

# Non-Cartesian MRI Scan Time Reduction through Sparse Sampling



# Non-Cartesian MRI Scan Time Reduction through Sparse Sampling

Proefschrift

ter verkrijging van de graad van doctor  
aan de Technische Universiteit Delft,  
op gezag van de Rector Magnificus prof. ir. K.F. Wakker,  
voorzitter van het College voor Promoties,  
in het openbaar te verdedigen op donderdag 7 juni 2001 om 10.30 uur

door

Frank Theo Anthonie Willem WAJER

natuurkundig ingenieur  
geboren te Amsterdam

Dit proefschrift is goedgekeurd door de promotor:

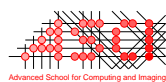
Prof.dr.ir. A.F. Mehlkopf

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof.dr.ir. A.F. Mehlkopf,	Technische Universiteit Delft, promotor
Dr.ir. D. van Ormondt,	Technische Universiteit Delft, toegevoegd promotor
Prof.dr.ir. P. Kruit,	Technische Universiteit Delft
Prof.dr.ir. C.P.A. Wapenaar,	Technische Universiteit Delft
Prof.dr.ir. M.A. Viergever,	Universiteit Utrecht
Dr. D. Graveron-Demilly,	Université Claude Bernard Lyon I, Frankrijk
Ir. M. Fuderer,	Philips Medical Systems, Best



This investigation is part of the research program of the Dutch Technology Foundation (STW).



This work was carried out in graduate school ASCI.  
ASCI dissertation series number 66.

ISBN 9064645787

Copyright © 2001 by F.T.A.W. Wajer

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the author.

Printed in The Netherlands

*Voor mijn ouders,  
voor Josephine*



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Magnetic Resonance Imaging . . . . .	1
1.2	Aim of this thesis . . . . .	2
1.3	Outline of this thesis . . . . .	3
<b>2</b>	<b>MRI scan techniques</b>	<b>5</b>
2.1	Introduction . . . . .	5
2.2	$k$ -Space trajectories . . . . .	8
2.2.1	Cartesian scan . . . . .	9
2.2.2	Radial scan . . . . .	9
2.2.3	Spiral scan . . . . .	11
2.3	Some other trajectory shapes . . . . .	13
2.4	Summary . . . . .	13
<b>3</b>	<b>Image reconstruction from complete data</b>	<b>15</b>
3.1	Introduction . . . . .	15
3.2	Discretization of the inverse Fourier transform . . . . .	16
3.2.1	Reconstruction by DFT . . . . .	16
3.2.2	Sampling density compensation . . . . .	16
3.3	Cartesian sampling . . . . .	18
3.4	Non-uniform sampling . . . . .	19
3.4.1	Introduction . . . . .	19
3.4.2	Image reconstruction for radial scans . . . . .	19
3.4.2.1	Projection reconstruction . . . . .	20
3.4.2.2	Linogram and the chirp-z reconstruction . . . . .	22
3.4.2.3	SqUbe reconstruction . . . . .	22
3.4.3	Gridding . . . . .	24
3.4.3.1	Introduction . . . . .	24
3.4.3.2	Pictorial Explanation of Gridding . . . . .	24
3.4.3.3	Algebraic explanation . . . . .	26
3.4.3.4	Interpolation error . . . . .	28
3.4.3.4.1	Error in approximating complex exponential . . . . .	28
3.4.3.4.2	Energy criterion . . . . .	28

3.4.3.5	Choice of convolution kernel . . . . .	29
3.4.3.6	Interpolation errors using the Kaiser-Bessel kernel . . . . .	31
3.4.3.7	Computational details . . . . .	34
3.4.3.8	Interpolation from uniform to non-uniform sample positions . . . . .	34
3.5	Maximum likelihood estimation . . . . .	36
3.6	Influence of the sample distribution on noise properties in the image . . . . .	37
3.7	Conclusions . . . . .	38
<b>4</b>	<b>Image reconstruction from incomplete data</b>	<b>39</b>
4.1	Introduction . . . . .	39
4.2	Scan-time reduction . . . . .	40
4.2.1	Cartesian scanning . . . . .	40
4.2.2	Radial and Spiral scanning . . . . .	41
4.3	Type 1: Edge-preserving filtering . . . . .	41
4.3.1	Non-linear Gaussian filter . . . . .	45
4.4	Type 2: Bayesian image estimation . . . . .	45
4.4.1	Introduction . . . . .	45
4.4.2	The most probable image . . . . .	46
4.4.3	The prior, $P(I)$ . . . . .	47
4.4.3.1	The 1D-Lorentz prior . . . . .	47
4.4.3.2	Alternative derivation of the 1D-Lorentz prior . . . . .	49
4.4.3.3	Properties of the 1D-Lorentz prior . . . . .	49
4.4.3.4	The 2D-Lorentz prior . . . . .	50
4.4.4	The likelihood . . . . .	53
4.4.4.1	Gradient of the log-likelihood . . . . .	54
4.4.5	Computation of the log-likelihood and its gradient . . . . .	54
4.4.5.1	Cartesian sampling . . . . .	54
4.4.5.2	Non-Cartesian sampling . . . . .	55
4.4.6	Choice of the hyperparameters [1] $\sigma_S$ and $a$ . . . . .	56
4.4.7	Posterior optimization . . . . .	57
4.4.8	Optimization method . . . . .	57
4.4.8.1	Bracketing the minimum . . . . .	58
4.4.8.2	Isolating the minimum along the search direction . . . . .	58
4.4.8.3	Number of function evaluations . . . . .	59
4.5	Summary and conclusions . . . . .	59
<b>5</b>	<b>Evaluation of image reconstruction from incomplete data</b>	<b>61</b>
5.1	Introduction . . . . .	61
5.2	Criteria . . . . .	62
5.2.1	Image performance measures . . . . .	62
5.2.2	Special performance measure for scan-time reduction . . . . .	65
5.2.2.1	Problems with applying $\text{perf}_4$ to non-Cartesian data-sets . . . . .	66
5.3	Summary and conclusions . . . . .	67



<b>6</b>	<b>Optimal sparse sample distributions</b>	<b>69</b>
6.1	Introduction . . . . .	69
6.2	Regular sparse spiral scanning . . . . .	69
6.3	Irregular sparse spiral scanning . . . . .	74
6.3.1	Introduction . . . . .	74
6.3.2	Best of random distributions . . . . .	75
6.3.3	Brute force search, by minimax . . . . .	76
6.4	Summary and conclusions . . . . .	77
<b>7</b>	<b>Applications</b>	<b>79</b>
7.1	Introduction . . . . .	79
7.2	Simulations . . . . .	81
7.2.1	Random scan . . . . .	81
7.2.2	Intra-scan rotation . . . . .	82
7.3	Real-world spiral scans . . . . .	82
7.3.1	Preamble . . . . .	82
7.3.2	Problems at $\vec{k} = \vec{0}$ . . . . .	82
7.3.3	Bayes reconstruction from sparse scans . . . . .	84
7.4	Summary and conclusions . . . . .	86
<b>8</b>	<b>Conclusions and Future Work</b>	<b>89</b>
8.1	Preamble . . . . .	89
8.2	General Overview . . . . .	89
8.3	Gridding . . . . .	90
8.4	Bayesian Reconstruction . . . . .	91
8.5	Suggestions for further research . . . . .	91
<b>A</b>	<b>Major Speedup of Reconstruction for Sensitivity Encoding with Arbitrary <math>k</math>-space Trajectories</b>	<b>93</b>
A.1	Introduction . . . . .	93
A.2	Method . . . . .	93
A.3	Results . . . . .	95
A.4	Conclusions . . . . .	95
<b>B</b>	<b>On the use of Gaussian prior knowledge</b>	<b>97</b>
	<b>Summary</b>	<b>99</b>
	<b>Samenvatting</b>	<b>101</b>
	<b>Bibliography</b>	<b>105</b>
	<b>List of symbols and abbreviations</b>	<b>115</b>
	<b>Dankwoord/Acknowledgement</b>	<b>117</b>



# Chapter 1

## Introduction

### 1.1 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is a technique that enables one to noninvasively view the interior of living beings [2–4]. It is based on the phenomenon of nuclear magnetism [5]. Many nuclei possess a magnetic moment  $\vec{\mu}_n$ , where the subscript n stands for a particular nuclear isotope. Examples of such nuclei are  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{31}\text{P}$ , with natural abundances of respectively 99.98%, 1.11%, and 100%. The magnitude of  $\vec{\mu}_n$  depends on the isotopic species at hand; that of  $^1\text{H}$  – the proton – is the largest of the mentioned three and therefore this isotope can emanate the strongest MRI signal. Another property favouring the strength of an MRI signal is mobility of the molecule hosting the nucleus. Since mobile water ( $\text{H}_2\text{O}$ ) is the most abundant molecule in living tissue, it is used as medium for MRI. Hence, a clinical MR image displays the distribution of water in a patient. Other substances, such as fat, can contribute too.

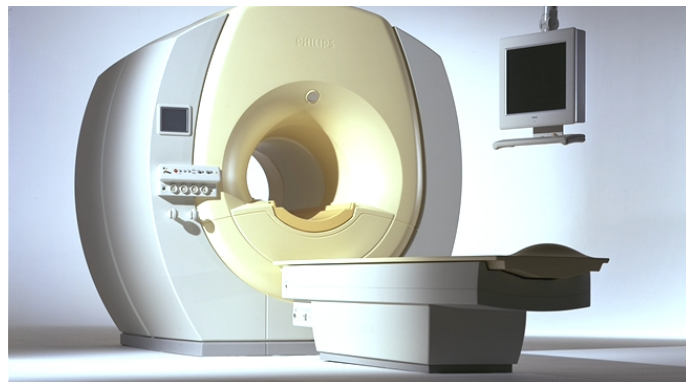


Figure 1.1: A clinical MRI scanner.

Fig. 1.1 shows a typical MRI scanner used in clinics. The bulk of the instrument consists of a strong electromagnet capable of sustaining a homogeneous magnetic field  $B_0$  of the order of 1 Tesla. The patient is slid into the bore of the magnet where his/her magnetic nuclei are aligned along  $B_0$ . Besides  $B_0$ , one needs additional magnetic fields to manipulate the nuclear moments once they are aligned. These fields are:

1. A radio-frequency field which is capable of tilting the aligned nuclei away from their preferred orientation along  $B_0$ .
2. Three so-called gradient fields which are parallel to  $B_0$  but whose magnitudes depend linearly on the three spatial coordinates  $x, y, z$  respectively. More on this in chapter 2.

Both the gradient fields and the radio-frequency field are computer-controlled, enabling one to adjust the nuclear magnetic moment as a function of time and spatial coordinates. I emphasize that the resulting MRI signal is not immediately suitable for use in the clinic. In fact, it is measured in the Fourier domain which implies that the image information is coded beyond recognition. Fourier transformation of the MRI signal to the spatial domain is required to obtain the desired image. The work described in this thesis has to do with estimating an image from a given MRI signal, rather than with MRI as such. This point is discussed in the next section.

## 1.2 Aim of this thesis

As mentioned in the previous section, the MRI signal is to be Fourier transformed from the measurement domain to the spatial domain in order to make the clinical information accessible to the human eye. Usually this poses no obstacle. However, increasingly one wishes to image (scan) time-dependent phenomena like bolus-spreading or movement of an object, *e.g.* of a heart. In those cases there is a dire need to omit part of the measurements (samples) because otherwise the scanner is unable to follow the events. Once this concession has been made, the Fourier transformation becomes inaccurate. In mathematics, this is called the ‘inverse problem’: How to estimate a function in one domain from incomplete (undersampled) data in another domain.

But there is more. Even if the data are complete, it may not have been possible to sample them on a regular rectangular (Cartesian) grid. In that case the Fourier transformation is time-consuming. It then becomes worthwhile to investigate whether resampling (regridding) the data – by computer – from the experimental sampling grid to a Cartesian sampling grid can be less time-consuming than Fourier transformation of non-Cartesian data.

After these preliminaries, it is now possible to state the aim of this thesis.

1. Investigation and improvement of resampling complete data from one sampling grid to another, also called regridding or gridding. I distinguish two kinds of resampling, namely from an arbitrary to a Cartesian grid, and vice versa. The usefulness of the first kind is evident: one wishes to make MRI data suitable for IFFT, enabling speedy reconstruction. The compensation for variation of the density of sampling

points requires special care. Kind 2) is useful for *e.g.* generation of radial or spiral data from Cartesian data. Such semi-real-world data are in turn useful for developing and testing image reconstruction algorithms.

2. Development of an algorithm for image reconstruction from MRI data that have been both arbitrarily sampled and locally undersampled. Note that arbitrary undersampling is not done intentionally, but rather is a consequence of unavoidable patient motion. Intentional undersampling occurs when omitting sample positions for the sake of scan time reduction. In both cases, the data contain insufficient information to reconstruct an image without additional knowledge of a general nature. Bayesian estimation is amenable to imposition of such knowledge and is therefore used in this thesis.
3. Testing of the new and existing image reconstruction algorithms with three kinds of data, namely
  - Simulated data
  - Semi-real-world data
  - Real-world data

### 1.3 Outline of this thesis

Chapter 2 describes background information deemed useful for signal processing professionals who wish to engage in image estimation from the raw MRI data which are acquired in the Fourier space (measurement space).

Chapter 3 deals with image estimation from complete raw data. Here, ‘complete’ pertains to the condition that the Nyquist sampling criterion is satisfied throughout the measurement space. In the case of Cartesian sampling, the estimator consists simply of IFFT. In other cases, two approaches are distinguished: 1) The sample positions are such that one can apply a dedicated estimator. 2) The sampling-position distribution is general, forcing one to interpolate the samples to a Cartesian grid (‘gridding’). After this interpolation, one can resort to IFFT. Much attention is given to the accuracy of interpolation. It turns out that in the case of general sample positions, the main source of error lies in properly accounting for the *varying density* of the sample positions.

Chapter 4 complements chapter 3 in the sense that it focuses on alleviating the effect of undersampling – *i.e.*, the case of local violation of the Nyquist criterion. Undersampling amounts to a fundamental lack of information. The effect of such a condition can be alleviated by invoking general prior knowledge about images. I impose this prior knowledge by applying Bayesian estimation which is particularly suitable for the task. Much attention is given to optimization of my algorithm. The computational cost per iteration has been reduced considerably. The number of iterations is yet to be limited, especially in the 3D case.

Chapter 5 concerns evaluation of image estimation from incomplete (undersampled) raw data. Issues are; image quality measures, reference images, image estimation (reconstruction) time.

Chapter 6 is devoted to the matter of how to omit spiral trajectories – *i.e.*, sparse sampling – with minimal loss of information. Ideally, the raw data satisfy Hermitian symmetry. This property can be exploited when designing and omitting trajectories. I investigate various symmetric cases.

Chapter 7 deals with applications. The methods alluded to above are tested on simulations and real-world data. The Bayesian estimator appears particularly suited for pseudo-random sample positions.

Chapter 8, finally, provides overall conclusions and dwells on recommendations for future research.

## Chapter 2

# MRI scan techniques

### 2.1 Introduction

In this section, I consider those aspects of sampling an MRI signal, deemed necessary to make signal processing professionals aware of some constraints attendant to MRI scanning.

In thermal equilibrium, the magnetic nuclei (spins) to be imaged are aligned along a strong magnetic field  $\vec{B}$ , applied along the  $z$ -axis [5]. This results in a macroscopic magnetic moment  $\vec{M}$ , parallel to  $z$ . When tilted,  $\vec{M}$  will precess around  $\vec{B}$  with the Larmor frequency, equal to

$$\omega = -\gamma|\vec{B}|, \quad (2.1)$$

where  $\gamma$  is the gyromagnetic ratio of the spins under investigation, usually protons (for protons  $\gamma/2\pi = 42.576$  MHz/T). The precessing magnetic moment induces the MRI signal in the receiver antenna. See Fig. 2.1.

As mentioned in chapter 1, the strength of  $\vec{B}$  in a scanner is made dependent on the spatial coordinates  $x, y, z$  and time  $t$  by special computer-controlled hardware [2, 3]. Hence we write  $\vec{B}(x, y, z, t)$ . As a consequence of Eq. 2.1,  $\omega$  undergoes the same spatial and time dependence as  $\vec{B}$ . We shall see below that this dependence serves to encode the MRI signal such that one can obtain an image through Fourier transformation.

Recapping, the magnetic field  $\vec{B}(x, y, z, t)$  comprises a static homogeneous part  $B_0$  and a space/time-dependent part. Note that both parts are assumed to be directed along the  $z$ -axis. As a consequence of the latter, we drop the vector bars in the corresponding equation,

$$B(x, y, z, t) = B_0 + xG_x(t) + yG_y(t) + zG_z(t), \quad (2.2)$$

where  $G_x(t)$ ,  $G_y(t)$ , and  $G_z(t)$  determine the strengths and time-dependences of linear gradients of  $B$ . In scanner jargon, one speaks of the  $x$ ,  $y$ , and  $z$  gradients respectively.

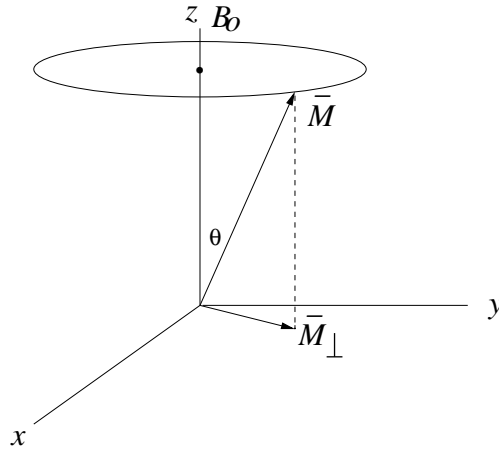


Figure 2.1: Precession of a macroscopic magnetization vector  $\vec{M}$ , tilted over an angle  $\theta$  with respect to a static homogeneous magnetic field  $B_0$  along the  $z$ -axis. The component of  $\vec{M}$  in the  $x, y$ -plane,  $\vec{M}_\perp$ , induces the MRI signal in the antenna of the scanner.

Assume now that  $\vec{M}$  is tilted away from the  $z$ -axis by a computer-controlled radio-frequency pulse and thus makes a component  $\vec{M}_\perp$  in the  $x, y$  plane. Fig. 2.1 shows an example where the tilt angle equals  $\theta$ . How to achieve this is beyond the scope of my thesis. See (*e.g.* [2]) for a good description. Also, relaxation affecting the magnitude of  $\vec{M}_\perp$  is ignored till further notice. Let us consider the precession of  $\vec{M}$  under this simplifying condition.

