are always willing to help. I am grateful for their kind support and willing attitude.

I would like to thank my colleagues from Swinburne University of Technology and Monash University who helped and accompanied me during my study, especially Zhenzhen Chen, Wen Hou, Jiaxi He, Ifat-Al Baqee, Sulaiman A Al Hasani, and Kamlesh Pawar.

I would like to express my special thanks to Dr Alex McKnight, who helped by proofreading the final draft for grammar and style.

Finally and most importantly, I would like to thank my parents, who encouraged and supported me to obtain my PhD degree. I could not have achieved my degree without their love.
Declaration

I declare that this thesis contains no material that has been accepted for the award of any other degree or diploma and to the best of my knowledge contains no material previously published or written by another person except where due reference is made in the text of this thesis.

Kai Zhu
2017
# Table of contents

**List of figures**

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1</td>
<td>Motivation and Objectives</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>Approach</td>
<td>4</td>
</tr>
<tr>
<td>1.3</td>
<td>Main Contributions</td>
<td>5</td>
</tr>
<tr>
<td>1.4</td>
<td>Thesis Outline</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>Fundamentals of Compressive Sensing</td>
<td>9</td>
</tr>
<tr>
<td>2.1</td>
<td>Introduction</td>
<td>9</td>
</tr>
<tr>
<td>2.2</td>
<td>Nyquist-Shannon Sampling Theorem</td>
<td>10</td>
</tr>
<tr>
<td>2.3</td>
<td>Compressive Sensing</td>
<td>11</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Sparse Representation</td>
<td>13</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Sensing System</td>
<td>16</td>
</tr>
<tr>
<td>2.3.3</td>
<td>$\ell_0$ Model</td>
<td>17</td>
</tr>
<tr>
<td>2.3.4</td>
<td>$\ell_1$ Model</td>
<td>18</td>
</tr>
<tr>
<td>2.3.5</td>
<td>Reconstruction Methods</td>
<td>19</td>
</tr>
<tr>
<td>2.4</td>
<td>Restricted Isometry Property</td>
<td>21</td>
</tr>
<tr>
<td>2.5</td>
<td>Coherence</td>
<td>22</td>
</tr>
<tr>
<td>2.6</td>
<td>Phase Transition</td>
<td>25</td>
</tr>
<tr>
<td>2.7</td>
<td>Conclusion</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td>Fundamentals of MRI Reconstruction</td>
<td>29</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>29</td>
</tr>
<tr>
<td>3.2</td>
<td>Physics of MRI</td>
<td>30</td>
</tr>
<tr>
<td>3.2.1</td>
<td>NMR Phenomenon</td>
<td>30</td>
</tr>
<tr>
<td>3.2.2</td>
<td>NMR Signal Excitation</td>
<td>34</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Imaging Sequences</td>
<td>37</td>
</tr>
<tr>
<td>3.3</td>
<td>Reconstruction Methods of MRI</td>
<td>40</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Reconstruction Problem</td>
<td>40</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Non-Cartesian Trajectory</td>
<td>41</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Data Sub-sampling in the Cartesian Coordinates</td>
<td>43</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Novel Sequence Design</td>
<td>54</td>
</tr>
<tr>
<td>3.4</td>
<td>Conclusion</td>
<td>56</td>
</tr>
<tr>
<td>4</td>
<td>Coherence Analysis of MRI Encoding</td>
<td>57</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>57</td>
</tr>
<tr>
<td>4.2</td>
<td>Non-Fourier Encoding in MRI</td>
<td>58</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Sequence Design</td>
<td>58</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Fourier Encoding</td>
<td>61</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Noiselet Encoding</td>
<td>62</td>
</tr>
<tr>
<td>4.2.4</td>
<td>Image Reconstruction Problem</td>
<td>63</td>
</tr>
<tr>
<td>4.3</td>
<td>Coherence Analysis</td>
<td>66</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Coherence Expression</td>
<td>66</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Sub-sampling</td>
<td>69</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Sensitivity Profile</td>
<td>71</td>
</tr>
<tr>
<td>4.4</td>
<td>Results</td>
<td>71</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Sparse Signal in Wavelet Domain</td>
<td>71</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Sparse Signal in Spatial Domain</td>
<td>80</td>
</tr>
<tr>
<td>4.5</td>
<td>Conclusion</td>
<td>82</td>
</tr>
<tr>
<td>5</td>
<td>Phase Transition of MRI Encoding</td>
<td>83</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>83</td>
</tr>
<tr>
<td>5.2</td>
<td>Phase Transition Analysis</td>
<td>84</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Theory</td>
<td>84</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Analysis Arrangements</td>
<td>87</td>
</tr>
<tr>
<td>5.3</td>
<td>Results</td>
<td>89</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Sparse Signal in Wavelet Domain</td>
<td>89</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Sharp Transition Phenomenon</td>
<td>95</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Sparse Signal in Spatial Domain</td>
<td>96</td>
</tr>
<tr>
<td>5.4</td>
<td>Conclusion</td>
<td>97</td>
</tr>
<tr>
<td>6</td>
<td>Image Reconstruction Simulation</td>
<td>99</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>99</td>
</tr>
<tr>
<td>6.2</td>
<td>2D Sub-sampling in Volume Imaging</td>
<td>100</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Volume Imaging</td>
<td>100</td>
</tr>
<tr>
<td>6.2.2</td>
<td>2D Sub-sampling</td>
<td>102</td>
</tr>
<tr>
<td>6.3</td>
<td>Image Reconstruction</td>
<td>107</td>
</tr>
</tbody>
</table>
List of figures

2.1 Comparison between the original image and the reconstructed image.  
(a) The original image. (b) The transformed image in the wavelet domain. (c) The reconstructed image. (d) Scaled difference between the original and reconstructed images. 14

2.2 The transformed image is considered sparse since only a small portion of the magnitude of pixels are significantly large. (a) Sorted magnitude values of all pixels in the original image. (b) Sorted magnitude value of all pixels in the wavelet transformed image. 15

2.3 A linear system coincides with a sparse solution. 18

2.4 Example of phase transition curve. 27

3.1 Spins in different environments. (a) Spins at random orientations. (b) Spins placed in a magnetic field. 30

3.2 $M_z$ as a function of time $t$. 32

3.3 $M_{xy}$ as a function of time $t$. 34

3.4 FID as a function of time $t$. 35

3.5 Slice selection process. 35

3.6 Timing diagram of the GRE sequence. 38

3.7 Timing diagram of a multi-slice imaging sequence. 39

3.8 Timing diagram of a volume imaging sequence. 40

3.9 Trajectory of a pulse sequence in the Cartesian coordinates. 42

3.10 Radial trajectory of a pulse sequence in the non-Cartesian coordinates. 43

3.11 Comparison between the original image and the reconstructed image with aliasing. (a) The original image. (b) The reconstructed image with aliasing. 44

3.12 GRAPPA reconstruction method. 47

3.13 TGRAPPA reconstruction method. 49

3.14 k-t GRAPPA reconstruction method. 50

3.15 Comparison between frequency encoding line in GRE and BPE. 54
4.1 Waveform of a sinc function in the time domain. .......................... 58
4.2 Timing diagram of the spin echo sequence. ................................. 59
4.3 Timing diagram of a non-Fourier pulse sequence based on the spin echo sequence. ................................................................. 60
4.4 The principle of Fourier encoding. (a) Phantom image. (b) Fourier-encoded image. ................................................................. 61
4.5 The principle of noiselet encoding. (a) Phantom image. (b) Noiselet-encoded image. ................................................................. 63
4.6 Example of artificial sensitivity profiles when the number of coils is 4. 64
4.7 Example of sensitivity encoded images when the number of coils is 4. 65
4.8 Sub-sampling techniques. (a) Random sub-sampling. (b), (c) and (d) Variable density sub-sampling. ................................. 70
4.9 Slowly and gradually changing sensitivity profiles. $N = 256$. 72
4.10 Quickly and suddenly changing sensitivity profiles. $N = 256$. 72
4.11 Gram Matrix of new dictionary constructed by different matrices. (a) The full Fourier matrix and the wavelet sparsifying matrix. (b) The full noiselet matrix and the wavelet sparsifying matrix. 73
4.12 Coherence results of $\mu\{\hat{T}_F\}$ with random sub-sampling when $N = 256$. (a) Map 1. (b) Map 2. ................................. 74
4.13 Coherence results of $\mu\{\hat{T}_F\}$ with random sub-sampling when $N = 1024$. (a) Map 1. (b) Map 2. ................................. 75
4.14 Coherence results of $\mu\{\hat{T}_F\}$ with variable density sub-sampling when $N = 256$. (a) Map 1. (b) Map 2. ................................. 75
4.15 Coherence results of $\mu\{\hat{T}_F\}$ with variable density sub-sampling when $N = 1024$. (a) Map 1. (b) Map 2. ................................. 76
4.16 Coherence results of $\mu\{\hat{T}_N\}$ with random sub-sampling when $N = 256$. (a) Map 1. (b) Map 2. ................................. 77
4.17 Coherence results of $\mu\{\hat{T}_N\}$ with random sub-sampling when $N = 1024$. (a) Map 1. (b) Map 2. ................................. 78
4.18 Coherence comparison of Fourier encoding with variable density sub-sampling, and noiselet encoding with random sub-sampling. (a) $N = 256$. (b) $N = 1024$. ................................. 79
4.19 Comparison between Fourier- and noiselet-transformed signal. (a) A randomly-generated sparse signal with length $N = 256$ and sparsity $K = 16$. (b) The Fourier transform of the original signal. (c) The noiselet transform of the original signal. 80
4.20 Coherence comparison of Fourier encoding and random sub-sampling, and noiselet encoding and random sub-sampling. (a) \( N = 256 \). (b) \( N = 1024 \). .......................................................... 81

5.1 Phase transition curves of Fourier encoding and noiselet encoding when \( N_c = 1 \) and \( N = 256 \). .......................................................... 89

5.2 Phase transition curves of Fourier encoding and noiselet encoding when \( N_c = 4 \) and \( N = 256 \). (a) Map 1. (b) Map 2. ....................... 90

5.3 Phase transition curves of Fourier encoding and noiselet encoding when \( N_c = 8 \) and \( N = 256 \). (a) Map 1. (b) Map 2. ....................... 91

5.4 Phase transition curves of Fourier encoding and noiselet encoding when \( N_c = 16 \) and \( N = 256 \). (a) Map 1. (b) Map 2. ....................... 92

5.5 Phase transition curves of Fourier encoding when \( N_c = 1 \) and \( N = 1024 \). (a) Map 1. (b) Map 2. ................................. 93

5.6 Phase transition curves of noiselet encoding when \( N = 256 \). (a) Map 1. (b) Map 2. ................................. 93

5.7 Phase transition curves of Fourier encoding and noiselet encoding when \( N_c = 1 \) and \( N = 1024 \). ................................. 94

5.8 Successful reconstructions at different \( \delta \) and \( \rho \) when \( N_c = 1 \) and \( N = 1024 \). (a) Noiselet. (b) Fourier. ....................... 95

5.9 Phase transition curves of Fourier encoding and noiselet encoding when \( N_c = 1 \) and \( N = 256 \). The original signal is sparse in the spatial domain. ....................... 96

6.1 Comparison between multi-slice 2D imaging and 3D volume imaging. (a) Multi-slice 2D imaging. (b) 3D volume imaging. ............. 101

6.2 The principle of 2D noiselet encoding. (a) Phantom image. (b) 2D noiselet-encoded data. ................................. 103

6.3 Comparison between 1D and 2D sub-sampling. (a) 1D variable density sub-sampling. (b) 2D variable density sub-sampling. (c) 1D random sub-sampling. (d) 2D random sub-sampling. ............. 104

6.4 Reconstructed images from sub-sampled Fourier-encoded data. (a) 1D variable density sub-sampling and \( 1/R = 30\% \). (c) 2D variable density sub-sampling and \( 1/R = 30\% \). (e) 1D variable density sub-sampling and \( 1/R = 10\% \). (g) 2D variable density sub-sampling and \( 1/R = 10\% \). (b), (d), (f) and (h) are differences between the original image and the reconstructed images (a), (c), (e) and (g), respectively. 105
6.5 Reconstructed images from sub-sampled noiselet-encoded data. (a) 1D random sub-sampling and \(1/R = 30\%\). (c) 2D random sub-sampling and \(1/R = 30\%\). (e) 1D random sub-sampling and \(1/R = 10\%\). (g) 2D random sub-sampling and \(1/R = 10\%\). (b), (d), (f) and (h) are differences between the original image and the reconstructed images (a), (c), (e) and (g), respectively. ............................ 106

6.6 Comparison between reconstructed images. (a) Reference phantom image. (b) Reconstructed image with noiselet encoding. (c) Difference between (a) and (b). (d) Reconstructed image with Fourier encoding. (e) Difference between (a) and (d). ............................. 108

6.7 Comparison between reconstructed images. (a) Part of reference phantom image. (b) Part of reconstructed image with noiselet encoding. (c) Part of reconstructed image with Fourier encoding. .............. 109

6.8 NMSE of reconstructed phantom images. ......................... 109

6.9 Comparison between reconstructed images. (a) Reference brain image. (b) Reconstructed image with noiselet encoding. (c) Difference between (a) and (b). (d) Reconstructed image with Fourier encoding. (e) Difference between (a) and (d). ...................... 111

6.10 Comparison between reconstructed images. (a) Part of reference brain image. (b) Part of reconstructed image with noiselet encoding. (c) Part of reconstructed image with Fourier encoding. ............ 112

6.11 NMSE of reconstructed brain images. .......................... 112

6.12 NMSE of reconstructed phantom images from data with noise. (a) Fourier encoding. (b) Noiselet encoding. ..................... 113

6.13 NMSE of reconstructed brain images from data with noise. (a) Fourier encoding. (b) Noiselet encoding. ..................... 113
Symbols

$B$ radius of weak-$\ell_p$ ball
$B_0$ external magnetic field
$B_1$ magnetic field caused by the RF pulse
$C$ a constant
$\mathbb{C}$ complex set
$d$ negative gradient of objective function
$E$ energy
$\Delta E$ energy difference
$f$ frequency
$f(x)$ regularized function
$g$ gradient of objective function
$G$ Gram Matrix
$G_x$ $x$ component of magnetic gradient
$G_y$ $y$ component of magnetic gradient
$G_z$ $z$ component of magnetic gradient
$G_s$ slice selection gradient
$G_\Phi$ phase encoding gradient
$H$ hard thresholding operation
$h$ the Planck’s constant
$I$ identity matrix
$i$ index
$j$ index
$K$ sparsity of the original signal
$k$ Boltzmann constant
$\ell_0$ 0 norm
$\ell_1$ 1 norm
$\ell_2$ 2 norm
$\ell_p$ $p$ norm
$\ell_\infty$ infinity norm
$M$ length of the measured data
$M_0$ net magnetization
$M_x$ $x$ component of the net magnetization
$M_y$ $y$ component of the net magnetization
$M_{xy}$ transverse component of the net magnetization due to $T_2$
$M^*_{xy}$ transverse component of the net magnetization due to $T_2^*$
\( \mathbf{M}_{xy0} \) at the start of the dephasing process
\( \mathbf{M}_z \) component of the net magnetization
\( N \) length of the original signal
\( N_c \) number of coils
\( [N] \) a set \( \{1, 2, \ldots, N\} \)
\( \text{card}(N) \) cardinality of set \( [N] \)
\( P_D \) proton density
\( p \) a constant
\( r \) residual
\( R \) reduction factor
\( R_{\text{net}} \) net reduction factor
\( R_{\text{nom}} \) nominal reduction factor
\( \mathbb{R} \) real set
\( s \) transformed compressible signal
\( s_i(j) \) acquired black data points in coil \( i \) at location \( j \)
\( s_{4,\text{ACS}} \) the gray data in the ACS line of coil 4
\( s(k_x, k_y) \) acquired signal in Fourier domain
\( S \) NMR signal
\( t \) time
\( t_p \) a time duration
\( T_1 \) spin-lattice relaxation time
\( T_2 \) spin-spin relaxation time
\( T_{2,\text{inhomo}} \) relaxation time caused by inhomogeneity
\( T^* \) combined relaxation time
\( T_E \) echo time
\( T_R \) repetition time
\( \mathbf{T} \) measurement matrix
\( \mathbf{T}_F \) partial Fourier encoding operator
\( \mathbf{T}_F \) full Fourier encoding operator
\( \hat{\mathbf{T}}_F \) Fourier encoding operator with wavelet transform
\( \mathbf{T}_N \) partial noiselet encoding operator
\( \mathbf{T}_N \) full noiselet encoding operator
\( \hat{\mathbf{T}}_N \) noiselet encoding operator with wavelet transform
\( \mathbf{T}_s \) selected encoding technique with sub-sampling applied
\( \mathbf{T}_s \) whole sub-sampling and encoding process
\( \mathbf{U} \) transform matrix
\( \mathbf{W} \) wavelet transform operator
\( w_i(j) \) GRAPPA weights in coil \( i \) at location \( j \)
\( x \) a direction
\( x_i \) the \( i \)th entry of \( x \)
\( \hat{x}_1 \) overlapped pixel
\( \hat{x}_2 \) overlapped pixel
\( x(1) \) pixel in a image
\( x(2) \) pixel in a image
\( \hat{x} \) reconstructed signal or image
\( y \) a direction
\( y \) measured data
\( y_i \) sub-sampled data from coil \( i \)
\( z \) a direction
$\alpha$ a constant
$\beta$ a constant
$\Gamma_1(1), \Gamma_1(2)$ sensitivity in the first coil
$\Gamma_2(1), \Gamma_2(2)$ sensitivity in the second coil
$\Gamma_i$ sensitivity profile of coil $i$
$\hat{\Gamma}_i$ transformed $\Gamma_i$
$\gamma$ the gyromagnetic ratio
$\delta$ ratio of $M$ to $N$
$\delta_K, \delta_{2K}$ RIP constant
$\epsilon$ parameters with relationship to noise
$\theta$ degree of RF pulse
$\Lambda$ support set
$\lambda$ regularization parameter
$\mu$ coherence value
$\nu$ frequency of proton
$\nu_0$ Larmor frequency
$\xi$ positive smoothing parameter
$\pi$ a constant
$\rho$ ratio of $K$ to $M$
$\rho(x, y)$ spin density at location $(x, y)$
$\sigma$ stop criterion
$\sigma_{\text{min}}$ minimum singular values
$\sigma_{\text{max}}$ maximum singular values
$\tau$ parameters with relationship to noise
$\Phi$ sensing matrix
$\hat{\Phi}$ a square sensing matrix
$\phi_i$ $i^{\text{th}}$ row or column of sensing matrix $\Phi$
$\Phi_K$ a matrix consisting of $K$ columns from $\Phi$
$\Phi_{\text{sub}}(K)$ sub-matrix formed from $K$ distinct columns of $\Phi$
$\chi(x)$ mother bases function of noiselets
$\Psi$ sparsifying matrix
$\psi_i$ $i^{\text{th}}$ row or column of sparsifying matrix $\Psi$
$\psi_i(x)$ $i^{\text{th}}$ entry of transformed vector $x$ under $\Psi$
$\hat{\Psi}$ $\Psi$ with each column normalized
$\hat{\psi}_i$ $i^{\text{th}}$ column of $\hat{\Psi}$
$\omega$ frequency of sine wave
$\emptyset$ empty set
$|\cdot|$ magnitude
$\|\cdot\|_0$ number of non-zero elements
$\|\cdot\|_1$ sum of the magnitude of all elements
$\|\cdot\|_2$ sum of the square of all elements
$\|\cdot\|_p$ $p^{\text{th}}$ root of sum of $p^{\text{th}}$ power of the magnitude of all elements
$\|\cdot\|_\infty$ maximum magnitude of a vector
$<\cdot, \cdot>$ inner product
$\infty$ infinity
$\odot$ element-wise multiplication
Abbreviations

1D 1-dimensional
2D 2-dimensional
3D 3-dimensional
ACS additional calibration signal
BP basis pursuit
BPE bunched phase encoding
CS compressive sensing
CT computational tomography
FID free induction decay
GRAPPA GeneRalized Autocalibrating Partially Parallel Acquisitions
GRE gradient-echo
GROG GRAPPA operator gridding
MHz millions of Hz
MR magnetic resonance
MRI magnetic resonance imaging
NMR nuclear magnetic resonance
NMSE normalized mean square error
NP non-deterministic polynomial-time
OECD Organization for Economic Cooperation and Development
OMP orthogonal matching pursuit
RF radio frequency
RIP restricted isometry property
SENSE SENSitivity Encoding
SNR signal-to-noise ratio
SOS sum of squares
Chapter 1

Introduction

1.1 Motivation and Objectives

Magnetic Resonance Imaging (MRI) [1] is a widely-used imaging technique in contemporary medical diagnosis. It is based on the nuclear magnetic resonance (NMR) phenomenon, which was first observed by Felix Bloch [2] and Edward Purcell [3] in 1946. They were awarded the Nobel Prize in 1952 for their contribution to discovering the NMR phenomenon.

MRI is a popular medical imaging technique due to its excellent resolution in tissue depiction, flexible slice orientation, and freedom from harmful radiation. According to Statista (https://www.statista.com/), a website which provides statistics results, there were 15.18 MRI units per million inhabitants in Australia in 2014, and 46.87 units per million inhabitants in Japan [4]. The number of examinations performed was 27.6 per 1000 population in Australia in 2013, and 51.7 on average in Organization for Economic Cooperation and Development (OECD) countries [5]. The average price of a diagnostic MRI was 215 U.S. dollars in Australia in 2014, and 1,119 U.S. dollars in the United States [6]. The price of an MRI unit may be up to millions of dollars [7]. Generally speaking, MRI is an expensive medical imaging technique, but it is widely used in medical diagnosis in Australia and worldwide.

One major problem of MRI is its relatively long data acquisition time, which may take up to tens of minutes or even hours. During the data acquisition process, patients are asked to stay still in the bore of the scanner. Although some equipment is used to stop patients from moving, this task still seems very difficult for healthy individuals, let alone senior people, young children and patients in great pain. Images will be corrupted with aliasing and artifacts if movements happen during the data acquisition process. Deteriorated image quality may greatly hamper the diagnosis of medical practitioners.
Research on data acquisition time reduction has been an active area for decades. By reducing the data acquisition time, it is anticipated that more scans can be finished in one unit of time period, leading to improved efficiency of scanners and decreased cost per scan. In addition, a short data acquisition time reduces the possibility of movement during scanning, leading to a higher success rate in acquiring a clear image. For applications like cardiac imaging, in which tissues or organs of interest move periodically, a short data acquisition time may determine the feasibility of application of MRI.

In traditional MRI, images of interest are encoded with Fourier encoding, so that the acquired data are stored in the Fourier domain, or k-space. Normally, k-space data are acquired row by row following a Cartesian trajectory. The data acquisition process means applying the 2-dimensional (2D) Fourier transform on the image of interest. In the data acquisition process, a row of data are collected in one phase encoding step within a very short time period, while the time between two consecutive phase encoding steps is much longer. The image of interest can be simply reconstructed by applying the 2D inverse Fourier transform on the acquired data. The image reconstruction process is completed in scanners within a very short time period, since the Fourier transform is built in scanners and has fast algorithms.

The reduction of data acquisition time is commonly achieved by sub-sampling the k-space data. In other words, some phase encoding steps are acquired while the remaining phase encoding steps are skipped during the data acquisition process. Data in one phase encoding step are always sampled or not sampled. Sub-sampling can only be applied in the phase encoding direction by skipping certain phase encoding steps. The reduction factor $R$ is defined as the ratio of the number of total phase encoding steps in theory to the number of scanned phase encoding steps. As a result, images of interest cannot be reconstructed by simply applying 2D inverse Fourier transform on the sub-sampled k-space data. Otherwise, reconstructed images will be corrupted by artifacts and aliasing. Specially designed reconstruction methods are required to reconstruct images from sub-sampled k-space data.

Many reconstruction methods have been proposed to reconstruct images from sub-sampled k-space data. For example, the keyhole method [8] was proposed to explore the redundancy in the temporal dimension in dynamic MRI, and the partial Fourier reconstruction method [9] was proposed to explore the conjugate symmetry property of k-space data. These methods were originally designed for preliminary MRI scanners with only one receiving coil.

In recent decades, phased array of coils are implemented in modern MRI scanners, providing the ability to simultaneously acquire data in each coil. Each coil in the
array has a distinctive sensitivity profile, which can be used as additional information in image reconstruction. The keyhole and partial Fourier reconstruction methods have been successfully extended to MRI scanners equipped with multiple receiving coils [10, 11].

Most recent image reconstruction methods exploit the additional information provided by the multiple-coil system. Currently, two major reconstruction methods built in MRI scanners are SENSitivity Encoding (SENSE) [12] and GeneRalized Autocalibrating Partially Parallel Acquisitions (GRAPPA) [13]. In SENSE, the sensitivity profiles of coils are required to be known as additional information. Normally, the sensitivity profiles are estimated before or during the actual scanning. Unfortunately, the estimated sensitivity profiles may be different from those in actual scanning due to movement, or in some applications, such as cardiac imaging, the sensitivity profiles are difficult to estimate. In contrast, in GRAPPA, the sensitivity profiles are not required to be known before reconstruction. Instead, additional calibration signal (ACS) lines are scanned in the low frequency part of k-space in order to interpolate the GRAPPA kernel. The ACS lines are extra phase encoding steps in the low frequency part of k-space. As a result, the net reduction factor $R_{net}$ is smaller than the nominal reduction factor $R_{nom}$, since more phase encoding steps are scanned than usual. Furthermore, when $R$ increases, the quality of GRAPPA reconstruction results decreases quickly.