Denoting the density of the magnetic moment at  $x, y, z$  and time  $t$  by the magnetization  $\vec{m}(x, y, z, t)$ , the precession angle – or phase – accumulated by  $\vec{m}(x, y, z, t)dV$  in a volume  $dV$  at  $x, y, z$  is

$$\phi(x, y, z, t) = - \int_{t'=0}^t \gamma |\vec{B}(x, y, z, t')| dt' = \omega_0 t - 2\pi (xk_x(t) + yk_y(t) + zk_z(t)), \quad (2.3)$$

with the definitions

$$\omega_0 = -\gamma |\vec{B}_0|, \quad (2.4)$$

$$k_x(t) = \frac{\gamma}{2\pi} \int_{t'=0}^t G_x(t') dt', \quad (2.5)$$

$$k_y(t) = \frac{\gamma}{2\pi} \int_{t'=0}^t G_y(t') dt', \quad (2.6)$$



$$k_z(t) = \frac{\gamma}{2\pi} \int_{t'=0}^t G_z(t') dt'. \quad (2.7)$$

Note that these definitions differ from those given in [2] by a factor of  $2\pi$ . With this choice of  $\vec{k}$  it is possible to get rid of any normalization constants in Fourier transforms. These formulae indicate that the accumulated precession angle at position  $x, y, z$  relative to that at the origin equals  $-2\pi$  times the inner product of the position vector  $\vec{r} = (x, y, z)$  and a vector  $\vec{k}(t) = (k_x(t), k_y(t), k_z(t))$ , defined in so-called  $k$ -space by Eqs. 2.5 - 2.7. The corresponding time dependence of the magnetization at  $(x, y, z)$  is

$$m_{\perp}(x, y, z, t) = m_{\perp}(x, y, z, 0) e^{i\phi(x, y, z, t)} = m_{\perp}(x, y, z, 0) e^{i\omega_0 t} e^{-2\pi i(xk_x(t) + yk_y(t) + zk_z(t))}, \quad (2.8)$$

where  $m_{\perp} = m_x + im_y$ . Integrating then over the field of view (FOV) gives the macroscopic perpendicular magnetic moment

$$M_{\perp}(t) = \iiint_{\text{FOV}} m_{\perp}(x, y, z, t) dV = e^{i\omega_0 t} \iiint_{\text{FOV}} m_{\perp}(x, y, z, 0) e^{-2\pi i(xk_x(t) + yk_y(t) + zk_z(t))} dV. \quad (2.9)$$

After removing the term  $e^{i\omega_0 t}$  in the detection stage, one finally arrives at the signal

$$S(\vec{k}(t)) \propto \iiint_{\text{FOV}} I(\vec{r}) e^{-2\pi i\vec{k}(t) \cdot \vec{r}} dV, \quad (2.10)$$

where  $m_{\perp}(x, y, z, 0)$  has been replaced by  $I(\vec{r})$ . The latter explicit replacement should make one aware of the fact that MRI images the nuclear magnetization rather than the nuclear density. Since the magnetization is influenced by relaxation, one possesses a handle to distinguish between tissues that have equal nuclear densities but are subject to different degrees of relaxation. This aspect is beyond the scope of this thesis; see Ref. [2] for details.

Eq. 2.10 states that the MRI signal is proportional to the Fourier transform of the image. As it stands, it implies that if one samples a signal in 3D  $k$ -space, subsequent inverse Fourier transformation yields a full, 3D image of the object under investigation. In general  $S(\vec{k}(t))$  is complex-valued. Fig. 2.2 shows an example of the absolute value of a 2D MRI signal on a logarithmic scale.

At this point, two comments are in order.

1. Often, one restricts a scan to two dimensions. To this end, one tilts (excites)  $\vec{M}$  only in a slice of interest and subsequently adapts the time dependence of  $G_x, G_y, G_z$  accordingly. In this way, one can produce 2D images of slices at various locations and in various orientations. As for slice orientation, one distinguishes transverse, coronal, and sagittal images. The corresponding slices are respectively

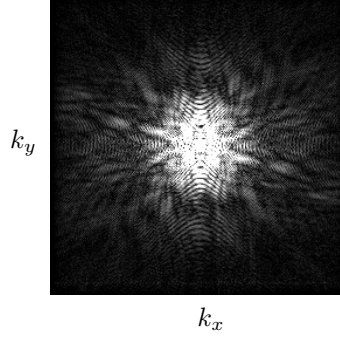


Figure 2.2: Absolute value of a 2D  $k$ -space signal on a logarithmic scale. Note that the signal is strong in the centre and weak at the edge.

perpendicular to the axis of a patient, from front to back along the axis, and from left to right along the axis. Fig. 2.3 shows an example of each mentioned orientation. Obviously, 2D scans require less time and less computing capacity than 3D scans.

2. Recall that I have ignored relaxation, which destroys  $M_{\perp}$  according to  $e^{-t/T_2^*}$ . When  $T_2^*$  is very long, it would suffice to tilt (excite) only once, and after this one could cover all  $k$ -space by proper control of  $G_x, G_y, G_z$  as a function of time. See *e.g.* Ref. [6–8] and section 2.2.3. This is the ideal case.

In reality, relaxation is fast ( $T_2^* \approx 20$  ms), causing rapid decay of the signal. In that case, one has to cover  $k$ -space piecewise, *i.e.* exciting the nuclear spins several or many times. Before each new excitation, one has to await restoration of thermal equilibrium to an extent that is dictated by the type of scan at hand. The path covered in  $k$ -space by  $\vec{k}(t)$  from one excitation till the next is called a trajectory. The time  $t$  in Eq. 2.10 is defined only during the course of a trajectory. The total scan time is determined mainly by the repeated waiting for (partial) thermal equilibrium before each new excitation. Hence, the fewer the number of excitations required, the shorter the total scan time.

## 2.2 $k$ -Space trajectories

From the previous section it is clear that MRI signals are measured in  $k$ -space. The gradients  $G_x, G_y, G_z$  determine the path or trajectory of the endpoint of  $\vec{k}$  through  $k$ -space [9, 10]. Along a trajectory the signal  $S(\vec{k}(t))$  is sampled at regular time intervals. The stronger the gradients, the shorter the time required to cover a trajectory. As mentioned in the previous section, ideally, all  $k$ -space is covered with a single, long trajectory. However, due to relaxation and inhomogeneity of  $B_0$ , the ideal case of only a single, long trajectory is often not possible. The total scan time is determined mainly by

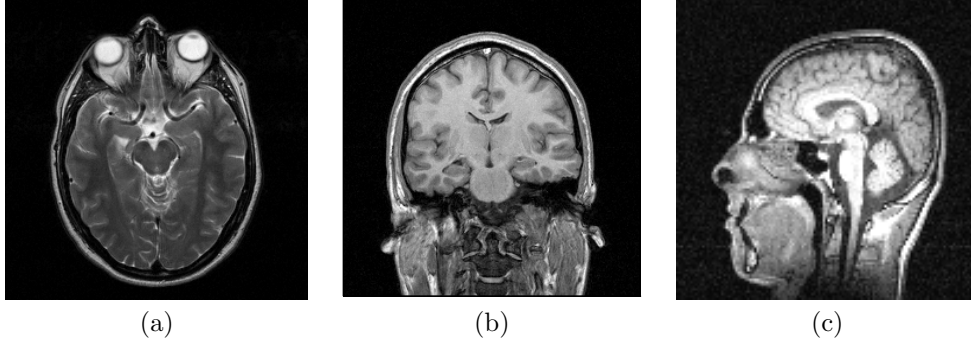


Figure 2.3: 2D images of slices of a human head in various orientations. (a) Transverse, (b) coronal, (c) sagittal.

the accumulated waiting times between successive trajectories used to cover all  $k$ -space (2D or 3D).

Many different trajectories have been conceived. I shall discuss mainly Cartesian, radial and spiral trajectories in the 2D case. More can be found in the review article by Hennig [10].

### 2.2.1 Cartesian scan

The most often used way of sampling  $k$ -space is Cartesian, depicted in Fig. 2.4. The preference for Cartesian has to do with two advantages. 1) To obtain an image, inverse FFT suffices. 2) Sampling starts at the edge of  $k$ -space where the signal is very weak. Hence, sampling errors which tend to occur in the initial part of a trajectory have little influence.

With Cartesian scanning,  $k$ -space is sampled along straight trajectories parallel to the  $k_x$ -axis. Along a trajectory, the sampling interval usually equals the distance between the successive trajectories, resulting in a square sampling pattern. Sampling along a trajectory starts at one edge of  $k$ -space and stops at the opposite edge. In this period,  $G_x$  is constant and  $G_y = 0$ , resulting in  $k_x(t) = (\gamma/(2\pi))G_x t$  and  $k_y(t)$  is constant. Prior to each trajectory, one needs gradients to go from  $\vec{k} = \vec{0}$  to the starting point of that trajectory. The corresponding paths are not shown in Fig. 2.4. If the gradient strength and relaxation allow, one can combine a number of straight trajectories shown in Fig. 2.4 into a single trajectory. This results in echo planar imaging (EPI), as discussed in Refs. [2, 6].

### 2.2.2 Radial scan

When  $T_2^*$  is short, there is no time to first go to the edge of  $k$ -space and subsequently start a Cartesian trajectory. In that case one resorts to radial scanning. In radial scanning

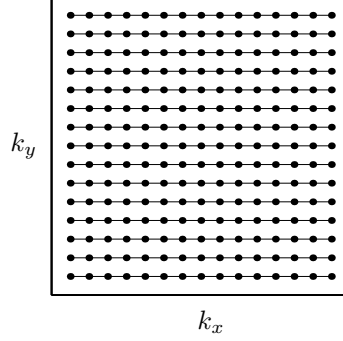


Figure 2.4: Cartesian  $k$ -space trajectories. The necessary paths from  $\vec{k} = \vec{0}$  to the starting point of each trajectory are not shown.

one starts at the centre of  $k$ -space and proceeds to the edge as fast as possible, *i.e.* along some straight line. In order to satisfy the Nyquist sampling criterion up to the edge of  $k$ -space, one needs ( $\pi$  times) more radial trajectories than Cartesian trajectories. On the other hand, artefacts caused by radial undersampling – in outer  $k$ -space – are less troublesome than artefacts caused by Cartesian undersampling.

Several radial scan alternatives suitable for 2D MRI exist. These are:

1. Projection Reconstruction (PR)

The oldest and most popular radial variant is the Projection Reconstruction (PR) method. In PR, all angles between successive radial trajectories are equal, requiring  $G_x = G_{\max} \cos(\alpha)$ ,  $G_y = G_{\max} \sin(\alpha)$ , with  $\alpha = n\Delta\alpha$ ,  $n = 0, \dots, N_{\text{rad}} - 1$ , and

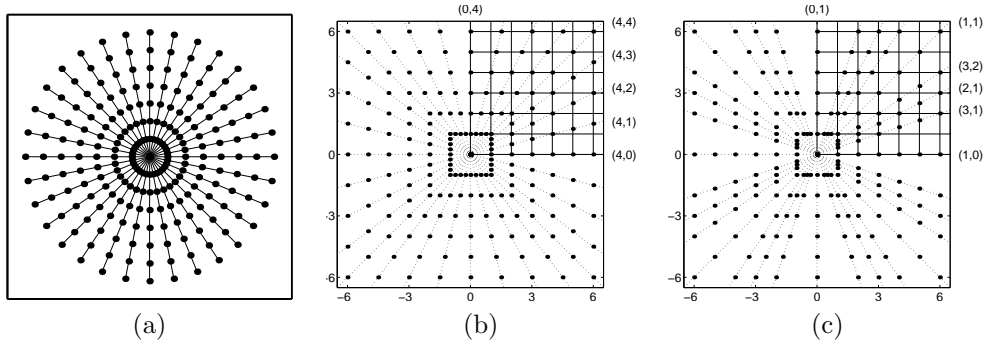


Figure 2.5: Radial  $k$ -space trajectories. (a) Projection Reconstruction (PR). (b) Lino-gram. (c) SRSFT. Taken from G.H.L.A. Stijnman's MSc Thesis [11].

$\Delta\alpha = 2\pi/N_{\text{rad}}$ , with  $N_{\text{rad}}$  the number of radials. In addition, the sampling interval

is independent of  $\alpha$ ; see Fig. 2.5a. Note that the sample positions reside on radial as well as circular lines. See Ref. [12] for recent applications in MRI.

After inverse FFT of every radial, the reconstruction becomes equal to that of X-ray tomography. The latter entails interpolation which causes some loss of accuracy. In addition, it requires more computation time. To alleviate these two drawbacks, Edholm *et al.* [13, 14] devised linogram scanning, described in the next point.

## 2. Linogram

With linogram scanning, the angles between successive radials, and the sample interval on each radial, vary such that the sample positions reside on radials as well as squares, see Fig. 2.5b. This makes reconstruction amenable to an FFT-based technique called chirp-z [11, 15–17], which is faster than projection reconstruction and obviates interpolation.

Linogram scanning is applied in X-ray tomography [13, 14, 18], in ultrasound tomography [19], and in MRI [20, 21].

## 3. SRSFT

The abbreviation SRSFT stands for Sparse Radial Scanning and Fourier Transform [22–24]. As with linogram, the angles between successive radials and the sample interval on each radial vary again such that the sample positions reside on radials as well as squares, see Fig. 2.5c. However, the number of sample positions that coincide with Cartesian grid points near the centre of  $k$ -space is maximized. An advantage of this sample distribution is that a preliminary fast (but low-resolution) reconstruction can be provided with inverse FFT. A price to pay is that some outer regions of  $k$ -space are undersampled, especially near the  $k_x$  and  $k_y$  axes.

### 2.2.3 Spiral scan

Like radial trajectories, spiral ones start at the center of  $k$ -space. However, these trajectories are not straight, but rather spiral around thus covering more of  $k$ -space. This allows a smaller number of trajectories to fulfill the Nyquist criterion, however at the expense of a longer acquisition time per trajectory [25–27].

A trajectory is an Archimedean spiral if described by the equation

$$k(t) = A\theta(t)e^{i(\theta(t)+\theta_0)}, \quad (2.11)$$

where  $k(t) = k_x(t) + ik_y(t)$ . The function  $\theta(t)$  does not determine the shape of the trajectory, it only determines the sample positions along the trajectory. The associated gradients are

$$G(t) = \frac{2\pi}{\gamma} A(1 + i\theta(t)) \frac{d\theta(t)}{dt} e^{i(\theta(t)+\theta_0)}, \quad (2.12)$$

where  $G(t) = G_x(t) + iG_y(t)$ . Note that for large  $\theta(t)$ ,  $|G(t)| \propto \theta(t) \frac{d\theta(t)}{dt}$ .

A nice property of an Archimedean spiral is the constant distance ( $2\pi A$ ) between windings. This helps to design the correct spiral trajectory for a given FOV. Ref. [28] provides a recent clinical application of spiral scanning.

Filling  $k$ -space with only one spiral trajectory is possible although acquisition time will be long, which can cause problems due to  $B_0$  inhomogeneities. The solution is to use more spiral trajectories with different starting angles  $\theta_0$ .

The function  $\theta(t)$  determines the sampling along the spiral trajectory. Several choices are possible. The use of velocity in the following descriptions comes from interpreting  $k(t)$  as a position vector and  $G(t)$  as its associated velocity.

1. constant angular velocity spiral

$\theta(t) \propto t$ . For large  $\theta(t)$  the angular velocity is constant, because  $|G(t)|/|k(t)| \propto \theta(t) \frac{d\theta(t)}{dt} / \theta(t) = \text{const.}$

2. constant linear velocity spiral

$\theta(t) \propto \sqrt{t}$ . This gives  $|G(t)| = \text{const.}$  for large  $\theta(t)$ . The expression is not valid close to the centre of  $k$ -space, since the gradient will become infinite. Therefore one starts with a constant angular velocity spiral which changes into a constant linear velocity spiral as soon as the gradients allow this.

3. optimal spiral

$\theta(t)$  is numerically optimized given the maximum gradient and slewrate ( $\frac{dG(t)}{dt}$ ) of the system [29, 30]. This is the preferred acquisition scheme, since it will give the shortest acquisition time with a given gradient system.

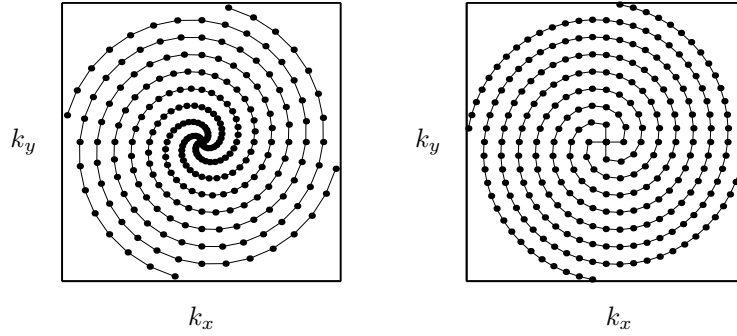


Figure 2.6: Constant angular velocity and constant linear velocity spiral trajectories.

Originally, the research for this thesis was aimed at sparse spiral scanning [31–33]. Unfortunately, economical conditions forced our industrial partner to abandon activities on spiral scanning halfway through my project. As a result, only a few preliminary spiral measurements could be put at my disposal. Only gradually, background information about the measurements turned out to be inadequate to warrant pursuit of my original subject.

## 2.3 Some other trajectory shapes

It should be mentioned that additional trajectory shapes have been reported. These are circular [34, 35], rosette [36–38], and stochastic trajectories [7, 39]. In addition, there is unintentional pseudo-random sampling, arising through correction of Cartesian scans for patient motion [40].

## 2.4 Summary

This chapter provided some background knowledge about MRI scanning deemed necessary for the signal processing community [41, 42]. Summarizing, the following points were dealt with.

- Sampling is done trajectory-wise in the Fourier domain, called  $k$ -space.
- Often, sample positions do not coincide with a rectangular (Cartesian) grid.
- The number of trajectories required to satisfy the Nyquist criterion everywhere in  $k$ -space, depends on the shape of the trajectories. For instance, one needs fewer spiral trajectories than radial trajectories.
- Each trajectory shape has its advantages and disadvantages. Technical realization of particular shapes is not a trivial matter.
- The measurement time is proportional to the number of trajectories. Hence, omission of trajectories results in reduction of measurement time.
- Omission of trajectories entails local undersampling, which in turn causes artefacts in the image.





## Chapter 3

# Image reconstruction from complete data

### 3.1 Introduction

Image reconstruction is the process of estimating an image from the raw data. This process is necessary because the raw data are acquired in the  $k$ -space and are therefore not directly interpretable as an image.

As already explained in chapter 2,  $k$ -space is the Fourier transform of the image space. The Fourier transform being invertible, one should in principle be able to estimate an image from the  $k$ -space data by applying the inverse Fourier-transform.

A number of aspects is to be dealt with:

- The raw data are sampled at discrete positions in  $k$ -space and are therefore not amenable to the continuous Fourier transform. Hence, one is forced to discretize the continuous Fourier transform which gives rise to various types of errors. This will be discussed in this chapter.
- An important aspect is the distribution of sample positions. First of all, this distribution should be sufficiently dense, implying that the Nyquist sampling criterion be satisfied everywhere in the covered  $k$ -space. If this requirement is met, we call the data complete or fully sampled. If the Nyquist sampling criterion is not met everywhere we speak of (local) undersampling. In this chapter, the data are complete.
- Complete data may have been sampled in various ways. One may distinguish the following types of sample position distributions
  - Cartesian; *i.e.* rectangular and uniform.
  - Regular nonuniform; *i.e.* governed by a (simple) expression, *e.g.* spiral.
  - Irregular or random; *e.g.* Cartesian, disturbed by random patient motion.

Various reconstruction methods exist. In this chapter, I distinguish

- Straightforward IFFT, for Cartesian sample position distributions.
- ‘Dedicated’ methods for some regular nonuniform sample position distributions such as linogram.
- Gridding, for arbitrary sample position distributions.
- Maximum likelihood reconstruction, which tries to estimate a desired set of parameters in such a way that the probability of the measured data given these parameters is maximized.

## 3.2 Discretization of the inverse Fourier transform

### 3.2.1 Reconstruction by DFT

The continuous inverse Fourier transform is defined as

$$I(\vec{r}) = \int S(\vec{k}) e^{2\pi i \vec{k} \cdot \vec{r}} d\vec{k}. \quad (3.1)$$

Reconstruction by Eq. 3.1 needs the data at all  $\vec{k}$ -positions. However,  $S(\vec{k})$  is sampled at discrete  $\vec{k}$ -positions,  $\vec{k}_n$ . The sampling forces us to discretize Eq. 3.1, which results in the following approximation:

$$I(\vec{r}) \simeq I_{\text{recon}}(\vec{r}) = \sum_n S(\vec{k}_n) e^{2\pi i \vec{k}_n \cdot \vec{r}} \Delta k_n \quad (3.2)$$

in which  $\Delta k_n$  represents a patch of  $k$ -space around the sample at  $\vec{k}_n$ . Finding a suitable size of  $\Delta k_n$  constitutes a major problem whenever the sample positions are irregular. The problem is referred to as density compensation because intuitively one may relate each  $\Delta k_n$  with the area/volume of the mentioned patch of  $k$ -space. See next section. Fig. 3.1 shows the discretization and integration process, which is actually an application of the trapezoidal rule known from numerical integration [43]. Applying the usually more accurate Simpson’s rule is not applicable here, because the complex exponential is not smoothly varying for large  $\vec{r}$ .

### 3.2.2 Sampling density compensation

Substituting Eq. 2.10 into Eq. 3.2, one can express  $I_{\text{recon}}(\vec{r})$  as the convolution of the true continuous image with a so-called point-spread function.

$$I_{\text{recon}}(\vec{r}) = \sum_n \left[ \int I(\vec{r}') e^{-2\pi i \vec{k}_n \cdot \vec{r}'} d\vec{r}' \right] e^{2\pi i \vec{k}_n \cdot \vec{r}} \Delta k_n \quad (3.3)$$

$$= \int \left[ \sum_n \Delta k_n e^{2\pi i \vec{k}_n \cdot (\vec{r} - \vec{r}')} \right] I(\vec{r}') d\vec{r}' \quad (3.4)$$

$$= \text{PSF}(\vec{r}) * I(\vec{r}), \quad (3.5)$$

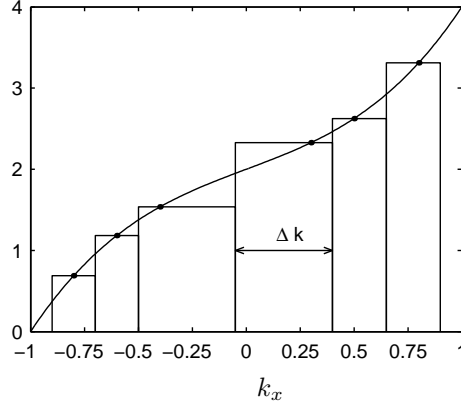


Figure 3.1: Discretization of integral. The value of the function is only known at the points denoted by a circle. This picture shows the necessity of including the  $\Delta k_n$  term in Eq. 3.2 to establish a correct integration.

in which  $*$  stands for convolution, and the point-spread function,  $\text{PSF}(\vec{r})$ , is defined in MRI as

$$\text{PSF}(\vec{r}) \stackrel{\text{def}}{=} \sum_n \Delta k_n e^{2\pi i \vec{k}_n \cdot \vec{r}}. \quad (3.6)$$

Eq. 3.5 shows that the reconstructed image approaches the true image when  $\text{PSF}(\vec{r})$  approaches a delta function. Note that in MRI, the point-spread function is not related to an experimental condition such as aperture in optics. Rather, the MRI point-spread function is a mathematical contraption that can be manipulated through the choice of the density compensation which is represented by the quantities  $\Delta k_n$ . Ideally, one should find a density compensation that makes  $\text{PSF}(\vec{r})$  approach a delta-function located at  $\vec{r} = \vec{0}$ . Degradation of image reconstruction caused by an inadequate  $\text{PSF}(\vec{r})$  – or inadequate sampling density compensation – can not be repaired later on.

One can distinguish the following alternative ways to obtain a sampling density compensation.

- Trivial case.  
When the sample distribution is Cartesian, all  $\Delta k_n$  are equal for reasons of symmetry. This allows replacement of the  $\Delta k_n$  by a constant in front of the summation operator.
- Analytical derivation [11].  
When the sample distribution is nonuniform but regular, derivation of an analytical expression for the density compensation is feasible. For example, consider a radial sample distribution. In this case one can rewrite Eq. 3.1 in polar coordinates and

then discretize. This results in

$$I_{\text{recon}}(\vec{r}) = \Delta k_r \Delta k_\theta \sum_n S(\vec{k}_n) e^{2\pi i \vec{k}_n \cdot \vec{r}} J(k_{r_n}, k_{\theta_n}), \quad (3.7)$$

where  $J(k_{r_n}, k_{\theta_n})$  is the Jacobian resulting from the transformation to polar coordinates.  $\Delta k_r$  and  $\Delta k_\theta$ , the step-sizes in radial and angular direction respectively, can be taken out of the summation because they are constants. Comparing Eq. 3.7 with Eq. 3.2 one can associate  $J(k_{r_n}, k_{\theta_n})$  with  $\Delta k_n$ , which means that we have found an analytical expression for the density compensation.

- The Voronoi algorithm [44, 45].  
This algorithm constructs a polygon (in 2D) or polyhedron (in 3D) around each sample position, in such a way that all points in the polygon/polyhedron are closest to the accompanying sample position. The area/volume of the polygon/polyhedron is substituted for  $\Delta k_n$ , on intuitive geometrical grounds. This algorithm is contained in MATLAB. It can be used when an analytical expression is not available [11, 46, 47]. However, the  $\Delta k_n$  found with this method can vary wildly from sample to sample when sampling is irregular, resulting in a PSF that does not resemble a delta-function.
- The Pipe-Menon algorithm [48].  
This algorithm has already been alluded to above. It is not based on geometric considerations. Rather, it varies all  $\Delta k_n$  in such a manner that the point-spread function approaches a delta function. This constitutes a superior philosophy. But the computational cost may be high. I have made two implementations using the C language [46, 47]. One of these constitutes an improvement of the original Pipe algorithm. The details of these implementations are beyond the scope of this thesis.

### 3.3 Cartesian sampling

When a sample position distribution is Cartesian – Cartesian sampling – the notion of Nyquist rate [49] applies. Consider the 1D case. A small drop of water at position  $x_0$  brings about an MRI signal proportional to  $\exp(-2\pi i x_0 k_x)$ . Suppose the drop is at the edge of the region where aliasing starts. Then, the sampling interval  $\Delta k$  is such that  $2\pi x_0 \Delta k = \pi$ , yielding  $\Delta k = 1/(2x_0)$ . Conversely, given a  $\Delta k$ , the greatest value of  $x_0$  at which no aliasing occurs is  $1/(2\Delta k)$ . With the given  $\Delta k$ , all drops in the interval  $[-x_0, x_0]$  will be properly imaged. This leads to a so-called **Field Of View** – FOV – equal to  $2x_0 = 1/\Delta k$ . In MRI practice,  $\Delta k$  is determined by the magnitude of the gradient field and the time domain sampling interval, as indicated in chapter 2.

The above considerations allow one to write for the 1D case  $k_n = n\Delta k = n/\text{FOV}$ . Using this simple relation, Eq. 3.2 becomes

$$I_{\text{recon}}(x) = \text{FOV}^{-1} \sum_n S(k_n) e^{2\pi i x n / \text{FOV}} \quad (3.8)$$

In practical cases  $I_{\text{recon}}(x)$  is calculated only at a finite number of points,  $x_m = m\text{FOV}/N$ ,  $m = -N/2, -N/2 + 1, \dots, N/2 - 1$ , where  $N$  equals the number of  $k$ -space samples, substituting gives

$$I_{\text{recon}}(x_m) = \text{FOV}^{-1} \sum_n S(k_n) e^{2\pi i m n / N} \quad (3.9)$$

which can be calculated very efficiently with the inverse fast Fourier transform (IFFT).

The above theory extends naturally to 2D and 3D MRI. It follows that Cartesian sampling enables one to use 2D or 3D IFFT for MR image reconstruction. The corresponding FOV's are determined by the magnitudes of the available gradient fields, through Eq. 2.5-2.7.

Finally, the possibility of uniform non-Cartesian sampling exists. For example, one may sample on a uniform hexagonal grid. This option is not considered in my thesis. Hence, in the sequel *non-uniform and non-Cartesian sampling mean the same*.

## 3.4 Non-uniform sampling

### 3.4.1 Introduction

In the previous section we saw that in case of Cartesian sampling the image can be reconstructed by applying 2D or 3D IFFT. However, to apply IFFT it is essential that  $k$ -space be sampled on a Cartesian grid. In case of a non-uniformly sampled  $k$ -space, one has to cope with Eq. 3.2 which is not amenable to IFFT. Unfortunately, the computational recourse – straightforward calculation of Eq. 3.2 – is very expensive. One way out is to first interpolate the non-Cartesian samples to a Cartesian grid and then use the IFFT. This results in the gridding algorithm which is explained in section 3.4.3.

Another aspect of non-uniform sampling is establishment of a FOV. This is problematic because the sampling interval is not constant. In practice, one uses the reciprocal of the largest or average sampling interval. Last but not least, non-uniform sampling can cause a major problem with respect to estimating the quantities  $\Delta k_n$  – or density compensation.

Before treating gridding, I shall first review some radial sampling distributions – non-uniform by definition, see section 2.2.2 – that can be handled with dedicated reconstruction algorithms.

### 3.4.2 Image reconstruction for radial scans

As mentioned in chapter 2 radial scanning is useful when the relaxation times are short. Because of their symmetry, radial sample distributions are amenable to dedicated reconstruction algorithms. Here, I distinguish two alternative radial implementations. 1) The samples reside on both straight lines and *circles* – referred to as Projection Reconstruction (PR). 2) The samples reside on both straight lines and *squares*. Alternative 2) in turn can be realized in several ways. Here, I shall dwell upon so-called Linogram and SRSFT, with respectively chirp-z and sqUbe as construction algorithms.

### 3.4.2.1 Projection reconstruction

Scanning along uniformly distributed radial trajectories is generally referred to as Projection Reconstruction (PR), which is actually a name of a reconstruction method borrowed by MRI in the past from the field of X-ray tomography. An alternative name of the reconstruction method is Filtered Back Projection. PR is MRI's first scan technique, proposed by Lauterbur [50] in 1973. The sample distribution is shown in Fig. 3.2. Note that the samples reside on straight lines as well as circles.

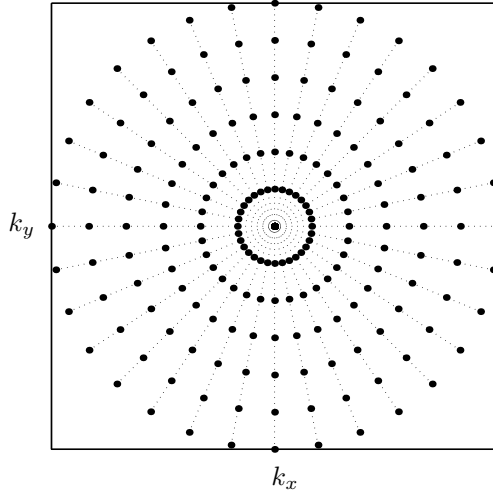


Figure 3.2: Samples on radial trajectories of a 2D PR sampling distribution. The samples are seen to reside on straight lines as well as circles.

A pictorial description of the principle of the PR reconstruction method is shown in Fig. 3.3, taken from Ref. [11]. See Refs. [15, 16] for details. One distinguishes the following steps.

1. Consider one radial trajectory only.
2. Apply 1D density compensation.
3. Apply 1D IFFT, resulting in a 1D projection image.
4. Spread the projection image over a 2D grid.
5. Rotate the spread projection image over the proper angle.
6. Interpolate the rotated spread projection image to the unrotated grid.
7. Repeat the above procedure for all trajectories and add all images. This yields the image of the scanned object.

The PR reconstruction method is still used by many. It can be made very fast by implementing dedicated hardware used in X-ray tomography. A disadvantage is that it uses interpolation, which causes some loss of accuracy.

Below, we shall see that by adapting the angles between adjacent radials and the sampling rate on each radial, such that the samples reside on straight lines and squares – rather than circles – one can resort to reconstruction algorithms that obviate interpolation. These reconstruction methods are called chirp-z and SqUbe.

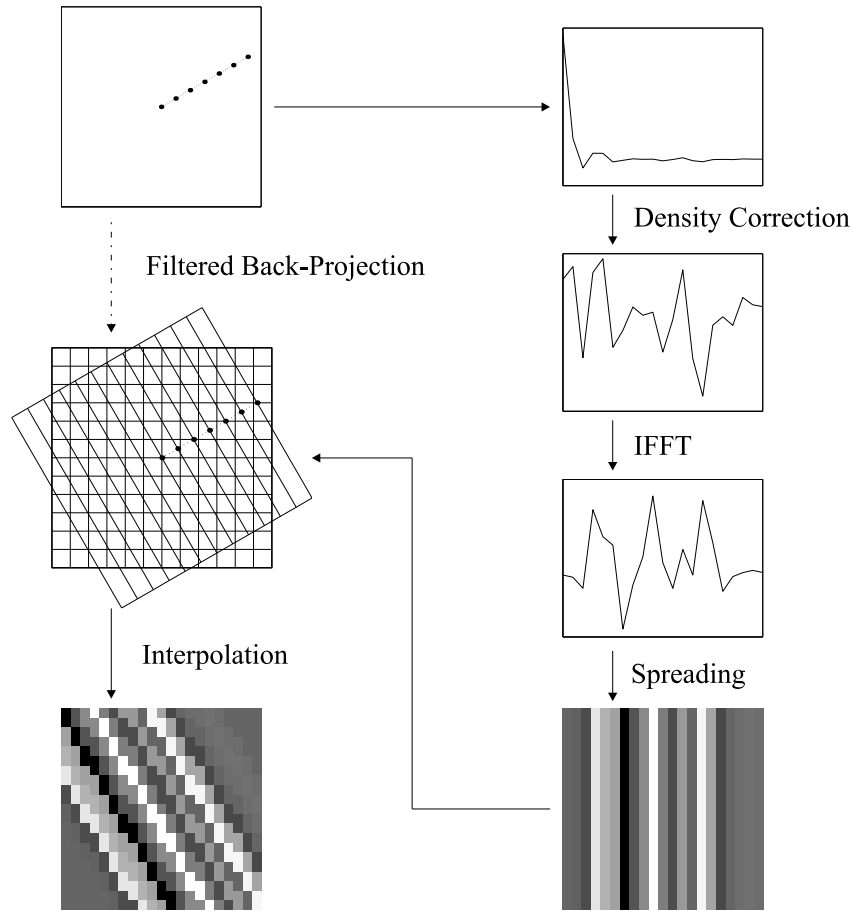


Figure 3.3: Pictorial description of filtered back-projection, applied to one radial of a PR distribution [11].

### 3.4.2.2 Linogram and the chirp-z reconstruction

Unlike PR, in linogram neither the angle between adjacent trajectories nor the radial sampling step are constant. In fact, these quantities are adapted such that the sample positions reside both on straight lines and on squares, see Fig. 3.4. This example shows a scan with  $N_{\text{rad}} = 32$  radial trajectories, while the number of samples on each trajectory is  $\frac{1}{2}N + 1 = 7$ . The  $k$ -space is partitioned into four regions, i)  $\phi \in (-\frac{1}{4}\pi, \frac{1}{4}\pi]$  and ii)  $\phi \in (\frac{1}{4}\pi, \frac{3}{4}\pi]$ , where  $\phi$  is the angle between a trajectory and the  $k_x$ -axis. Regions iii and iv are obtained from respectively regions i and ii by reflection with respect to the centre of  $k$ -space. In region i, the trajectories pass through the points  $(\frac{1}{8}N_{\text{rad}}, m)$ , where  $m$  is an integer in the region  $-\frac{1}{8}N_{\text{rad}} < m \leq \frac{1}{8}N_{\text{rad}}$ . Combining these preliminaries, the sample positions in regions i and iii of Fig. 3.4 are determined by the integers  $k$  and  $m$  through the expressions

$$\begin{aligned} Nk_x &= n, & -\frac{1}{2}N < n \leq \frac{1}{2}N, \\ Nk_y &= n \frac{8m}{N_{\text{rad}}}, & -\frac{1}{8}N_{\text{rad}} < m \leq \frac{1}{8}N_{\text{rad}}. \end{aligned} \quad (3.10)$$

In regions ii and iv, the roles of  $k_x$  and  $k_y$  are interchanged with respect to regions i and iii, respectively.

Substituting the  $k$ -space coordinates of Eq. 3.10, the argument of the 2D IDFT exponential,  $2\pi i(xk_x + yk_y)$ , becomes  $2\pi i(xn + yn(8m/N_{\text{rad}}))/N$ . The term in  $x$  is amenable to IFFT, but the term in  $y$  is not. However, chirp-z can handle the term in  $y$  by subsequently two FFTs, an inner product, and an IFFT and is therefore still fast and accurate; see Ref. [11, 51, 52] and references therein. Interpolation is obviated. A similar approach applies to 3D linogram [11].

For more details, the reader is referred to the forthcoming PhD thesis of R. Lethmate (UCB Lyon I, december 2001). In the present thesis, I limit myself to mentioning that for linogram scanning chirp-z is faster and more accurate than PR and is a good alternative to gridding.

### 3.4.2.3 SqUbe reconstruction

Linogram scanning is useful in that it enables one to perform fast and accurate image reconstruction. Note however, that the indices determining the directions of 2D linogram trajectories possess only a single variable. This limits the degrees of freedom in the design of radial/square (2D). (A similar restriction occurs in 3D radial/cubic (3D) sample distributions.) For instance, linogram does not provide the possibility to maximize the number of sample positions coinciding with a Cartesian grid. The latter circumstance prompted Graveron-Demilly to use two indices  $l$  and  $m$  for determining the directions of 2D radial/square and three for radial/cubic trajectories [22]. This enabled her to maximize the number of Cartesian sample positions in a square/cubic region around  $\vec{k} = \vec{0}$  where the signal is strongest. This distribution is indicated by the name SRSFT. See Fig. 2.5c.