A newly emerged and promising approach to reconstruct images from sub-sampled k-space data is by using compressive sensing (CS) [14, 15]. In CS theory, a sparse signal can be recovered from many fewer randomly-selected measurements than the length of the original signal by solving a convex optimization problem. Two prerequisites of CS theory are that (1) the original signal is sparse, either naturally or after certain kind of transformation, and (2) the signal is acquired by a random sensing system. To guarantee successful reconstruction, the random sensing matrix must satisfy certain rules, such as the restricted isometry property (RIP) [14, 16, 17], and coherence [18, 19].

CS has been successfully applied to the reconstruction problem of MRI from sub-sampled k-space data [20]. Magnetic resonance (MR) images are commonly assumed to be sparse after wavelet transform. Data acquired in traditional MRI scanners are encoded with Fourier encoding. As a result, most of the energy of data is stored in the low frequency part, or the center of k-space. The low frequency part of k-space controls the main structure and brightness of the image of interest, while the high frequency part determines the details. Data in different phase encoding steps carry different importances in k-space. In order to keep the main structure
and brightness of the image, more phase encoding steps should be scanned in the low frequency part of k-space, while fewer should be sampled in the high frequency part of k-space. This kind of sub-sampling is called variable density sub-sampling, since the sampling density varies according to location. The consequence of variable density sub-sampling is that the details of image are sacrificed in the data acquisition process. In addition, variable density sub-sampling violates the randomness prerequisite of the sensing matrix to some extent, leading to suboptimal reconstruction performance.

To fulfill random sub-sampling, data in each phase encoding step should carry the same importance in k-space. Noiselet encoding was proposed recently to spread the energy of k-space data evenly across all phase encoding steps [21]. The energy of noiselet transformed data is naturally spread evenly in all phase encoding steps. In the MRI literature, the k-space is referred to as the Fourier domain, since data are encoded by Fourier encoding in traditional MRI scanners. In this thesis, the k-space is referred to as the transformed domain. It is the Fourier domain if Fourier encoding is applied. Similarly, it is the noiselet domain if noiselet encoding is applied.

Since the energy of k-space data is spread evenly in k-space when noiselet encoding is applied, random sub-sampling can be chosen to collect data during the data acquisition process. It was demonstrated in [21] that images reconstructed from sub-sampled data encoded with noiselet encoding have better quality than those reconstructed from sub-sampled data encoded with traditional Fourier encoding in 2D imaging.

Unfortunately, the question why such encoding methods improve reconstruction results has not been well analyzed and understood.

1.2 Approach

In this thesis, the difference between Fourier encoding and noiselet encoding is analyzed in the following three aspects: coherence, phase transition and image reconstruction.

Coherence serves as an important quality measure of the sensing matrix, or from another perspective, the quality measure of the sparsifying dictionary of CS. It can be defined as the maximum inner product of two different columns of the normalized sparsifying dictionary. Combinations of different encoding and sub-sampling techniques lead to different sparsifying dictionaries and coherence values. In CS, the smaller the coherence value, the better. By comparing the coherence values of the corresponding encoding and sub-sampling techniques, the performance
of CS reconstruction can be analyzed. Different sensitivity profiles, sub-sampling patterns, and coil numbers are used in this simulation in order to cover as many cases as possible in real scanning.

The relationship between coherence value and reconstruction quality is not linear. In order to quantitatively compare the reconstruction performance, phase encoding analysis is further studied. Let \( \delta = M/N \), and \( \rho = K/M \), where \( K, M, \) and \( N \) stand for the sparsity (number of non-zero or significant entries) of the original signal, the length of measured data and the length of the original signal, respectively. Phase transition states that in the plane of \((\delta, \rho)\), there is a curve which divides the plane into two phases. Below the curve the original signal can be successfully reconstructed, while above the curve the original signal cannot be successfully reconstructed, both with an overwhelming probability. Combinations of different encoding and sub-sampling techniques lead to different reconstruction results and phase transition curve positions. By comparing the positions of the phase transition curves of corresponding encoding and sub-sampling techniques, the quantitative performance of CS reconstruction can be analyzed. As with coherence analysis, different sensitivity profiles, sub-sampling patterns, and coil numbers are used in this simulation in order to cover as many cases as possible in real scanning.

Image reconstruction from sub-sampled 3-dimensional (3D) volume imaging data encoded with two different encoding methods is also simulated. Images are assumed to be chosen from a slice of 3D data, of which both directions are phase encoding directions, so that sub-sampling can be applied in two directions. Phantom and MR images are used in the simulation. The number of coils and the reduction factor vary, in order to demonstrate the performance of CS-based reconstruction in different situations.

### 1.3 Main Contributions

The main contributions of this thesis are summarized as follows.

- In order to analyze the difference between Fourier encoding and noiselet encoding in CS theory, the first coherence analysis is conducted in this thesis. Simulation results show that the coherence value of partial noiselet encoding sub-sampled with random sub-sampling is smaller than that of partial Fourier encoding sub-sampled with variable density sub-sampling at the same \( R \). This part of the content can be found in Chapter 4.
• Phase transition analysis is further conducted in order to compare the quantitative performance of CS-based MRI reconstruction from sub-sampled data encoded by two different encoding techniques. Phase transition results also show that the performance of CS-based MRI reconstruction from sub-sampled data encoded with partial noiselet encoding is better than that encoded with partial Fourier encoding. This part of the content can be found in Chapter 5.

• Image reconstruction simulation is finally conducted in order to compare the reconstruction difference from sub-sampled data of simulated 3D volume imaging using two different encoding techniques. Reconstruction results show that images recovered from data encoded with partial noiselet encoding are of better quality. The results of this thesis provide guidance for pulse sequence design for 3D volume imaging using novel encoding techniques. This part of content can be found in Chapter 6.

• Although the main focus of this research is MR images which are sparse in the wavelet domain, signals which are sparse in the spatial domain are also considered in order to make a comprehensive comparison. Similar coherence and phase transition analyses are conducted in order to compare the two encoding techniques. This part of the content can be found in Chapter 4 and 5.

1.4 Thesis Outline

This thesis studies the coherence analysis of CS-based MRI reconstruction from sub-sampled k-space data encoded with Fourier encoding and noiselet encoding. In order to quantitatively compare the reconstruction performance, phase encoding analysis is further studied. Finally, image reconstruction from sub-sampled 3D volume imaging data encoded with two different encoding methods is simulated. This thesis is organized as follows.

• In Chapter 1, first the motivation and objectives of this thesis are described. A general background of the research is provided to emphasize the importance of this research. The main contributions of this thesis are then summarized. Finally, the thesis outline is given.

• In Chapter 2, the fundamentals of CS are discussed. First, the Nyquist-Shannon sampling theorem is reviewed to introduce the limitation of conventional sampling theorem and the necessity of CS. The theory of CS is then described,
including the sparse representation of signals, the sensing system, \( \ell_0 \) and \( \ell_1 \) models, and basic reconstruction methods. To guarantee successful reconstruction, the RIP and coherence of CS are further introduced. Phase transition is then described to show the performance of CS reconstruction. Finally, applications of CS are briefly reviewed.

- In Chapter 3, the fundamentals of MRI reconstruction are summarized. First, the physics of MRI is reviewed, including the NMR phenomenon, NMR signal excitation and the MR imaging sequences. Second, the image reconstruction problem for sub-sampled k-space data is illustrated, and some typical methods to reduce data acquisition time are discussed, such as following non-Cartesian trajectories, sub-sampling k-space data in Cartesian coordinates, and designing novel sequences.

- In Chapter 4, the coherence analysis results of two different encoding techniques are shown. First, the basics of non-Fourier encoding in MRI are introduced. In this section, the sequence design for non-Fourier encoding based on spin echo sequence is described. The theory of noiselet encoding and its properties are illustrated. Second, the coherence analysis arrangements are introduced. In this section, the coherence expression of the CS-based MRI reconstruction problem is given, and the sub-sampling pattern and sensitivity profiles used in the simulation are illustrated. Thirdly, the coherence results are provided to demonstrate the difference in reconstruction results from sub-sampled data encoded by Fourier encoding and noiselet encoding. Finally, the conclusion of this chapter is presented.

- In Chapter 5, the phase transition analysis results are given. First, the phase transition analysis for the CS-based MRI reconstruction problem is described. Second, the phase transition analysis arrangements are introduced. Then, the phase transition results are shown to demonstrate the difference between reconstruction results from sub-sampled data encoded with Fourier encoding and noiselet encoding.

- In Chapter 6, image reconstruction simulation is conducted to compare the quality difference of reconstructed images from sub-sampled data encoded with two different encoding techniques. First, volume imaging by Fourier encoding and noiselet encoding are introduced. Second, the image reconstruction problem from sub-sampled data of volume imaging is described. Thirdly, both phantom and MR images are used to test the performance of two different encoding techniques.
• In Chapter 7, the conclusion and discussion of this thesis are highlighted. Recommendations for future research are listed.
Chapter 2

Fundamentals of Compressive Sensing

2.1 Introduction

Sampling is the process of converting a continuous signal to its discrete form, so that it can be processed by modern digital computing devices. The Nyquist-Shannon Sampling Theorem has served as the foundation of modern sampling technologies for decades. It requires that the sampling rate is no less than twice the bandwidth of the continuous signal. In applications, discrete signals are compressed without sacrificing much of their fidelity in order to save space. In sum, a large amount of data is sampled while most are discarded, which creates a paradox.

In this chapter, first, the Nyquist-Shannon Sampling Theorem and its advantages and disadvantages are discussed. Second, the theory of CS is introduced. CS was proposed by Candès, Romberg, Tao [14] and by Donoho [15] in 2006 to tackle this paradox. It has drawn considerable attention in many areas, such as applied mathematics, signal processing, and computer science. Many signals can be represented by using a few non-zero coefficients in a suitable basis or dictionary. In CS, a sparse signal can be recovered from many fewer randomly-selected measurements than the length of the original signal. The CS theory, including sparse representation, the sensing system, $\ell_0$ and $\ell_1$ models, and reconstruction methods, is presented. Restricted Isometry Property (RIP), coherence and phase transition, which are used to determine the performance of CS reconstruction, are also discussed.
2.2 Nyquist-Shannon Sampling Theorem

Signals generated in natural environments, such as sounds, images, and videos, are continuous in the time or spatial domains. These continuous signals need to be sampled and recorded in discrete form in order to be analyzed and stored in modern digital systems. The discrete signals must contain all the information of the continuous signals. That is, continuous signals can be reconstructed from discrete signals with minimal error.

Sampling is the process of converting a continuous signal to its discrete form. Suppose a continuous signal is of finite bandwidth. It can be recovered exactly from a series of equally-spaced discrete samples taken at a frequency no less than twice the bandwidth of the continuous signal. This is called the Nyquist-Shannon Sampling Theorem [22, 23]. This sampling theorem acts as a fundamental bridge between analog signals and digital signals. It provides a sufficient condition that a discrete signal collected above a sampling rate contains all the information of the continuous signal. However, it does not mean that perfect reconstruction does not exist from data sampled below the Nyquist rate. In contrast, perfect reconstruction does exist under some constraints, which is the foundation stone of compressive sensing.

The Nyquist-Shannon Sampling Theorem established the basis of modern digital signal processing theories. Continuous signals can be sampled by obeying the sampling theorem and transformed to discrete signals, which then are digitized into digital signals. Digital signals can be easily processed and stored by digital devices, and transferred between devices at high speed. Digital devices are more robust, more flexible, as well as cheaper and more widely used than their analog counterparts.

Although digital signal processing offers dramatic advantages over continuous signal processing, it also has several problems. The first problem is that the sampling frequency tends to become higher and higher with the unceasing requirement of better fidelity. The amount of data has grown from a small trickle to a torrent. For example, in MRI, the Larmor frequency is normally up to the scale of millions of Hz (MHz). The minimum sampling frequency may be even higher in other applications. Rapidly increasing sampling frequencies require faster analog-digital converters, more powerful processors, transmission media with wider bandwidth, and storage products with higher capacity. For example, the frequency of modern CPUs is easily up to GHz. To increase efficiency, 10nm technology has been announced to shrink the size of CPUs, which is approaching the limit of physics [24]. These requirements challenge the development of contemporary technologies and tightened financial budgets, since better devices commonly mean higher prices.
Another problem of digital signal processing is that not all sampled data need to be stored. In other words, measured data can be compressed without sacrificing much information, or some data are redundant. Compression techniques aim to find the most concise representation of a signal which is capable of achieving a target level of acceptable distortion. One popular solution of compression is by transforming signals to another domain, in which the transformed signal is sparse or compressible. Therefore, a signal with considerable length can be represented by only a few non-zero coefficients. For example, instead of recording a sine wave with infinite length, it is better to store its amplitude, frequency and phase. For some images, it is known that they have only a few elements with significant values, while the rest can be ignored after certain kinds of transform, such as Fourier transform and wavelet transform. Therefore, instead of recording images with high dimensions, it is better to store the significant values and their locations and the transforming contract. Many lossy compression techniques for images, sounds or videos, such as JPEG, JPEG2000, MPEG, and MP3, comply with this principle.

In sum, when the Nyquist-Shannon Sampling Theorem is used, a large amount of data need to be sampled but most are discarded, which creates paradox.

### 2.3 Compressive Sensing

The paradox posed by the Nyquist-Shannon Sampling Theorem and compression technology leads to a simple idea: is it possible to directly measure compressed data and reconstruct the original signal? Since a small portion of the transformed data are capable of representing the original signal, we are interested in directly measuring data in the compressed form, instead of first sampling at high frequency and then compressing the sampled data.

The newly-emerged compressive sensing (or compressed sensing, compressive sampling) breaks the twice-bandwidth boundary set by the Nyquist-Shannon Sampling Theorem. By applying CS, data can be obtained directly in compressed form, and perfect reconstruction can be achieved from many fewer measurements than usually required by the Nyquist-Shannon Sampling Theorem. CS enables a potentially large reduction in the sampling and computation costs for sensing signals which have a sparse or compressible representation.

CS was proposed by Candès, Romberg, Tao [14] and by Donoho [15] in 2006. These researchers assert that signals which are sparse or compressible can be reconstructed with an overwhelming probability from many fewer measurements than the length of the original signal. The number of non-zero or significant elements
of the sparse signal is called the sparsity, normally denoted as $K$. It is claimed that
the length of measurements is proportional to the sparsity of the original signal. The
design of sensing schemes and their extensions to real world data acquisition systems
are central challenges in CS.

The CS theory is different from the Nyquist-Shannon Sampling Theorem in
several aspects. First, CS mainly focuses on sensing and recovering discrete signals
with finite length, while the Nyquist-Shannon Sampling Theorem typically considers
continuous signals with infinite length. In addition, CS measures data in the form
of inner products between the original signal and a sensing matrix, while in the
Nyquist-Shannon Sampling Theorem the continuous signal is sampled at equally-
spaced time points. In CS, the randomness of the sensing matrix plays a vital role,
which will be analyzed in this thesis. Finally, the two frameworks differ in how the
original signal is reconstructed. In CS, signal reconstruction is achieved by solving
nonlinear convex optimization problems, which may be hardware-demanding and
time-consuming. In the Nyquist-Shannon framework, signal recovery is achieved
through sinc interpolation, which is a linear process that requires little computation
and has a simple interpretation. Although the reconstruction process of the original
signal by CS appears to be a problem, it is not a major concern in CS, since the aim
of CS focuses on acquiring less data.

Two general requirements of CS are that first the signal is sparse or compressible,
either naturally or after certain kinds of transformation, and second the sensing matrix
obeys certain principles. Many signals are either naturally sparse, or compressible
after certain kinds of transformation, for example, images and videos. The sensing
matrix must obey the restricted isometry property (RIP), initially named the uniform
uncertainty principle. Generally speaking, RIP requires that the sensing matrix
should be incoherent with the sparsifying matrix. Otherwise, reconstruction from
measured data will fail. It has been reported that Gaussian, Bernoulli, and partial
random Fourier matrices obey RIP, and randomness is exploited to show optimal
or near-optimal conditions on the number of measurements in terms of sparsity
[25–27]. Unfortunately, RIP is non-deterministic polynomial-time (NP) hard.

Another important measure of quality of the sensing matrix, or from another
perspective, the quality of the sparsifying matrix, is coherence. It can be defined as
the maximum inner product of two different columns of the normalized sparsifying
dictionary. In CS, the smaller the coherence value, the better.

Furthermore, in CS, by approaching the problem via polytope geometry, or more
precisely the notion of $K$-neighbourliness, Donoho observed sharp phase transition
curves for the random Gaussian matrix, separating regions where recovery fails or succeeds with high probability [28–30].

CS has been successfully applied in many areas. In this thesis, we mainly focus on the application of CS in MRI [20]. The application of CS has been extended to other areas, such as imaging [31, 32], A/D conversion [33], radar [34], wireless communication [35, 36], and compressive sensor networks [37, 38].

2.3.1 Sparse Representation

A signal is sparse if only a few entries are non-zero, while the rest are zero. The number of non-zero entries is named the sparsity, often denoted as $K$ [14, 15]. Let $[N]$ denote a set $\{1, 2, \cdots, N\}$, and card$(N)$ denote the cardinality of set $[N]$. The support of a signal vector $x \in \mathbb{C}^N$ is the set consisting of the non-zero entries. That is,

$$\text{supp}(x) := \{i \in [N] : x_i \neq 0\}, \quad (2.1)$$

where $x_i$ is the $i^{\text{th}}$ entry of $x$. The vector $x$ is $K$-sparse if at most $K$ of its entries are non-zero. That is,

$$||x||_0 = \text{card}(\text{supp}(x)) \leq K, \quad (2.2)$$

where $|| \cdot ||_0$ denotes the $\ell_0$ norm. $|| \cdot ||_p$ is called the $p$-norm, which is defined as

$$||x||_p := \left( \sum_{i=1}^{N} |x_i|^p \right)^{1/p}, 1 \leq p < \infty. \quad (2.3)$$

Specifically, when $p = 0$,

$$||x||_0 := \text{card}(\text{supp}(x)), \quad (2.4)$$

which denotes the number of non-zero elements. When $p = 1$,

$$||x||_1 = \sum_{i=1}^{N} |x_i|, \quad (2.5)$$

which means that $\ell_1$ norm is the sum of the magnitude of all entries of vector $x$. When $p = \infty$,

$$||x||_{\infty} := \max \{|x_i|, i = 1, \cdots, N\}, \quad (2.6)$$

which denotes the maximum magnitude of vector $x$.

Not all signals are naturally sparse. However, a wide range of signals can be considered approximately sparse, or compressible, after certain kinds of transformation.
The sorted entries of transformed signals take the form of a rapid decay, typically like a power law. Only a small number of entries are significantly big, while the rest can be ignored. For example, in Fig. 2.1 (a), an $256 \times 256$ image of a cameraman is shown. This image contains many details, such as a person, buildings, a camera, and tripod. The sorted magnitude of all pixels of the original image is shown in Fig. 2.2 (a). It is easily observed that about 80% of the pixels have a value between 0.1 to 0.8, and almost no pixels have a value of zero. In other words, this image is not naturally sparse.

![Image comparison](image_url)

Fig. 2.1 Comparison between the original image and the reconstructed image. (a) The original image. (b) The transformed image in the wavelet domain. (c) The reconstructed image. (d) Scaled difference between the original and reconstructed images.

The 2D wavelet transform was applied on the original image, and the transformed image is shown in Fig. 2.1 (b). It is seen that only a small fraction of the pixels are bright, or non-zero, while the rest are zero. The sorted magnitude of all pixels of the
transformed image is shown in Fig. 2.2 (b). Only a very small portion of the data are significantly large, while the others are almost zero. Choosing a threshold of 0.1 and setting those below the threshold as zero, about 7,000 are kept unchanged, which is about 10%. The truncated transformed image is considered sparse. If the location and value of the significant entries are stored, the space needed is about 20% of that of the original image.

Fig. 2.2 The transformed image is considered sparse since only a small portion of the magnitude of pixels are significantly large. (a) Sorted magnitude values of all pixels in the original image. (b) Sorted magnitude value of all pixels in the wavelet transformed image.

The reconstructed image from the truncated transformed data is shown in Fig. 2.1 (c), and the scaled difference between the original and the reconstructed images is shown in Fig. 2.1 (d). The reconstructed image also keeps rich details as in the original image. The difference between the original and the reconstructed images is so small that it is almost indistinguishable by the naked eye if not scaled. This shows that 10% of significant entries of the transformed image can reconstruct the image with unnoticeable differences.

This process can be expressed using mathematical language. Let \( \Psi \in \mathbb{C}^{N \times N} \) denote a transform matrix with orthonormal rows \( \psi_i \). The entries of transformed vector \( x \) are denoted as \( \psi_i x \), which can be arranged in a descending order according to magnitude, i.e., \( |\psi_{(1)} x| > |\psi_{(2)} x| > \cdots > |\psi_{(N)} x| \). \( |\psi_{(i)} x| \) obeys a weak-\( \ell_p \) ball of
radius $B$ for $0 < p < \infty$ if

$$|\psi(i)x| \leq B \cdot i^{-\frac{1}{p}}, i = 1, 2, \cdots, N.$$  \hspace{1cm} (2.7)

The parameter $p$ determines the descending speed of decay. It is easily appreciated that the smaller $p$, the faster the decay, and the smaller the sparsity $K$.

By recording the values and locations of the significant entries, only $2K$ instead of $N$ data points are needed to reconstruct the original image with acceptable sacrifice. In the sparse or compressible case, $2K \ll N$. If the values and locations of the significant entries are easily obtained, the reconstruction process is of trivial difficulty. Unfortunately, determination of the values and locations of the significant entries is not trivial. An extra sensing step is required in order to collect enough data for reconstruction.

### 2.3.2 Sensing System

In CS, instead of sampling the original signal directly, data are measured by a sensing matrix $\Phi \in \mathbb{C}^{M \times N}$ with $M \ll N$ [14, 15]. Let $y \in \mathbb{C}^M$ denote the measured data. The data acquisition process is expressed as

$$y = \Phi x. \hspace{1cm} (2.8)$$

Since $M \ll N$, there are more unknown variables than equations. Eq. 2.8 is under-determined, which means it cannot be directly reconstructed by inverse process.

Suppose $x$ is sparse after a certain kind of orthonormal transform, $\Psi$. Let $s \in \mathbb{C}^N$ denote the transformed compressible signal, which is given as

$$s = \Psi^T x, \text{ or } x = \Psi s. \hspace{1cm} (2.9)$$

The data acquisition process is then expressed as

$$y = \Phi x = \Phi \Psi s. \hspace{1cm} (2.10)$$

The matrix $\Phi$ works as dimension reduction, i.e., it maps a signal with length $N$ to measurements with length $M$, where $M \ll N$. CS aims to recover the original signal with length $N$ from many fewer measurements with length $M$. In addition, it is generally assumed that the sensing matrix is non-adaptive, which means that the matrix $\Phi$ is fixed and does not change depending on the previously acquired measurements. In some applications, adaptive sensing may lead to significant
performance improvements. In this thesis, we focus only on the non-adaptive sensing matrix.

The data acquisition process in CS is accomplished by introducing a sensing matrix. There are two intuitive but important theoretical questions in CS. The first one is how to design the sensing matrix so that the measured data preserve all the information in the original signal. A random sensing matrix is commonly used in CS, and randomness plays a vital role in the construction of the sensing matrix. In the following chapters, the importance of the randomness of the sensing matrix in the image reconstruction problem of MRI by CS is thoroughly explored. The benefits of the random sensing matrix are demonstrated by the image reconstruction results. The second question is how the original signal can be reconstructed from many fewer measurements. This is accomplished by solving a nonlinear convex optimization problem. We focus on the second question first.

### 2.3.3 $\ell_0$ Model

An intuitive idea to reconstruct the sparse signal $x$ is based on the principle that if two signals $x_1$ and $x_2$ with sparsity $K$, and $y = \Phi x_1 = \Phi x_2$, then $x_1 = x_2$. In other words, there is only one unique sparse signal which corresponds to $y = \Phi x$. Therefore, to reconstruct the sparse signal $x$ is to find its non-zero entries, which is expressed as

$$
\min ||x||_0 \quad \text{s.t.:} \quad y = \Phi x, \quad (2.11)
$$

where $|| \cdot ||_0$ denotes the number of non-zero entries. If the locations and values of the non-zero entries are determined, the reconstruction problem can be easily solved.

Unfortunately, the $\ell_0$ minimization problem is NP hard [39, 40] and therefore computationally intractable. A polynomial-time algorithm can be performed in a number of steps bounded by a polynomial expression in the size of its input. It is illustrated in more detail as follows. Since signal $x$ has sparsity $K$, it is straightforward to consider all possible subsets of $[N]$ consisting of $K$ entries. As a result, it is possible to solve every rectangular system $y = \Phi x$, or its equivalent square system $\Phi_K^* y = \Phi_K^* \Phi_K x$, where $\Phi_K$ denotes a matrix consisting of all possible combinations of $K$ columns from $\Phi$. To select $K$ out of $N$ is a simple problem in permutation and combination, which has $\binom{N}{K}$ possible combinations. Suppose $N = 1000$, and $K = 10$, there are $\binom{1000}{10}$, which is $\binom{1000}{10} = 10^{20}$ possible combinations. Since $K = 10$, the minimum $M$ is also 10 to have just enough equations for unknown variables. If one linear system with size $10 \times 10$ can be solved within $10^{-10}$ seconds, the total time needed to iterate through all possible combinations would be $10^{10}$ seconds, or more than
300 years. As a result, the $\ell_0$ minimization problem is NP hard. It cannot be solved within acceptable polynomial-time.

Furthermore, the $\ell_0$ minimization problem 2.11 is a non-convex optimization problem. Even in the very lucky case when the right $\Phi_K$ is found at the beginning of trial and error, the search for global minimization may be trapped in local minimum.