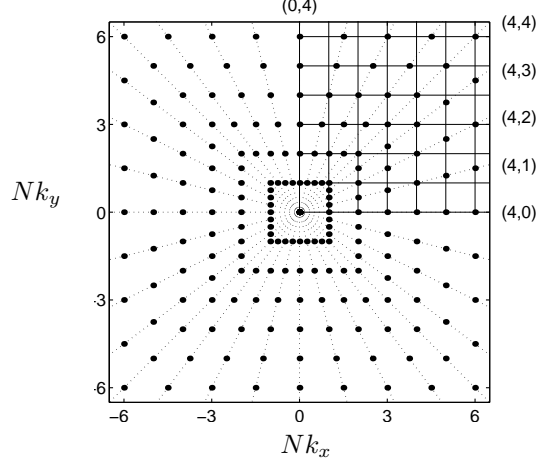


Figure 3.4: Radial trajectories in  $k$ -space of a 2D linogram distribution. Indicated are the indices  $(\frac{1}{8}N_{\text{rad}}, m)$  or  $(m, \frac{1}{8}N_{\text{rad}})$  of several directions. In this example,  $N_{\text{rad}} = 32$  and  $N = 12$  (ignoring points at  $Nk_x$  or  $Nk_y = -6$ ). Regions i, ii, iii, iv, mentioned in the text, range from respectively NE to SE, NW to NE, NW to SW, and SW to SE.

The rationale for applying SRSFT is that it allows one to very quickly reconstruct a low-resolution previewing image. The remaining points can be used later on to improve the resolution. For the latter purpose, Stijnman devised the reconstruction algorithm sqUbe (square  $\cup$  cubic) [11]. Like chirp-z, sqUbe obviates interpolation.

As mentioned above, the task is to accommodate two directional indices in the reconstruction algorithm. For the 2D case, the argument of the IDFT exponential,  $2\pi i(xk_x + yk_y)$ , becomes  $2\pi i(xk + yk(l/(mN)))$ , with  $-\frac{1}{2}N < k \leq \frac{1}{2}N$ . In SRSFT,  $l$  and  $m$  are to satisfy the relations  $|l| \leq N_s$  and  $|m| \leq N_s$ , where  $2N_s + 1$  is the size of the square around  $\vec{k} = \vec{0}$  not lacking any Cartesian positions. As a rule,  $N_s \ll N$ .

As with 2D linogram, the term involving  $y$  is again computationally expensive. However, one can easily check that

$$\exp[2\pi i(xk + (y + m)k(l/m))/N] = \exp[2\pi i((x + l)k + yk(l/m))/N],$$

which serves to greatly reduce the computational burden of the sqUbe reconstruction.

In conclusion, sqUbe is a reconstruction algorithm for general square/cubic radial sample distributions that obviates interpolation and exploits a symmetry relation to reduce the computation time.

### 3.4.3 Gridding

#### 3.4.3.1 Introduction

Gridding is an algorithm for computation of Eq. 3.2 in the case of an arbitrarily non-uniformly sampled  $k$ -space. Non-uniformity of sample positions makes it impossible to use IFFT and straightforward computation of Eq. 3.2 is time-consuming. The main feature of gridding is fast and accurate interpolation of the data from the given non-uniform grid to a uniform one, restoring the possibility of using IFFT. See Fig. 3.5 which is based on work by Zwaga [53]. Overall, gridding consists of four steps:

1. Density compensation.
2. Interpolation to a Cartesian grid.
3. Inverse fast Fourier transform.
4. Roll-off correction.

The first step of gridding, the density compensation, is common to most reconstruction methods. A notable exception is Bayesian reconstruction, treated in the next chapter. Finding an optimal density can be truly difficult, especially when the non-uniformity is severe. Recently, an important contribution to optimizing the density compensation was made by Pipe and Menon [48]. Jointly with Woudenberg and de Jong, I have made an implementation of their method. For this aspect, see [41, 46, 47, 54].

For this thesis, I have spent considerable time on the interpolation step resulting in 1) improved formulation, 2) increased accuracy, 3) reduction of computation time.

#### 3.4.3.2 Pictorial Explanation of Gridding

In gridding, interpolation is accomplished by convolving the samples with a kernel. One can compare this to sinc-convolution, the classical way to interpolate a uniformly sampled bandlimited signal [55]. What type of kernel to use will be explained in section 3.4.3.5. Note that this type of interpolation is different from normal interpolation, because in the former a sample at a Cartesian point can be affected by neighbouring samples, while in the latter it keeps its value. Moreover, the entire  $k$ -space is treated the same way.

Writing out the convolution with the kernel,  $C(\vec{k})$ , and discretizing like in Eq. 3.2

$$S'(\vec{k}) = \int S(\vec{k}')C(\vec{k} - \vec{k}')d\vec{k}' \quad (3.11)$$

$$\simeq \sum_n S(\vec{k}_n)C(\vec{k} - \vec{k}_n)\Delta k_n \quad (3.12)$$

$$= \sum_n [S(\vec{k}_n)\Delta k_n]C(\vec{k} - \vec{k}_n) \quad (3.13)$$

Eq. 3.13 shows that one has to convolve the density-compensated signal,  $S(\vec{k}_n)\Delta k_n$ , with the kernel  $C(\vec{k})$ . At first, it looks like we have not gained anything, because we have

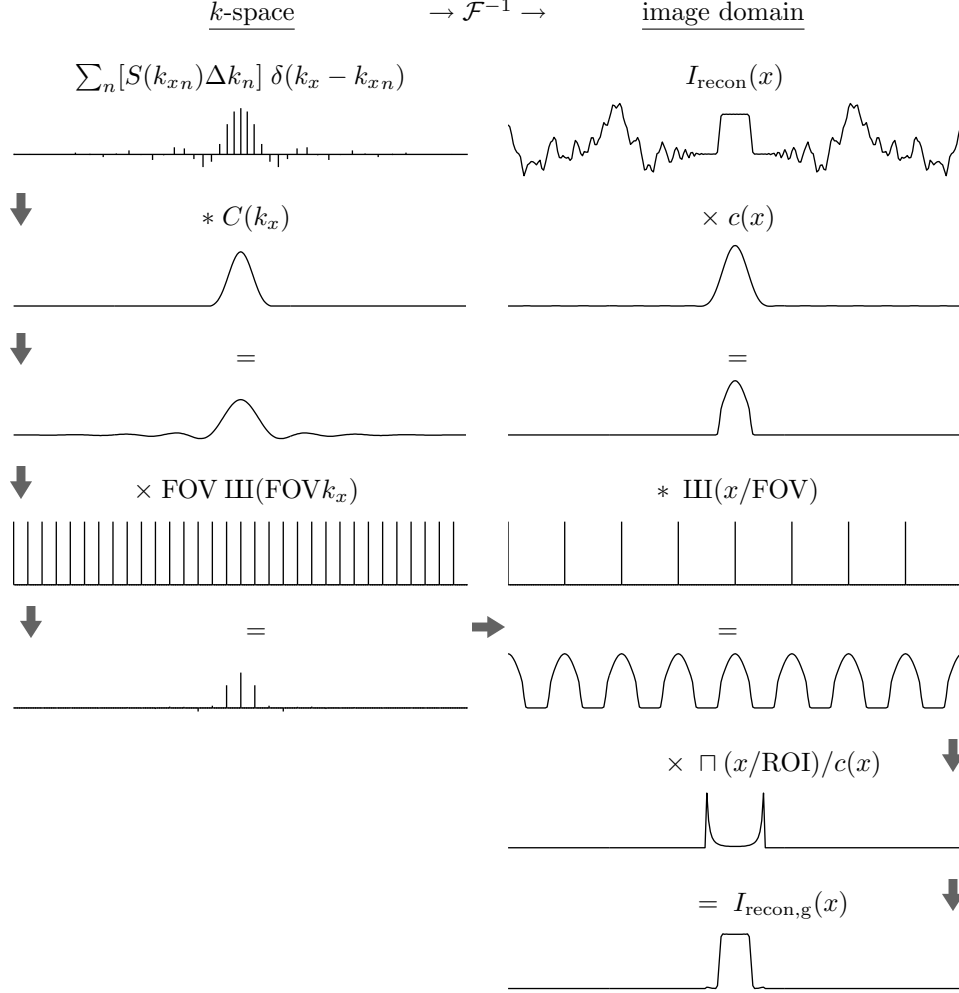


Figure 3.5: Nonuniform sampling and convolution gridding image reconstruction. The left column shows the sampling and reconstruction process in  $k$ -space, the right column shows the corresponding signals and operations in the image domain. Based on a Figure in Ref. [53]. The arrows show the actual computational route.

to sum over all  $n$  in Eq. 3.13 to get one value of  $S'(\vec{k})$ . However,  $S'(\vec{k})$  can be sampled on a Cartesian grid allowing IFFT to be used. To reduce the computation time of the convolution we use a finite-sized kernel, which means that the kernel will be zero outside a certain limited interval. This means that the summation in Eq. 3.13 is reduced to a summation over the points where we know that the kernel is non-zero. Choosing the

size of the kernel many times smaller than the total size of  $k$ -space, results in a drastic reduction of calculation time for the evaluation of the convolution in Eq. 3.13.

The signal resulting from the convolution,  $S'(\vec{k})$ , is sampled on a Cartesian grid with sampling interval  $\Delta k_{\text{Cart}}$ . This interval should be small enough to support a FOV of the same or larger size as the region of interest (ROI). This is the region where we want to know  $I_{\text{recon}}(\vec{r})$ . Unfortunately, a finite-sized kernel in  $k$ -space has infinite extent in image space. This infinite extent causes aliasing artefacts, whenever the image of the density-compensated signal  $S(\vec{k})\Delta k$  is non-zero outside the FOV defined by the resampling interval,  $\Delta k_{\text{Cart}}$ , of the Cartesian grid we are convolving to. Fortunately, these artefacts can be made arbitrarily small by choosing the kernel appropriately and making  $\Delta k_{\text{Cart}}$  small enough. This will be explained in section 3.4.3.6. Fig. 3.6 shows that the part of the signal that will alias into the ROI. It is contained in ROI-wide regions centered around multiples of FOV. The inverse Fourier transform of the kernel should have the property that it suppresses any signal present in these regions. Regions that do not alias into the ROI are indicated by “don’t care”. Finally, I dwell briefly on steps 3 and 4 of gridding. In step 3, the resampled result of the convolution is IFFT’ed to the image domain. When we choose  $\Delta k_{\text{Cart}}$  smaller than required to support the ROI, we cut off the outer part. Step 4 corrects for the shape of the inverse of the convolution kernel. The convolution with the kernel in  $k$ -space is equivalent to a multiplication with the inverse Fourier transform of the kernel in the image domain,  $c(x)$ . To correct for the non-rectangular shape of  $c(x)$ , the final step consists of division by this function. This is called the roll-off correction.

### 3.4.3.3 Algebraic explanation

This section will give a more formal derivation of the gridding algorithm with an arbitrary number of dimensions.

Writing down the relation between the kernel  $C(\vec{k})$  and its inverse Fourier transform  $c(\vec{r})$  gives

$$c(\vec{r}) = \int C(\vec{k}) e^{2\pi i \vec{k} \cdot \vec{r}} d\vec{k}. \quad (3.14)$$

Multiplying both sides with  $e^{2\pi i \vec{k}_n \cdot \vec{r}}$  one gets

$$e^{2\pi i \vec{k}_n \cdot \vec{r}} c(\vec{r}) = \int C(\vec{k}) e^{2\pi i (\vec{k} + \vec{k}_n) \cdot \vec{r}} d\vec{k} = \int C(\vec{k} - \vec{k}_n) e^{2\pi i \vec{k} \cdot \vec{r}} d\vec{k}. \quad (3.15)$$

Discretizing the integral on the right hand side of Eq. 3.15 on a Cartesian grid with interval  $\Delta k_{\text{Cart}}$  gives

$$e^{2\pi i \vec{k}_n \cdot \vec{r}} c(\vec{r}) \simeq \Delta k_{\text{Cart}} \sum_m C(\vec{k}_m - \vec{k}_n) e^{2\pi i \vec{k}_m \cdot \vec{r}}. \quad (3.16)$$

The step-size of this grid,  $\Delta k_{\text{Cart}}$ , is chosen smaller than the step-size needed to support the desired ROI.

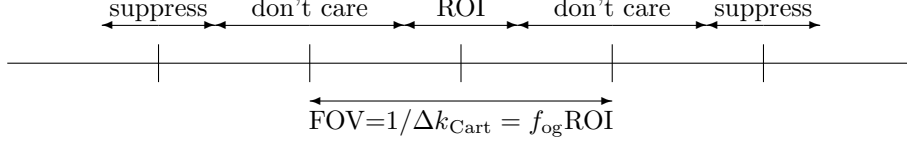


Figure 3.6: Necessary properties of the convolution kernel. By choosing  $\Delta k_{\text{Cart}}$  small enough, FOV will be larger than the ROI. This results in don't care regions that will alias into the FOV but not into the ROI. The suppress regions are where the inverse Fourier transform of the convolution kernel should be ideally zero.

The discretization causes parts of  $e^{2\pi i \vec{k}_n \cdot \vec{r}} c(\vec{r})$  outside the FOV,  $x, y, \dots \in [-1/(2\Delta k_{\text{Cart}}), 1/(2\Delta k_{\text{Cart}})]$  to alias into the ROI. The error in this approximation is dependent on the choice of  $C(\vec{k})$  and will be elaborated in section 3.4.3.6.

We are only interested in a limited region of the image domain, *i.e.* the ROI. Suppose  $c(\vec{r})$  is non-zero inside this region so we can safely divide by it, giving

$$e^{2\pi i \vec{k}_n \cdot \vec{r}} \simeq \frac{\Delta k_{\text{Cart}}}{c(\vec{r})} \sum_m C(\vec{k}_m - \vec{k}_n) e^{2\pi i \vec{k}_m \cdot \vec{r}}. \quad (3.17)$$

Eq. 3.17 gives us a recipe how to replace a complex exponential  $e^{2\pi i \vec{k}_n \cdot \vec{r}}$  with a  $\vec{k}_n$  not residing on a Cartesian grid with a linear combination of complex exponentials with their  $\vec{k}_m$  on a Cartesian grid. The weights of the linear combination are determined by the convolution kernel  $C(\vec{k})$ .

Using Eq. 3.17, Eq. 3.2 becomes

$$I_{\text{recon}}(\vec{r}) \simeq I_{\text{recon,g}} = \sum_n S(\vec{k}_n) \left( \frac{\Delta k_{\text{Cart}}}{c(\vec{r})} \sum_m C(\vec{k}_m - \vec{k}_n) e^{2\pi i \vec{k}_m \cdot \vec{r}} \right) \Delta k_n. \quad (3.18)$$

Finally, exchanging the summations gives

$$I_{\text{recon,g}}(\vec{r}) = \frac{\Delta k_{\text{Cart}}}{c(\vec{r})} \sum_m \left( \sum_n [S(\vec{k}_n) \Delta k_n] C(\vec{k}_m - \vec{k}_n) \right) e^{2\pi i \vec{k}_m \cdot \vec{r}}. \quad (3.19)$$

The inner summation is the convolution of the kernel with the non-uniform samples. The outer summation is an IFFT. And the division by  $c(\vec{r})$  is the roll-off correction. So Eq. 3.19 describes the gridding algorithm.

The above derivation shows exactly which approximations occur in the gridding algorithm due to the interpolation. There is only one, namely Eq. 3.17. The error in approximating the continuous inverse Fourier transform by Eq. 3.2 – which includes the density compensation – is shared with other reconstruction methods and can therefore not be blamed only on gridding. As I will show in the following sections, the error in Eq. 3.17 can be made arbitrarily small.

### 3.4.3.4 Interpolation error

We will consider convolution kernels that are separable in the  $k$ -coordinates, *i.e.*  $C(k_x, k_y, \dots) = C_1(k_x)C_1(k_y)\dots$ . This means that the inverse Fourier transform of the kernel is also separable, *i.e.*  $c(x, y, \dots) = c_1(x)c_1(y)\dots$ . This allows us to reduce the examination to 1D, as is easily shown by expanding Eq. 3.17. In the following sections the subscript on  $c_1$  is dropped, for reasons of clarity.

#### 3.4.3.4.1 Error in approximating complex exponential

In the previous section it was shown that gridding involves only one error, *i.e.* the approximation in Eq. 3.17. Minimizing the absolute value of this error will minimize the gridding error. Therefore, an appropriate criterion for selection of a convolution kernel is the minimization of the maximum of

$$E_1(x, k_x) = \left| e^{2\pi i x k_x} - \frac{\Delta k_{\text{Cart}}}{c(x)} \sum_m C(k_{xm} - k_x) e^{2\pi i x k_{xm}} \right| \quad (3.20)$$

$$= \frac{1}{c(x)} \left| c(x) - \Delta k_{\text{Cart}} \sum_m C(k_{xm} - k_x) e^{2\pi i x (k_{xm} - k_x)} \right|, \quad (3.21)$$

with  $x \in [-\text{ROI}/2, \text{ROI}/2]$ .

$E_1$  is periodic in  $k_x$  with period  $\Delta k_{\text{Cart}}$ , so it is sufficient to use  $|k_x| < \Delta k_{\text{Cart}}/2$ .

#### 3.4.3.4.2 Energy criterion

Writing out the complete gridding algorithm gives

$$I_{\text{recon,g}}(x) = \frac{\square(x/\text{ROI})}{c(x)} \{ \text{III}(x/\text{FOV}) * (c(x)I_{\text{recon}}(x)) \} \quad (3.22)$$

and rearranging results in

$$I_{\text{recon,g}}(x) = \square(x/\text{ROI}) \left\{ \text{III}(x/\text{FOV}) * \left( \left[ \frac{\square(x/\text{ROI})}{c(x)} * \text{III}(x/\text{FOV}) \right] c(x)I_{\text{recon}}(x) \right) \right\}. \quad (3.23)$$

The convolution of the left shah function with the term to the right causes aliasing. The amount of energy that aliases into the ROI equals

$$\int \left| \left[ \frac{\square(x/\text{ROI})}{c(x)} * \text{III}(x/\text{FOV}) \right] c(x)I_{\text{recon}}(x) \right|^2 dx. \quad (3.24)$$

The integral should be evaluated over all  $x$  that will alias into the ROI, *i.e.*  $|x - n\text{FOV}| < \text{ROI}/2, n \neq 0$ , see Fig. 3.6.

However, the above equation depends on  $I_{\text{recon}}(x)$ . Clearly any criterion should be independent of the image to be reconstructed. Because we have no prior knowledge on what  $I_{\text{recon}}(x)$  should be, we assume it constant, resulting in

$$E_2 = \int \left| \left[ \frac{\square(x/\text{ROI})}{c(x)} * \text{III}(x/\text{FOV}) \right] c(x) \right|^2 dx. \quad (3.25)$$

Minimizing  $E_2$  gives us a second criterion for selecting a convolution kernel.

The term between brackets in Eq. 3.25 describes the influence of the roll-off correction. Without this term the criterion becomes the minimization of

$$E_3 = \int |c(x)|^2 dx. \quad (3.26)$$

This is the least squares criterion, and finding a kernel  $c(x)$  based on this criterion is well known [56]. In case  $\text{FOV} = \text{ROI}$ , the solution is the zero order Prolate Spheroidal Wavefunction (PSWF).

#### 3.4.3.5 Choice of convolution kernel

To the author's knowledge there is no way to optimize  $c(x)$  using either Eq. 3.21 or Eq. 3.25 as the criterion. However, when  $c(x)$  is more or less constant for  $|x| < \text{ROI}/2$ , we expect the solution to approach the kernel found using the least squares criterion. This will be the case for  $\text{FOV} \gg \text{ROI}$ .

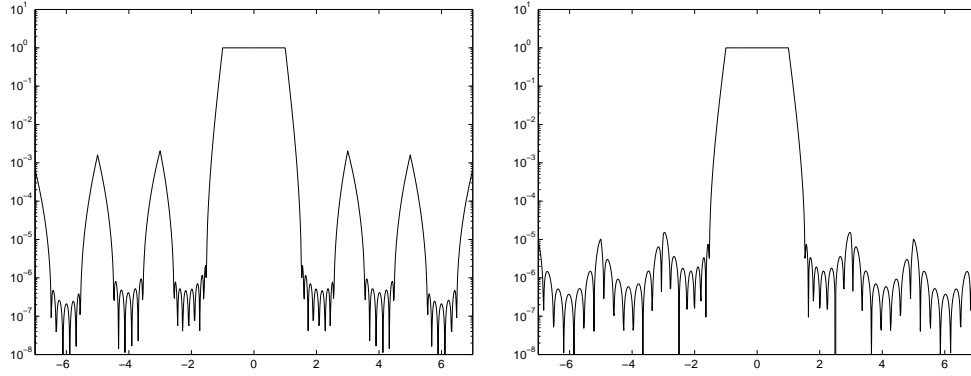


Figure 3.7: Inverse Fourier transform of two possible convolution kernels with width 3 plotted on a logarithmic scale, with  $\text{ROI} = 1$  and  $\text{FOV} = 2$ , including roll-off correction. Left: FirLS kernel. The locations where the kernel is large,  $x = \dots, -5, -3, 3, 5, \dots$ , are the don't care regions of Fig. 3.6. Right: Kaiser-Bessel kernel.

Fig. 3.7 on the left shows the inverse Fourier transform of an optimized kernel using the least squares criterion in case  $\text{ROI} = 1$  and  $\text{FOV} = 2$ , including the roll-off correction. I will call this kernel the FirLS kernel [46]. The locations where the FirLS kernel is large,  $x = \dots, -5, -3, 3, 5, \dots$ , are exactly the don't care regions of Fig. 3.6, because this is the way the kernel was designed. Note that the shape of the FirLS kernel will depend on its width  $L$ , the size of the FOV relative to the ROI. If any of these parameters change we need to recalculate the FirLS kernel. However, the kernels could be precalculated and stored in a lookup table.

Fig. 3.7 on the right shows the inverse Fourier transform of the Kaiser-Bessel kernel, again including the roll-off correction. The Kaiser-Bessel kernel [57] is equal to

$$C(k_x) = \begin{cases} \frac{1}{L} I_0(B\sqrt{1 - (2k_x/L)^2}) & |k_x| \leq \frac{L}{2} \\ 0 & \text{otherwise} \end{cases}, \quad (3.27)$$

in which  $I_0$  is the zero-order modified Bessel function of the first kind,  $L$  is the width of the kernel and  $B$  is a shape-parameter.

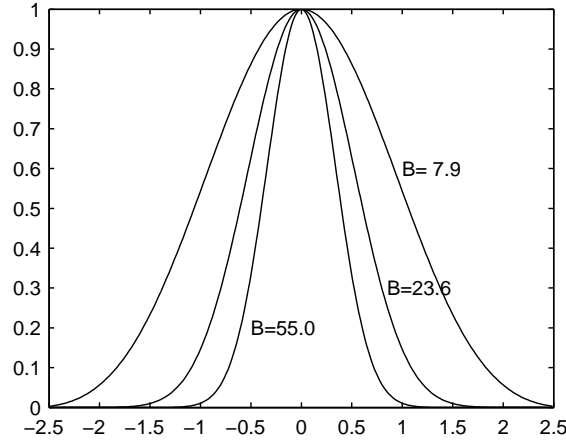


Figure 3.8: Kaiser Bessel kernel for different values of  $B$  for  $L = 5$ . The value at  $k_x = 0$  has been normalized to 1. The width of the kernel is inversely proportional to  $B$ .

The inverse Fourier transform of the Kaiser-Bessel kernel is given by

$$c(x) = \frac{\sin \sqrt{\pi^2 L^2 x^2 - B^2}}{\sqrt{\pi^2 L^2 x^2 - B^2}}. \quad (3.28)$$

$E_2$  of the Kaiser-Bessel kernel is slightly higher than that of the FirLS kernel [58, 59]. However, by increasing the width of the Kaiser-Bessel kernel or FOV just a little we get similar values of  $E_2$  as that of the FirLS kernel. This means that with a small increase in computation time we get similar performance. In MRI practice the Kaiser-Bessel kernel is now widely used [26, 59, 60].

A benefit of using the Kaiser-Bessel kernel is that we have an analytic expression for the kernel and its inverse Fourier transform. This allows easy adaptation when using different FOV and/or ROI, or different levels of desired accuracy.



### 3.4.3.6 Interpolation errors using the Kaiser-Bessel kernel

Using the Kaiser-Bessel kernel as convolution kernel, the interpolation error will depend on the three variables:  $B$ ,  $L$  and  $f_{\text{og}}$ . The latter two determine the computation time of the gridding algorithm. Hence, at fixed computation time,  $B$  is left over as a free parameter for adjusting the accuracy of interpolation. In this section it is assumed that  $\text{ROI} = \text{FOV}/f_{\text{og}} = 1$ . If this is not the case, all variables should be scaled appropriately.

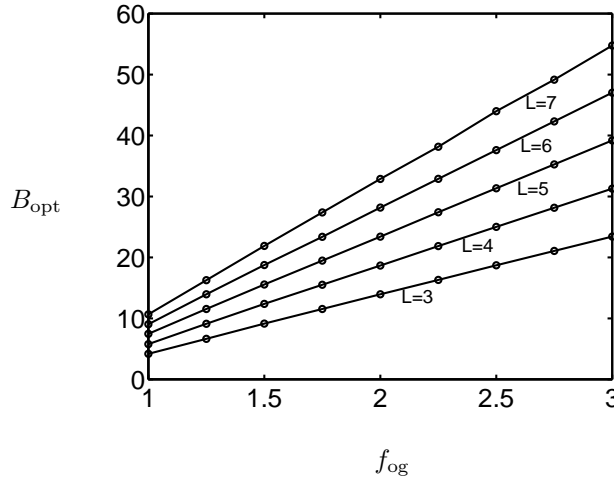


Figure 3.9: Optimal values for  $B$  found by minimizing criterion  $E_2$  of Eq. 3.25 for different values of  $L$  and  $f_{\text{og}}$ . The lines are linear fits.

An optimal value for  $B$  was found by minimizing criterion  $E_2$  of Eq. 3.25. Fig. 3.9 shows the result. The optimal  $B$  appear to lie on straight lines. The lines in Fig. 3.9 are linear fits and the fit is very accurate. Moreover, the linear fits suggested a very simple relation between the optimal value for  $B$  and the other parameters  $f_{\text{og}}$  and  $L$ . In fact, through numerical analysis of the graphs we found [46, 61])

$$B_{\text{opt}} = (f_{\text{og}} - \frac{1}{2})\pi L \quad (3.29)$$

Fig. 3.10 shows the criterion  $E_2$  as a function of  $L$  and  $f_{\text{og}}$  using the optimal value for  $B$ . Note that the  $x$ -axis is equal to  $f_{\text{og}}L$ , which determines the computation time of the convolution. We see that the accuracy increases as both  $L$  and  $f_{\text{og}}$  increase, but that for a fixed computation time, the gain of increasing  $f_{\text{og}}$  diminishes. Keep in mind that increasing  $f_{\text{og}}$  also increases the time necessary for computing the subsequent IFFT.

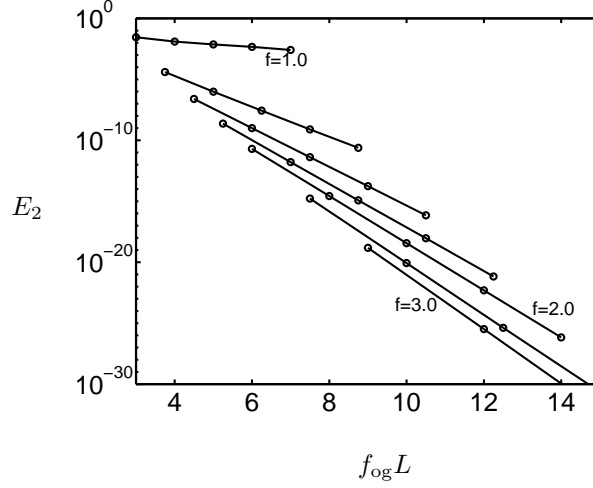


Figure 3.10: Value of the criterion  $E_2$  for different values of  $L$  and  $f_{\text{og}}$  using the optimal value for  $B$ . Note that the  $x$ -axis is  $f_{\text{og}}L$ , which determines the computation time of the convolution. In the figure  $f$  stands for  $f_{\text{og}}$ .

Using criterion  $E_1$  for finding an optimal value for  $B$  gives a slightly smaller value for  $B$ . Such a value for  $B$  would increase accuracy according to criterion  $E_1$ . However, Eq. 3.29 allows a very simple expression for the accuracy.

Fig 3.11 shows the absolute value of criterion  $E_1(x, k_x)$  for  $L = 3$  and  $f_{\text{og}} = 2$ , for all  $x$  inside the ROI. The values of  $k_x$  are limited to  $0 < k_x < 1$ . This is allowed, because of the periodicity of  $E_1(x, k_x)$  in  $k_x$ . The period is equal to  $\Delta k_{\text{Cart}} = 1/f_{\text{og}}$ . The picture is representative for typical values of  $L$  and  $f_{\text{og}}$ . The largest error occurs at the edge of the ROI, *i.e.*  $x = -1/2$ , and it is independent of  $k_x$ . Remember that the error is an aliasing error. The aliasing at  $x = -1/2$  comes from  $x = mf_{\text{og}} - 1/2$ ,  $|m| = 1, 2, \dots, \infty$ , as indicated in Fig. 3.6. Because  $c(x)$  decreases for increasing  $x$ , we only take the largest contribution, coming from  $x = f_{\text{og}} - 1/2$ , since this one is closest to  $x = -1/2$ . Therefore the error is approximately equal to

$$E_1 \simeq E'_1 = \left| \frac{c(f_{\text{og}} - 1/2)}{c(1/2)} e^{2\pi i(f_{\text{og}} - 1/2)k_x} \right| = \frac{\pi L \sqrt{f_{\text{og}}(f_{\text{og}} - 1)}}{\sinh(\pi L \sqrt{f_{\text{og}}(f_{\text{og}} - 1)})}, \quad (3.30)$$

As an example, let us calculate  $E'_1$  for the values used in Fig. 3.11.  $E'_1 = 4.34 \times 10^{-5}$ , which compares well with the true value of  $E_1 = 4.50 \times 10^{-5}$ , see Fig. 3.11.

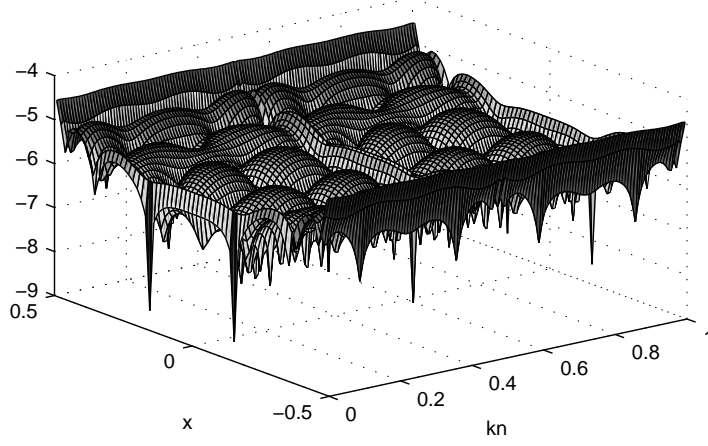


Figure 3.11: Absolute value, on a log scale, of  $E_1(x, k_x)$  for  $L = 3$  and  $f_{\text{og}} = 2$ , for all  $x$  inside the ROI. The values of  $k_x$  are limited to  $0 < k_x < 1$ . This is allowed, because of the periodicity in  $k_x$ . The period is equal to  $\Delta k_{\text{Cart}} = 1/f_{\text{og}}$ . The largest error ( $= 4.50 \times 10^{-5}$ ) occurs at  $x = -1/2$ , and it is nearly independent of  $k_x$ .

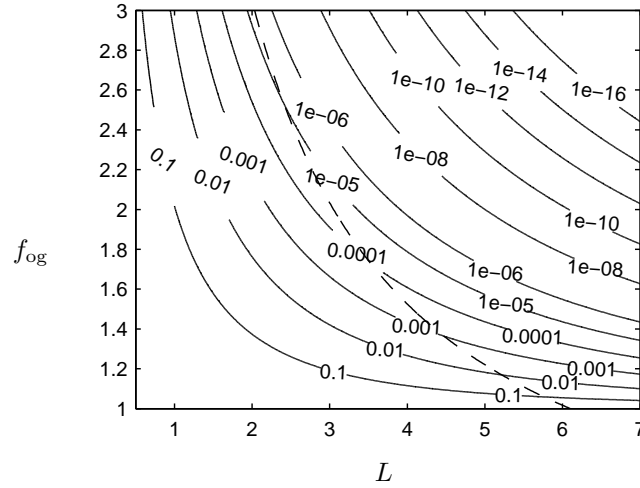


Figure 3.12: Contour lines of  $E'_1$ , Eq. 3.30. The dashed line is a line of constant convolution computation time ( $f_{\text{og}}L = \text{const.}$ ). For a fixed convolution computation time it is more accurate to use larger  $f_{\text{og}}$  and a smaller  $L$ .

Fig. 3.12 shows the contour lines of  $E'_1$  using Eq. 3.30. The dashed line is a line of constant convolution computation time, given by  $f_{\text{og}}L = \text{const.}$  For a fixed convolution computation time, it is more accurate to use larger  $f_{\text{og}}$  and a smaller  $L$ . Or, for a fixed accuracy, the calculation will be faster with a larger  $f_{\text{og}}$  and smaller  $L$ . Note that the increase in accuracy at fixed computation time diminishes for larger  $f_{\text{og}}$ . We have seen the same behaviour with the  $E_2$  criterion.

Increasing  $f_{\text{og}}$  also increases the calculation time for the subsequent IFFT. Therefore, the procedure for finding optimal gridding parameters for  $L$  and  $f_{\text{og}}$  is

1. choose appropriate level of accuracy
2. increase  $f_{\text{og}}$  and decrease  $L$ , to keep the accuracy at the same level
3. continue up to the point where the increase in computation time for the IFFT cancels the decrease in computation time for the convolution.

Finally, I conclude that at a certain choice of  $L$  and  $f_{\text{og}}$  one can make  $E_1$  (ie. the error in approximating the complex exponential by Eq. 3.17) smaller than machine accuracy. This conclusion is supported by both criteria  $E_1$  and  $E_2$ .

### 3.4.3.7 Computational details

I have reduced the computation of interpolation by taking the following measures.

- Making use of separability of Kaiser-Bessel kernel, *i.e.* in the 2D case  $C(k_x, k_y) = C(k_x)C(k_y)$ . This means that one has to evaluate the Kaiser-Bessel kernel only at  $2f_{\text{og}}L$  instead of  $2f_{\text{og}}^2L^2$  positions.
- The calculation of the Kaiser-Bessel kernel is expensive, due to the Bessel function. Instead of evaluating  $C(k_x)$  during the convolution, I precalculate and store it in a table. Values between points in the table are interpolated.
- The array storing the result of the convolution should be large enough to store values produced by the samples at the edge of  $k$ -space. This enables elimination of tests (if-statements) that check whether the convolution produces results that are outside the array that should be wrapped around. The wrapping around is done in a separate routine.

### 3.4.3.8 Interpolation from uniform to non-uniform sample positions

Interpolation from uniform to non-uniform sample positions is often necessary. For instance, one may need to validate an image with respect to the non-uniformly sampled  $k$ -space data from which it was reconstructed. A way to do this is to FFT the image to a Cartesian  $k$ -space grid and to then interpolate to the original sample positions. This application is used in chapter 4. Another application is to generate samples on arbitrary  $k$ -space positions from Cartesian data, for the purpose of simulations.

Sampling theory tells us that this interpolation is in principle possible by convolving the uniform samples with a sinc-function. In practice, this convolution is impractical and

we need to resort to a convolution with a finite-sized kernel. Using the same approach as in gridding, but now in the reverse direction we write:

$$c(\vec{r}) = \int C(\vec{k}) e^{2\pi i \vec{k} \cdot \vec{r}} d\vec{k}. \quad (3.31)$$

Multiplying both sides with  $e^{-2\pi i \vec{k}_n \cdot \vec{r}}$  one gets

$$e^{-2\pi i \vec{k}_n \cdot \vec{r}} c(\vec{r}) = \int C(\vec{k}) e^{2\pi i (\vec{k} - \vec{k}_n) \cdot \vec{r}} d\vec{k} = \int C(\vec{k}_n - \vec{k}) e^{-2\pi i \vec{k} \cdot \vec{r}} d\vec{k}. \quad (3.32)$$

The right-hand side of Eq. 3.32 is discretized using Cartesian positions  $\vec{k}_m$ . Correspondingly,  $\Delta k_m = \Delta k_{\text{Cart}}$ .

$$e^{-2\pi i \vec{k}_n \cdot \vec{r}} c(\vec{r}) \simeq \Delta k_{\text{Cart}} \sum_m C(\vec{k}_n - \vec{k}_m) e^{-2\pi i \vec{k}_m \cdot \vec{r}}. \quad (3.33)$$

Since  $c(\vec{r})$  is greater than zero everywhere in the ROI, division by  $c(\vec{r})$  is allowed, giving

$$e^{-2\pi i \vec{k}_n \cdot \vec{r}} \simeq \frac{\Delta k_{\text{Cart}}}{c(\vec{r})} \sum_m C(\vec{k}_n - \vec{k}_m) e^{-2\pi i \vec{k}_m \cdot \vec{r}} \quad (3.34)$$

Next, I discretize the Fourier integral relating the  $k$ -space signal and its corresponding image, Eq. 2.10, and substitute Eq. 3.34.

$$S(\vec{k}_n) = \int I(\vec{r}) e^{-2\pi i \vec{k}_n \cdot \vec{r}} d\vec{r} \simeq \Delta r_{\text{Cart}} \sum_p I(\vec{r}_p) e^{-2\pi i \vec{k}_n \cdot \vec{r}_p} \simeq \quad (3.35)$$

$$\Delta k_{\text{Cart}} \Delta r_{\text{Cart}} \sum_p I(\vec{r}_p) \left( \frac{1}{c(\vec{r}_p)} \sum_m C(\vec{k}_n - \vec{k}_m) e^{-2\pi i \vec{k}_m \cdot \vec{r}_p} \right). \quad (3.36)$$

Exchanging then the summations, one finds for interpolation of  $I(\vec{k}_m)$  to arbitrary positions  $\vec{k}_n$

$$S(\vec{k}_n) \simeq \Delta k_{\text{Cart}} \Delta r_{\text{Cart}} \sum_m \left( \sum_p \frac{I(\vec{r}_p)}{c(\vec{r}_p)} e^{-2\pi i \vec{k}_m \cdot \vec{r}_p} \right) C(\vec{k}_n - \vec{k}_m). \quad (3.37)$$

Sinc interpolation is a special case of Eq. 3.37. Note that density compensation is not necessary because both  $\vec{k}_m$  and  $\vec{r}_p$  reside on Cartesian grids. By proper choice of the kernel parameters and  $\Delta k_{\text{Cart}}$ , one can make the error (which is equivalent to the error in Eq. 3.17) in Eq. 3.37 as small as necessary for the application at hand.