### 2.3.4 $\ell_1$ Model

The $\ell_0$ minimization problem 2.11 is not a tractable problem. Generally, $\ell_1$ norm instead of $\ell_0$ norm is used in CS since it is more tractable. The mathematical model is then expressed as

$$\min ||x||_1 \quad \text{s.t.: } y = \Phi x,$$

which is usually called $\ell_1$ minimization or basis pursuit (BP). The $\ell_1$ minimization problem and the $\ell_0$ minimization problem share the same solution. However, problem 2.12 is a convex problem, and can be solved by many convex optimization methods.

The sparsity of the solution is illustrated in Fig. 2.3. The square in the figure is the $\ell_1$ ball. It is seen that the linear system first coincides with a vertex of the $\ell_1$ ball, no matter from which direction the line moves towards the square.

![Fig. 2.3 A linear system coincides with a sparse solution.](image)

Noise-free data hardly exist in real measurements. In contrast, it is quite common that data are corrupted by noise during capture, storage, transmission, processing or conversion. In the presence of noise, the linear constraint in 2.12 is not met. The mathematical model is then expressed as

$$\min ||x||_1 \quad \text{s.t.: } ||y - \Phi x||_2 \leq \epsilon,$$
or in a unconstrained version as

\[
\min ||x||_1 + \frac{\tau}{2}||y - \Phi x||^2_2, \tag{2.14}
\]

where both \(\varepsilon\) and \(\tau\) are parameters with relationship to noise. With careful choice of \(\varepsilon\) and \(\tau\), these two problems are the same.

### 2.3.5 Reconstruction Methods

The \(\ell_1\) minimization problem of CS is convex, and can be easily solved. Many reconstruction methods have been proposed to solve 2.12 or 2.13. Here two groups of algorithms based on different reconstruction principles are introduced: greedy algorithms and iterative thresholding algorithms.

#### Greedy Algorithms

Greedy algorithms rely on iterative approximation of the signal coefficients and support. This can be achieved, either by iteratively identifying the support of the signal until a convergence criterion is met, or alternatively by obtaining an improved estimate of the sparse signal at each iteration that attempts to account for the mismatch to the measured data.

One commonly used greedy algorithm is called orthogonal matching pursuit (OMP) [41, 42]. OMP tries to search for a sparse minimizer one entry at a time. Therefore, the original signal with sparsity \(K\) can be recovered exactly in \(K\) iterations. OMP starts with finding the column of the sensing matrix \(\Phi\) most correlated to the measured data \(y\), then repeats this step by correlating the columns with the signal residual, which is obtained by subtracting the contribution of a partial estimate of the signal from the original measurement vector.

The OMP algorithm is briefly described in Algorithm 1. Many modifications based on OMP have also been proposed, such as CoSaMP [43], regularized OMP [44], and stagewise OMP [45].

#### Iterative Thresholding Algorithms

The iterative thresholding method [46] is straightforward. The intuition for iterative thresholding algorithms relies on the approximate inversion of the sensing matrix \(\Phi\) by multiplying its adjoint \(\Phi^*\). Instead of solving the constrained \(\ell_1\) minimization (2.12 or 2.13), the iterative thresholding method solves the unconstrained \(\ell_1\) minimization 2.14. Starting from an initial signal estimate \(x^0\), the algorithm iterates a
Algorithm 1 OMP

Input: Sensing matrix $\Phi$, measured data $y$
Output: Sparse signal $x$

Initialization: Support set $\Lambda^0 = \emptyset$, sparse signal $x^0 = 0$, residual $r^0 = 0$, stop criterion $\sigma$

Iteration:
1. search for one column of sensing matrix which is most relevant with residual $j = \arg \min_{j \in |N|} |r^{i-1}, \phi_j|$, where $\phi_j$ is the $j^{\text{th}}$ column of $\Phi$
2. update support set $\Lambda^i = \Lambda^{i-1} \cup \{ j \}$
3. update sparse signal by least square method $x^i = \arg \min ||y - \Phi_{\Lambda^i} x||_2$
4. update residual $r^i = y - \Phi_{\Lambda^i} x^i$
5. check stop criterion
   - if $||r^i||_2^2 < \sigma$, stop
   - otherwise, repeat from step 1

Gradient descent step followed by hard thresholding until a convergence criterion is met. The iterative hard thresholding algorithm is to solve $y = \Phi x$ with the knowledge that the original signal is $K$ sparse. First, the gradient of 2.14 is calculated. Instead of solving $y = \Phi x$ directly, $\Phi^* y = \Phi^* \Phi x$ is solved. Then, the minimum value of the function is obtained by an iterative process following the gradient direction. A thresholding procedure is used during this process. The search for a minimum value can be interpreted as the fixed point equation $x^{i+1} = (I - \Phi^* \Phi) x^i + \Phi^* y$. Only the $K$ largest absolute entries of $x^{i+1}$ are kept since the original signal is $K$ sparse. Once the minimum value is obtained, the gradient at this point is again calculated. The process is repeated until the decrease between two iterations is smaller than a stop criterion.

The iterative thresholding algorithm is briefly described in Algorithm 2.

Algorithm 2 Iterative Hard Thresholding Algorithm

Input: Sensing matrix $\Phi$, measured data $y$, sparsity $K$
Output: Sparse signal $x$

Initialization: signal $x^0 = 0$

Iteration:
repeat until a stopping criterion is met
$x^{i+1} = H(x^i + \Phi^* (y - \Phi x^i))$, $H$ is a hard thresholding operation
2.4 Restricted Isometry Property

In CS, one important theoretical question is how to design a sensing matrix so that successful recovery can be obtained from many fewer measurements. To guarantee successful recovery, first the original signal should be sparse, and second the sensing matrix should satisfy certain rules.

An important sufficient condition for exact reconstruction is the restricted isometry property (RIP) \[14, 16, 17\], also known as the uniform uncertainty principle. For a sensing matrix \(\Phi\) with unit column norms, the RIP constant \(\delta_K\) is defined as the smallest value which satisfies

\[(1 - \delta_K)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta_K)\|x\|_2^2,\] (2.15)

for all sparse signals \(x\) with sparsity \(K\).

In the definition of the RIP in 2.15, the bounds are assumed to be symmetric about 1. However, the bounds can be arbitrary and the symmetry is just for notational convenience. That is, the RIP can be denoted as

\[\alpha\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq \beta\|x\|_2^2,\] (2.16)

where \(0 < \alpha \leq \beta < \infty\). The definition 2.15 and 2.16 are essentially the same. By scaling \(\Phi\) in 2.16, the bounds can satisfy the symmetry property in 2.15. For example, if multiplying \(\Phi\) in 2.16 by a factor \(\frac{1}{\beta + \alpha}\), the \(\delta_K\) in 2.15 can be denoted as \(\frac{\beta - \alpha}{\beta + \alpha}\).

In [17], it is shown that a \(K\)-sparse signal can be exactly reconstructed from the measurement of \(\Phi\) if \(\delta_{2K} < 1\). It has been proven that if the sensing matrix \(\Phi\) satisfies the RIP of order \(2K\) with constant \(\delta_{2K} \in (0, 1/2]\), the length of measurements

\[M \geq C \cdot K \cdot \log\left(\frac{N}{K}\right),\] (2.17)

where \(C = \frac{1}{2\log(\sqrt{24} + 1)} \approx 0.28\).

In addition to determining the length of measurements and guaranteeing successful reconstruction, the RIP constant also determines the stability of reconstruction. In the presence of noise, i.e., \(y = \Phi x + \varepsilon\), the reconstructed signal \(\hat{x}\) satisfies

\[\|x - \hat{x}\|_2^2 \leq \frac{4\varepsilon^2}{1 - \delta_{2K}},\] (2.18)
As a result, the reconstruction error becomes smaller when $\delta_{2^K}$ becomes smaller, and vice versa. Ideally, $\delta_K$ should be as small as possible.

The RIP constant not only determines the reconstruction ability of CS, but also the reconstruction error. It is an important indicator in CS. However, the computation of $\delta_K$ is also NP hard, since it requires iteration through all possible sparse signals with sparsity $K$ and length $N$.

In [14, 47, 48], it is shown that the RIP is equivalent to

$$(1 - \delta_K) \leq \sigma^2_{\text{min}}[\Phi_{\text{sub}}(K)] \leq \sigma^2_{\text{max}}[\Phi_{\text{sub}}(K)] \leq (1 + \delta_K),$$

where $\Phi_{\text{sub}}(K)$ is a $M \times K$ sub-matrix formed from $K$ distinct columns of the sensing matrix $\Phi$, $\sigma_{\text{min}}[\Phi_{\text{sub}}(K)]$ and $\sigma_{\text{max}}[\Phi_{\text{sub}}(K)]$ are the minimum and maximum singular values of $\Phi_{\text{sub}}(K)$, respectively. Therefore, instead of calculating $\delta_K$, it is assessed by the distance between 1 and $\sigma_{\text{min}}[\Phi_{\text{sub}}(K)]$ and between 1 and $\sigma_{\text{max}}[\Phi_{\text{sub}}(K)]$. The smaller the distance, the smaller the $\delta_K$, and hence the better the performance of $\Phi$.

In [14, 47, 48], the eigenvalue $\lambda[\Phi_{\text{sub}}(K)^T \Phi_{\text{sub}}(K)]$ is used to calculate singular values, where $\lambda[\Phi_{\text{sub}}(K)^T \Phi_{\text{sub}}(K)] = \sigma^2[\Phi_{\text{sub}}(K)]$. To estimate $\delta_K$, first $K$ out of $N$ columns of $\Phi$ are chosen to form $\Phi_{\text{sub}}(K)$, then the eigenvalues $\lambda[\Phi_{\text{sub}}(K)^T \Phi_{\text{sub}}(K)]$ are calculated. This process should be repeated enough times to make a fair estimate of $\delta_K$.

### 2.5 Coherence

To verify that the sensing matrix $\Phi$ satisfies RIP is cumbersome, since all $\binom{N}{K}$ sub-matrices should be considered. Apart from RIP, another important measure of the quality of the sensing matrix is coherence [18, 19, 94]. Similar to the RIP constant, the smaller the coherence, the better the recovery algorithms perform.

In CS, instead of observing the original signal $x \in \mathbb{C}^N$ directly, a small number $M$ of transform coefficients of $x$ are collected. Suppose an orthogonal matrix $U \in \mathbb{C}^{N \times N}$ is the transform matrix, with the property that

$$U^*U = NI,$$

where $I$ is an identity matrix. The transform process is denoted as

$$y = Ux.$$
If all transform coefficients are observed, the reconstruction process is trivial. The original signal \( x \) can be reconstructed by multiplying \( \frac{1}{N} U^* \) on both sides of 2.21. However, only a small portion of transform coefficients are sampled in CS, which makes the reconstruction process very complicated. To answer the question of how many samples are needed to successfully reconstruct the original signal from the \( \ell_1 \) minimization problem, a simple relationship is obtained in [18], that is

\[
M \geq C \cdot \mu^2 \{ U \} \cdot K \cdot \log(N),
\]

where \( C \) is a small constant, \( K \) is the sparsity of the original signal, and \( \mu \{ U \} \) is defined as the largest magnitude among the entries of \( U \):

\[
\mu \{ U \} = \max |U_{i,j}|, 1 \leq i, j \leq N.
\]

The meaning of \( \mu \{ U \} \) can be regarded as a rough measure of how concentrated the rows of \( U \) are. Since \( U \) is an orthogonal matrix and each row or column of \( U \) has an \( \ell_2 \) norm equal to \( \sqrt{N} \), the value of \( \mu \) is in the range between 1 and \( \sqrt{N} \). \( \mu = 1 \) means that the rows of \( U \) are perfectly flat, which requires the least number of observed samples. In contrast, \( \mu = \sqrt{N} \) means that a row of \( U \) is maximally concentrated while the other rows vanish, which offers no guarantee for recovery.

In CS, the transform matrix \( U \) is commonly decomposed as a product of a sparsifying matrix \( \Psi \in \mathbb{C}^{N \times N} \) and an orthogonal sensing matrix \( \hat{\Phi} \in \mathbb{C}^{N \times N} \), as denoted in 2.8, 2.9 and 2.10. The only difference here is that the sensing matrix \( \hat{\Phi} \) is of size \( N \times N \) instead of \( M \times N \). Therefore, the length of measured data is \( N \), while only \( M \) out of \( N \) is used to reconstruct the original signal. Suppose that

\[
U = \hat{\Phi} \Psi, \quad \hat{\Phi}^* \hat{\Phi} = NI, \quad \Psi^* \Psi = I,
\]

then the parameter \( \mu \) is denoted as

\[
\mu \{ \hat{\Phi} \Psi \} = \max |< \phi_i, \psi_j >|, 1 \leq i, j, \leq N,
\]

where \( < \cdot, \cdot > \) denotes the inner product, and \( \phi_i \) and \( \psi_j \) denote the \( i \)th row of \( \hat{\Phi} \) and \( j \)th column of \( \Psi \), respectively. Here, \( \mu \{ \hat{\Phi} \Psi \} \) serves as a rough characterization of the degree of similarity between the sparsifying and the sensing matrices, and \( \mu \{ \hat{\Phi} \Psi \} \) is referred to as the mutual coherence between \( \hat{\Phi} \) and \( \Psi \). When \( \mu \{ \hat{\Phi} \Psi \} \) is close to its minimum value 1, the rows of the sensing matrix spread in the sparsifying
domain. The relationship 2.22 still holds and becomes

\[ M \geq C \cdot \mu^2 \{ \hat{\Phi} \Psi \} \cdot K \cdot \log(N). \] (2.26)

In 2.26, the minimum number of samples is given according to the mutual coherence of the sensing and the sparsifying matrices, the sparsity of the original signal and the length of the original signal. However, as we may easily anticipate, reconstruction with sampled data from certain locations may perform better or worse than that with data from other locations. However, this is not shown in 2.26. To take the sub-sampling process into consideration, it is proposed to define the coherence in a dictionary perspective.

Suppose \( x \in \mathbb{C}^N \) is sparse on dictionary \( \Psi \in \mathbb{C}^{N\times N} \). If \( x \) is naturally sparse, \( \Psi \) is the identity matrix. Mathematically,

\[ x = \Psi s. \] (2.27)

The coherence of \( \Psi \) is defined as

\[ \mu \{ \Psi \} = \max_{1 \leq i \neq j \leq N} \frac{|\psi_i^T \psi_j|}{||\psi_i|| \cdot ||\psi_j||}, \] (2.28)

where \( \psi_i \) is the \( i \)th column of \( \Psi \). Coherence can also be obtained by calculating the Gram Matrix, which is defined as

\[ G = \tilde{\Psi}^T \tilde{\Psi}, \] (2.29)

where \( G \in \mathbb{C}^{N\times N} \) is the Gram Matrix, and \( \tilde{\Psi} \in \mathbb{C}^{N\times N} \) is \( \Psi \) with each column normalized. The maximum magnitude of the off-diagonal elements of \( G \) is defined as the coherence. If \( \mu \{ \Psi \} = 1 \), at least two columns of the dictionary are the same or proportional, which means there are redundant columns in the dictionary. In contrast, if \( \mu \{ \Psi \} = 0 \), it means that any two columns of the dictionary are mutually incoherent.

Suppose that the original signal \( x \) has been reconstructed by \( x = \Psi s \) with a sparse representation, and the following inequality is satisfied

\[ ||s||_0 < \frac{1}{2} \left( 1 + \frac{1}{\mu \{ \Psi \}} \right). \] (2.30)

The sparse signal \( s \) is the solution of

\[ \min ||s||_1 \quad \text{s.t.:} \quad x = \Psi s, \] (2.31)
which can be recovered by the OMP method [49, 50]. Once the sparse signal $s$ is obtained, the original signal $x$ is easily recovered.

Suppose $x$ is measured by a sensing matrix $\Phi \in \mathbb{C}^{M \times N}$, in which sub-sampling has already been involved. Then 2.31 becomes

$$\min ||s||_1 \quad \text{s.t.: } y = \Phi \Psi s.$$ (2.32)

If the following inequality is satisfied

$$||s||_0 < \frac{1}{2}(1 + \frac{1}{\mu \{\Phi \Psi\}}),$$ (2.33)

the sparse signal $s$ can be recovered and so can the original signal $x$. $\mu \{\Phi \Psi\}$ is calculated by using the Gram Matrix method.

The inequality 2.33 shows the relationship between the sparsity of the original signal and the coherence between the sparsifying and the sensing matrices. It is clear that by choosing different sensing matrices, the coherence between the sparsifying and the sensing matrices may change, which further influences the recovery of the sparse signal. In other words, since the sparsity of a signal is fixed when the signal and the sparsifying transform are chosen, certain choices of the sensing matrix may satisfy the inequality 2.33, while some may not. By calculating the coherence between the sparsifying and the sensing matrices, it is possible to find a suitable sensing matrix, since the sparsifying matrix is commonly fixed.

### 2.6 Phase Transition

The theory of CS asserts that a sparse signal can be faithfully recovered from sub-sampled data, the length of which is far less than that of the original signal. One important aspect of the CS theory is how aggressively a sparse signal can be sub-sampled while preserving all information. An important result in this aspect is called phase transition [30, 51].

Let $\delta = M / N$, and $\rho = K / M$, where $K, M$, and $N$ stand for the sparsity of the original signal, the length of the measured data and the length of the original signal, respectively. Phase transition states that in the plane of $(\delta, \rho)$, there is a curve which divides the plane into two phases. Below the curve the original signal can be successfully reconstructed, while above the curve the original signal cannot be successfully reconstructed, both with an overwhelming probability.

In [30], the phase transition of real valued CS using the sensing matrix with i.i.d. random Gaussian entries was studied. The Gaussian condition of the sensing matrix
can be considerably relaxed, which results in the well-known observed universality of phase transition [51]. In [52], the phase transition of complex valued CS was studied. The results show that the same phase transition property exists for complexed valued CS, and the phase transition of complexed valued CS presents higher successful phase than that of real valued CS.

How to plot the phase transition curve is described as follows. Suppose \( x \in \mathbb{C}^N \) is a sparse signal, where \( N \) is the length of the original signal. The sparsity of the original signal \( x \) is denoted as \( K \), which means only \( K \) out of \( N \) entries are non-zero. Measured data \( y \in \mathbb{C}^M \) are acquired by a sensing matrix \( \Phi \in \mathbb{C}^{M \times N} \), where \( M \) is the length of the measured data. The sparse signal is reconstructed by solving the \( \ell_1 \) minimization problem 2.12. As defined above, \( \delta = \frac{M}{N} \), and \( \rho = \frac{K}{M} \). Generally speaking, in CS, \( M \approx 3 \sim 5K \), and \( M \ll N \), so \( K < M \ll N \). In the plane of \( (\delta, \rho) \), for a fixed \( M \), it is easy to appreciate that the original signal can be reconstructed if \( \rho \) is below a certain value, which means \( K \) is sufficiently small. In contrast, the reconstruction will fail if \( \rho \) is above a certain value.

The phase transition property is denoted as an ascending curve in the plane of \( (\delta, \rho) \). An example is shown in Fig. 2.4. The horizontal axis denotes \( \delta \), while the vertical axis denotes \( \rho \). Both \( \delta \) and \( \rho \) are from 0 to 1. For a specific pair of \( (\delta, \rho) \), such as \((0.5, 0.5)\), it means \( M = 0.5 \times N \), and \( K = 0.5 \times M \). If \( N = 256 \), then \( M = 128 \), and \( K = 64 \). A sparse signal with sparsity \( K \) is randomly generated. The locations where the entries are non-zero are randomly determined, so as the values of the entries. In this case, the sensing matrix \( \Phi \) is a partial Fourier transform, which is obtained by randomly sub-sampling the full discrete Fourier matrix. A conjugate gradient method [20] is applied to solve 2.12.

The normalized mean square error (NMSE) of the reconstructed signal \( \hat{x} \) is obtained by

\[
\text{NMSE} = \frac{||\hat{x} - x||_2^2}{||x||_2^2}.
\]

(2.34)

A threshold of 0.1 is used to determine the performance of the reconstruction. If \( \text{NMSE} \leq 0.1 \), the reconstruction is regarded as successful. Otherwise, the reconstruction is regarded as a failure. At each pair of \( (\delta, \rho) \), this reconstruction process is repeated 500 times, in order to minimize the influence of randomness. The probability of successful reconstruction is calculated by dividing the number of successful reconstructions by the number of total repeated times. Phase transition is defined as the value of \( \rho \) at which the original signal is successfully recovered with a probability of 50%.
2.6 Phase Transition

While the phase transition curve is continuous between 0 to 1, it is only possible to choose discrete locations in the plane of \((\delta, \rho)\) in simulation. In simulation, \(\delta\) is from 0 to 1 with a step of 0.04, so as \(\rho\). A polynomial fitting process is applied after the reconstruction process to generate a continuous curve.

![Phase Transition Curve](image)

**Fig. 2.4 Example of phase transition curve.**

The phase transition curve may be affected by several different factors, such as the choice of the sensing matrix and the reconstruction method, the NMSE threshold, and the repeated times. Despite these factors, the general trend of the phase transition curve stays unchanged. Above the curve, the probability of successful reconstruction is less than 50%, while below the curve, the probability of successful reconstruction is higher than 50%. In fact, the performance will decrease/increase so quickly that it will soon reach 0 or 1 not far away from the phase transition curve.

The phase transition curve moves upwards or downwards when different sensing matrices are used. As depicted in [52], the phase transition curve will move upwards if complex valued CS instead of real valued CS is chosen, which means complex valued CS enables fewer measurements \(M\) when the sparsity \(K\) is fixed, or larger \(K\) when the length of the measured data \(M\) is fixed. In the following chapters, we will show that different phase transition curves are obtained if partial Fourier matrix or partial Noiselet matrix is selected. Phase transition serves as a good quantitative measurement of the performance of CS reconstruction.
2.7 Conclusion

In this chapter, the fundamentals of CS are reviewed. First, the Nyquist-Shannon Sampling Theorem is introduced to show the challenges we face in modern digital sampling systems. Second, to overcome these challenges, CS was proposed in order to faithfully recover the original signal from many fewer measurements. The CS theory, including the sparse representation, the sensing system, \( \ell_0 \) and \( \ell_1 \) models, and reconstruction methods, is described in this chapter. To guarantee successful reconstruction, RIP and coherence analysis are also presented. A quantitative measure of how aggressively a sparse signal can be recovered, which is named phase transition, is finally introduced.
Chapter 3

Fundamentals of MRI Reconstruction

3.1 Introduction

Magnetic resonance imaging (MRI) is an imaging technique which produces high quality images of the inside of the human body. It is widely used in contemporary clinical practices due to its excellent resolution in tissue depiction, flexible slice orientation, and freedom from harmful radiation [1]. In this chapter, first, the physical background of MRI is introduced, including the nuclear magnetic resonance (NMR) phenomenon, NMR signal excitation, and imaging sequences.

One major problem of MRI is its relatively long data acquisition time, which may lead to blur and artifacts in magnetic resonance (MR) images, and restrict its application in certain areas. It is a challenging research question as to what would be the minimum number of samples required for MR images. Intensive research activities have been carried out in order to answer this question, which aims to shorten the data acquisition time, since it is proportional to the number of samples. In traditional MRI, images of interest are encoded with Fourier encoding, so that the acquired data are stored in the Fourier domain, or k-space. A common practice to shorten the data acquisition time is to sub-sample the k-space data by deliberately skipping some data points. In the second part of this chapter, the problem of reconstruction from sub-sampled k-space data is introduced, followed by data acquisition time reduction techniques, such as following non-Cartesian trajectories, sub-sampling data in Cartesian coordinates, or designing novel sequences.
3.2 Physics of MRI

3.2.1 NMR Phenomenon

Spin

MRI is based on the nuclear magnetic resonance (NMR) phenomenon, which was first observed by Bloch [2] and Purcell [3] in 1946. NMR arises in atoms possessing a property known as spin [1]. Spin is a fundamental property of nature, denoting the intrinsic angular momentum of atomic and sub-atomic particles, for example, electrons, protons, and whole nuclei. It is multiples of 1/2 and can be positive or negative. Particles with spins with opposite signs can pair up and cancel out the manifestation of spin. In NMR, only nuclei with unpaired spins are of importance. As nuclei rotate, they act as small magnetic dipoles with random orientations. Matter even of a small size contains a huge number of nuclei. However, when a matter is in the natural environment, the net magnetization $M_0$ of nuclei is undetectable since the dipoles are at random orientations and they cancel out each other’s magnetic effect. Fig. 3.1 (a) shows spins at random orientations.

![Fig. 3.1 Spins in different environments. (a) Spins at random orientations. (b) Spins placed in a magnetic field.](image)

When placed in an external magnetic field $B_0$, the dipoles show distinctively different characteristics. The dipoles align with the direction of $B_0$. Some are parallel to it, while the others are anti-parallel. The two spin states are referred to as "spin-up" or "parallel", and "spin-down" or "anti-parallel", respectively. Fig. 3.1 (b) shows spins placed in a magnetic field $B_0$. Dipoles in parallel state are of low energy, while dipoles in anti-parallel state are of high energy. The difference in numbers of dipoles...
in parallel and anti-parallel states leads to the NMR signal. The energy difference between the two states is defined as,

$$\Delta E = \frac{h \gamma |B_0|}{2\pi}.$$  \hspace{1cm} (3.1)

where $h$ is the Planck’s constant, $\gamma$ is the gyromagnetic ratio, and $|B_0|$ is the magnitude of $B_0$.

Spins can undergo a transition between the two energy states by absorbing a photon with a specific frequency. Since the energy of a photon is related to its frequency, $\nu$, the energy of the photon is written as

$$E = h\nu.$$ \hspace{1cm} (3.2)

The energy of the photon should match the energy difference between the two energy states in order to ignite the transition, which means

$$E = \Delta E.$$ \hspace{1cm} (3.3)

Therefore, the following equation should be valid,

$$h\nu = \frac{h \gamma |B_0|}{2\pi}.$$ \hspace{1cm} (3.4)

The frequency of the photon is then given as

$$\nu = \frac{\gamma |B_0|}{2\pi}.$$ \hspace{1cm} (3.5)

This frequency is named the Larmor frequency.

**$T_1$ Relaxation**

After being placed in a magnetic field for a sufficient length of time, the net magnetization $M_0$ reaches equilibrium. The direction of the net magnetization $M_0$ is parallel to that of the magnetic field $B_0$. The direction of the magnetic field $B_0$ is generally referred to as the $z$ direction, while the transverse plane perpendicular to the $z$ direction is composed of the $x$ direction and the $y$ direction. At equilibrium, the $z$ component of the net magnetization, $M_z$, equals $M_0$, while the $x$ and the $y$ components of the net magnetization, $M_x$ and $M_y$, respectively, are zero. The net magnetization is linear with the magnitude of the magnetic field $B_0$. It can also be
changed by exposure to a system containing a sequence at Larmor frequency, which will ignite the transition of spins between energy states.

The net magnetization cannot reach its equilibrium instantly. In fact, the process may take up to seconds. The time constant which describes how it reaches its equilibrium is called the spin-lattice relaxation time, or $T_1$. The net magnetization grows to its equilibrium obeying the following expression,

$$M_z = M_0 \left(1 - e^{-t/T_1}\right).$$

(3.6)

Fig. 3.2 shows $M_z$ as a function of $t$ when $M_0 = 1$ and $T_1 = 256$.

Since spins favor the low energy state, the energy of the spin system decreases as $M_z$ grows to its equilibrium. The decreased energy is emitted to nearby nuclei as a transfer of heat through collisions, rotations or electromagnetic interactions.

$T_2$ Relaxation

At equilibrium, the net magnetization $M_0$ has no $x$ or $y$ components. It can be tipped into the transverse plane by applying a radio frequency (RF) pulse which is perpendicular to the $z$ direction. The frequency of the RF pulse is equal to (or at least is very close to) the Larmor frequency, and the magnetic field caused by the RF pulse is named the $B_1$ magnetic field. The flip angle $\alpha$ is proportional to the
frequency $f$ of the $B_1$ field and the time duration $t_p$, that is
\begin{equation}
\alpha = ft_p = \gamma B_1 t_p. \tag{3.7}
\end{equation}

The flip angle can be between 0 and 180°. At 180°, the net magnetization also aligns with $B_0$, but in the opposite direction. At any other degrees, the net magnetization will rotate around the $z$ direction. This process is called precession. The rotation frequency equals the Larmor frequency. After the RF pulse is turned off, the $z$ component of net magnetization, $M_z$, will experience the $T_1$ relaxation, while the transverse component of the net magnetization, $M_{xy}$, will start a different relaxation process.

The transverse magnetization $M_{xy}$ dephases because the spins start to experience slightly different magnetic fields. The difference of magnetic fields comes from molecular interaction. The time constant which defines the decay of the transverse magnetization is called the spin-spin relaxation time, or $T_2$. The decay of the net magnetization is expressed as
\begin{equation}
M_{xy} = M_{xy0} e^{-t/T_2}, \tag{3.8}
\end{equation}

where $M_{xy0}$ is the transverse component of net magnetization at the start of the dephasing process. In practice, $T_2$ is always smaller than $T_1$.

Two factors contribute to the decay of the net magnetization in the transverse plane. As mentioned above, the decay results from the differences in local magnetic fields. First, molecular interactions lead to the pure $T_2$ described above. Second, although the magnetic field $B_0$ is assumed to be constant everywhere, there is always inhomogeneity. the variations in $B_0$ lead to $T_{2,inhomo}$. The time constant resulting from the combination of these two factors is called $T_2^*$, which is defined as
\begin{equation}
\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_{2,inhomo}}. \tag{3.9}
\end{equation}

$T_2^*$ is always smaller than $T_2$. The decay of the net magnetization caused by these two factors is expressed as
\begin{equation}
M_{xy}^* = M_{xy0} e^{-t/T_2^*}. \tag{3.10}
\end{equation}

Fig. 3.3 shows $M_{xy}^*$ as a function of $t$ when $M_{xy0} = 1$ and $T_2 = 128$, $T_2^* = 64$, respectively.
3.2.2 NMR Signal Excitation

Free Induction Decay

As the net magnetization rotates around the $z$ direction, its $z$ component experiences the $T_1$ relaxation, and its $xy$ component experiences the $T_2^*$ relaxation. NMR signals are acquired with coils of wire placed in the transverse plane. The $xy$ component induces a current in the coils. As a function of time, the current gives a sine wave. Furthermore, the current does not persist for ever. It decays with the time constant $T_2^*$. This signal is called a free induction decay (FID), obeying the following form: $\sin(\omega t) e^{-t/T_2^*}$, where $\omega$ is the frequency of the sine wave. Fig. 3.4 shows an FID as a function of $t$ when $T_2^* = 512$, and the frequency of the sine wave is 1/256.

Slice Selection

The main magnetic field $B_0$ in an MR scanner is assumed to be homogeneous, or constant everywhere. If an RF pulse is applied to tip the net magnetization into the transverse plane, the FID recorded will be from all the spins placed in the main magnetic field, since all of them have the same Larmor frequency. A magnetic gradient allows us to achieve locate positions. A magnetic gradient is a variation in the magnetic field with respect to location. The gradients in $x$, $y$, and $z$ directions are
denoted as $G_x$, $G_y$ and $G_z$, respectively. Magnetic gradients are created by gradient coils in MR scanners.

2D MR images depict a slice of a volume with a specified thickness. To achieve slice selection, a 1-dimensional (1D) linear magnetic gradient along $z$ direction is applied during the period of the RF pulse. During scanning, an RF pulse together with a slice selection gradient tips the net magnetization in a specified slice into the transverse plane. Spins outside this slice are not be influenced. After slice selection, the RF pulse and the slice selection gradient are turned off, and the main magnetic field in the scanner becomes homogeneous again at this time. For example, by applying an RF pulse and a slice selection gradient to Fig. 3.1 (b), the slice selection process is displayed in Fig. 3.5.
Fundamentals of MRI Reconstruction

Frequency Encoding
The slice selection gradient is used to select a specific slice with thickness along the $z$ direction. To locate a position in the transverse plane of the slice, the $x$ and the $y$ values are determined by the frequency encoding gradient and the phase encoding gradient, respectively.

The center of the transverse plane where $(x, y) = (0, 0)$ is called the isocenter of the magnet. At the isocenter, the magnetic field is $B_0$ and the corresponding Larmor frequency is $\nu_0$. A 1D linear magnetic gradient is applied in the $x$ (or $y$) direction, causing the magnetic field in this direction to vary linearly as a function of position. The magnetic field along this direction is presented as

$$B(x) = B_0 + xG_x. \quad (3.11)$$

Correspondingly, the Larmor frequency is presented as

$$\nu(x) = \gamma B(x) = \gamma B_0 + \gamma xG_x = \nu_0 + \gamma xG_x. \quad (3.12)$$

The Larmor frequency also varies linearly as a function of position. As a result, the excited NMR signal has more than one frequency. By measuring the frequencies of the NMR signal, the location along the $x$ direction can be determined. This process is called frequency encoding, since the Larmor frequency is proportional to the position of spin.

Phase Encoding
Frequency encoding alone can only determine the position along one direction. This direction is called the frequency encoding direction. To determine the location along the other direction, phase encoding is used. Likewise, this direction is called the phase encoding direction. The location along the frequency encoding direction is determined by its frequency, and the location along the phase encoding direction is determined by its phase.

Phase encoding is accomplished by introducing a 1D linear magnetic gradient along the phase encoding direction. The magnetic field along the phase encoding direction is presented as

$$B(y) = B_0 + yG_y. \quad (3.13)$$

Correspondingly, the Larmor frequency is presented as

$$\nu(y) = \gamma B(y) = \gamma B_0 + \gamma yG_y = \nu_0 + \gamma yG_y. \quad (3.14)$$
Up to this step, phase encoding is the same as frequency encoding. Since the Larmor frequencies differ from each other, spin will rotate at its own unique Larmor frequency. If the phase encoding gradient is turned off some time after it was turned on, spins at different locations will experience the same \( B_0 \) again and will rotate at the same Larmor frequency. Although the frequencies are the same, the phase is different, since the spins have been rotating with different frequencies for some time. By measuring the phase, the location along the phase encoding direction can be determined.

MR Image Formulation

The \((x,y)\) locations of spins in the transverse plane are determined by measuring the frequency and the phase of the FID, as illustrated above. The mathematical expression of this process is given as

\[
s(k_x, k_y) = \int_{x} \int_{y} \rho(x, y) e^{-i2\pi(k_xx + k_yy)} \, dx \, dy,
\]

(3.15)

where \( s(k_x, k_y) \) denotes the acquired signal in Fourier domain, or k-space, and \( \rho(x, y) \) denotes the spin density at location \((x, y)\). Generally speaking, the signal acquisition process is equivalent to applying 2D Fourier transform to the spin density. The acquired signal is sampled and presented in discrete form. The discrete signal is expressed as

\[
s[k_x, k_y] = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \rho[x,y] e^{-i2\pi(k_xx + k_yy)},
\]

(3.16)

where \( k_x = 0, 1, \ldots, N - 1 \), and \( k_y = 0, 1, \ldots, N - 1 \).

3.2.3 Imaging Sequences

NMR signals are obtained by applying a series of pulse sequences in order. A timing diagram is a multi-axis plot of the pulse sequences versus time. Generally, a timing diagram contains an RF pulse, a slice selection gradient pulse, a phase encoding gradient pulse, a frequency encoding pulse, and a signal. In this section, some basic sequences are introduced.

Gradient-Echo Sequence

The gradient-echo (GRE) sequence is a commonly used pulse sequence in modern clinical practice. Its timing diagram is shown in Fig. 3.6.
First, an RF pulse is applied together with a slice selection gradient \((G_s)\) in order to flip the net magnetization in the aiming slice to the transverse plane. The flip angle produced by the RF pulse is usually between 0 and 90°. A small flip angle leads to short acquisition time, but a weak signal, while a large flip angle results in long acquisition time, but a strong signal. The flip angle is determined by the \(B_1\) field caused by the RF pulse and its time duration \(t_p\).

Following the RF pulse and the slice selection gradient is the phase encoding gradient \((G_\Phi)\). The phase encoding gradient varies from the minimum to the maximum value with 128 or 256 steps. Spins at different positions in the phase encoding direction possess different phases after the phase encoding gradient is turned off.

At the same time as the phase encoding gradient, a dephasing frequency encoding gradient is applied, so the excited NMR signal is in phase at the center of the acquisition period. The dephase frequency encoding gradient has the opposite sign to the frequency encoding gradient during the acquisition period. The dephasing gradient causes the spins to dephase, while the frequency encoding gradient refocuses the spins, to generate an echo at the center of the acquisition period.

The time from the beginning of the RF pulse to the center of the acquisition period is called the echo time \((T_E)\), while the time from the beginning of the RF pulse to the beginning of the next RF pulse is called the repetition time \((T_R)\). The total time consumption for a complete scan using the GRE sequence is the multiplication of \(T_R\) and the phase encoding steps. The signal from the GRE sequence is expressed as

\[
S = k\rho (1 - e^{-T_R/T_1})e^{-T_E/T_2^*}. \tag{3.17}
\]
Multi-slice Imaging

As shown in Fig. 3.6, $T_R$ is much greater than $T_E$. In other words, most of the time in $T_R$ is wasted. To improve time efficiency, a straightforward approach is to excite multiple slices in one $T_R$. A timing diagram of 3 RF pulses in one $T_R$ is shown in Fig. 3.7.

![Timing diagram of a multi-slice imaging sequence.](image)

With the application of this sequence, three slices can be acquired within the time required for only one slice in the GRE sequence. That is, although the time needed for each slice is the same as usual, the total time efficiency is improved by three times.

The only restriction of the multi-slice GRE sequence is that the acquisition of one slice should not affect the acquisition of other slices. To achieve this goal, the distance between each excited slices should be sufficiently far. Since the slice selection gradient is linear as a function of the location, the total magnetization is also linear, as is the Larmor frequency. As long as the bandwidth of the Larmor frequency at each location does not overlap with that of others, the acquisition of each slice is independent.

Volume Imaging

Although multi-slice imaging is capable of acquiring multiple slices and shortening the scan time, the resolution in the slice direction is often worse than that in the frequency encoding and the phase encoding directions. In other words, the voxels are not isotropic. To improve resolution in the slice direction, or if isotropic voxels are desired, a volume imaging sequence is preferred compared with single tomographic sequences. Volume imaging acquires NMR signals from a volume rather than a
single tomographic slice. Unlike multi-slice imaging, volume imaging acquires data from contiguous slices through a region of the imaged object. Fig. 3.8 displays the timing diagram of a volume imaging sequence.

![Timing diagram of a volume imaging sequence.](image)

The RF pulse and the slice selection gradient are the same as in the GRE sequence. However, unlike the GRE sequence, the forthcoming phase encoding gradient is conducted in both the phase encoding direction and the slice direction. In the multi-slice imaging sequence, the slice direction is in the spatial domain, while in the volume imaging sequence, the slice direction is in k-space. Both of the two phase encoding gradients vary from the minimum to the maximum value, and they cycle through all possible combinations in order to scan the whole volume. The frequency encoding gradient is the same as in the GRE sequence, and an echo is generated and recorded at the center of the acquisition period.

### 3.3 Reconstruction Methods of MRI

#### 3.3.1 Reconstruction Problem

Compared with other contemporary imaging techniques, such as computational tomography (CT) or ultrasound, MRI offers many advantages, including but not limited to high resolution, freedom from radiation, and flexible slice orientation [1]. One major problem of MRI is its relatively long data acquisition time. For example, if $T_R$ is one second and there is a total of 256 phase encoding gradient steps, the scan time is 256 seconds, or 4 minutes and 16 seconds. To improve the signal-to-noise ratio (SNR), in some applications, each phase encoding gradient step is repeated
several times and the signals are averaged. If each phase encoding gradient step is repeated twice, the total scan time for one image will be 8 minutes and 32 seconds. In multi-slice imaging or volume imaging, more than one slice is acquired, and the total scan time may take hours.

Long scan time brings about many restrictions to MRI. For some patients, it is very hard to stay still for a long time, especially seniors and young children. Movement during scanning leads to misrepresentation of protons and artifacts [53]. In some applications like chest imaging or cardiac imaging, the long scan time causes many restrictions for MRI because of the movement of the human body. In addition, MR scanners are very expensive to purchase and to maintain. Reducing scan time can improve the efficiency of diagnosis, which further lowers the cost of each scan. For these reasons, research in the area of scan time reduction in MRI has been very active in recent decades.

Generally, novel scan time reduction methods take advantage of the multi-coil technology implemented in modern MR scanners. These scanners are equipped with phased array of coils, and they have the capability to simultaneously collect data from multiple coils. Each coil in the array has its distinctive sensitivity profile. Ideally, the sum of squares (SOS) of the sensitivity profiles from all coils equals one. An image from each coil emphasizes only one part of the proton density, and the SOS of images from all coils displays the proton density with uniform distribution weights.

To shorten the data acquisition time, many methods have been proposed, including but not limited to following non-Cartesian trajectories, sub-sampling k-space data in Cartesian coordinates, using novel pulse sequences, and combinations of these methods.

### 3.3.2 Non-Cartesian Trajectory

The GRE sequence, multi-slice imaging, and volume imaging described in previous sections all collect data in the Cartesian coordinates. The frequency encoding lines are parallel to each other, and the distances between adjacent frequency encoding lines are the same. The phase encoding direction is perpendicular to the frequency encoding direction. Fig. 3.9 shows the trajectory of a pulse sequence in the Cartesian coordinates.

Images are reconstructed from data acquired with trajectory in the Cartesian coordinate by applying the 2D Fourier transform, which is easy and fast to implement. However, trajectories in the Cartesian coordinates suffer from long scan time equal to the multiplication of the $T_R$ and the total phase encoding steps.
For trajectories in the Cartesian coordinates, frequency encoding lines at the center of k-space determine the brightness and main structure of the image, while frequency encoding lines at the peripheral of k-space define the details of the image. In other words, the importance of different frequency encoding lines across the k-space is not the same. If some frequency encoding lines are dropped out at the periphery of k-space, the quality of the image will not be heavily influenced. However, if some frequency encoding lines are dropped at the center of k-space, the quality of the image will be greatly changed.

Trajectories in the non-Cartesian coordinates normally focus more on the center of k-space than the periphery [54, 55]. Fig. 3.10 shows a radial trajectory of a pulse sequence in the non-Cartesian coordinates. All radial lines pass the center of k-space, and extend to the periphery. The angle between adjacent radial lines is the same across the k-space. As a result, all radial lines are of the same importance in the reconstruction.

Trajectories in the non-Cartesian coordinates require less time to scan k-space data compared with that in the Cartesian coordinates. However, the reconstruction process is more complicated due to its irregular sampling pattern. A gridding method is required in order to interpolate the data point to the Cartesian coordinates before the 2D Fourier transform can be applied [56, 57]. When scanning with trajectories in the non-Cartesian coordinates, more data are acquired in the low frequency part, since all the lines pass through the center of the k-space, while less data are acquired in the high frequency part, because the distance between adjacent lines at the edge is much greater.
3.3 Reconstruction Methods of MRI

Fig. 3.10 Radial trajectory of a pulse sequence in the non-Cartesian coordinates.

Although the time consumption with non-Cartesian trajectories is economical, the reconstruction process is not straightforward. Additional steps are required in order to obtain images from data acquired with non-Cartesian trajectories. As a result, the application of non-Cartesian trajectories is restricted to limited practices. The majority of scans are completed with trajectories in the Cartesian coordinates.

3.3.3 Data Sub-sampling in the Cartesian Coordinates

In the Cartesian coordinates, 128 or 256 phase encoding steps are usually used to form an MR image. To shorten the data acquisition time, some frequency encoding lines are skipped deliberately. The factor by which the number of frequency encoding lines is reduced is referred to as the reduction factor $R$, which is defined by dividing the number of total frequency encoding lines by the number of actual frequency encoding lines in the actual scan. Reconstructed images will have artifacts by directly applying 2D inverse Fourier transform to the sub-sampled k-space data. Specially designed reconstruction methods are required to reconstruct images with acceptable quality from sub-sampled k-space data. SENSE [12], GRAPPA [13], CS-based optimization [20] and their extensions are the main reconstruction methods in MRI.

SENSE

SENSE was proposed to address the problem of reconstructing full FOV images from sub-sampled k-space data of multiple coils by exploiting the sensitivity profile of coils [12]. If the k-space data are sub-sampled, the reconstructed images have aliasing. For example, the image shown in Fig. 3.11 (a) is first transformed to
k-space, and the k-space data are sub-sampled by collecting one out of every two frequency encoding lines. The image reconstructed by directly applying 2D inverse Fourier transform is shown in Fig. 3.11 (b). In the original image, two bright pixels are symmetrical vertically. In the reconstructed image, the two bright points overlap each other and become indistinguishable. In fact, the top and bottom halves are added together to form an 1/2 FOV image.

![Fig. 3.11 Comparison between the original image and the reconstructed image with aliasing. (a) The original image. (b) The reconstructed image with aliasing.](image)

SENSE consists of two steps to unfold the reconstructed image with aliasing. First, the aliased image of each coil is reconstructed by directly applying 2D inverse Fourier transform to the sub-sampled k-space data. Second, the full FOV image is created from these intermediate aliased images.

The process of SENSE is illustrated in the following example. Suppose the two bright pixels in the original image shown in Fig. 3.11 (a) are denoted by \( x(1) \) and \( x(2) \). Also suppose two coils are used in the scanning, and their sensitivities at these two bright pixel locations are \( \Gamma_1(1), \Gamma_1(2) \) for the first coil, and \( \Gamma_2(1), \Gamma_2(2) \) for the second coil, respectively. Let \( \hat{x}_1 \) and \( \hat{x}_2 \) denote the overlapped pixels in the two intermediate aliased images. The aliasing process is described by the following equations:

\[
\Gamma_1(1)x(1) + \Gamma_1(2)x(2) = \hat{x}_1, \tag{3.18}
\]
\[
\Gamma_2(1)x(1) + \Gamma_2(2)x(2) = \hat{x}_2. \tag{3.19}
\]

These equations can be written in matrix form,

\[
\begin{bmatrix}
\Gamma_1(1) & \Gamma_1(2) \\
\Gamma_2(1) & \Gamma_2(2)
\end{bmatrix}
\begin{bmatrix}
x(1) \\
x(2)
\end{bmatrix}
= \begin{bmatrix}
\hat{x}_1 \\
\hat{x}_2
\end{bmatrix}. \tag{3.20}
\]
In 3.20, there are two unknown variables and two linear equations. This problem can be easily solved by multiplying the inverse matrix of the first matrix (if it exists) on the left of Eq. 3.20 to both sides of the equation. The number of coils, $N_c$, should be no less than the reduction factor $R$. If $N_c \geq R$, there are more equations than or the same number of equations as the unknown variables. The unknown variables can be solved by a pseudo-inverse procedure. However, if $N_c < R$, the reconstruction problem becomes under-determined. To solve the equation, the inverse or pseudo-inverse matrix must exist.

The prerequisite of SENSE for successful reconstruction is the accurate sensitivity profiles of coils. The inversion cannot be performed if the sensitivity profiles are unknown. Sensitivity profiles can be obtained by a prescan during which data with a smaller FOV are acquired [12]. However, sensitivity profiles during the actual scan may be different from the prescan to some extent due to many factors, such as patient movement. Therefore, in [58] the authors proposed to obtain the sensitivity profiles during the actual scan, by adopting a variable density sub-sampling pattern, which means collecting more frequency encoding lines the low frequency part of k-space, instead of sub-sampling them the same way as in the high frequency part. Since the calibration data are acquired simultaneously with the data to be reconstructed, errors due to sensitivity mis-calibration are eliminated.

Although these two methods differ in the time when the scanning for sensitivity information is conducted, both share a similar sensitivity estimation procedure. First, a low resolution image for each coil is reconstructed from k-space data obtained by the prescan or the low frequency part. Then the reference image is obtained by the SOS of the sensitivity encoded images from all coils. The sensitivity profile of each coil is calculated by dividing the sensitivity encoded image by the reference image. An interpolation step is finally adopted to increase the size of the sensitivity profiles.

SENSE has been applied successfully in cardiac real-time imaging [59] and other clinical practice [60, 61]. However, the requirement of accurate sensitivity profiles limits its applications where sensitivity profiles are difficult to estimate.

**Extensions of SENSE**

Ever since SENSE was proposed, there have been many extensions reported. In this section, some typical extensions are introduced.

In [62], the authors propose to extend SENSE to 2D sub-sampling, named 2D SENSE. In volume imaging, since there are two phase encoding directions, sub-sampling can be conducted in both directions. As a result, aliasing occurs in two directions. 2D SENSE consists of two consecutive SENSE steps. Each SENSE step
unfolds aliasing in one direction. The term 2D is given as sub-sampling is applied in two dimensions.

Auto-SENSE was introduced in [63] and has been applied in dynamic MRI, such as cardiac imaging. In dynamic MRI, sequential frames are obtained in a time series. In adjacent odd and even frames, one frequency encoding line is shifted when sub-sampling k-space data. As mentioned previously, the low frequency part of k-space data determines the main structure of the image, while the high frequency part of k-space data controls the details. In dynamic MRI, it is assumed that the change between two adjacent frames are considerably small. Therefore, the low frequency parts of k-space data from adjacent frames are considered to be identical or very close, while only the high frequency part of k-space data change greatly. The sub-sampled low frequency parts of adjacent even and odd frames can be combined to extract the sensitivity profiles of coils. Once sensitivity profiles are estimated, SENSE is applied in both of the even and odd frames to reconstruct images. In Auto-SENSE, prescan is avoided, which saves data acquisition time, and sensitivity profiles can be dynamically updated, which leads to fewer artifacts and less noise.

A reconstruction algorithm based on Tikhonov regularization is presented in [64], which reduces the SNR loss due to geometric correlations in the spatial information from the array coil elements. In SENSE with Tikhonov regularization, a reference image is used as prior information about the reconstructed image in order to provide regularized estimates for the reconstruction. The reference image can be obtained by a prescan. It is asserted that this method can reduce g-factors and improve image qualities in SENSE reconstruction.

In typical SENSE, accurate sensitivity profiles must be provided before reconstruction takes place. In [65], the authors propose JSENSE to jointly estimate sensitivity profiles and reconstruct images. The sensitivity profiles are represented by a model with unknown parameters, and the reconstruction problem is further formulated as a nonlinear optimization problem. In JSENSE, a polynomial function is chosen to model the sensitivity profiles. As a result, the estimated sensitivity profiles from ACS lines have a degree of uncertainty, and they are iteratively updated during the reconstruction process.

The combination of SENSE and the CS reconstruction method (introduced later) was introduced in [66], and named CS-SENSE. CS-SENSE asserts that the aliased image is also sparse, like that shown in Fig. 3.11 (b). As a result, the aliased image can be further sub-sampled and reconstructed by CS. CS-SENSE contains two sub-sampling and corresponding reconstruction steps. In the sub-sampling process, first, k-space data is sub-sampled according to a regular sub-sampling pattern. That
is, one of the $R$ frequency encoding lines are acquired. Second, the sub-sampled k-space data is further sub-sampled according to a variable density sub-sampling pattern. That is, the low frequency part of k-space is more densely sub-sampled than the high frequency part of k-space. In the reconstruction process, first, a CS reconstruction method is applied to reconstruct reduced FOV image for each coil. Second, the reduced FOV images from all coils are used as input for the SENSE reconstruction step. In CS-SENSE, the net reduction factor is the product of the reduction factor of two sub-sampling steps.