### 3.5 Maximum likelihood estimation

Maximum likelihood (ML) tries to estimate a desired parameter, or set of parameters, in such a way that the probability of the measured data given these parameters is maximized [62–64]. In MRI, the image is to be interpreted as the set of parameters and the sampled  $k$ -space as the data-set.

The model function I use for the ML estimation is a discretized version of the imaging equation Eq. 2.10:

$$S(\vec{k}_n) = \Delta r_{\text{Cart}} \sum_{p=1}^{N_x N_y} I(\vec{r}_p) e^{-2\pi i \vec{k}_n \cdot \vec{r}_p} + \mathcal{N}_S(\vec{k}_n) \quad (3.38)$$

Eq. 3.38 amounts to modelling the object as  $N_x \times N_y$  drops of water each placed in the centre of a pixel on a Cartesian grid. Drop No.  $p$  gives rise to a  $k$ -space signal = amplitude  $\times e^{-2\pi i \vec{k}_n \cdot \vec{r}_p}$ . The number of water molecules in each drop – proportional to the corresponding pixel intensity  $I(\vec{r}_p)$  – is to be estimated by ML. The pixels residing on known Cartesian positions, a density compensation is obviated, and  $\Delta r_p = \Delta r_{\text{Cart}}$  can be put in front of the summation operator.

Eq. 3.38 can be written compactly as

$$\underline{\mathbf{s}} = T \underline{\mathbf{i}} + \underline{\mathbf{n}}, \quad (3.39)$$

by stacking the respective columns of  $S$ ,  $I$  and  $\mathcal{N}_S$  on top of each other and  $T_{n,p} = \Delta r_{\text{Cart}} e^{-2\pi i \vec{k}_n \cdot \vec{r}_p}$ . The measurement noise,  $\mathcal{N}_S$  (or  $\underline{\mathbf{n}}$ ), is assumed to be independently and Gaussian distributed [65]. This means that the probability density function (PDF) of the noise is given by

$$\mathbf{P}(\underline{\mathbf{n}}) = \frac{1}{(2\pi\sigma_S^2)^{N_{\text{samp}}}} e^{-(\underline{\mathbf{n}}^\dagger \underline{\mathbf{n}})/2\sigma_S^2} = \frac{1}{(2\pi\sigma_S^2)^{N_{\text{samp}}}} e^{-((\underline{\mathbf{s}} - T \underline{\mathbf{i}})^\dagger (\underline{\mathbf{s}} - T \underline{\mathbf{i}}))/2\sigma_S^2} \quad (3.40)$$

with  $N_{\text{samp}}$  the number of samples in  $k$ -space and  $\sigma_S$  the standard deviation of the measurement noise. Maximizing Eq. 3.40 results in the following well-known solution [66, 67].

$$\underline{\mathbf{i}}_{\text{ML}} = (T^\dagger T)^{-1} T^\dagger \underline{\mathbf{s}}. \quad (3.41)$$

The following comments on Eq. 3.41 are in order.

- Eq. 3.41 is equivalent to a *linear least squares fit* of the image pixel intensities to the signal sampled at arbitrary positions  $k_n$ . This simplification of the ML approach is a consequence of the assumed Gaussian nature of the measurement noise.
- Numerical computation of Eq. 3.41, although obviating the density compensation, is prohibitively expensive: If the sizes of the image and  $k$ -space data matrices are both  $N_x \times N_y$ , the size of  $T$  is  $N_x N_y \times N_x N_y$  [53]. Ill-conditioning [66, 67] of  $T^\dagger T$  may give numerical instabilities. Rosenfeld [68] used SVD to tackle these instabilities.

- If  $k$ -space sampling is Cartesian, and the sizes of the image and  $k$ -space data matrices are both  $N_x \times N_y$ , multiplication by  $T$  is equivalent to an  $N_x \times N_y$  FFT, and  $T^\dagger T$  is equal to the  $N_x N_y \times N_x N_y$  unit matrix. In that case, Eq. 3.41 is equivalent to reconstruction by IFFT.
- The maximum likelihood solution is closely related to Eq. 3.2. To see this, rewrite the ML solution as  $\mathbf{i}_{\text{ML}} = T^\dagger [T(T^\dagger T)^{-2} T^\dagger] \mathbf{s}$ . Eq. 3.2 can be written as  $\mathbf{i}_{\text{recon}} = T(\Delta k) \mathbf{s}$ . Here,  $\Delta k$  is a diagonal matrix with the density compensation factors on the diagonal. Therefore, Eq. 3.2 and Eq. 3.41 give the same result, whenever  $T(T^\dagger T)^{-2} T^\dagger$  is a diagonal matrix. For non-uniform sampling this would be a coincidence [53]. The above equation does show an alternative to finding a density compensation, *i.e.* find a diagonal matrix that ‘best’ fits  $T(T^\dagger T)^{-2} T^\dagger$ , see [69] for details. Again, the inversion of  $T^\dagger T$  may lead to instabilities.
- The maximum likelihood solution is equal to deconvolution of the PSF, for  $\Delta k_n = 1$ , in Eq. 3.5. Eq. 3.41 can be rewritten as  $(T^\dagger T) \mathbf{i}_{\text{ML}} = T^\dagger \mathbf{s}$ . Now,  $T^\dagger T$  can be shown to be equivalent to the PSF in case of using a density compensation of  $\Delta k_n = 1$ , see Eq. 4.32.  $(T^\dagger T) \mathbf{i}_{\text{ML}}$  is equal to a convolution of the PSF and  $\mathbf{i}_{\text{ML}}$ , see 4.4.5.2, and  $T^\dagger \mathbf{s}$  is equal to  $\mathbf{i}_{\text{recon}}$ , for  $\Delta k_n = 1$ , Eq. 3.2. Therefore, Eq. 3.41 is equivalent to a deconvolution of the PSF in Eq. 3.5.
- The maximum likelihood approach plays an important role in Bayesian reconstruction, which is the subject of the next chapter.

### 3.6 Influence of the sample distribution on noise properties in the image

If sampling of noise-corrupted  $k$ -space signal is Cartesian, the correlation between the noise in individual pixels of the corresponding image domain is zero. This favours correct interpretation of the image. The present section investigates the possible detrimental influence of non-uniform sampling on the image noise properties.

$$S(\vec{k}_n) = S'(\vec{k}_n) + \mathcal{N}_S(\vec{k}_n), \quad (3.42)$$

where  $S'(\vec{k}_n)$  is the  $k$ -space data devoid of noise, and  $\mathcal{N}_S(\vec{k}_n)$  is the measurement noise.

$$I(\vec{r}) = \sum_n S(\vec{k}_n) e^{2\pi i \vec{k}_n \cdot \vec{r}} \Delta k_n = I'(\vec{r}) + \sum_n \mathcal{N}_S(\vec{k}_n) e^{2\pi i \vec{k}_n \cdot \vec{r}} \Delta k_n = I'(\vec{r}) + \mathcal{N}_I(\vec{r}), \quad (3.43)$$

with  $I'(\vec{r})$  and  $\mathcal{N}_I(\vec{r})$  the images reconstructed from the noiseless data and from the measurement noise, respectively.

Next, we calculate the expectation of  $I(\vec{r})$ , and use  $\text{E}[\mathcal{N}_S(\vec{k}_n)] = 0$ .

$$\text{E}[I(\vec{r})] = I'(\vec{r}) + \sum_n \text{E}[\mathcal{N}_S(\vec{k}_n)] e^{2\pi i \vec{k}_n \cdot \vec{r}} \Delta k_n = I'(\vec{r}). \quad (3.44)$$

Eq. 3.44 enables one to derive the covariance of the noise contained in image pixels.

$$\begin{aligned}\text{Cov}[\text{Re } \mathcal{N}_I(\vec{r}_p), \text{Re } \mathcal{N}_I(\vec{r}_q)] &= \text{E}[(\text{Re } \mathcal{N}_I(\vec{r}_p) - \text{E}[\text{Re } \mathcal{N}_I(\vec{r}_p)])(\text{Re } \mathcal{N}_I(\vec{r}_q) - \text{E}[\text{Re } \mathcal{N}_I(\vec{r}_q)])] \\ &= \sigma_S^2 \sum_n \cos(2\pi \vec{k}_n \cdot (\vec{r}_p - \vec{r}_q)) \Delta k_n^2,\end{aligned}\quad (3.45)$$

where we have used the fact that the  $k$ -space noise is uncorrelated and has equal variance  $\sigma_S^2$  at each sample position for both the real and imaginary part. Similarly we find

$$\text{Cov}[\text{Im } \mathcal{N}_I(\vec{r}_p), \text{Im } \mathcal{N}_I(\vec{r}_q)] = \sigma_S^2 \sum_n \cos(2\pi \vec{k}_n \cdot (\vec{r}_p - \vec{r}_q)) \Delta k_n^2 \quad (3.46)$$

and

$$\text{Cov}[\text{Re } \mathcal{N}_I(\vec{r}_p), \text{Im } \mathcal{N}_I(\vec{r}_q)] = \sigma_S^2 \sum_n \sin(2\pi \vec{k}_n \cdot (\vec{r}_p - \vec{r}_q)) \Delta k_n^2 \quad (3.47)$$

Eqs. 3.45 - 3.46 are the desired noise covariances. However, in order to better appreciate the expression, I return to a continuous integral, replacing one power of  $\Delta k_n$  by  $\Delta k(\vec{k})$  and the other by  $d\vec{k}$ , and the summation by an integral. Eq. 3.45 then becomes

$$\text{Cov}[\text{Re } \mathcal{N}_I(\vec{r}_p), \text{Re } \mathcal{N}_I(\vec{r}_q)] \simeq \sigma_S^2 \text{Re} \int \Delta k(\vec{k}) e^{2\pi i \vec{k} \cdot (\vec{r}_p - \vec{r}_q)} d\vec{k} \quad (3.48)$$

Eq. 3.48 says that the noise covariance in the image domain is proportional to the Fourier transform of the sampling density compensation  $\Delta k(\vec{k})$ . This result implies that the flatter a sample distribution, the more the noise covariance is proportional to a delta function  $\delta(\vec{r}_p - \vec{r}_q)$ . The sampling density compensation being flat in the Cartesian case, the latter result concurs with the first sentence of this section. See also Ref. [70].

### 3.7 Conclusions

The work described in this chapter allows the following conclusions.

- For accuracy of reconstruction, Cartesian sampling is best.
- Physical conditions may demand non-uniform sampling, like radial or spiral. For linogram sampling, a dedicated fast and accurate reconstruction algorithm is available.
- Gridding is a convenient and accurate image reconstruction method for any non-uniformly, but sufficiently densely, sampled signal.
- The sampling density compensation is the main source of image reconstruction errors. The more a sample distribution is irregular, the larger the errors.
- Irregular sample positions, locally violating the Nyquist sampling criterion, arise through intrascan motion correction in  $k$ -space. This is detrimental to functional imaging.
- The noise covariance in the image domain is proportional to the Fourier transform of the sampling density compensation.



## Chapter 4

# Image reconstruction from incomplete data

### 4.1 Introduction

Incomplete or sparse sampling of  $k$ -space arises in several cases. For instance, this will happen if one deliberately omits a number of trajectories to reduce the measurement time. Secondly, there may be hardware imperfections or motion of the patient, which may cause certain parts of  $k$ -space to be undersampled.

In this chapter, I will describe several methods that are used for reconstruction of images from sparsely sampled  $k$ -space. A sparse scan means that locally the Nyquist criterion may be violated – *i.e.* local undersampling. Images reconstructed from such sparse datasets by an inverse Fourier transform will suffer from undersampling-related artefacts. The algorithms employed should be able to reconstruct an artefact-free image from such sparse data. Two types of algorithms can be distinguished.

The first type tries to estimate the omitted data. Together with the measured data this should result in an artefact-free image. However, it is crucial that one should be able to pinpoint which at which  $k$ -space positions data have been omitted. In the case of scan-time reduction this task is usually possible since the reduction requires omission of trajectories from a full scan. On the other hand, if the sample positions are irregular – as is the case *e.g.* after motion correction – it is difficult to define used and omitted trajectories.

The second type tries to estimate the image directly, given the *available* data. The FT of the estimated image should agree with the available data to within the ‘noise band’. This approach is more flexible.

## 4.2 Scan-time reduction

### 4.2.1 Cartesian scanning

In Cartesian scanning, the trajectories are not equivalent to each other. One can distinguish them from each other by their shortest distance to the origin of  $k$ -space. Those trajectories that pass the origin at short distance – also called low-frequency encodings – are usually all scanned because they contain the highest signal energy. Omission of these encodings is therefore highly detrimental. In fact, scan-time reduction is accomplished rather by omitting high-frequency encodings; this results in ringing artefacts. Furthermore, the omission can be done either by truncating beyond a certain  $k_y$ -value or by irregular omission of  $k_y$ -values. Fig. 4.1 shows the two approaches for 30% scan-time reduction.

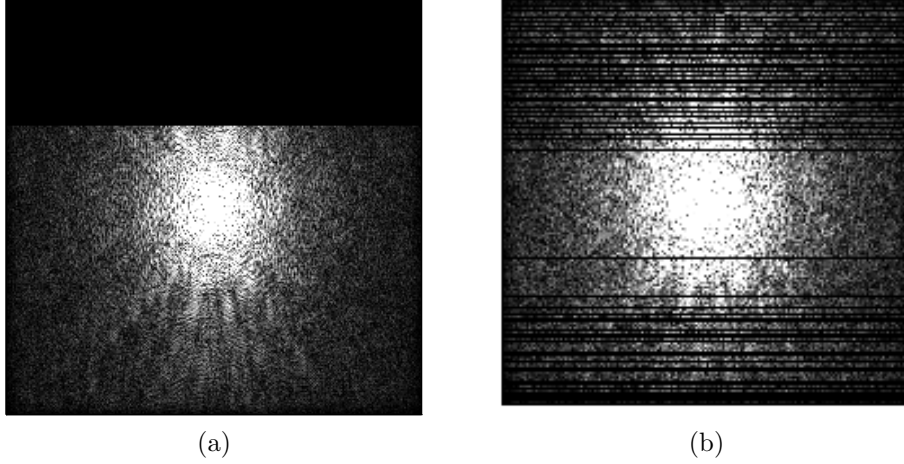


Figure 4.1: Examples of 30% Cartesian scan-time reduction. a) Scan-time reduction by one-sided  $k$ -space truncation. b) Scan-time reduction by omitting encodings irregularly.

Truncating  $k$ -space as in Fig. 4.1a brings about a loss of resolution. Furthermore, if the signal has not decayed completely at the point of truncation, the reconstructed image will contain ringing artefacts. These can be diminished by windowing, but at the cost of even more loss of resolution. To regain full resolution the measured data would have to be *extrapolated*.

Omitting high-frequency encodings irregularly, as in Fig. 4.1b, will lead to ringing artefacts in reconstructed images too. The difference with truncation is that to regain full resolution, the measured data can now be *interpolated* between measured data which do contain high-frequency information.

Measurement-time reduction of Cartesian scans has been investigated extensively by Marseille [71].

### 4.2.2 Radial and Spiral scanning

As explained in chapter 2, radial and spiral trajectories share the property that each passes through the origin of  $k$ -space. Thus, unlike Cartesian trajectories, each radial/spiral trajectory is equally important; the notion of high- and low-frequency encodings does not apply. In addition, omission of various adjacent trajectories causes significant loss of signal energy. See Fig. 4.2. The left-hand part of Fig. 4.2, a, shows two aspects 1) As mentioned in chapter 2, SRSFT (radial) indeed possesses many Cartesian samples positions near the origin of  $k$ -space. 2) Wide gaps occur beyond a certain value of  $|k|$ . As for spiral omissions, shown in Fig. 4.2b, one can see significant loss of signal intensity near the origin of  $k$ -space. In the following, I shall discuss the two types

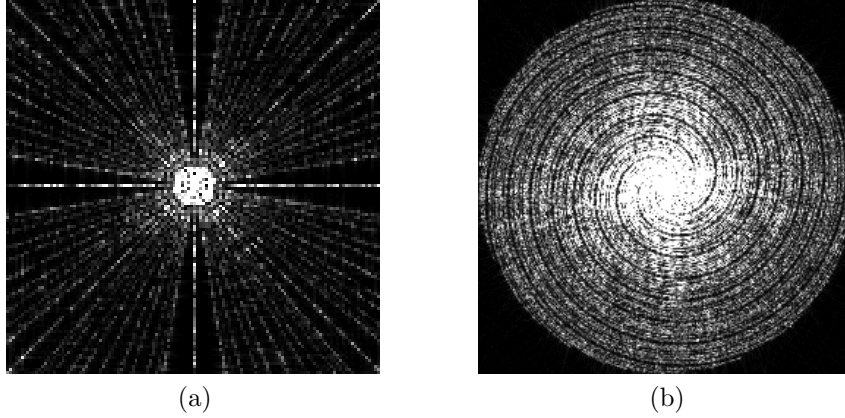


Figure 4.2: Examples of radial and spiral scan-time reduction [72]. a) Radial, SRSFT; see text. b) Spiral.

of algorithms to sparse  $k$ -space data. These are edge-preserving filtering (type 1) and Bayesian image estimation (type 2).

### 4.3 Type 1: Edge-preserving filtering

Edge-preserving filtering is a method that has been used to eliminate artefacts caused by the omission of encodings in Cartesian scans [71, 73]. Before describing how this method can also be applied to non-Cartesian scans I will first review the method.

As described above, the image reconstructed from a sparse Cartesian scan will contain ringing-like artefacts. Smoothing the image will reduce these artefacts. However, the smoothing must be carried out by a non-linear filter, because a linear filter can not interpolate or extrapolate data in  $k$ -space. The reason being that linear filtering is equal to a multiplication in  $k$ -space, and the omission of data, resulting in zeroes at certain locations in  $k$ -space will remain zero after the multiplication.

The non-linear filtering will accomplish the necessary interpolation or extrapolation in  $k$ -space. However, it will also affect the measured data, which is not desired, since this is a Type 1 algorithm, so it should keep the measured data intact. Therefore the data at measured locations are should be constrained or reset after FT to their measured values.

The above process is repeated until convergence. This method was successfully applied to sparse Cartesian scans [71, 73]. In this section I will extend this technique to non-Cartesian scans.