**GRAPPA**

Another major branch of reconstruction methods from sub-sampled k-space data is GRAPPA [13]. Unlike SENSE, GRAPPA does not require exact information on sensitivity profiles. Instead, GRAPPA requires that the low frequency part of k-space is fully sampled. Another major difference is that SENSE reconstruction is mainly completed in the image space, while GRAPPA reconstruction is mainly finished in the Fourier space.

The reconstruction process of GRAPPA is illustrated in Fig 3.12. The horizontal axis denotes the phase encoding direction, and the vertical axis denotes different coils. The frequency encoding direction is perpendicular to the plane. Hollow circles in the figure indicate frequency encoding lines which are not sampled, while solid (black or gray) circles denote frequency encoding lines which are sampled. The black ones are sampled in the same way as SENSE. That is, one of the $R$ lines is acquired. The gray ones are additional calibration signal (ACS) lines, which are intentionally acquired in GRAPPA. In this example, $R = 2$ and only one ACS line is acquired. In real applications, $R$ can be chosen greater than 2 and normally more than $R - 1$ ACS lines are acquired in order to improve the performance of GRAPPA. More data acquisition time is needed when more ACS lines are acquired.

![Fig. 3.12 GRAPPA reconstruction method.](image-url)
Generally speaking, the GRAPPA reconstruction method contains two steps. In the first step, the GRAPPA weights (or kernels) are determined by an interpolation process. In the example shown in Fig. 3.12, the black data points acquired from all coils in adjacent phase encoding steps are used to interpolate the gray data point in the ACS line of coil 4. The interpolation process is given as

\[ \sum_{i=1}^{4} \sum_{j=1}^{4} s_i(j)w_i(j) = s_{4,ACS}, \]  

(3.21)

where \( i \) iterates coils, \( j \) iterates phase encoding steps, \( s_i(j) \) denotes the acquired black data points in coil \( i \) at location \( j \), \( s_{4,ACS} \) denotes the gray data in the ACS line of coil 4, and \( w_i(j) \) denotes GRAPPA weights. Each black data point is given a weight to balance the left- and right-hand side of the interpolation equation. In this case, there are a total of 16 weights to be determined but only one equation. There are 15 free variables, which means the equation has infinite solutions. The interpolation process is repeated for all data points along the frequency encoding direction. Suppose there are 256 data points in the frequency encoding direction, meaning that there are 256 equations but only 16 unknown variables, which makes the problem of finding GRAPPA weights over-determined. In GRAPPA, it is assumed that the GRAPPA weights across the k-space are the same. In addition to moving along the frequency encoding direction, the same interpolation process is repeated along the phase encoding direction in the fully-sampled low frequency part of k-space. Since there are more equations than unknown variables, the GRAPPA weights can be obtained by a pseudo-inverse step, as in SENSE. With the same procedure, the GRAPPA weights for interpolating the gray data points in the other 3 coils are obtained.

After completing the interpolation step, the missing data points are reconstructed. The missing data points are placed on the right-hand side of Eq. 3.21 instead of the gray data points in the ACS lines. Since the weights and black data points are known, the missing data points can be easily calculated by a pure arithmetic process. Finally, 2D inverse Fourier transform is applied to convert the reconstructed k-space data into an image.

The biggest advantage of GRAPPA compared with SENSE is that GRAPPA does not need the accurate sensitivity profiles of coils. Instead, GRAPPA adopts a variable density sub-sampling pattern which acquires ACS lines in the low frequency part of k-space. It is recommended that in situations where sensitivity profiles can be estimated, SENSE should be considered first.
Extensions of GRAPPA

Since GRAPPA was proposed, many extensions have been reported. In this section, some typical extensions are introduced.

TGRAPPA [67] was proposed to apply GRAPPA in dynamic MRI. In dynamic MRI, k-space data are acquired at different times, which can be regarded as being acquired in a higher dimensional k-t space. In TGRAPPA, no additional ACS lines are acquired. The ACS lines are obtained by combining the acquired frequency encoding lines in the low frequency part of k-space in every $R$ frames. When a new frame is scanned, the oldest frame is dropped. As a result, the ACS lines update as time progresses. With updated ACS lines, GRAPPA weights can be calculated and applied in each frame. The reconstruction process of TGRAPPA is illustrated in Fig. 3.13.

![TGRAPPA reconstruction method](image)

A method called k-t GRAPPA is proposed in [68]. Unlike TGRAPPA, k-t GRAPPA acquires ACS lines in each frame. k-t GRAPPA uses data from adjacent time frames, together with data from adjacent blocks, to interpolate GRAPPA weights. The reconstruction process of k-t GRAPPA is illustrated in Fig. 3.14. Both TGRAPPA and k-t GRAPPA can be used in either a single coil or multiple coils, since they also explore information in time space.

In GRAPPA, acquired data in adjacent phase encoding steps of all coils are used to interpolate one data point in the ACS line of a coil. In [69], the authors propose that not only acquired data in adjacent phase encoding steps of all coils should be used, but also acquired data in the same frequency encoding lines, since they are also very close to the data point in the ACS line. By introducing more data points in the interpolation process, the GRAPPA weights are more accurate.

In GRAPPA, the weights are assumed to be homogeneous everywhere. Once the weights are calculated, they are applied across the whole k-space. In [70], the authors claim that for low and high frequency parts of k-space, the difference
between them should be considered, which means different GRAPPA weights should be calculated and applied in these areas. To obtain better reconstruction results, ACS lines should be segmented along the frequency encoding direction, and GRAPPA weights in different segments should be calculated separately. In addition, out of the ACS area, k-space should be sampled with different reduction factors according to locations.

In [71], GRAPPA is applied to 3D k-space data, and named 2D GRAPPA or 2D GRAPPA Operator, depending on the way of reconstructing image from sub-sampled k-space data in two phase encoding directions. For 3D k-space data, as mentioned previously, sub-sampling along two phase encoding directions can be used to improve the net reduction rate. Like the original GRAPPA, 2D GRAPPA weights can be obtained by interpolating the data points in ACS lines with acquired data in all three directions. This process is named 2D GRAPPA, since it is quite similar to the reconstruction process from sub-sampled k-space in one direction. Alternatively, the reconstruction process can be split into two separate consecutive 1D GRAPPA operations. Each operation reconstructs missing data in one direction. This process is named 2D GRAPPA Operator.

To date, no methods based on GRAPPA have considered the problem of noise. In real scans, acquired data are corrupted by noise, which leads to inaccurate GRAPPA weights. To pick out the outliers, the authors in [72] propose a method called Robust GRAPPA. In this method, the outliers are multiplied with a very small constant or zero, so the weights can be calculated with higher accuracy.

Since GRAPPA weights are assumed to be homogeneous across the whole k-space, it is possible to calculate GRAPPA weights from reconstructed missing lines. In [73], the authors propose that GRAPPA weights should be calculated with the consideration of reconstructed missing data. First, GRAPPA weights are calculated from a few ACS lines, and missing frequency encoding lines are reconstructed. Then
GRAPPA weights are also calculated from the reconstructed missing data. To ensure convergence, the authors also proposed a special method to determine the weights between new and old GRAPPA weights. Since GRAPPA weights are iteratively updated, this method is named iGRAPPA.

The interpolation process of determining GRAPPA weights is time-consuming and computationally heavy. In [74], a random matrix is introduced to multiply on both sides of the interpolation equation. The random matrix has fewer rows than columns, leading to reduced computational cost of the interpolation process. Furthermore, the random matrix preserves the accuracy of GRAPPA weights.

**CS-Based Reconstruction Method**

According to the recently developed CS theory, a signal can be reconstructed with an overwhelming probability from randomly sampled data if the original signal is sparse. In MRI, images are intrinsically sparse, such as angiography, or compressible after certain kinds of transformation, such as brain images. Therefore, it is straightforward to apply CS to the image reconstruction problem from sub-sampled k-space data in MRI.

In [20], the authors propose to reconstruct images by minimizing the $\ell_1$ norm of the transformed image, subject to data fidelity constraints. In the CS-based image reconstruction method, the k-space data is sub-sampled by a variable density sub-sampling technique. Mathematically, the reconstruction problem is expressed as

$$\min \| \Psi x \|_1 \quad \text{s.t.} \quad \| T_F x - y \|_2 < \epsilon, \quad (3.22)$$

where $\Psi$ stands for the wavelet transform operator, $x$ the objective image, $T_F$ the partial Fourier transform operator obtained by sub-sampling the full discrete Fourier matrix with variable density sub-sampling technique, and $y$ the measured data. $\epsilon$ is a parameter relating to noise.

To solve Eq. 3.22, first the constrained problem is converted to the unconstrained Lagrangian form:

$$\min \lambda \| \Psi x \|_1 + \| T_F x - y \|_2^2, \quad (3.23)$$

where $\lambda$ is a regularization parameter to determine the trade-off between sparsity and data fidelity. If appropriately chosen, the solution of Eq. 3.23 can be the same as the solution of Eq. 3.22. $\lambda$ is determined by solving Eq. 3.23 with different values, and then choosing $\lambda$ so that the data fidelity constraint in Eq. 3.22 is satisfied.

The authors propose to solve 3.23 by using a conjugate gradient descent algorithm with backtracking line search. The function $f(x)$ is defined as in Eq. 3.23. The
conjugate gradient of \( f(x) \) is

\[
\nabla f(x) = \lambda \nabla ||\Psi x||_1 + 2T^*_F (T_F x - y). \tag{3.24}
\]

The \( \ell_1 \) norm is defined as the sum of absolute values of entries, which is not smooth for all values of \( x \). Therefore, the \( \ell_1 \) norm is approximated by a smooth function, 

\[
|x| = \sqrt{x^*x + \xi},
\]

where \( \xi \) is a positive smoothing parameter. With this approximation, the conjugate gradient of \( \ell_1 \) norm is expressed as

\[
\nabla ||x||_1 = \frac{x}{\sqrt{x^*x + \xi}}. \tag{3.25}
\]

As a result, the conjugate gradient of \( f(x) \) is further expressed as

\[
\nabla f(x) = \lambda \frac{\Psi^*\Psi x}{\sqrt{(\Psi x)^*(\Psi x) + \xi}} + 2T^*_F (T_F x - y). \tag{3.26}
\]

The conjugate gradient descent algorithm with backtracking line search is briefly described in Algorithm 3 [20].

In [20], it is demonstrated that improved spatial resolution and accelerated acquisition for multi-slice fast spin echo brain imaging and 3D contrast enhanced angiography are achieved.

**Extension of CS-Based Reconstruction Method**

Ever since Lustig’s work [20], there have been many extensions reported. In this section, some typical extensions are introduced.

In [75], the authors review the requirements for successful CS reconstruction. They describe the natural fit of CS to MRI, and give an intuitive understanding of CS reconstruction by describing it as a process of interference cancellation. Four example applications of CS in MRI are given in the paper. They also introduce factors influencing the application of CS in MRI, such as limitations imposed by MR hardware, the characteristics of different types of images, and clinical concerns.

Similar to SENSE and GRAPPA, it is straightforward to apply CS in dynamic MRI. Dynamic MRI offers more sparsity than conventional 2D imaging, by applying Fourier transform along the temporal dimension, assuming that only parts of FOV change at high temporal rate, while other parts remain stationary or change slowly. Random sub-sampling for CS can be realized by randomly skipping frequency encoding lines in each dynamic frame. In [76], the feasibility of CS for accelerating dynamic MRI is assessed.
Generally speaking, CS-based image reconstruction methods from sub-sampled k-space data in MRI are solved by iterative optimization algorithms, such as conjugate gradient descent algorithm [20], and the alternating direction method of multipliers [77]. One major problem of iterative optimization algorithms is the long computational time and heavy computational cost. In [78], the authors propose multi-core CPUs to accelerate computational efficiency. By dispatching multi-channel data to multi-core CPUs, the optimization process is parallelized. In [79], GPUs are used to accelerate the computational efficiency, since modern GPUs have more cores than CPUs. Parallel computing is a promising technique to reduce the reconstruction time of CS-based reconstruction methods when multiple coils are used in MRI.

**Algorithm 3** Conjugate Gradient Descent Algorithm with Backtracking Line Search

**Input:**
- $\lambda$ - regularization parameter
- $\Psi$ - sparsifying transform operator
- $y$ - measured data
- $T_F$ - partial Fourier transform operator
- $\xi$ - positive smoothing parameter
- TolGrad - stopping criteria by gradient magnitude (default $10^{-4}$)
- MaxIter - stopping criteria by number of iterations (default 100)
- $\alpha$ - line search parameter (default $\alpha = 0.05$)
- $\beta$ - line search parameter (default $\beta = 0.6$)
- $i$ - iterator

**Output:**
- $x$ - objective image

**Initialization:**
- $i = 0, x = 0, g^0 = \nabla f(x), d^0 = -g^0$

**Iteration:**
- while ($||g^i|| < $ TolGrad && $i < $ MaxIter) % backtracking line search
  - $t = 1$
  - while ($f(x^i + t d^i) > f(x^i) + \alpha t \text{Real}(g^i \ast d^i)$)
    - $t = \beta t$
    - $x^{i+1} = x^i + t d^i$
    - $g^{i+1} = \nabla f(x^{i+1})$
    - $r = ||g^{i+1}||_2^2$
    - $d^{i+1} = -g^{i+1} + r d^i$
    - $i = i + 1$
3.3.4 Novel Sequence Design

In the above section, the trajectories used are in the Cartesian coordinates, and the encoding method is Fourier transform. Some authors have proposed pulse sequences with trajectories which oscillate from the Cartesian coordinates, and sequences with encoding methods other than the traditional Fourier transform. In this section, several novel sequence designs are reviewed.

Bunched Phase Encoding

In conventional 2D sequences, such as GRE, in order to form an image of size $256 \times 256$, 256 phase encoding steps are implemented. During each $T_R$, 256 data points in the frequency encoding direction are acquired. In [80], the authors propose to acquire data in the frequency encoding direction at a higher rate and along a zigzag/oscillating trajectory. This kind of oscillating trajectory is equivalent to acquiring data in a bunch of phase encoding steps simultaneously, and it is named bunched phase encoding (BPE). The comparison between a straight frequency encoding line in GRE and an oscillating frequency encoding line in BPE is given in Fig. 3.15. It is noted that the black dots in this figure denote separate data points instead of frequency encoding lines as in Fig. 3.12.

![Straight Frequency Encoding Line in GRE](image1)

![Oscillating Frequency Encoding Line in BPE](image2)

Fig. 3.15 Comparison between frequency encoding line in GRE and BPE.

In BPE, more data are acquired in the frequency encoding direction while $T_R$ is not changed. As a result, phase encoding steps can be reduced, which leads to reduced data acquisition time. Normally $R$ times more data are acquired in the frequency encoding direction in one $T_R$. The oscillating trajectory is created by a gradient oscillation, which enables data to be acquired at locations which are offset some distance from the integer locations in the Cartesian coordinates. Normally, the offset is smaller than the distance between two consecutive phase encoding steps. By location oscillation, the BPE sequence equals data acquisition in multiple phase encoding steps in one single phase encoding step. In BPE, one coil is able to offer
3.3 Reconstruction Methods of MRI

reduced data acquisition time and accurate image reconstruction. As a result, it does not require the accurate sensitivity profiles of coils, which are difficult to retrieve in some applications. The image reconstruction process is simple and free from cumbersome iterations.

Although BPE requires only one coil, it can be combined with a phased array of coils to provide even higher reduction factor. In [81], BPE with SENSE is proposed to take advantage of multiple coils. In the first step, the aliased image is constructed from each receiver coil by BPE reconstruction. Then the unaliased image is created by SENSE using the sensitivity profiles of coils.

In BPE, modified sequences and rapidly changing gradients are required to acquire bunched data. These requirements limit its applications in standard MR scanners. In [82], the authors propose the self-calibrating GRAPPA operator gridding (GROG) method to generate bunched data, which can be achieved on MR scanners with a phased array of coils. The GROG method, first proposed in [57], is developed to grid data sampled along a non-Cartesian trajectory. Since GROG is based on GRAPPA, it can only be implemented on scanners with multiple coils. GROG shifts acquired non-Cartesian data to the nearest Cartesian location, thereby converting non-Cartesian data to Cartesian data. In contrast, in GROG-BPE, GROG is used to interpolate data in non-Cartesian locations from data in Cartesian locations, in order to simulate the effect of an oscillating gradient. As a result, a standard straight frequency encoding line is capable of producing bunched data with GROG.

Rotating RF Coil

The RF pulse to flip the net magnetization to the transverse plane is generated by an RF coil, which is usually fixed stationary in the scanner. In [83–85], the authors propose a rotating RF coil system. The RF coil rotates around the object being scanned for spin excitation and signal acquisition. The rotating RF coil system may obviate the need for multiple channels and complex RF decoupling of many coils. The images reconstructed from data with rotating RF coil are comparable to those from data with the conventional RF system.

Wave-CAIPI

In volume imaging, the sensitivity profiles of a particular coil in consecutive slices are quite similar. As a result, the information provided by the sensitivity profiles of coils is not orthogonal, which limits the degree of freedom. A high reduction factor is difficult to achieve in conventional parallel imaging. In [86], the authors propose wave-CAIPI, which introduces sinusoidal phase encoding and partition
encoding gradients simultaneously during the readout of frequency encoding lines. Therefore, inter-slice shifts are created by modifying the phase encoding and the partition encoding strategy. Using this method, highly efficient k-space sampling, which spreads the aliasing evenly in all three dimensions can be achieved. Highly accelerated volume imaging with low artifacts and low SNR penalties is obtained because wave-CAIPI fully uses the spatial variation in coil sensitivities.

### 3.4 Conclusion

In this chapter, the physical background of MRI is first introduced, including the NMR phenomenon, NMR signal excitation, and MR imaging sequences. Then, the reconstruction problem from sub-sampled k-space data is introduced, followed by data acquisition time reduction techniques, such as following the non-Cartesian trajectory, sub-sampling data in the Cartesian coordinates, or designing novel sequences.
Chapter 4

Coherence Analysis of MRI Encoding

4.1 Introduction

In traditional MRI, images of spin density are encoded with Fourier encoding. It is known that most of the energy of the scanned data is located in the low-frequency part of k-space. Therefore, phase encoding steps in the low frequency part of k-space are of more importance than that in the high frequency part of k-space. It is common practice to use variable density sub-sampling together with CS-based reconstruction methods to recover images from sub-sampled k-space data [20]. By using variable density sub-sampling, more phase encoding steps are selected from the low frequency part of k-space, while fewer are selected from the high frequency part of k-space. One major problem of variable density sub-sampling is that, although most of the energy is conserved, the details of images are lost. The reason for this problem is that data from the low frequency part of k-space determine the major structure and brightness of images, while data from the high frequency part of k-space control the details of images.

In [21], the authors propose the use of noiselet encoding to replace the traditional Fourier encoding in MRI. The k-space data is flattened when noiselet encoding is used, which means that each frequency encoding line has the same importance in k-space. In other words, each frequency encoding line has the same amount of information about the structure and details of images. It is suggested to randomly select frequency encoding lines, and to reconstruct images using CS-based reconstruction methods.

In this chapter, non-Fourier encoding in MRI is first reviewed. Then the coherence analysis is presented. Finally, the coherence results are presented and discussed.
4.2 Non-Fourier Encoding in MRI

4.2.1 Sequence Design

As demonstrated in Chapter 3, a pulse sequence contains an RF pulse, a slice selection gradient, a phase encoding gradient, a frequency encoding gradient and a signal. The RF pulse together with the slice selection gradient is used to select spins in a slice with a specific thickness from a volume. To avoid interference from adjacent slices, the RF pulse is designed to be a sinc function, since it has a square frequency distribution. A square frequency distribution has very sharp slopes, which prevent interference from adjacent slices with similar Larmor frequencies. The sinc function is defined as

\[
sinc(t) = \begin{cases} 
1 & t = 0, \\
\frac{\sin(\pi t)}{\pi t} & \text{otherwise.}
\end{cases}
\] (4.1)

The waveform of a sinc function in the time domain is shown in Fig. 4.1.

![Fig. 4.1 Waveform of a sinc function in the time domain.](image)

It has been proposed that by introducing a spatially selective RF excitation, encoding methods other than Fourier encoding can be achieved, such as SVD encoding [87], wavelet encoding [88], non-Fourier encoding [89], random encoding [90], and noiselet encoding [21]. In theory, the spatial profile of arbitrary shape can be excited by modulating the envelope of the RF excitation pulse [91, 92]. If these
spatial excitation profiles are functions of a non-Fourier orthogonal basis, such as those just cited, the field of view along the encoding direction is encoded on that basis, and can be reconstructed by applying the appropriate inverse transformation. When the flip angle is small, designing the excitation profile is as simple as making the RF pulse envelope equal to the Fourier transform of the desired profile [91, 92]. Therefore, to encode by the noiselet transform in MRI is to make the envelope of the RF pulse equal to the Fourier transform of its corresponding matrices.

Fig. 4.2 shows a timing diagram of the spin echo sequence, which is commonly used in scanning. One advantage of the spin echo sequence is that it produces $T_2$-dependent signal. In the human body, some tissues have similar $T_1$ values but different $T_2$ values. As a result, it is advantageous to have an imaging sequence which produces images with $T_2$-dependence, so that different tissues can be differentiated.

In the timing diagram shown in Fig. 4.2, first a slice selective RF pulse with $\theta$ degree is applied together with a slice selection gradient to select spins from a slice from a volume. $\theta$ can be chosen from a very small angle to 90°. At the time $T_E/2$, a slice selective 180° RF pulse is applied together with a slice selection gradient to select the slice again.

Between the $\theta$ and 180° RF pulses, the phase encoding gradient is applied, which varies between the maximum and minimum value in 128 or 256 steps. Although the phase encoding gradient can be applied after the 180° RF pulse, it is applied before in order to minimize the $T_E$ time period.

The frequency encoding gradient is applied after the 180° RF pulse, during which the echo is generated and collected. An additional frequency encoding gradient is applied between the $\theta$ and 180° RF pulses, which is used to dephase the spins so
that they will rephase and generate an echo at the center of the readout window. The FID generated after the first RF pulse is discarded. This process is repeated every $T_R$ seconds until all the phase encoding steps have been recorded.

The pulse sequence of noiselet encoding is designed based on the spin echo sequence. As mentioned previously, non-Fourier encoding techniques are achieved by modulating the envelope of the first RF pulse. In Fig. 4.3, the timing diagram of a non-Fourier pulse sequence based on the spin echo sequence is shown.

![Fig. 4.3 Timing diagram of a non-Fourier pulse sequence based on the spin echo sequence.](image)

In the timing diagram shown in Fig. 4.3, first a spatially selective RF pulse is applied together with the phase encoding gradient to encode the phase encoding direction into the desired encoding domain. The envelope of the spatially selective RF pulse is designed by taking the Fourier transform of the desired encoding matrix. The spatially selective RF pulses differ from each other in different phase encoding steps. To achieve a specific small flip angle, the integral of each individual RF pulse over time must equal that of the sinc RF pulse used in the traditional Fourier encoding for the same flip angle. The phase encoding gradient varies between the minimum and maximum value in 128 or 256 steps.

Unlike in the spin echo sequence, the spatially selective RF pulse does not select a slice in the non-Fourier encoding technique introduced in this section. The slice selection process is achieved by the $180^\circ$ RF pulse and the slice selection gradient applied at $T_E/2$.

The frequency encoding gradient is applied after the $180^\circ$ RF pulse, during which the echo is generated and collected. An additional frequency encoding gradient is applied between the spatially selective RF pulse and the $180^\circ$ RF pulse, which is used to dephase the spins so that they rephase and generate an echo at the center
4.2 Non-Fourier Encoding in MRI

of the readout window. The FID generated after the spatially selective RF pulse is
discarded. This process is repeated every \( T_R \) seconds until all the phase encoding
steps have been recorded.

4.2.2 Fourier Encoding

In traditional MRI, data are collected in the Fourier domain, or k-space. The process
of data acquisition equals applying a 2D Fourier transform on an image representing
the spin density in the field of view. The 1D Fourier transform applied on the left-
hand side of the image matrix encodes phase, while that applied on the right-hand
side of the image encodes frequency. The conventional Fourier encoding is expressed
as

\[
y = \tilde{T}_F x, \tag{4.2}
\]

where \( \tilde{T}_F \) stands for the full Fourier encoding operator, \( x \) is the image representing
the spin density, and \( y \) is the scanned data. Most of the energy of the k-space data is
in the low frequency part, while the high frequency part contains very limited energy.
For example, a phantom image is shown in Fig. 4.4 (a), and the Fourier-encoded
image is shown in Fig. 4.4 (b).

![Fig. 4.4 The principle of Fourier encoding. (a) Phantom image. (b) Fourier-encoded image.](image)

Fig. 4.4 (b) shows that the low frequency part of k-space is bright while the rest
is dark, which means most of the energy is constrained at the center. As a result, the
CS-based image reconstruction method requires that more frequency encoding lines
are sampled in the low frequency part, while fewer are sampled in the high frequency
part. Therefore, variable density sub-sampling is applied in the traditional MRI with
Fourier encoding. One major problem of variable density sub-sampling is that the
details of images are lost although the main structures are conserved.
4.2.3 Noiselet Encoding

To overcome this disadvantage, it is advised to spread the energy evenly across all phase encoding steps in k-space. This can be achieved by replacing the traditional Fourier encoding with noiselet encoding.

Noislets [93] are "noise-like" functions which are totally uncompressible by orthogonal wavelets. They are constructed via a multiscale iteration in exactly the same way as wavelets, but with a difference. Wavelets are constructed by translates and dilates of the mother bases function, while noiselets are constructed by twisting the translates and dilates of the mother bases function. The mother bases function $\chi(x)$ of noiselets is defined as

$$\chi(x) = \begin{cases} 1, & x \in [0,1) \\ 0, & \text{otherwise.} \end{cases} \quad (4.3)$$

The family of noiselet basis functions is generated in the interval $[0,1)$ as

$$f_1(x) = \chi(x),$$

$$f_{2n}(x) = (1-i)f_n(2x) + (1+i)f_n(2x-1), \quad (4.4)$$

$$f_{2n+1}(x) = (1+i)f_n(2x) + (1-i)f_n(2x-1),$$

where $i = \sqrt{-1}$ and $f_{2n}, f_{2n+1}$ form the unitary basis for vector space $V_n$. A $4 \times 4$ noiselet matrix is shown below

$$\frac{1}{2} \begin{bmatrix} 0-1i & 1+0i & 1-0i & 0+1i \\ 1+0i & 0+1i & 0-1i & 1-0i \\ 1-0i & 0-1i & 0+1i & 1+0i \\ 0+1i & 1-0i & 1+0i & 0-1i \end{bmatrix}. \quad (4.5)$$

By replacing Fourier encoding with noiselet encoding, it is possible to spread the energy of k-space evenly across all phase encoding steps. The noiselet encoding process can be written as

$$y = \tilde{T}_N x, \quad (4.6)$$

where $\tilde{T}_N$ stands for the full noiselet encoding operator.

For example, a phantom image is shown in Fig. 4.5 (a), and the noiselet-encoded image is shown in Fig. 4.5 (b). Noiselet transform is applied on the vertical direction, while Fourier transform is applied on the horizontal direction. This kind of encoding is named 1D noiselet encoding. The figure shows that the energy is spread evenly across all phase encoding steps in k-space by incorporating noiselet encoding.
4.2 Non-Fourier Encoding in MRI

Fig. 4.5 The principle of noiselet encoding. (a) Phantom image. (b) Noiselet-encoded image.

4.2.4 Image Reconstruction Problem

As demonstrated in Chapter 3, when using CS-based reconstruction methods to recover images from sub-sampled k-space data, Problem 3.22 is chosen to be solved when the number of coils is 1. Phased array of coils are commonly provided in modern MRI, which enable simultaneous data acquisition from multiple coils. The sensitivity profiles of coils are different from each other, and can be used as additional information to facilitate the reconstruction of images from sub-sampled k-space data. Ideally, the sum of squares of sensitivity profiles of all coils at any location equals one. An example of artificial sensitivity profiles when the number of coils $N_c = 4$ is shown in Fig. 4.6.

In Fig. 4.6, the sensitivity profiles of coils change gradually and are different from each other. The bright part of the sensitivity profile indicates the region which is emphasized in the sensitivity encoded image, while the clarity is reduced in the sensitivity encoded image in the region where the sensitivity profile is dark. Since the sensitivity profiles of coils are different from each other, different parts of the image are emphasized when phased array of coils are applied. For example, if coils with the sensitivity profiles shown in Fig. 4.6 are applied when scanning the phantom image, the sensitivity encoded images are shown in Fig. 4.7. It is obvious that different parts of the image are emphasized in different sensitivity encoded images.
Coherence Analysis of MRI Encoding

The data acquisition process of multiple coils is denoted by

\[ y = \begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_{N_c} \end{pmatrix} = \begin{pmatrix} T_s(\Gamma_1 \odot x) \\ \vdots \\ T_s(\Gamma_i \odot x) \\ \vdots \\ T_s(\Gamma_{N_c} \odot x) \end{pmatrix} = \hat{T}_s x, \quad i = 1, 2, \ldots, N_c, \quad (4.7) \]

where \( y_i \) stands for the sub-sampled data from coil \( i \), \( \Gamma_i \) the sensitivity profile of coil \( i \), \( T_s \) a selected encoding technique with the sub-sampling process applied, \( \odot \) the element-wise multiplication, and \( \hat{T}_s \) the whole sub-sampling and encoding process in the second brackets. Therefore, the reconstruction problem when multiple coils are used is expressed as

\[ \min \| \Psi^T x \|_1 \quad \text{s.t.} \quad \| \hat{T}_s x - y \|_2 < \epsilon. \quad (4.8) \]
This problem can be solved via the conjugate gradient descent algorithm with backtracking line search illustrated in Chapter 3.

In [21], it is demonstrated that by using noiselet encoding, better reconstruction results are obtained compared with the conventional Fourier encoding in 2D MRI. Coherence and phase transition analysis are conducted in this and the following chapters in order to better understand the benefits of spreading the energy of the k-space data across all phase encoding steps.
4.3 Coherence Analysis

4.3.1 Coherence Expression

In this section, the coherence of CS-based reconstruction methods from sub-sampled k-space data encoded by Fourier encoding and noiselet encoding is analyzed.

In 2D MRI, only the phase encoding direction can be sub-sampled, which is accomplished by skipping some phase encoding steps during the data acquisition process. In contrast, the frequency encoding direction is always fully sampled. Therefore, it is common practice to first apply an inverse Fourier transform to the sub-sampled k-space data to transform the frequency encoding direction into the image domain. As a result, the k-space data are transformed to hybrid-space data. Since sub-sampling is only along the phase encoding direction, the 2D image reconstruction problem can be decomposed into many 1D rows of images (or columns of images, depending on the phase encoding direction) reconstruction problem, although sometimes the correlations between rows (or columns) are considered in order to improve the reconstruction results. Therefore, in the coherence analysis, only 1D sparse signal is considered.

Suppose $x \in \mathbb{C}^N$ is a 1D signal and is sparse on dictionary $\Psi \in \mathbb{C}^{N \times N}$. Mathematically,

$$x = \Psi s,$$

where $s \in \mathbb{C}^N$ is the transformed sparse signal. In $s$, only $K$ out of $N$ entries are non-zero, while the rest are zero. $K$ is called the sparsity of $s$. $s$ is considered compressible if only $K$ out of $N$ entries are significantly large, while the rest are negligible. CS asserts that if $x$ is sparse or compressible, naturally or after a specified transformation, it can be reconstructed with an overwhelming probability from a limited number of measurements by solving the following problem:

$$\min \|\Psi^T x\|_1 \quad \text{s.t.} \quad y = \Phi x = \Phi \Psi s,$$

where $y \in \mathbb{C}^M$ is the measured signal, and $\Phi \in \mathbb{C}^{M \times N}$ is the sensing matrix, with $M \ll N$.

Coherence serves as a qualitative characterization of the degree of similarity between each column of the dictionary [18, 19]. In 4.10, $\Phi \Psi$ serves as a new dictionary, which is denoted later by $\hat{\Psi}$. Similarly as denoted in 2.28, the coherence
of the new dictionary \( \hat{\Psi} \) is defined as

\[
\mu\{\hat{\Psi}\} = \max_{1 \leq i, j \leq N, i \neq j} \frac{|\hat{\psi}_i^T \hat{\psi}_j|}{||\hat{\psi}_i|| \cdot ||\hat{\psi}_j||},
\]

(4.11)

where \( \hat{\psi}_i \) is the \( i \)th column of \( \hat{\Psi} \). Coherence can also be obtained by calculating the Gram Matrix, which is defined as

\[
G = \tilde{\hat{\Psi}}^T \tilde{\hat{\Psi}},
\]

(4.12)

where \( G \in \mathbb{C}^{N \times N} \) is the Gram Matrix, and \( \tilde{\hat{\Psi}} \in \mathbb{C}^{N \times N} \) is \( \hat{\Psi} \) with each column normalized. The maximum magnitude of the off-diagonal elements of \( G \) is defined as the coherence. If \( \mu\{\hat{\Psi}\} = 1 \), at least two columns of the dictionary are the same or proportional, which means that there are redundant columns in the dictionary. In contrast, if \( \mu\{\hat{\Psi}\} = 0 \), it means that any two columns of the dictionary are mutually incoherent.

If the following inequality is satisfied

\[
||s||_0 < \frac{1}{2} \left(1 + \frac{1}{\mu\{\hat{\Psi}\}}\right),
\]

(4.13)

the sparse signal \( s \) can be recovered from the measured data \( y \), as can the original signal \( x \). Therefore, in the problem of recovering a sparse signal \( s \) with sparsity \( K \) from the measured data \( y \) with length \( M \), a smaller \( \mu\{\hat{\Psi}\} \) will lead to larger \( K \). As a result, \( \mu\{\hat{\Psi}\} \) should be as small as possible.

In MRI, images often contain complicated patterns and great detail. When CS-based reconstruction methods are used, these images are assumed to be sparse in the wavelet domain, as shown in Fig. 2.1. Therefore, in the coherence analysis simulation, we assume that the original 1D signal \( x \) is also sparse in the wavelet domain. That is,

\[
x = Ws, \text{ or } s = W^T x.
\]

(4.14)

Here, \( W \) denotes the wavelet transform.

In modern MRI scanners, data are simultaneously collected from a phased array of coils, each with a specific and distinctive sensitivity profile. In the coherence analysis simulation, we also suppose that there are \( N_c \) coils, and the sensitivity profile of coil \( i \) is denoted by a 1D vector \( \Gamma_i \in \mathbb{C}^N, i = 1, 2, \cdots, N_c \). Fourier encoding is commonly used in modern MRI scanners. With Fourier encoding, the sub-sampled
Coherence Analysis of MRI Encoding

k-space data of coil $i$ is written as

$$y_i = \mathbf{T}_F(\Gamma_i \odot \mathbf{x}) = \mathbf{T}_F \hat{\Gamma}_i \mathbf{x} = \mathbf{T}_F \hat{\Gamma}_i \mathbf{W}s,$$

where $\odot$ denotes entry-wise multiplication, and the diagonal of matrix $\hat{\Gamma}_i \in \mathbb{C}^{N \times N}$ is $\Gamma_i$. The sub-sampled k-space data from all coils is written as

$$
\begin{pmatrix}
y_1 \\
\vdots \\
y_i \\
\vdots \\
y_{N_c}
\end{pmatrix} =
\begin{pmatrix}
\mathbf{T}_F \hat{\Gamma}_1 \mathbf{W} \\
\vdots \\
\mathbf{T}_F \hat{\Gamma}_i \mathbf{W} \\
\vdots \\
\mathbf{T}_F \hat{\Gamma}_{N_c} \mathbf{W}
\end{pmatrix} \mathbf{s} = \hat{\mathbf{T}}_F \mathbf{s},
$$

where $\hat{T}_F$ denotes the operation in the second bracket.

Similarly, if noiselet encoding is selected, the sub-sampled k-space data from all coils is written as

$$
\begin{pmatrix}
y_1 \\
\vdots \\
y_i \\
\vdots \\
y_{N_c}
\end{pmatrix} =
\begin{pmatrix}
\mathbf{T}_N \hat{\Gamma}_1 \mathbf{W} \\
\vdots \\
\mathbf{T}_N \hat{\Gamma}_i \mathbf{W} \\
\vdots \\
\mathbf{T}_N \hat{\Gamma}_{N_c} \mathbf{W}
\end{pmatrix} \mathbf{s} = \hat{\mathbf{T}}_N \mathbf{s},
$$

where $\hat{T}_N$ denotes the operation in the second bracket. In the special case when $N_c = 1$, Eq. 4.16 is reduced to

$$y = \mathbf{T}_F \mathbf{x} = \mathbf{T}_F \mathbf{W}s = \hat{\mathbf{T}}_F \mathbf{s},$$

where $\hat{T}_F = \mathbf{T}_F \mathbf{W}$. And Eq. 4.17 is reduced to

$$y = \mathbf{T}_N \mathbf{x} = \mathbf{T}_N \mathbf{W}s = \hat{\mathbf{T}}_N \mathbf{s},$$

where $\hat{T}_N = \mathbf{T}_N \mathbf{W}$.

In order to determine which encoding method is better, $\mu\{\hat{T}_F\}$ is compared with $\mu\{\hat{T}_N\}$. Here, Haar wavelet is chosen since it has been proven that noiselets and Haar wavelets are perfectly incoherent [94].

For signals which are naturally sparse, the dictionary $\Psi$ is an identity matrix, i.e., $\mathbf{x} = \mathbf{s}$. In this case, normally partial Fourier matrix is used as the sensing matrix in
CS. Therefore, the measured data $y$ is written as

$$y = T_F x = T_F s,$$  \hspace{1cm} (4.20)

where $T_F \in \mathbb{C}^{M \times N}$ is the partial Fourier matrix, consisting of $M$ selected rows of the full DFT matrix. When a partial noiselet matrix is used as the sensing matrix, the measured data $y$ is written as

$$y = T_N x = T_N s,$$  \hspace{1cm} (4.21)

where $T_N \in \mathbb{C}^{M \times N}$ is the partial noiselet matrix, consisting of $M$ selected rows of the full noiselet matrix. To determine which sensing matrix is better, $\mu \{ T_F \}$ is compared with $\mu \{ T_N \}.$

### 4.3.2 Sub-sampling

In MRI, an important challenge is the relatively long data acquisition time. One solution to this challenge is to sub-sample the k-space data, which is accomplished by skipping some phase encoding lines deliberately during the data acquisition process, and then reconstructing images from sub-sampled k-space data using specially-designed reconstruction methods. To reconstruct images by CS-based reconstruction methods from sub-sampled k-space data encoded by Fourier encoding, variable density sub-sampling is chosen in order to select more frequency encoding lines in the low frequency part of k-space, and fewer in the high frequency part of k-space. When the k-space data are encoded by noiselet encoding, random sub-sampling is chosen, since the energy in k-space is flattened evenly across all phase encoding steps. In random sub-sampling, each frequency encoding line is collected with a fixed probability.

In coherence analysis, both of the sub-sampling patterns are chosen to select from the fully sampled data. For random sub-sampling, each phase encoding step is collected with a fixed probability. For variable density sub-sampling, the phase encoding steps in the low frequency part are sub-sampled with a high probability, while the phase encoding steps in the high frequency part are sub-sampled with a low probability. It is easy to understand that variable density sub-sampling may have different forms according to how much is defined as the low frequency part, and how much is sampled in the low/high frequency part. The ratio between the length of the original signal, $N$, to the length of the sub-sampled data, $M$, is named the reduction factor, $R$. An example of random sub-sampling and variable density sub-sampling
Coherence Analysis of MRI Encoding

is shown in Fig. 4.8, where white rows denote selected frequency encoding lines, while black rows denote skipped frequency encoding lines. In all four sub-figures, 128 out of 256 lines are selected.

(a) (b) (c) (d)

Fig. 4.8 Sub-sampling techniques. (a) Random sub-sampling. (b), (c) and (d) Variable density sub-sampling.

In Fig. 4.8 (a), each row is selected with a fixed probability 0.5. The probability is fixed across all rows from top to bottom. In Fig. 4.8 (b)-(d), rows at the center are sub-sampled more densely than those at the top or bottom, which is the reason why it is named variable density sub-sampling. In these three sub-figures, 128 out of 256 lines are selected, according to the different forms. The ratio of the low frequency part in Fig. 4.8 (b) is the same as that in Fig. 4.8 (c). However, they differ in how much is sampled in the low/high frequency part. In Fig. 4.8 (b) and (c), the ratio of the low frequency part is higher than that in Fig. 4.8 (d). At least two parameters determine the behavior of variable density sub-sampling. Selection of the appropriate variable density sub-sampling parameters is a cumbersome task which requires trial and error.
4.4 Results

4.3.3 Sensitivity Profile

Modern MRI scanners are commonly equipped with a phased array of coils to simultaneously collect data in k-space. Each coil has its own sensitivity profile, providing additional information for reconstruction. These sensitivity profiles are different from each other and gradually change according to their spatial locations. Ideally, the sum of squares of these sensitivity profiles at any location equals one.

In coherence analysis, two cases of sensitivity profiles are used in the simulation, including one that changes slowly and gradually, and one which changes quickly and suddenly. When $N_c = 4$, the two cases are shown in Fig. 4.9 and 4.10. In both cases, the sum of squares of all $N_c$ coils at any location equals one.

4.4 Results

4.4.1 Sparse Signal in Wavelet Domain

Coherence serves as a qualitative measure of the sensing matrix, or from another perspective, a qualitative measure of the sparsifying dictionary. In this section, the simulation results of $\mu \{\hat{T}_F\}$ and $\mu \{\hat{T}_N\}$ are presented. Signals are assumed to be sparse in the wavelet domain, such as MR images.

The coherence between the full encoding matrix and the sparsifying matrix is calculated before sub-sampling is applied. The multiplication of the full encoding matrix and the sparsifying matrix serves as a new dictionary. The coherence of the new dictionary is calculated by the Gram Matrix. The Gram Matrices of the new dictionary constructed by the full Fourier matrix and the wavelet sparsifying matrix, and by the full noiselet matrix and the wavelet sparsifying matrix are shown in Fig. 4.11. In both of the sub-figures in Fig. 4.11, the diagonal entries are 1 while the off-diagonal entries are 0. Since coherence is defined as the maximum absolute value of the non-diagonal entries of the Gram Matrix, the coherence values of the two encoding methods are 0. This means that the columns of the new dictionary are mutually incoherent, which makes the dictionary suitable for CS reconstruction.

When sub-sampling is applied during the data acquisition process, specified rows from the full encoding matrix are chosen to form the partial encoding matrix, leading to different coherence performances. In the following section, sub-sampling is considered in coherence simulation.
Fig. 4.9 Slowly and gradually changing sensitivity profiles. \( N = 256. \)

Fig. 4.10 Quickly and suddenly changing sensitivity profiles. \( N = 256. \)
4.4 Results

Fig. 4.11 Gram Matrix of new dictionary constructed by different matrices. (a) The full Fourier matrix and the wavelet sparsifying matrix. (b) The full noiselet matrix and the wavelet sparsifying matrix.

Fourier Encoding

In the case of Fourier encoding, variable density sub-sampling is normally chosen to sub-sample phase encoding steps, despite the difficulties to determine the best parameters. In contrast, random sub-sampling is seldom considered in Fourier encoding. First, a comparison of the coherence of Fourier encoding with variable density sub-sampling and that with random sub-sampling is made.

Fig. 4.12 shows the coherence $\mu \{ \hat{T}_F \}$ with random sub-sampling when $N = 256$. The horizontal axis denotes the ratio between the length of the sub-sampled data, $M$, to the length of the original signal, $N$, i.e., $1/R$. The ratio is from 0.04 to 0.96, with a step of 0.04. The vertical axis denotes the coherence, with 0 meaning totally incoherent, and 1 meaning maximally coherent. At each ratio along the horizontal axis, sub-sampling is randomly generated 500 times, and the corresponding coherence values are calculated. The average value of 500 results is displayed in the figure. In the simulation, the number of coils are chosen to be 1, 4, 8, and 16. For the latter three cases, two different coil sensitivity profiles are used for each case. When the coil number is 1, the sensitivity profile is assumed to be even across all locations.

Two obvious observations can be obtained from Fig. 4.12. First, the coherence value decreases when the reduction factor $R$ decreases, or $1/R$ increases. When $1/R$ approaches close to 1, the coherence value drops close to 0. As illustrated above, the coherence value is 0 if no sub-sampling is applied. This simply implies that the reconstruction results are better if more data are sampled. In contrast, if $1/R$ decreases close to 0, the coherence value will be close to 1, which means
the columns of the new dictionary are coherent. The reconstruction result will be greatly influenced. Second, by comparing Fig. 4.12 (a) and (b), it is seen that the coherence value of multiple coils with quickly changing sensitivity profiles are smaller than those with slowly changing sensitivity profiles, in the middle range of $1/R$. This indicates that there is room for improvement in the coherence value and reconstruction results for partial Fourier encoding with random sub-sampling.

The same simulation was repeated, except that the length of signal $N = 1024$. The results are shown in Fig. 4.13. The same observations can be made as for Fig. 4.12, since the coherence lines in the two figures are almost the same. The coherence results imply that random sub-sampling may not be the optimal sub-sampling choice for Fourier encoding.

Variable density sub-sampling is commonly selected to collect phase encoding steps from the fully sampled data encoded with Fourier encoding. It has been demonstrated in many publications that successful reconstructions can be obtained from sub-sampled k-space data encoded with Fourier encoding and sub-sampled by variable density sub-sampling. Fig. 4.14 shows the coherence $\mu\{\hat{T}_F\}$ with variable density sub-sampling when $N = 256$. 

---

Fig. 4.12 Coherence results of $\mu\{\hat{T}_F\}$ with random sub-sampling when $N = 256$. 
(a) Map 1. (b) Map 2.
Fig. 4.13 Coherence results of $\mu\{\hat{T}_F\}$ with random sub-sampling when $N = 1024$. (a) Map 1. (b) Map 2.

Fig. 4.14 Coherence results of $\mu\{\hat{T}_F\}$ with variable density sub-sampling when $N = 256$. (a) Map 1. (b) Map 2.
The same simulation arrangements as in random sub-sampling were implemented. As mentioned previously, there are different forms of variable density sub-sampling patterns depending on the definition of the low frequency part and how much is sampled in the low/high frequency part. All possible parameter combinations are tried and the pair which produces the lowest coherence value is chosen in the simulation. The optimal coherence values are presented in Fig. 4.14.

A comparison of Fig. 4.14 with Fig. 4.12 makes it clear that the coherence values of Fourier encoding with optimal variable density sub-sampling are much smaller than that with random sub-sampling, which means that reconstruction results will be better when variable density sub-sampling is used in Fourier encoding. The coherence values of Fourier encoding with sub-optimal variable density sub-sampling will locate between the two groups of lines in the two figures, indicating a performance change between the two extreme cases.

The same simulation as in Fig. 4.14, except that $N = 1024$, was conducted, and the results are shown in Fig. 4.15. The same conclusion can be drawn by comparing Fig. 4.15 with Fig. 4.13. In sum, the coherence between partial Fourier encoding and wavelet sparsifying matrix is smaller when variable density sub-sampling is used other than random sub-sampling.

Fig. 4.15 Coherence results of $\mu \{ \hat{T}_F \}$ with variable density sub-sampling when $N = 1024$. (a) Map 1. (b) Map 2.
4.4 Results

**Noiselet Encoding**

Most of the energy of the measured data encoded with Fourier encoding is in the low frequency part of k-space. As a result, variable density sub-sampling is required in order to collect more phase encoding steps in the low frequency part of k-space. Noiselet encoding is applied in order to spread the energy evenly across all phase encoding steps. As a result, each frequency encoding line has the same importance. Random sub-sampling instead of variable density sub-sampling is chosen to select phase encoding steps from the fully sampled data. The same simulation arrangements as in coherence simulation of Fourier encoding are used in the coherence simulation of noiselet encoding with random sub-sampling. The coherence results are shown in Fig. 4.16 when the length of signal \( N = 256 \).

![Coherence Results](image)

**Fig. 4.16** Coherence results of \( \mu \{ \hat{T}_N \} \) with random sub-sampling when \( N = 256 \). (a) Map 1. (b) Map 2.

By comparing coherence results in Fig. 4.12 with those in Fig. 4.16, it is obvious that the coherence lines for noiselet encoding with random sub-sampling are much lower than those for Fourier encoding with random sub-sampling. It is clear that by spreading the energy of the k-space data across all phase encoding steps, the coherence of the dictionary is dramatically decreased, providing a better guarantee of successful CS reconstruction.
A further comparison was made of the coherence results in Fig. 4.14 and in Fig. 4.16. All coherence lines share the same trend, which indicates that similar reconstruction results can be obtained for both cases. A quantitative coherence comparison is made later in this chapter. Phase transition and reconstruction results are given in following chapters.

The coherence results of noiselet encoding and random sub-sampling are shown in Fig. 4.17 when the length of signal $N = 1024$. By comparing coherence values in Fig. 4.13, Fig. 4.15 and Fig. 4.17, the same conclusion can be obtained as in the case when $N = 256$.

Fig. 4.17 Coherence results of $\mu \{ \tilde{T}_N \}$ with random sub-sampling when $N = 1024$. (a) Map 1. (b) Map 2.

**Coherence Comparison**

In the above sections, the simulation results demonstrate that

- for Fourier encoding, the coherence value is much smaller when variable density sub-sampling is used instead of random sub-sampling;

- for noiselet encoding, random sub-sampling is capable of producing similar coherence values as the optimal variable density sub-sampling for Fourier encoding.
In this section, a quantitative coherence comparison is made between the two encoding methods with their corresponding optimal sub-sampling methods. When the length of signal $N = 256$, the coherence lines of different coils are close to each other in Fig. 4.14 and Fig. 4.16. Therefore, the coherence when the number of coils $N_c = 1$ was chosen for comparison. The same is true when $N = 1024$. Fig. 4.18 shows the coherence comparison of the two encoding methods with their corresponding optimal sub-sampling methods.

![Coherence comparison](image)

Fig. 4.18 Coherence comparison of Fourier encoding with variable density sub-sampling, and noiselet encoding with random sub-sampling. (a) $N = 256$. (b) $N = 1024$.

First, a comparison the coherence results of Fourier encoding and noiselet encoding indicates that the coherence value of noiselet encoding is smaller than that of Fourier encoding at the same $1/R$. As mentioned previously, the "optimal" variable density sub-sampling for Fourier encoding is chosen by trial and error. Unfortunately, due to the limited number of trials for each parameter pair and the limited number of parameter pairs tried, the term "optimal" is difficult, and theoretically impossible, to obtain. Although this may be taken as reducing the validity of the coherence results, the truth is that optimal variable density sub-sampling is also difficult to find and implement in real scanning. This difficulty in turn makes the simulation valid for comparison. Optimal variable density sub-sampling tries to balance the
low frequency part and the high frequency part, in order to preserve both the main structure and the details of images. However, the balance is difficult to maintain.

In noiselet encoding, the situation is dramatically simplified. Since the energy of the k-space data is spread evenly across all phase encoding steps, each frequency encoding line has the same importance. In other words, each frequency encoding line contains the same amount of information about the main structure and details of images. Therefore, random sub-sampling is applied for noiselet encoding. The coherence value is even smaller than that of Fourier encoding with optimal variable density sub-sampling, which means that noiselet encoding with random sub-sampling is more suitable for CS reconstruction.