If possible, the data on omitted trajectories are initially estimated by averaging the data on neighbouring trajectories. This suffices, for radial and spiral measurements, for the centre of  $k$ -space, which is oversampled. However, the edges of  $k$ -space are estimated poorly. Subsequently, the initial estimates are improved in an iterative manner by shuttling back and forth between the measurement domain and the image domain [71]. In each iteration, the ringing artefacts in the image are smoothed by a non-linear Gaussian filter, while fine details are subsequently restored in the measurement domain by resetting the data on the measured trajectories to their measured values. A short description of the non-linear Gaussian filter will be presented later in this section.

The flow diagram in Fig. 4.3 describes the core of the estimation procedure. Therefore only the most important parts of the algorithm are displayed. A detailed flow diagram is presented at the end of this section.

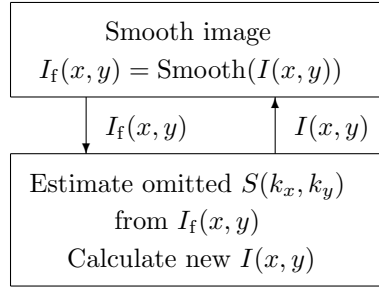


Figure 4.3: Basic flow diagram of the estimation procedure. It shows the shuttling between  $k$ -space and the image domain.

However, a complication arises whenever  $k$ -space sampling is non-Cartesian. As a result, the transformation between the two domains in each iteration can not be accomplished by the 2D-(I)FFT. Moreover, Type 1 algorithms require density compensation with its accompanying problems.

The transformation from  $k$ -space to image domain is accomplished by the gridding algorithm, described in chapter 3. In principle, the reverse transformation could also be accomplished by gridding, but this is not optimal since I wish to constrain the measured data to their original values.

The solution is to estimate the omitted trajectories in such a way, that the resulting image (after gridding reconstruction) matches the smoothed image best in least squares

sense, while constraining the measured data to their measured values. This is described by the following equation

$$\underset{\underline{\mathbf{s}}(k_x, k_y)}{\text{MIN}}' \sum_{x, y} |G\underline{\mathbf{s}}(k_x, k_y) - \underline{\mathbf{i}}_f(x, y)|^2. \quad (4.1)$$

$\underline{\mathbf{s}}(k_x, k_y)$  is the data written as a vector.  $I_f(x, y)$  is the smoothed image also written as a vector, see section 3.5.  $G$  is a matrix representing the gridding reconstruction. Gridding can be written in this way because it is a linear operator. The prime on the MIN signifies a constrained minimization.

Next, I decompose  $\underline{\mathbf{s}}(k_x, k_y)$  into a measured part,  $\underline{\mathbf{s}}_m(k_x, k_y)$ , and an omitted part,  $\underline{\mathbf{s}}_o(k_x, k_y)$

$$\underline{\mathbf{s}} = (\underline{\mathbf{s}}_m^T \mid \underline{\mathbf{s}}_o^T)^T. \quad (4.2)$$

The gridding matrix  $G$  can be written as  $G = R^{-1}F^{-1}C\Delta K$ , see 3.4.3.1, with  $\Delta K$ ,  $C$ ,  $F^{-1}$  and  $R^{-1}$  representing density compensation, convolution with the kernel, IFFT and roll-off correction, respectively. All these operations are linear so they can be written in matrix form.

The decomposition in measured and omitted part is also applied on  $G$ , giving

$$\begin{aligned} G &= (G_m \mid G_o) \\ &= R^{-1}F^{-1} (C_m \mid C_o) \left( \begin{array}{c|c} \Delta K_m & 0 \\ \hline 0 & \Delta K_o \end{array} \right) \\ &= R^{-1}F^{-1} (C_m \Delta K_m \mid C_o \Delta K_o). \end{aligned} \quad (4.3)$$

Combining Eqs. 4.2 and 4.3, one can derive:

$$\begin{aligned} \underline{\mathbf{i}} &= G\underline{\mathbf{s}} \\ &= \underline{\mathbf{i}}_m + G_o\underline{\mathbf{s}}_o. \end{aligned} \quad (4.4)$$

$\underline{\mathbf{i}}_m(x, y)$  is the reconstructed image using only the measured samples, and is fixed given the measured data  $\underline{\mathbf{s}}_m(k_x, k_y)$ .

Eq. 4.4 allows one to eliminate the constraint on the minimization in Eq. 4.1 to get

$$\underset{\underline{\mathbf{s}}_o(k_x, k_y)}{\text{MIN}} \sum_{x, y} |G_o\underline{\mathbf{s}}_o - (\underline{\mathbf{i}}_f - \underline{\mathbf{i}}_m)|^2. \quad (4.5)$$

Solving Eq. 4.5 for  $\underline{\mathbf{s}}_o(k_x, k_y)$  gives

$$\underline{\mathbf{s}}_o = (G_o^\dagger G_o)^{-1} G_o^\dagger (\underline{\mathbf{i}}_f - \underline{\mathbf{i}}_m). \quad (4.6)$$

$G_o^\dagger G_o$  does not have a nice structure to allow fast inversion. Moreover, the size of  $G_{nm}$  is typically on the order of  $60000 \times 20000$ , so direct evaluation of Eq. 4.6 is impossible. I therefore use an optimization method, such as the conjugate gradient method, to find  $\underline{\mathbf{s}}_o(k_x, k_y)$  in an iterative manner.

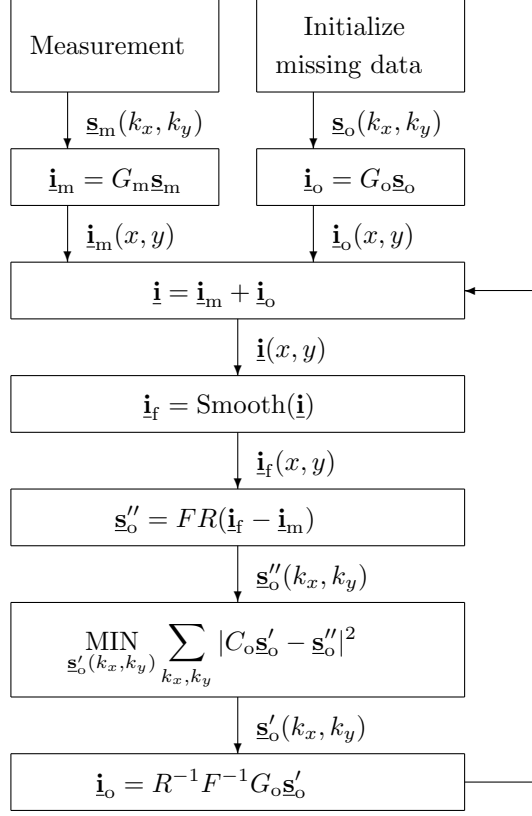


Figure 4.4: Detailed flow diagram of Type 1 algorithm using edge-preserving filtering.

Note that if one measures on a uniform Cartesian grid,  $G_o = F_o^{-1}$ . And because  $F_o^{-1} F_o = I$ , there is no need for any matrix inversion in Eq. 4.6. This means that Eq. 4.6 can be used directly, and actually reduces to a partial 2D-FFT.

One can speed up the minimization of Eq. 4.5 by weighting the summation with  $R^\dagger R$ . This enables one to eliminate the 2D-IFFT during the minimization

$$\begin{aligned}
 \text{MIN}_{\mathbf{s}_o'(k_x, k_y)} \sum_{x, y} |R(G_o \mathbf{s}_o - (\mathbf{i}_f - \mathbf{i}_m))|^2 = \\
 \text{MIN}_{\mathbf{s}_o'(k_x, k_y)} \sum_{k_x, k_y} |C_o \Delta K_o \mathbf{s}_o - \mathbf{s}_o''|^2,
 \end{aligned} \tag{4.7}$$

with  $\mathbf{s}_o''(k_x, k_y) = FR(\mathbf{i}_f(x, y) - \mathbf{i}_m(x, y))$ . This simplification was possible due to the fact that the Fourier matrix  $F$  is a unitary matrix. The minimization of Eq. 4.7 is, however,

not equivalent to the minimization in Eq. 4.5. Rather, the minimization in Eq. 4.7 is equivalent to

$$\underset{\underline{\mathbf{s}}'_o(k_x, k_y)}{\text{MIN}} \sum_{k_x, k_y} |C_o \underline{\mathbf{s}}'_o - \underline{\mathbf{s}}''|^2, \quad (4.8)$$

with  $\underline{\mathbf{s}}'_o(k_x, k_y) = \Delta K_o \underline{\mathbf{s}}_o(k_x, k_y)$ .

Combining everything results in the algorithm described in the flow diagram in Fig. 4.4.

### 4.3.1 Non-linear Gaussian filter

The nonlinear Gaussian filter I use to smooth the image during each iteration of our estimation algorithm, is defined by the following equation [65]:

$$I_f(k, l) = \frac{\sum_{m, n=-D/2}^{D/2} I(k+m, l+n) w(k, l, m, n)}{\sum_{m, n=-D/2}^{D/2} w(k, l, m, n)}. \quad (4.9)$$

$w(k, l, m, n)$  is given by

$$w(k, l, m, n) = e^{-\frac{1}{4}(I(k+m, l+n) - I(k, l))^2 / \kappa^2}, \quad (4.10)$$

where  $I_f(k, l)$  is the filtered pixel value at position  $(k, l)$ ,  $I(k, l)$  is the pixel value in the original image at position  $(k, l)$ .  $D$  is the ‘size’ of the filter.

The non-linear Gaussian filter is an edge-preserving filter, because only those pixels with  $(I(k+m, l+n) - I(k, l))^2 < \kappa^2$  will contribute significantly to the summation in Eq. 4.9.

## 4.4 Type 2: Bayesian image estimation

### 4.4.1 Introduction

With Type 1 image reconstruction from sparse data, one estimates the omitted  $k$ -space samples. However, with the alternative reconstruction of Type 2, one adjusts the image iteratively until it is compatible with the *available*  $k$ -space data; the latter may be sampled at *arbitrary* positions in  $k$ -space. Two important aspects in this approach are: 1) The notion of omitted samples ceases to exist, especially in the case of motion-corrected sample positions. 2) I chose to evaluate the compatibility between image and data in the  $k$ -space rather than in the image domain. In section 3.4.3.8, I pointed out that this choice avoids the density compensation because it involves transformation from a Cartesian to an arbitrary grid instead of the reverse.

As a result of undersampling, information is irretrievably lost. To alleviate this loss, one may invoke prior knowledge. Following Fuderer, Marseille, Lettington and Hong, and McNally [71, 73–76], I make use of the fact that the histogram of differences of intensity between nearest-neighbour image pixels has Lorentzian shape. Following Marseille [71], I use Bayesian estimation to impose this property. Alternatively, one could treat the Lorentzian property – called Cauchy in mathematics – as regularization imposed on maximum likelihood estimation [77, 78].

#### 4.4.2 The most probable image

My reconstruction algorithm aims at finding the image  $I_{\text{MAP}}$ , that maximizes the conditional probability density function [79, 80]  $P(I|S)$ , of the image  $I$ , given the measured  $k$ -space data  $S$ , and ignoring omitted samples. In the context of Bayesian reconstruction, this conditional probability density function is called the posterior [81].  $P(I|S)$  is a shorthand notation for  $P(I_{1,1}, I_{1,2}, \dots, I_{N_y, N_x} | S_1, \dots, S_{N_{\text{samp}}})$ . I will use this notation in all subsequent equations.

$I_{1,1}$	$I_{2,1}$	$\cdots$	$I_{n_x,1}$	$\cdots$	$I_{N_x,1}$
$I_{1,2}$	$I_{2,2}$	$\cdots$	$I_{n_x,2}$	$\cdots$	$I_{N_x,2}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
$I_{1,n_y}$	$I_{2,n_y}$	$\cdots$	$I_{n_x,n_y}$	$\cdots$	$I_{N_x,n_y}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
$I_{1,N_y}$	$I_{2,N_y}$	$\cdots$	$I_{n_x,N_y}$	$\cdots$	$I_{N_x,N_y}$

Figure 4.5: The image,  $I$ , written as a matrix. The matrix elements represent the pixels in the image.

Using Bayes-theorem [62, 81–85]

$$P(I|S) = \frac{P(S|I)P(I)}{P(S)}, \quad (4.11)$$

and minimizing  $-\log P(I|S)$  instead of maximizing  $P(I|S)$ , one arrives at the following problem

$$\min_I \left\{ -\log P(S|I) - \log P(I) \right\}. \quad (4.12)$$

In deriving Eq. 4.12, the evidence  $P(S)$  was treated as a constant, because it does not depend on the image  $I$ .  $P(S|I)$  is the likelihood. It is exactly the same term as in maximum likelihood estimation, see section 3.5. The likelihood will be further elaborated on in section 4.4.4. The prior,  $P(I)$ , which describes any *a priori* knowledge one has about the image, will be explained in the next section.



### 4.4.3 The prior, $P(I)$

$P(I)$  is the prior, which allows incorporation of *a priori* knowledge about the image into the reconstruction algorithm. The prior depends on the number of dimensions of the image. In case of 2D Cartesian scanning one dimension, for example  $k_y$ , is fully sampled. This means that we can already apply IFFT in the  $k_y$ -direction. The 2D reconstruction problem then reduces to a 1D problem. In this case all rows in the image can be reconstructed separately using a 1D-prior, see [71] for details. So in the Cartesian case no 2D-prior is necessary. In the non-Cartesian no dimension is fully sampled. Therefore, the reconstruction problem is 2D and we want a 2D-prior.

I will now start discussing the one dimensional case and then extend to two dimensions.

#### 4.4.3.1 The 1D-Lorentz prior

Although in one dimension the image will be a vector, see Fig. 4.6, I will still call it an image. As explained above, in 2D-Cartesian scanning it will be a single row in the 2D-image.

$I_1$	$I_2$	$\cdots$	$I_{n_x}$	$\cdots$	$I_{N_x}$
-------	-------	----------	-----------	----------	-----------

Figure 4.6: The image,  $I$ , in the one-dimensional case. In this case the image will be a vector. The vector-elements represent the pixels in the image. In 2D-Cartesian scanning the picture represents a row in the 2D-image.

I start off by expanding the prior as follows [71, 86, 87]

$$P(I) = P(I_1, I_2, \dots, I_{n_x}, \dots, I_{N_x}) = P(I_1)P(I_2|I_1) \cdots P(I_{n_x}|I_1, I_2, \dots, I_{n_x-1}) \cdots P(I_{N_x}|I_1, \dots, I_{N_x-1}). \quad (4.13)$$

Every term describes the probability of finding a pixel value given the values of all the pixels to the left of it. Note that there are no terms with dependency on a right neighbour.

I now assume that  $I_{n_x}$  depends (directly) only on its left neighbour,  $I_{n_x-1}$ . This corresponds to the image being a Markov sequence [82, 83, 86], with the following Markov property

$$P(I_{n_x}|I_{n_x-1}, I_{n_x-2}, \dots) = P(I_{n_x}|I_{n_x-1}). \quad (4.14)$$

Substituting Eq. 4.14 into Eq. 4.13, and ignoring the boundaries, gives

$$P(I) = \prod_{n_x=1}^{N_x} P(I_{n_x}|I_{n_x-1}). \quad (4.15)$$

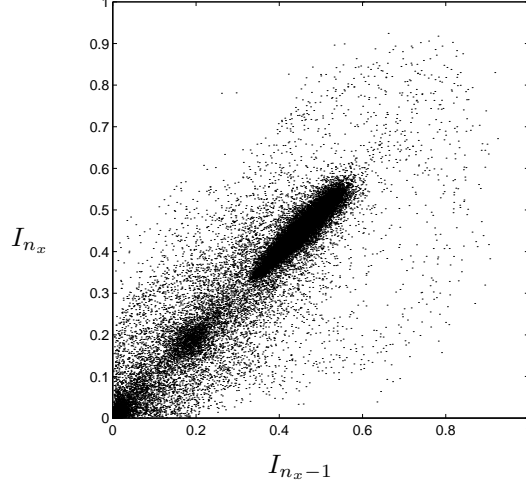


Figure 4.7: Scatterplot of the value of a pixel against the value of its left neighbour pixel. There is a large concentration of points along the diagonal. Further away from the diagonal are steadily fewer points. This is valid for all points along the diagonal.

We know that  $P(I_{n_x}|I_{n_x-1}) = \frac{P(I_{n_x}, I_{n_x-1})}{P(I_{n_x-1})}$ . To simplify, I will assume  $P(I_{n_x})$  uniformly distributed. This is equal to having no prior knowledge on the pixel values themselves. This results in  $P(I_{n_x}|I_{n_x-1}) \propto P(I_{n_x}, I_{n_x-1})$ . Fig. 4.7 shows a typical scatterplot of  $I_{n_x}$  against its left neighbour  $I_{n_x-1}$ . Most points are near the diagonal. Further away from the diagonal are steadily fewer points. Since most points are along the diagonal and the behaviour perpendicular to the diagonal is more or less the same for all points along the diagonal, I write

$$P(I_{n_x}, I_{n_x-1}) = P\left(\frac{I_{n_x} - I_{n_x-1}}{\sqrt{2}}, \frac{I_{n_x} + I_{n_x-1}}{\sqrt{2}}\right) = P\left(\frac{I_{n_x} - I_{n_x-1}}{\sqrt{2}}\right)P\left(\frac{I_{n_x} + I_{n_x-1}}{\sqrt{2}}\right), \quad (4.16)$$

by rotating the ‘coordinates’, where the last step follows from the assumed independency of the behaviour perpendicular to the diagonal and along the diagonal. I again use the fact that I have no prior knowledge of the pixels themselves, amounting to  $P(\frac{I_{n_x} + I_{n_x-1}}{\sqrt{2}})$  being constant. Finally, this results in  $P(I_{n_x}|I_{n_x-1}) \propto P(I_{n_x} - I_{n_x-1}) = P(\xi_{n_x})$ , where  $\xi_{n_x} \stackrel{\text{def}}{=} I_{n_x} - I_{n_x-1}$ . I have already left out the factor  $\sqrt{2}$ , because this will only scale the  $a$ -parameter in the prior, see Eq. 4.18, it does not change the prior itself. Substitution in Eq. 4.15 gives

$$P(I) \propto \prod_{n_x=1}^{N_x} P(\xi_{n_x}). \quad (4.17)$$

In [74], Fuderer found that the probability density function of  $\xi_{n_x}$  is Lorentzian-shaped (see also [71, 75, 76, 79, 88]). This leads to the following expression for the 1D-prior

$$P(I) \propto \prod_{n_x=1}^{N_x} \frac{a}{\pi(a^2 + \xi_{n_x}^2)}. \quad (4.18)$$

#### 4.4.3.2 Alternative derivation of the 1D-Lorentz prior

In this section I will describe an alternative derivation of the 1D-Lorentz prior. We start off directly with the differences of the pixel-values,  $\xi_{n_x}$ , see Fig. 4.8. There is no loss

$\xi_1$	$\xi_2$	$\cdots$	$\xi_{n_x}$	$\cdots$	$\xi_{N_x}$
---------	---------	----------	-------------	----------	-------------

Figure 4.8: The difference image in the one-dimensional case. The elements are constructed from the image by  $\xi_{n_x} = I_{n_x} - I_{n_x-1}$ .

of information (except for the DC-component), because we can reconstruct the pixel-values from the differences by applying  $I_{n_x+1} = I_{n_x} + \xi_{n_x}$ . Therefore the differences and the DC-component together contain the same information as the image itself. It should have the same prior since both descriptions represent the same thing. First we shall assume that the DC-component is uniformly distributed, effectively removing it from the prior. Second, assume the pixel differences  $\xi_{n_x}$  are independent. These two assumptions directly lead to the 1D-prior

$$P(I) \propto P(\xi_1, \xi_2, \dots, \xi_{N_x}) = \prod_{n_x=1}^{N_x} P(\xi_{n_x}), \quad (4.19)$$

which is equivalent to Eq. 4.17.