4.4.2 Sparse Signal in Spatial Domain

In this section, coherence simulation when the original signal is sparse in the spatial domain is conducted.

![Graphs showing comparison between Fourier- and noiselet-transformed signal.](image)

Fig. 4.19 Comparison between Fourier- and noiselet-transformed signal. (a) A randomly-generated sparse signal with length $N = 256$ and sparsity $K = 16$. (b) The Fourier transform of the original signal. (c) The noiselet transform of the original signal.

Since the original signal is naturally sparse, there is no need to sparsify it in another dictionary. Alternatively, the sparsifying matrix can be chosen as the identity
matrix. A randomly-generated sparse signal is shown in Fig. 4.19 (a) with length $N = 256$ and sparsity $K = 16$. The sub-figures (b) and (c) show data after Fourier transform and noiselet transform, respectively.

If the original signal is sparse in the spatial domain, both the real and the imaginary parts of its Fourier transformed data fall in a limited range, except that the central datum of the real part contains the 0 frequency of the original signal, which normally has the maximum magnitude. Similarly, both the real and the imaginary parts of its noiselet transformed data fall in a limited range. Variable density sub-sampling should be chosen if most of energy are stored in the center of the transformed domain. Therefore, random sub-sampling is applied in both of the two sensing matrices.

The coherence simulation arrangements are the same as before, except that the sparsifying matrix is the identity matrix. The number of coils $N_c = 1$ since the original signal is sparse in the spatial domain. The coherence results are shown in Fig. 4.20. Unlike the cases shown in the examples in which the signals are sparse in the wavelet domain, the coherence value of the partial noiselet matrix is larger than that of the partial Fourier matrix at the same $1/R$.

Fig. 4.20 Coherence comparison of Fourier encoding and random sub-sampling, and noiselet encoding and random sub-sampling. (a) $N = 256$. (b) $N = 1024$. 
The coherence results indicate that for signals which are sparse in the spatial domain, the partial Fourier matrix is better than the partial noiselet matrix as the sensing matrix for CS reconstruction. Therefore, the random partial Fourier matrix is recommended for CS reconstruction when signals are sparse in the spatial domain.

4.5 Conclusion

In this chapter, coherence analysis is conducted for Fourier encoding and noiselet encoding. Coherence results show that for signals which are sparse in the wavelet domain, such as MR images, the combination of noiselet encoding and random sub-sampling is the better choice. However, for signals which are naturally sparse, the combination of the Fourier matrix and random sub-sampling is recommended.
Chapter 5

Phase Transition of MRI Encoding

5.1 Introduction

Coherence serves as a quality measure of the sensing matrix, or from another perspective, the quality of the sparsifying dictionary in CS. However, the relationship between the coherence value and the quality of the reconstruction result is not linear. For example, if the coherence value of a sensing matrix is two times smaller than that of another sensing matrix, we can only predict that the quality of the reconstruction result should be better, but it does not mean that the quality of the reconstruction result is two times better.

In order to compare the quantitative performance of CS-based image reconstruction from sub-sampled k-space data encoded with two different encoding techniques, phase transition analysis of CS was conducted. Let $\delta = M/N$, and $\rho = K/M$, where $K, M,$ and $N$ stand for the sparsity of the original signal, the length of the measured data and the length of the original signal, respectively. Phase transition states that in the plane of $(\delta, \rho)$, there is a curve which divides the plane into two phases. Below the curve the original signal can be successfully reconstructed, while above the curve the original signal cannot be successfully reconstructed, both with an overwhelming probability. The phase transition curve depends on many factors, including but not limited to the choice of sensing matrix, and different sparsifying dictionaries. As a result, the phase transition curve can be used to quantitatively compare the performance of CS reconstruction.

In this chapter, phase transition analysis is carried out in order to compare the reconstruction results from sub-sampled k-space data encoded with two different encoding techniques.
5.2 Phase Transition Analysis

5.2.1 Theory

Suppose that there is a sparse signal $x \in \mathbb{C}^N$. In the sparse signal, only $K$ out of $N$ entries are non-zero, and the number of non-zero entries is named the sparsity. In CS theory, a sparse signal can be reconstructed from many fewer randomly-selected measurements than the length of the original signal. The signal is measured by a random matrix $\Phi \in \mathbb{C}^{M \times N}$, and the measured data is denoted as $y \in \mathbb{C}^M$. This process is expressed as

$$y = \Phi x.$$  \hspace{1cm} (5.1)

The original signal is reconstructed by solving the following problem

$$\min ||x||_1 \quad \text{s.t.:} \quad y = \Phi x.$$  \hspace{1cm} (5.2)

Some signals are not naturally sparse but compressible after certain kinds of transformation. Let $\Psi \in \mathbb{C}^{N \times N}$ denote the sparsifying transform matrix, and the compressible signal $s \in \mathbb{C}^N$ is denoted as

$$x = \Psi s.$$  \hspace{1cm} (5.3)

Then the CS reconstruction problem is denoted as

$$\min ||\Psi^T x||_1 \quad \text{s.t.:} \quad y = \Phi x,$$  \hspace{1cm} (5.4)

or

$$\min ||s||_1 \quad \text{s.t.:} \quad y = \Phi \Psi s.$$  \hspace{1cm} (5.5)

According to coherence analysis, if the following inequality is satisfied,

$$||s||_0 < \frac{1}{2} \left( \frac{1}{\mu} + \frac{1}{\mu(\Phi \Psi)} \right),$$  \hspace{1cm} (5.6)

the compressible signal $s$ can be successfully recovered, as can the original signal $x$. It is easily appreciated and has been demonstrated in Chapter 4 that different sensing matrices and sparsifying matrices lead to different coherence values and reconstruction performances. Unfortunately, the relationship between the coherence value and the quality of reconstruction result is not linear.

Phase transition serves as a quantitative measure of CS reconstruction. It was first observed in [30] when the authors tried to develop asymptotic methods to count faces
of random high-dimensional polytopes. It was shown that the conclusions drawn in
the paper have surprising implications in statistics, probability, information theory
and signal processing, and potential impacts in practical subjects such as medical
imaging and digital communications. The authors introduced three implications
in the paper, convex hull of Gaussian point clouds, signal recovery from random
projections, and gross error corrections, of which we are interested in the second
implication.

In the implication of signal recovery from random projections, we intend to
recover a $K$-sparse signal $x \in \mathbb{R}^N$ from its measurements. The original signal is first
proposed to be in the real set $\mathbb{R}$ instead of the complex set $\mathbb{C}$. The measuring process
is conducted by asking $M$ questions, and each question is in the form of $y_i = \langle \phi_i, x \rangle$
on a vector $\phi_i \in \mathbb{R}^N$. Here $\phi_i$ is also in the real set $\mathbb{R}$ instead of the complex set $\mathbb{C}$. The problem is "How many questions should be asked?" or mathematically "How
big should $M$ be?" in order to fully reconstruct the original signal $x$.

If $\phi_i$ is the standard unit vector basis, $N$ measurements are obviously enough
to recover the original signal $x$. However, considering the sparsity of the original
signal, fewer questions will be enough. If $\phi_i$ is independent random vectors with
i.i.d. Gaussian entries, $M = 2K + 1$ should suffice. However, in order to recover the
original signal, all possible $K$ combinations of $\phi_i$ should be tried out, which is not a
tractable task.

Therefore, a more practical question is "How many questions are needed to
allow computationally tractable recovery of the original signal?". Mathematically,
the sensing process is denoted as $y = \Phi x$ as in CS, where $\phi_i$ is the rows of $\Phi$ with
i.i.d. Gaussian entries. Problem 5.2 is used to recover the original signal from
measurements. In Section 7.3 of [30], it is shown that for $N$ much larger than $K$, the
following inequality should be satisfied in order to successfully recover the original
signal

$$M \geq 2 \cdot K \cdot \log(N/M)(1 + o_p(1)),$$

(5.7)

where $o_p(1)$ denotes a sequence of random variables $(V_M)$ having $P\{|V_M| > \varepsilon\} = o(1)$
as $N \to \infty$ for each $\varepsilon > 0$. Therefore, instead of acquiring data with length $2K + 1$, roughly $2 \cdot K \cdot \log(N/M)$ is enough to efficiently reconstruct the $K$-sparse signal, which is still far less than the length of the original signal.

In [30], $K, M$ and $N$ are large but also comparable in size. That is,

$$\delta = M/N, \quad \rho = K/M, \quad N \to \infty,$$

(5.8)
where $\delta$ measures the sub-sampling ratio, that is how many fewer measurements are acquired than the length of the original signal, and $\rho$ measures to what extent the effective number of degrees of freedom $K$ is smaller than the number of measurements.

In the plane of $(\delta, \rho)$, a phase transition phenomenon is observed, like Fig. 2.4. The phase transition curve divides the plane into two phases. Below the curve the original signal can be successfully reconstructed, while above the curve the original signal cannot be successfully reconstructed, both with an overwhelming probability. The phase transition curve passes the location $(\delta, \rho)$ in the plane at which the successful recovery rate is 50% of total repeated reconstructions. Away from the phase transition curve, the successful recovery rate drops to zero or increases to one very quickly. In CS, phase transition works as a performance measure of the sparsity-undersampling trade-off.

In [30], phase transition of real valued CS (the original signal is in the real set $\mathbb{R}$) using the sensing matrix with i.i.d. random Gaussian entries was studied. However, in real applications, it often seems that Gaussianity is not required. In fact, the Gaussian condition of the sensing matrix can be considerably relaxed, which results in the well-known observed universality of phase transition [51].

In [51], the authors reviewed connections between phase transitions in high-dimensional combinatorial geometry and phase transitions occurring in modern high-dimensional data analysis and signal processing. In combinatorial geometry, phase transition appears as abrupt changes in the properties of face counts of convex polytopes when the dimensions are varied. In data analysis, phase transition arises as the abrupt breakdown of linear model selection, robust data fitting or compressive sensing reconstructions, when the complexity of the model or the number of outliers increases beyond a limit. However, the thresholds in these different problems appear in the same critical locations after appropriate calibration of variables.

The authors conducted extensive computational experiments and formal inferential analysis to test the hypothesis that the phase transition phenomenon is universal across a range of underlying matrix ensembles. Several different matrix ensembles and problem sizes were tried and visually the empirical phase transition does not depend on the ensemble. The phase transition agrees well with the asymptotic theory assuming Gaussianity in [30].

In [51], instead of using i.i.d. Gaussian matrix as the sensing matrix, Bernoulli, Fourier, Ternary, Hadamard, Ecpander and Rademacher ensembles are used as the sensing matrix. Phase transition of the Fourier matrix when $N = 200$ is shown in the paper, and similar results are obtained to the i.i.d. Gaussian matrix. The authors
claimed that empirical results for both Gaussian and non-Gaussian ensembles show finite $N$ transition bands centered around the asymptotic phase transition derived from an i.i.d. Gaussian assumption. Since the phase transition curves are different from each other when different ensembles are used, the reconstruction results of CS when using different sensing matrices can be compared.

In [52], phase transition of complex valued CS was studied. The original signal $x$ is assumed to be in the complex set $\mathbb{C}$ instead of the real set $\mathbb{R}$ in this paper. Results show that the same phase transition property exists for complexed valued CS, and phase transition of complexed valued CS presents a higher phase than that of real valued CS.

As discussed in [30, 51, 52], phase transition is universal across a range of underlying matrix ensembles, and different ensembles lead to different phase transition results which can be compared quantitatively. Therefore, the phase transition curve provides a quantitative measure of one important aspect of CS, that is how aggressively a sparse signal can be sub-sampled.

In traditional MRI, Fourier encoding is chosen to encode images into k-space data. In [21], noiselet encoding was proposed to encoded image into k-space data. The image is reconstructed by a CS-based method from sub-sampled k-space data. In CS, combinations of different sensing and sparsifying matrices lead to different reconstruction results. Therefore, it is possible to plot the phase transition curve of Fourier encoding and noiselet encoding, and compare the results. In the following sections, the phase transition curves of Fourier encoding and noiselet encoding under different situations are simulated.

### 5.2.2 Analysis Arrangements

The phase transition curve is defined in the plane of $(\delta, \rho)$. $\delta = M/N$, and $\rho = K/M$, where $K, M$ and $N$ stand for the sparsity of the original signal, the length of the measured data, and the length of the original signal. The range of both $\delta$ and $\rho$ is from 0 to 1. For a location $(\delta_i, \rho_j)$ in the plane, the length of the measured data $M = \delta_i \cdot N$ and the sparsity of the original signal $K = \rho_j \cdot M = \rho_j \cdot \delta_i \cdot N$. A sparse signal $x$ with length $N$ and sparsity $K$ is randomly generated, and the measured data $y$ with length $M$ are acquired by applying a random sensing matrix $\Phi \in \mathbb{C}^{M \times N}$. Here, the partial Fourier operator $T_F$ or the partial noiselet operator $T_N$ is used. The number of rows of the partial Fourier or noiselet operator is $M$. The original signal is reconstructed by solving problem 5.2 or 5.4 depending on the sparsifying matrix.

The NMSE of the reconstructed signal is defined in Eq. 2.34. If NMSE is less than a threshold, it is regarded as a successful reconstruction. Otherwise, it is
regarded as a failure. Multiple tries of reconstruction are repeated. If the number of successful reconstructions is more than half of the total reconstruction tries, the location \((\delta_i, \rho_j)\) is under the phase transition curve. In contrast, if the number of successful reconstructions is less than half of the total reconstruction tries, the location \((\delta_i, \rho_j)\) is above the phase transition curve. The phase transition curve passes all the locations at which the number of successful reconstructions is equal to half of the total reconstruction tries.

MR images are commonly assumed to be sparse after wavelet transform. Therefore, the sparse representation \(s \in \mathbb{C}^N\) with sparsity \(K\) is first randomly generated in the wavelet domain. The original signal \(x \in \mathbb{C}^N\) is the inverse wavelet transform of the sparse representation \(s\). One difference we should note is that the significant values of a wavelet-transformed image are mostly located in the upper left corner in the wavelet domain, while in the simulation the significant values are randomly located. It is easily appreciated that the real case is just one special situation of the simulation. In the coherence analysis, there is no other requirement of the original signal except that it should be sparse in the wavelet domain.

The sub-sampling techniques and sensitivity profiles used in the phase transition analysis are the same as in Chapter 4. Random sub-sampling is applied when a partial noiselet encoding matrix is used as the sensing matrix, and variable density sub-sampling is applied when a partial Fourier encoding matrix is used as the sensing matrix.

For signals which are sparse in the spatial domain, the original signal \(x \in \mathbb{C}^N\) is randomly generated with sparsity \(K\) in the spatial domain, as in Section 4.4. Random sub-sampling is applied in both of the two encoding methods.

As demonstrated in [52], the phase transition curve of complex valued CS is higher than that of real valued CS. It is anticipated that the phase transition curve of Fourier encoding and that of noiselet encoding will be different from each other.
5.3 Results

5.3.1 Sparse Signal in Wavelet Domain

In this section, phase transition simulation is conducted when the original signal is sparse in the wavelet domain. Different numbers of coils and sensitivity profiles are used in the simulation in order to cover as many cases as possible.

First, phase transition when $N = 256$ and $N_c = 1$ is conducted. The simulation results are shown in Fig. 5.1. It is seen that the phase transition curve of noiselet encoding is generally higher than that of Fourier encoding, which indicates that better reconstruction results can be obtained by CS from sub-sampled k-space data encoded with noiselet encoding than Fourier encoding.

![Phase transition curves of Fourier encoding and noiselet encoding when $N_c = 1$ and $N = 256$.](image)

Another interesting phenomenon in Fig. 5.1 is that when $\delta$ is in the range of 0.8 to 1, the phase transition curve of Fourier encoding is higher than that of noiselet encoding. One possible reason for this phenomenon can be found in the coherence analysis results. In Fig. 4.18 when $1/R$ (actually $1/R$ is the same as $\delta$, since both are defined as $M/N$) approaches 1, the coherence line of Fourier encoding and that of noiselet encoding are very close to each other. In addition, the coherence values of both Fourier encoding and noiselet encoding are very close to 0, which means that
the columns of the new dictionaries are almost incoherent with each other, leading to very good reconstruction results. In an ideal case, better reconstruction results will be obtained by CS from sub-sampled k-space data encoded with Fourier encoding than noiselet encoding. This will be further discussed when the original signal is sparse in the spatial domain.

In reality, we mainly focus on the left half of the phase transition plane, since in CS the length of measured data $M$ is assumed to be far less than the length of the original signal $N$, while in the right half of the phase transition plane, $M$ is more than half of $N$, which contradicts the assumptions of CS theory.

The simulation results of phase transition analysis are shown in Fig. 5.2 when $N = 256$ and $N_c = 4$. In this figure, the first observation is that the phase transition curve of noiselet encoding is generally higher than that of Fourier encoding, which indicates that better reconstruction performance can be obtained by CS from sub-sampled k-space data encoded with noiselet encoding than Fourier encoding. More detailed discussions are provided below.

![Fig. 5.2 Phase transition curves of Fourier encoding and noiselet encoding when $N_c = 4$ and $N = 256$. (a) Map 1. (b) Map 2.](image)

For the phase transition simulation results with sensitivity profile case 1, which changes slowly and gradually, it is noted that the phase transition curve of noiselet encoding is generally higher than that of Fourier encoding, except that the phase
transition curve of Fourier encoding surpasses a little bit or is very close to that of noiselet encoding when $\delta$ is close to 1. This is similar as $N_c = 1$.

For the phase transition simulation results with sensitivity profile case 2, which changes quickly and suddenly, it is noted that the phase transition curves of both Fourier encoding and noiselet encoding reach 1 within a short range in the horizontal direction. This phenomenon shows that the reconstruction results benefit if the sensitivity profile of coils changes dramatically. The phase transition curve of noiselet encoding is higher than that of Fourier encoding in the left half of the phase transition plane before the phase transition curve reaches 1.

Another important feature of the phase transition simulation results of sensitivity profile case 2 is that the distance between two phase transition curves is larger compared with that in sensitivity profile case 1 at the same $\delta$ in the left half of the phase transition plane. Although the two curves are close to each other at first sight, the vertical gap is large since the slope of both is steep.

The same phase transition simulation is conducted when $N = 256$, $N_c = 8$ and $N_c = 16$. The simulation results are shown in Fig. 5.3 and 5.4, respectively. The same conclusions can be drawn from these two figures as in Fig. 5.2.

![Fig. 5.3 Phase transition curves of Fourier encoding and noiselet encoding when $N_c = 8$ and $N = 256$. (a) Map 1. (b) Map 2.](image-url)
Fig. 5.4 Phase transition curves of Fourier encoding and noiselet encoding when $N_c = 16$ and $N = 256$. (a) Map 1. (b) Map 2.

The phase transition results for Fourier encoding when $N_c = 1, 4, 8$ and 16 are shown in Fig. 5.5. It is clear that when multiple coils are used, the reconstruction results are dramatically improved. This is not clearly revealed in the coherence analysis results. It seems that the introduction of multiple coils creates a great difference in reconstruction by CS.

In addition, there is a difference when the number of coils changes. Although the coherence results seem very close to each other, they are not identical, especially in sensitivity profile case 2. Due to the width of the lines and the area of the figures, the difference between coherence lines are not clearly shown. However, careful examination reveals that in Fig. 4.14 (b) the coherence line of $N_c = 4$ is the lowest, followed by that of $N_c = 8$ and that of $N_c = 16$. The phase transition curves in Fig. 5.5 clearly show this trend.

Another interesting phenomenon is that the coherence value increases and the phase transition curve drops when the number of coils increases from 4 to 8 and 16. This is due to the design of the sensitivity profiles. When the number of coils increases, the range of magnitude of the designed sensitivity profiles shrinks. As a result, the difference in sensitivity profiles becomes less significant, and the effect on coherence becomes worse when the number of coils increases.
Fig. 5.5 Phase transition curves of Fourier encoding when $N = 256$. (a) Map 1. (b) Map 2.

Fig. 5.6 Phase transition curves of noiselet encoding when $N = 256$. (a) Map 1. (b) Map 2.
In reality, sensitivity profiles of coils in the same MRI scanner change in different scans for many reasons, such as the influence of the human body. Therefore, sensitivity profiles of coils should be estimated before or during scanning in each scan. Unfortunately, sometimes the sensitivity profiles cannot be well estimated due to motion. Furthermore, in some modern MRI scanners, the number of coils can be up to 32 or even higher. During scanning, not all coils can acquire data with acceptable fidelity, because some coils may be too far away from the slice of interest. As a result, an increase in the number of the coils cannot guarantee that the reconstruction results will be better. A conclusion drawn from this phenomenon is that if the differences between the sensitivity profiles of coils are not large enough, an increase in the number of coils may not improve the performance of reconstruction by CS. In some cases, the performance may even be reduced.

The phase transition results for noiselet encoding when $N_c = 1, 4, 8$ and 16 are shown in Fig. 5.6. The same conclusion as for the phase transition simulation of Fourier encoding can be drawn.

The simulation results of phase transition analysis are shown in Fig. 5.7 when $N = 1024$. Same conclusions can be drawn from this figure as in Fig. 5.1 when $N = 256$.

![Fig. 5.7 Phase transition curves of Fourier encoding and noiselet encoding when $N_c = 1$ and $N = 1024$.](image-url)
5.3 Results

5.3.2 Sharp Transition Phenomenon

The shape and location of the phase transition curve may be affected by several factors, such as the choice of sparsifying matrices, differences in sensing matrices, and accuracy of reconstruction methods. Another important factor influencing the shape and location of the phase transition curve is the threshold which is used to define a location, whether it is above or below the phase transition curve. For example, in the phase transition figures above, the phase transition curve is assumed to pass all the locations at which the number of successful reconstructions is equal to half of the total reconstruction tries. If the threshold is defined as another ratio, for example, 60% of the total tries, the phase transition shape and location will change accordingly. However, plotting the phase transition curve of different thresholds is cumbersome. Therefore, we next consider the successful reconstructions at different $\delta$ and $\rho$.

![Phase Transition Curves](image)

Fig. 5.8 Successful reconstructions at different $\delta$ and $\rho$ when $N_c = 1$ and $N = 1024$. (a) Noiselet. (b) Fourier.

Fig. 5.8 shows successful reconstructions at different $\delta$ and $\rho$ when $N_c = 1$ and $N = 1024$. In the figure, the horizontal axis denotes $\rho$, and the vertical axis denotes the number of successful tries. Lines with different colors denote the reconstruction results of Fourier encoding or noiselet encoding at different $\delta$. It is obvious that the slope of the lines of the noiselet simulation is much greater than that of the Fourier
simulation. At the same $\delta$, when $\rho$ increases, the boundary between successful reconstructions and unsuccessful reconstructions for noiselet-encoded data is much clearer than that for Fourier-encoded data. The reason for this phenomenon can be found in the difference between random sub-sampling and variable density sub-sampling. For variable density sub-sampling, it is hard to maintain the balance between data acquired from the low frequency part and data from the high frequency part. At certain $\rho$, it is possible that some variable density sub-sampling choices may be much worse than the optimal choice, while for random sub-sampling, although each repetition is different, the chance of extremely bad cases is rare.

### 5.3.3 Sparse Signal in Spatial Domain

In this section, phase transition simulation when the original signal is sparse in the spatial domain is simulated. The length of the original signal is 256, and phased array of coils are not used. Random sub-sampling is used in both Fourier and noiselet simulations. Phase transition simulation results are shown in Fig. 5.9.

![Phase transition curve](image)

Fig. 5.9 Phase transition curves of Fourier encoding and noiselet encoding when $N_c = 1$ and $N = 256$. The original signal is sparse in the spatial domain.

Fig. 5.9 shows that when signals are sparse in the spatial domain, the phase transition curve of Fourier encoding is higher than that of noiselet encoding. This means that at the same $\delta$, a sparse signal with larger $K$, which is less sparse, can be
reconstructed from the same $M$ measurements when noiselet encoding is used. The phase transition analysis and the coherence analysis agree well with each other when the original signal is sparse in the spatial domain.

5.4 Conclusion

Phase transition serves as a quantitative measure of how aggressively a sparse signal can be recovered by CS. In this chapter, first the phase transition theory is reviewed. Following that, phase transition simulations when the original signal is sparse in the wavelet domain and the spatial domain are conducted. Simulation results show that the phase transition analysis and the coherence analysis agree well with each other. When the original signal is sparse in the spatial domain, Fourier encoding with random sub-sampling is a better choice for CS, while when the original signal is sparse in the wavelet domain, noiselet encoding with random sub-sampling offers better reconstruction results.
Chapter 6

Image Reconstruction Simulation

6.1 Introduction

MRI plays a vital role in contemporary medical imaging diagnosis, by providing images of human tissues and organs with high resolution. 2D imaging sequences can acquire a set of 2D data, which can be reconstructed to form a 2D image. The two dimensions of the 2D data are the phase encoding and the frequency encoding directions. Sub-sampling can only be applied in the phase encoding direction, since the frequency encoding direction is always fully sampled.

Volume imaging is also widely used in MRI. In volume imaging, a 3D data set can be acquired and then reconstructed to form a 3D image. Compared with 2D images, 3D images are capable of providing more information about the human body. Of the three dimensions, two are phase encoding directions, and one is the frequency encoding direction. As a result, sub-sampling can be applied in two phase encoding directions, in order to provide more flexibility when sub-sampling data. When reconstructing images, the frequency encoding direction can be first transformed to the spatial domain by applying the 1D inverse Fourier transform, since the frequency encoding direction is always fully sampled. Therefore, the 3D image reconstruction problem can be decomposed to many 2D image reconstruction problems. In the decomposed 2D image reconstruction problems, both directions of the acquired data are phase encoding directions, and sub-sampling can be applied in two dimensions.

In this chapter, a simulation of image reconstruction from sub-sampled 3D volume imaging data encoded with two encoding techniques is conducted. Both phantom and MR images are used in the simulation.
6.2 2D Sub-sampling in Volume Imaging

6.2.1 Volume Imaging

In MRI, 2D imaging sequences, such as GRE and the spin echo sequence introduced in Chapter 3 and 4, are used to acquire data from a slice of the human body. The RF pulse and the slice selective gradient are designed to select spins in a slice with a specific thickness and to avoid interference from adjacent slices. A multi-slice 2D imaging sequence is capable of collecting several slices in one TR. The 3D data obtained by a multi-slice 2D imaging sequence are in the hybrid domain. In the frequency encoding direction, data are Fourier encoded and fully sampled. In the phase encoding direction, data are also Fourier encoded but can be sub-sampled by deliberately skipping some phase encoding steps. In the slice direction, data are in the spatial domain. The 2D inverse Fourier transform can be applied on the data in one slice to reconstruct a 2D image. All slices are stacked together to form a 3D image.

As described in Chapter 3, volume imaging sequences can be designed by modifying 2D imaging sequences, such as GRE and the spin echo sequence. Volume imaging acquires NMR signals from a volume of the object rather than a single tomographic slice. In volume imaging, a phase encoding gradient is applied in the slice direction. As a result, all three dimensions are Fourier encoded, of which two are phase encoding directions, and one is the frequency encoding direction. Sub-sampling can be applied in two phase encoding directions by deliberately skipping some phase encoding steps, while data are always fully sampled in the frequency encoding direction. 1D inverse Fourier transform can be applied on each of the three dimensions to reconstruct a 3D image. Fig. 6.1 displays multi-slice 2D imaging and 3D volume imaging.

In Fig. 6.1 (a), the vertical direction is the phase encoding direction, and the horizontal direction is the frequency encoding direction. A slice is selected in the slice direction, indicated by a red square. In Fig. 6.1 (b), both the vertical and the horizontal directions are phase encoding directions, and the direction perpendicular to the plane is the frequency encoding direction. The encoding dimensions can be manipulated by operators. If 1D inverse Fourier transform is applied in the frequency encoding direction, the 3D data are transformed to the hybrid space, as in multi-slice imaging. As a result, a slice can be selected in the transformed frequency encoding direction, indicated by a red square, like the slice in Fig. 6.1 (a). This is the reason why these three encoding dimensions are selected.
6.2 2D Sub-sampling in Volume Imaging

Fig. 6.1 Comparison between multi-slice 2D imaging and 3D volume imaging. (a) Multi-slice 2D imaging. (b) 3D volume imaging.

The difference between the two slices in Fig. 6.1 (a) and (b) is obvious. In Fig. 6.1 (a), the two dimensions of the slice are the phase encoding and the frequency encoding directions, of which sub-sampling can be applied in the phase encoding direction only. In contrast, in Fig. 6.1 (b), both of the two dimensions of the slice are phase encoding directions, so that sub-sampling can be applied in two dimensions. One advantage of volume imaging is that it provides more flexibility when sub-sampling data. Another advantage is that isotropic voxels can be obtained in volume imaging.

Let \( y \) denote the decomposed sub-sampled 2D k-space data, which is defined as

\[
y = \begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_{N_c} \end{pmatrix} = \begin{pmatrix} T_s(\Gamma_1 \odot x) \\ \vdots \\ T_s(\Gamma_i \odot x) \\ \vdots \\ T_s(\Gamma_{N_c} \odot x) \end{pmatrix} = \hat{T}_s x, \quad i = 1, 2, \cdots, N_c, \tag{6.1} \]

where \( x \) is the 2D image of interest, \( y_i \) is the sub-sampled data of coil \( i \), \( \Gamma_i \) the sensitivity profile of coil \( i \), \( T_s \) a selected encoding technique with the sub-sampling process applied, \( \odot \) the element-wise multiplication, and \( \hat{T}_s \) the whole sub-sampling and encoding process in the second brackets. If without noise, the following problem is solved to reconstruct the image of interest:

\[
\min ||\Psi^T x||_1 \quad \text{s.t.} \quad \hat{T}_s x = y, \quad \tag{6.2}
\]
where $\Psi$ is the wavelet transform operator. If the data are corrupted by noise, which is very common in real data acquisition, the following problem is solved to reconstruct the image of interest:

$$\min ||\Psi^T x||_1 \quad \text{s.t.} \quad ||\hat{T}_s x - y||_2 < \varepsilon,$$

(6.3)

where $\varepsilon$ is a parameter related to noise. This problem can be solved via the conjugate gradient descent algorithm with backtracking line search illustrated in Chapter 3. In Chapter 4, the 1D sub-sampling process involved in $T_s$ is conducted in one phase encoding direction. This problem can also be used to reconstruct an image from sub-sampled volume imaging data by involving 2D sub-sampling in $T_s$ similar to the 1D case.

If noiselet encoding can be applied in volume imaging, both of the two phase encoding directions will be encoded by noiselet encoding, while the frequency encoding direction will be encoded by traditional Fourier encoding. As mentioned above, 1D inverse Fourier transform can be applied in the frequency encoding direction since data in this direction are always fully sampled. As a result, the image space image reconstruction problem can be decomposed as a series of 2D image reconstruction problems.

### 6.2.2 2D Sub-sampling

A phantom image is shown in Fig. 6.2 (a), and the 2D noiselet-encoded data are shown in Fig. 6.2 (b). It is seen that the energy is spread across all phase encoding steps by incorporating noiselet encoding in two dimensions. Therefore, 2D random sub-sampling should be applied for noiselet-encoded volume imaging data. For Fourier-encoded volume imaging, the Fourier data of a slice is the same as that displayed in Fig. 4.4. Since sub-sampling is applied in two phase encoding dimensions, 2D variable density sub-sampling should be applied for Fourier-encoded volume imaging data.

Fig. 6.3 shows the comparison between 1D and 2D sub-sampling. For 2D variable density sub-sampling, the low frequency part of k-space is sampled densely, while the high frequency part is sampled sparsely. As in the 1D case, there are different forms, depending on how much is defined as the low frequency part and how much is sampled in the low/high frequency part. For random sub-sampling, in the 1D case only the rows are randomly selected, while in the 2D case all locations are randomly selected. 2D variable density sub-sampling is used for Fourier-encoded
Fig. 6.2 The principle of 2D noiselet encoding. (a) Phantom image. (b) 2D noiselet-encoded data.

volume imaging data, while 2D random sub-sampling is used for noiselet-encoded volume imaging data.

The benefits of 2D sub-sampling in the MRI reconstruction problem are substantial. First, by dividing the sub-sampling process in two dimensions, better reconstruction results can be obtained at the same reduction factor compared with 1D sub-sampling. Second, higher reduction factors can be achieved in 2D sub-sampling than in 1D sub-sampling. In the following example, a phantom is used to test the performance of CS-based MRI reconstruction when 1D and 2D sub-sampling are used.

In Fig. 6.4, reconstructed images from sub-sampled Fourier-encoded data are shown. When 30% of data are sampled, reconstructed images from data with 1D and 2D variable density sub-sampling are displayed in sub-figures (a) and (c), and their differences with the original image are shown in sub-figures (b) and (d). The results when 10% of data are sampled are shown in sub-figures (e)-(h). The difference images are scaled in order to clearly show the errors.

In Fig. 6.5, reconstructed images from sub-sampled noiselet-encoded data are shown. When 30% of data are sampled, reconstructed images from data with 1D and 2D random sub-sampling are displayed in sub-figures (a) and (c), and their differences with the original image are shown in sub-figures (b) and (d). The results when 10% of data are sampled are shown in sub-figures (e)-(h). The difference images are scaled in order to clearly show the errors.

From Fig. 6.4 and Fig. 6.5, it is obvious that when 30% of data are sampled, clear images can be reconstructed from sub-sampled data with 1D sub-sampling, in both Fourier encoding and noiselet encoding. In the 2D sub-sampling case, the reconstructed images have less error than the 1D sub-sampling. Therefore,
better reconstruction results can be obtained at the same reduction factor in 2D sub-sampling than in 1D sub-sampling. When only 10% of data are sampled, the reconstructed images from both encoding methods with 1D sub-sampling are not acceptable. In contrast, reconstructed images from data with 2D sub-sampling still possess clear details, and the differences between reconstructed images and the original image are minor.

2D sub-sampling is more flexible than 1D sub-sampling. Reconstructed images from 2D sub-sampled data have better qualities than those from 1D sub-sampled data. Therefore, applying noiselet encoding in volume imaging not only provides a new encoding method, but also offers opportunities to achieve a higher reduction factor without sacrificing much information. In the following sections, comparisons between Fourier encoding and noiselet encoding in volume imaging are made.
Fig. 6.4 Reconstructed images from sub-sampled Fourier-encoded data. (a) 1D variable density sub-sampling and $1/R = 30\%$. (c) 2D variable density sub-sampling and $1/R = 30\%$. (e) 1D variable density sub-sampling and $1/R = 10\%$. (g) 2D variable density sub-sampling and $1/R = 10\%$. (b), (d), (f) and (h) are differences between the original image and the reconstructed images (a), (c), (e) and (g), respectively.
Fig. 6.5 Reconstructed images from sub-sampled noiselet-encoded data. (a) 1D random sub-sampling and $1/R = 30\%$. (c) 2D random sub-sampling and $1/R = 30\%$. (e) 1D random sub-sampling and $1/R = 10\%$. (g) 2D random sub-sampling and $1/R = 10\%$. (b), (d), (f) and (h) are differences between the original image and the reconstructed images (a), (c), (e) and (g), respectively.
6.3 Image Reconstruction

In this and the following sections, image reconstruction simulation from 2D sub-sampled volume imaging data is conducted. Data are acquired from a Siemens Skyra 3T scanner with Fourier encoding. Sub-sampling is not applied during the scanning. The sensitivity encoded images are reconstructed by applying 2D inverse Fourier transform on the sliced data from all coils. The image of interest is obtained by the sum of the squares of the sensitivity encoded images. Images reconstructed from the fully sampled data are used as the reference images. Noiselet transform is applied on both directions of the sensitivity encoded images to simulate the noiselet-encoded data.

Sub-sampling is applied manually after k-space data are fully acquired. 2D random sub-sampling is applied when noiselet encoding is used, and 2D variable density sub-sampling is used when Fourier encoding is chosen. Phantom and brain image are used to test the performance of the reconstruction method based on CS. The size of the original image is $256 \times 256$, as is the size of the reconstructed images.

Since measured data are often corrupted by noise, noise is also considered in the image reconstruction simulation. White Gaussian noise is added to the sub-sampled k-space data manually after the sub-sampling process is completed. Three levels of noise are added to the sub-sampled k-space data, which are 20dB, 30dB and 40dB. It is noted that the higher the number, the less noise is added. Without noise, Problem 6.2 is solved via the conjugate gradient descent method with backtracking line search. When noise is considered, Problem 6.3 is solved also via the conjugate gradient descent method with backtracking line search.

6.3.1 Phantom

First, a phantom is used to compare the image quality from the 2D sub-sampled Fourier-encoded data with that from the 2D sub-sampled noiselet-encoded data. In total, 8 coils are used in the scanning. Both of the dimensions are assumed to be phase encoding directions, encoded with Fourier encoding or noiselet encoding. Only 5% of the fully-sampled data from each coil are collected. The reconstruction process is repeated 100 times in order to minimize the influence of randomness. To evaluate the performance, the NMSE of the reconstructed images is evaluated and compared. Reconstructed images with the lowest NMSE from both of the two encoding methods are shown in the following figures.

The reference phantom image is shown in Fig. 6.6 (a). It has sharp transitions at boundaries and great detail. A reconstructed image from the 2D sub-sampled
噪声编码的数据如图6.6 (b) 所示。原始图像与重构图像的差异如图6.6 (c) 所示。几乎没有任何视觉差异可以被肉眼识别。所有差异图像（包括这个和接下来的图像）都以 (0,0.5) 的尺度显示，因为原始图像的尺度为 (0,1)，并且通常差异图像的最大幅度小于 0.5。重构的仿生影像的NMSE为 0.030。
reconstructed image from the 2D sub-sampled Fourier-encoded data is shown in Fig. 6.6 (d). The difference between the original image and the reconstructed image is shown in Fig. 6.6 (e). It is obvious that the details in the phantom are not well reconstructed. The NMSE of the reconstructed phantom is 0.049.

To make a clear visual comparison, parts of the original image and the reconstructed images are shown. The size of the selected part is $128 \times 128$. The selected parts are shown in Fig. 6.7. The small squares in the reconstructed image with noiselet encoding are clearly shown, while they are obscure in the reconstructed image with Fourier encoding.

Fig. 6.7 Comparison between reconstructed images. (a) Part of reference phantom image. (b) Part of reconstructed image with noiselet encoding. (c) Part of reconstructed image with Fourier encoding.

Fig. 6.8 NMSE of reconstructed phantom images.
The sorted NMSE of all 100 times are shown in Fig. 6.8. The NMSE of reconstructed images from the noiselet-encoded data is much smaller than that from the Fourier-encoded data. The range of the NMSE of reconstructed images from the noiselet-encoded data is from 0.03 to 0.05, while the range of the NMSE of reconstructed images from the Fourier-encoded data is from 0.05 to 0.09. The highest NMSE of reconstructed images from the noiselet-encoded data is almost the same as the lowest NMSE of reconstructed images from the Fourier-encoded data.

6.3.2 Brain Image

Second, a brain image is used to compare the image quality from the 2D sub-sampled Fourier-encoded data with that from the 2D sub-sampled noiselet-encoded data. In total 8 coils are used in the scanning. Both of the two dimensions are assumed to be phase encoding directions, encoded with Fourier encoding or noiselet encoding. Only 5% of the fully sampled data from each coil are collected. The reconstruction process is repeated 100 times in order to minimize the influence of randomness. To evaluate the performance, the NMSE of the reconstructed images is evaluated and compared. Reconstructed images with the lowest NMSE from both of the two encoding methods are shown in the following figures.

The reference brain image is shown in Fig. 6.9 (a). It clearly shows the structure of the human brain, containing many gyrus and the skeleton of the head. A reconstructed image from the 2D sub-sampled noiselet-encoded data is shown in 6.9 (b). The difference between the original image and the reconstructed image is shown in 6.6 (c). Although there are some differences, the reconstructed image maintains the main structure and detail of the original image. The NMSE of the reconstructed image is 0.052. A reconstructed image from the 2D sub-sampled Fourier-encoded data is shown in 6.9 (d). The difference between the original image and the reconstructed image is shown in 6.9 (e). The differences are larger compared to its noiselet counterpart, and the reconstructed image is not as clear as its noiselet counterpart. The NMSE of the reconstructed phantom is 0.061.

The selected parts of the original brain image and the reconstructed images are shown in Fig. 6.10 in order to make a further visual comparison. The reconstructed image with Fourier encoding is more obscure than the reconstructed image with noiselet encoding.

The sorted NMSE of all 100 times are shown in Fig. 6.11. The NMSE of reconstructed images from the noiselet-encoded data is smaller than that from the Fourier-encoded data. The range of the NMSE of reconstructed images from the noiselet-encoded data is from 0.05 to 0.07, while the range of the NMSE of re-
constructed images from the Fourier-encoded data is from 0.06 to 0.09. Generally speaking, there are 15% improvements in NMSE when noiselet encoding is used instead of Fourier encoding.

Fig. 6.9 Comparison between reconstructed images. (a) Reference brain image. (b) Reconstructed image with noiselet encoding. (c) Difference between (a) and (b). (d) Reconstructed image with Fourier encoding. (e) Difference between (a) and (d).
Fig. 6.10 Comparison between reconstructed images. (a) Part of reference brain image. (b) Part of reconstructed image with noiselet encoding. (c) Part of reconstructed image with Fourier encoding.

Fig. 6.11 NMSE of reconstructed brain images.

6.3.3 Noise

Finally, the impact of noise in the CS-based reconstruction method from the 2D sub-sampled k-space data is simulated. 20dB, 30dB and 40dB of noise are added to the sub-sampled data. When the SNR is 20dB, generally speaking the original data are badly corrupted, while when the SNR is 40dB, the effect of noise is limited. The NMSE of reconstructed images from noise-corrupted 2D sub-sampled k-space data are shown in Fig. 6.12 and Fig. 6.13 for the phantom and brain image, respectively.

In Fig. 6.12 and Fig. 6.13, for the noiselet encoding cases, the NMSE of reconstructed images from data with 40dB of noise is almost the same as the NMSE
Fig. 6.12 NMSE of reconstructed phantom images from data with noise. (a) Fourier encoding. (b) Noiselet encoding.

Fig. 6.13 NMSE of reconstructed brain images from data with noise. (a) Fourier encoding. (b) Noiselet encoding.
of reconstructed images from data without noise. The NMSE of reconstructed images from data with 30dB of noise is larger than the NMSE of reconstructed images from data without noise. For the brain image, the NMSE difference is even larger than that for the phantom, since the brain image is more complicated. When 20dB of noise is applied, the NMSE of both cases is over 0.2 and the quality of reconstructed images is badly corrupted.

For the Fourier encoding cases, the NMSE of reconstructed images from data with 40dB of noise is similar to that from data with 30dB of noise in the noiselet encoding cases. When 30dB of noise is added to the sub-sampled data, the NMSE of reconstructed images is already much larger than that from data with 20dB of noise in the noiselet encoding cases. When 40dB of noise is applied, the NMSE is above 0.4 and the reconstructed images are badly corrupted.

6.4 Conclusion

In this chapter, image reconstruction simulation is conducted in order to compare the reconstructed image quality from data encoded with Fourier encoding with that with noiselet encoding. First, the differences between multi-slice imaging and volume imaging are introduced and the benefits of 2D sub-sampling are demonstrated by single coil image reconstruction. Better reconstruction results and higher reduction factor can be obtained with 2D sub-sampling compared with 1D sub-sampling. Second, phantom and brain image are used to compare the different performance of CS-based reconstruction method between Fourier encoding and noiselet encoding. It is shown that the reconstructed image quality from the 2D sub-sampled noiselet-encoded data is better than that from the 2D sub-sampled Fourier-encoded data, and the NMSE is lower. When noise is considered, the NMSE of reconstructed images from noiselet-encoded data with a certain degree of noise is lower than that from Fourier-encoded data with the same degree of noise.
Chapter 7

Conclusion and Future Work

In this chapter, we summarize the main contributions of this thesis, and list some possible directions for future work.

7.1 Conclusion

MRI is a widely-used imaging technique in contemporary medical diagnosis. It is capable of producing high-resolution images of the internal human body in any direction without involving invasive operation. A major problem of MRI is its relatively long data acquisition time due to physical constraints. In order to shorten the data acquisition time, sub-sampling is applied in the phase encoding direction. A consequence of sub-sampling is that the inverse 2D Fourier transform cannot be directly applied to transform the sub-sampled k-space data to the image domain. Otherwise, aliasing will appear in the reconstructed images.

In CS theory, a sparse signal can be faithfully reconstructed with negligible error from many fewer randomly-selected measurements. CS has been successfully applied to the MRI reconstruction problem from sub-sampled k-space data. In conventional MRI with Fourier encoding, most of the energy of the measured data is located in the low frequency part of k-space. Therefore, variable density sub-sampling is commonly used in CS-based reconstruction, which acquires more data from the low frequency part of k-space and less data from the high frequency part of k-space. Variable density sub-sampling violates the randomness requirement of the sensing matrix to some extent, leading to sub-optimal reconstruction results.

Noiselet encoding is proposed to overcome this disadvantage. It is capable of spreading the energy of k-space data evenly across all phase encoding steps. As a result, random sub-sampling instead of variable density sub-sampling can be applied during the data acquisition of the CS-based image reconstruction problem of MRI.
2D image reconstruction results show that a reconstructed image from randomly sub-sampled noiselet-encoded data is of better quality than that from variable densely sub-sampled Fourier-encoded data. In this thesis, we have carried out work on the coherence analysis of MRI encoding and CS-based MRI reconstruction in order to gain understanding of the performance of CS-based MRI reconstruction using sub-sampled data from noiselet and Fourier encoding.

In Chapter 4, first, non-Fourier encoding was first reviewed and coherence analysis for CS-based MRI reconstruction was conducted. Coherence serves as a measure of the quality of the sensing matrix, or from another perspective, the quality of the sparsifying matrix of CS. First, when the original signal is sparse in the wavelet domain, the coherence value of Fourier encoding with variable density sub-sampling is much smaller that that of Fourier encoding with random sub-sampling at the same reduction rate. This means that the quality of the new dictionary is much better when variable density sub-sampling is applied in Fourier encoding, leading to reconstructed images of better quality by CS-based reconstruction. The coherence simulation results are consistent with common practices in CS-based reconstruction. Second, random sub-sampling is chosen when noiselet encoding is applied. The coherence value of noiselet encoding with random sub-sampling is smaller than that of Fourier encoding with optimal variable density sub-sampling at the same reduction rate. This means that the quality of the new dictionary is better when random sub-sampling is applied in noiselet encoding, leading to reconstructed images of better quality by CS-based reconstruction. In addition, random sub-sampling is easy to generate. In contrast, to find optimal variable density sub-sampling is a cumbersome task. Finally, when the original signal is sparse in the spatial domain, random sub-sampling is chosen for both Fourier encoding and noiselet encoding. The coherence value of Fourier encoding with random sub-sampling is smaller than that of noiselet encoding with random sub-sampling, indicating that Fourier encoding is a better option when the original signal is sparse in the spatial domain.

The relationship between the coherence value and the reconstruction results is not linear. Phase transition serves as a quantitative measure of the reconstruction results of CS-based reconstruction. In Chapter 5, phase transition analysis was carried out in order to quantitatively compare the performance of CS-based MRI reconstruction from sub-sampled Fourier-encoded k-space data with that from sub-sampled noiselet-encoded k-space data. Simulation results show that phase transition analysis and coherence analysis are consistent with each other. Specifically, when the original signal is sparse in the spatial domain, Fourier encoding with random sub-sampling is a better choice for CS-based reconstruction, while when the original signal is sparse
in the wavelet domain, noiselet encoding with random sub-sampling offers better reconstruction results.

After demonstrating the benefits of noiselet encoding with random sub-sampling in CS-based MRI reconstruction, image reconstruction simulations were conducted in order to compare the reconstructed image quality from the 2D sub-sampled Fourier-encoded volume imaging data with that from the 2D sub-sampled noiselet-encoded volume imaging data in Chapter 6. First, differences between multi-slice imaging and volume imaging were introduced and the benefits of 2D sub-sampling were demonstrated by single coil image reconstruction. Better reconstruction results and a higher reduction factor can be obtained using 2D sub-sampling compared with 1D sub-sampling. Second, phantom and brain images were used to compare the performance differences of CS-based MRI reconstruction methods between Fourier encoding and noiselet encoding. The reconstructed image quality from 2D sub-sampled noiselet-encoded data is better than that from 2D sub-sampled Fourier-encoded data, and the reconstructed image from the 2D sub-sampled noiselet-encoded data has a lower NMSE value. When noise is considered, the NMSE of reconstructed images from 2D sub-sampled noiselet-encoded data with a certain degree of noise is lower than that from 2D sub-sampled Fourier-encoded data with the same degree of noise.

In sum, combinations of different encoding and sub-sampling techniques result in different coherence properties of the sparsifying matrix. In CS-based MRI reconstruction, the coherence value of noiselet encoding with random sub-sampling is smaller than that of Fourier encoding with variable density sub-sampling at the same reduction factor, leading to better image reconstruction results.

7.2 Future Work

In this section, some possible directions for future research are listed.

- Noiselet encoding sequence design in volume imaging
  Simulation results demonstrate that images reconstructed from 2D sub-sampled noiselet-encoded data are of better quality than those from 2D sub-sampled Fourier-encoded data. Intuitively, the next step is to design the pulse sequence of noiselet encoding in volume imaging. After that, phantom and human body scanning should be conducted in order to test the performance of the pulse sequence. Image reconstruction from the fully sampled volume imaging data could be followed to determine whether the expected image can be obtained by directly applying the inverse Fourier transform and inverse noiselet transform.
on the volume imaging data. If successful reconstruction can be achieved, 2D sub-sampling should be manually applied to the fully-sampled volume imaging data. Image reconstruction from the 2D sub-sampled volume imaging data could be conducted to finally test the reconstruction performance of CS-based MRI reconstruction when in vivo volume imaging data are used.

• Noiselet encoding in non-Cartesian coordinates
As discussed in Chapter 3, pulse sequences following the non-Cartesian coordinates offer unique advantages in Fourier encoding, since more data are collected naturally from the low frequency part of k-space. Another interesting advantage of pulse sequences in the non-Cartesian coordinates is that sub-sampling can be randomly selected, since each frequency encoding line passes the center of k-space and carries the same amount of information. It is possible to compare the quality of reconstructed images from the randomly sub-sampled noiselet-encoded data in the Cartesian coordinates with that from randomly sub-sampled Fourier-encoded data in the non-Cartesian coordinates with the same reduction factor. Although it may seem unnecessary to apply noiselet encoding in the non-Cartesian coordinates since the energy of k-space data is spread across all phase encoding steps, it is worth trying since pulse sequences in the non-Cartesian coordinates are normally faster than those in the Cartesian coordinates. However, one important issue of noiselet encoding in the non-Cartesian coordinates is that in the gridding process, linear interpolation may not work since the data are "noise-like".

• Analysis of other non-Fourier encoding techniques
As discussed in Chapter 4, other non-Fourier encoding techniques, such as SVD encoding, and wavelet encoding, have been proposed. Each has unique properties. Similar analysis of them can be conducted to compare the performance differences when CS-based MRI reconstruction is applied.
References


Publication