#### 4.4.3.3 Properties of the 1D-Lorentz prior

First, let us consider the following configuration of image pixels.

$\cdots$	$I_{n_x}$	$I_{n_x+1}$	$I_{n_x+2}$	$\cdots$
----------	-----------	-------------	-------------	----------

We know the probability of finding a certain pixel value given the pixel value of its left neighbour. Now let us calculate the probability of finding a certain pixel value given *only* the second left neighbour, *i.e.*

$$P(I_{n_x+2}|I_{n_x}) = \int P(I_{n_x+2}|I_{n_x+1})P(I_{n_x+1}|I_{n_x})dI_{n_x+1} = \frac{2a}{\pi \left[ (2a)^2 + (I_{n_x+2} - I_{n_x})^2 \right]} \quad (4.20)$$

In this derivation it is assumed that the integration is from  $I_{n_x+1} = -\infty$  to  $I_{n_x+1} = \infty$ . In practice  $I_{n_x+1}$  will have only a finite range of values. However, when  $a$  is small enough

the integration is still effectively from  $-\infty$  to  $\infty$ . Eq. 4.20 is a nice property: The prior of an image with twice as low a resolution has the same form as the original image. The width of the Lorentzian,  $a$ , is larger by a factor of two, showing the information loss of omitting pixel  $I_{n_x+1}$ . This example can also be done for the third left neighbour etc., in all cases the form of the prior will stay the same, only the  $a$ -parameter increases linearly with the distance between the pixels.

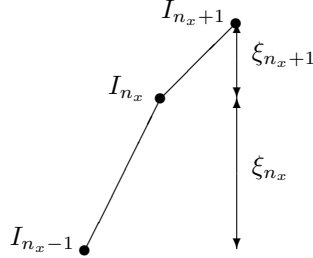


Figure 4.9: Possible intensity-steps for going from one intensity-plateau of an image to an adjacent one in two steps.

The second property concerns the probability of intensity-steps in the image from one intensity-plateau to another. Fig. 4.9 shows possible intensity-steps when going from intensity-plateau to the next in two steps. The joint probability of the two steps is

$$P(\xi_1, \xi_2) = \frac{a^2}{\pi^2(a^2 + \xi_1^2)(a^2 + \xi_2^2)}, \quad (4.21)$$

which is plotted in Fig. 4.10. The figure implies that the 1D-Lorentz prior prefers

- two equal steps, *i.e.*  $\xi_1 = \xi_2$ , if the total step  $(\xi_1 + \xi_2)$  is smaller than  $2a$
- one big step and a small step, if the total step is larger than  $2a$

This preference means that the 1D-Lorentz prior has an edge-enhancing property for large edges and a smoothing property for small ones. The parameter  $a$  gives us control at what level smoothing or enhanced should occur.

#### 4.4.3.4 The 2D-Lorentz prior

The 1D-prior described in section 4.4.3.1 ignores the dependence of a pixel on the pixel ‘above’. This was allowed in the 1D case, because in 1D a pixel has only left and right neighbours. To derive the 2D-prior I will use the approach used in section 4.4.3.2. Fig. 4.11 is the equivalent of Fig. 4.8 for the 2D case. Again,  $\xi_{n_x, n_y} = I_{n_x, n_y} - I_{n_x-1, n_y}$  and  $\eta_{n_x, n_y} \stackrel{\text{def}}{=} I_{n_x, n_y} - I_{n_x, n_y-1}$ . Although not valid in this case, for example  $I_{2,2} - I_{1,1} = \xi_{2,1} + \eta_{2,2} = \eta_{1,2} + \xi_{2,2}$ , I still assume independent pixel differences at different locations

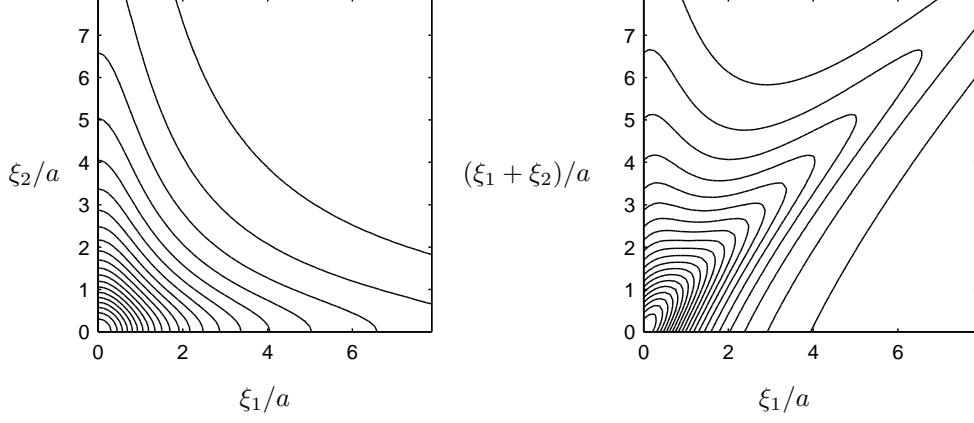


Figure 4.10: Contour curves of  $P(\xi_1, \xi_2)$ . Left: For  $\xi_1/a$  against  $\xi_2/a$ . Right: For  $\xi_1/a$  against  $(\xi_1 + \xi_2)/a$ .

in the image, resulting in

$$P(I) \propto P((\xi_{1,1}, \eta_{1,1}), \dots, (\xi_{n_x, n_y}, \eta_{n_x, n_y}), \dots, (\xi_{N_x, N_y}, \eta_{N_x, N_y})) = \prod_{n_x=1, n_y=1}^{N_x, N_y} P(\xi_{n_x, n_y}, \eta_{n_x, n_y}). \quad (4.22)$$

$\xi_{1,1}, \eta_{1,1}$	$\xi_{2,1}, \eta_{2,1}$	$\cdots$	$\xi_{n_x,1}, \eta_{n_x,1}$	$\cdots$	$\xi_{N_x,1}, \eta_{N_x,1}$
$\xi_{1,2}, \eta_{1,2}$	$\xi_{2,2}, \eta_{2,2}$	$\cdots$	$\xi_{n_x,2}, \eta_{n_x,2}$	$\cdots$	$\xi_{N_x,2}, \eta_{N_x,2}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
$\xi_{1,n_y}, \eta_{1,n_y}$	$\xi_{2,n_y}, \eta_{2,n_y}$	$\cdots$	$\xi_{n_x,n_y}, \eta_{n_x,n_y}$	$\cdots$	$\xi_{N_x,n_y}, \eta_{N_x,n_y}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
$\xi_{1,N_y}, \eta_{1,N_y}$	$\xi_{2,N_y}, \eta_{2,N_y}$	$\cdots$	$\xi_{n_x,N_y}, \eta_{n_x,N_y}$	$\cdots$	$\xi_{N_x,N_y}, \eta_{N_x,N_y}$

Figure 4.11: Matrix of the pixel value differences  $\xi_{n_x, n_y} = I_{n_x, n_y} - I_{n_x-1, n_y}$  and  $\eta_{n_x, n_y} = I_{n_x, n_y} - I_{n_x, n_y-1}$ .

The problem of finding the 2D-prior has now been reduced to finding  $P(\xi_{n_x, n_y}, \eta_{n_x, n_y})$ . Fig. 4.12 shows a contour plot of the 2D-histogram of the pixel value differences  $\xi_{x,y}$  and  $\eta_{x,y}$  of an image of a human brain. It appeared representative for a whole class of MRI images. The contours of the histogram are approximately circular, which means that the histogram is an isotropic function of  $\xi_{x,y}$  and  $\eta_{x,y}$ . To model this histogram, one could

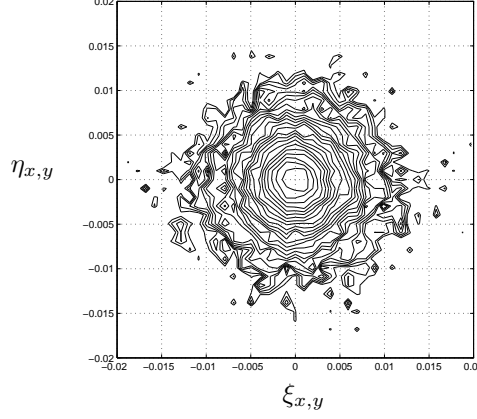


Figure 4.12: Contour plot of the 2D-histogram of  $\xi_{x,y}$  and  $\eta_{x,y}$  of an image of a human brain. The contours are approximately circular, which means that the histogram is an isotropic function of  $\xi_{x,y}$  and  $\eta_{x,y}$ . This histogram is representative for a whole class of MRI images.

simply multiply the 1D-priors of the  $x$ -direction and the  $y$ -direction, as in

$$\frac{a^2}{\pi^2(a^2 + \xi_{n_x, n_y}^2)(a^2 + \eta_{n_x, n_y}^2)}. \quad (4.23)$$

However, this function is not isotropic. To remedy this, I propose the following function [89, 90]

$$P(\xi_{n_x, n_y}) = \frac{a}{2\pi(a^2 + \xi_{n_x, n_y}^2 + \eta_{n_x, n_y}^2)^{1.5}}. \quad (4.24)$$

This expression is isotropic. Furthermore, this function is consistent with the 1D-prior, since integrating over  $\xi_{x,y}$  or  $\eta_{x,y}$  results in the correct 1D-prior.

Fig. 4.13 shows the central slice through the histogram in Fig. 4.12 and the fit to it of the function of Eq. 4.24. The quality of the fit appears good. Note that the power of 1.5 in the denominator of Eq. 4.24 is essential in fitting the histogram. In Eq. 4.23, the corresponding power is only 1, which is too small to follow the descent of the histogram at large  $\xi_{x,y}$ . This constitutes another reason to reject Eq. 4.23 as a model for the histogram in Fig. 4.12. Using Eq. 4.22 and Eq. 4.24, I end up with the following equation, which I use as 2D-prior:

$$P(I) = \prod_{n_x=1, n_y=1}^{N_x, N_y} \frac{a}{2\pi(a^2 + \xi_{n_x, n_y}^2 + \eta_{n_x, n_y}^2)^{1.5}}. \quad (4.25)$$

No further prior knowledge, see [71], was used.

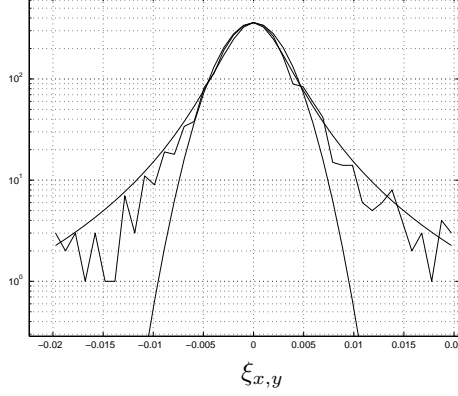


Figure 4.13: Plot of the central slice ( $\eta_{x,y} = 0$ ) through the histogram in Fig. 4.12 and the fit to it of the function of Eq. (4.24). The highest values of the histogram are in the centre. This means that the most probable values are  $\xi_{x,y} = 0$  and  $\eta_{x,y} = 0$ . These values correspond to the homogeneous regions in the image. Also shown is a fit of a Gaussian function. Apparently, the tails of the histogram are wider than that of a Gaussian, which means that  $\xi_{x,y}$  or  $\eta_{x,y}$  have a substantial probability of being large. These quantities correspond to the edges in the image.

#### 4.4.4 The likelihood

In the following equations I will write the image as a vector,  $\mathbf{i}$ , by stacking all columns of the image on top of each other, as in

$$\mathbf{i} = (I_{1,1}, I_{1,2}, \dots, I_{N_x, N_y})^T \quad (4.26)$$

As already mentioned in section 3.5, the  $k$ -space signal and image are related to each other through the 2D Fourier-transform. In addition, the samples in  $k$ -space are contaminated by additive, white, Gaussian noise with zero mean and standard deviation  $\sigma_S$  [65]. I therefore model the data as

$$\underline{\mathbf{s}} = T\mathbf{i} + \underline{\mathbf{n}}, \quad (4.27)$$

in which  $T_{n,p} = \Delta r_{\text{Cart}} e^{-2\pi i \vec{k}_n \cdot \vec{r}_p}$  and  $\underline{\mathbf{n}}$  is the measurement noise.  $\vec{r}_p$  are the locations of the pixels, which are assumed to be distributed on a Cartesian grid. Arbitrary, irregular sample positions,  $\vec{k}_n$ , are allowed. When the  $\vec{k}_n$  are indeed irregular, straight-forward calculation of  $T\mathbf{i}$  is very expensive. A solution to this problem will be given below. Should the  $\vec{k}_n$  coincide with a rectangular grid,  $T$  reduces to a 2D-FFT (actually cutting out part of the FFT of the image).

Using the Gaussian nature of the measurement noise results in the following expression

for the log-likelihood term

$$\log P(S|I) = -\frac{|\underline{\mathbf{s}} - T\underline{\mathbf{i}}|^2}{2\sigma_S^2} + \text{const.} \quad (4.28)$$

The log-likelihood is, up to an unimportant constant, equal to an unweighted least-squares term. It is unweighted, because we assumed uncorrelated noise with equal variance at each sampling point. However, it may be beneficial to ignore these assumptions and include some weighting to speed up convergence. In [58,91] it is suggested to use the density compensation as weighting. This indeed has the nice property that the resulting image after the first iteration in the optimization process is equal to the image obtained using the gridding algorithm. However, finding a suitable density compensation, especially for sparse sample distributions, is difficult. In the following, we therefore use the unweighted approach.

The usual least-squares solution of Eq. 4.28, which in this case is equivalent to the solution of the maximum likelihood technique, is impossible if the number of  $k$ -space data is smaller than the number of pixels in the image. This happens when deliberately omitting  $k$ -space samples. As already explained in section 3.5, even if we have enough samples  $T$  can be ill-conditioned, effectively making the inversion of  $T^\dagger T$  in Eq. 3.41 impossible.

However, maximizing the posterior Eq. 4.12, allows incorporation of *a priori* information (the  $P(I)$  term) about the image. This will regularize the reconstruction and compensate any ill-conditioning of  $T$ .

#### 4.4.4.1 Gradient of the log-likelihood

In following paragraphs we will need the gradient of the log-likelihood, which is equal to

$$\nabla \log P(S|I) = \frac{\partial}{\partial \underline{\mathbf{i}}} \log P(S|I) = \frac{T^\dagger (\underline{\mathbf{s}} - T\underline{\mathbf{i}})}{\sigma_S^2}. \quad (4.29)$$

Note that the maximum likelihood estimator Eq. 3.41 can be derived by setting the gradient of the log-likelihood to zero.

### 4.4.5 Computation of the log-likelihood and its gradient

#### 4.4.5.1 Cartesian sampling

In this case  $T\underline{\mathbf{i}}$  is equivalent to cutting out part of the FFT of the image. Calculation of the log likelihood can be done straightforward by applying Eq. 4.28. This requires only one FFT.

The gradient of the log-likelihood is also calculated in a straightforward way by applying Eq. 4.29. Since the gradient includes the same factor,  $\underline{\mathbf{s}} - T\underline{\mathbf{i}}$ , as the log-likelihood this needs to be calculated only once. The multiplication with  $T^\dagger$  is equivalent to an IFFT. This means that the gradient can be calculated using one IFFT.



#### 4.4.5.2 Non-Cartesian sampling

In the case of non-Cartesian sampling  $T$  will be equal to a DFT matrix and straightforward calculation of the matrix-vector product  $T\mathbf{i}$  is very expensive (order of  $N^4$  operations as opposed to 2D-FFT which requires  $N^2 \log(N)$  operations, when  $N_x = N_y = N$ ).

To speed up the calculation one can apply the interpolation method described in section 3.4.3.8 for a multiplication by  $T$  and the method described in section 3.4.3.3 for a multiplication by  $T^\dagger$ . This will drastically reduce the computation cost. The interpolations, however, have to be carried out for every iteration during the minimization. Luckily this can be circumvented using the following approach.

First of all, I expand the likelihood term as follows:

$$|\underline{\mathbf{s}} - T\mathbf{i}|^2 = \underline{\mathbf{s}}^\dagger \underline{\mathbf{s}} - 2\text{Re}(\mathbf{i}^\dagger T^\dagger \underline{\mathbf{s}}) + \mathbf{i}^\dagger T^\dagger T \mathbf{i}, \quad (4.30)$$

where the division by  $2\sigma_S^2$  has been dropped for reasons of clarity.

The first term in Eq. 4.30 can be completely dropped since it does not depend on the image, while  $\mathbf{i}^\dagger T^\dagger \underline{\mathbf{s}}$  is equal to

$$\mathbf{i}^\dagger T^\dagger \underline{\mathbf{s}} = \Delta r_{\text{Cart}} \sum_p I^*(\vec{r}_p) \left( \sum_n S(\vec{k}_n) e^{2\pi i \vec{k}_n \cdot \vec{r}_p} \right), \quad (4.31)$$

the term in parentheses equals a non-uniform 2D-IDFT as described in 3.2. It will be approximated using the technique implied by Eq. 3.18. Note that there is no density compensation in Eq. 4.31.

The right part of the third term of Eq. 4.30 can be rewritten as

$$\begin{aligned} (T^\dagger T \mathbf{i})(\vec{r}_p) &= \Delta r_{\text{Cart}}^2 \sum_n e^{2\pi i \vec{k}_n \cdot \vec{r}_p} \left( \sum_q e^{-2\pi i \vec{k}_n \cdot \vec{r}_q} I(\vec{r}_q) \right) = \\ &= \sum_q \left( \sum_n e^{2\pi i \vec{k}_n \cdot (\vec{r}_p - \vec{r}_q)} \right) I(\vec{r}_q) = \sum_q Q(\vec{r}_p - \vec{r}_q) I(\vec{r}_q), \end{aligned} \quad (4.32)$$

where  $Q(\vec{r}) \stackrel{\text{def}}{=} \sum_n e^{2\pi i \vec{k}_n \cdot \vec{r}}$  has exactly the same form as the term in parentheses in Eq. 4.31, so I calculate it using the same approximation technique. Actually they are calculated at the same time, because the only difference is the  $S(\vec{k}_n)$  term. Note, however, that the size of  $Q(\vec{r})$  should equal two times the size (in each dimension) of the image, since  $\vec{r}_p - \vec{r}_q$  has twice the range of  $\vec{r}_p$ . This forms no obstacle in applying Eq. 3.18 and Eq. 4.31 and  $Q(\vec{r})$  can still be calculated at the same time.

The last term in Eq. 4.32 represents a 2D convolution of  $Q(\vec{r})$  with the image  $I(\vec{r})$ , which can be calculated efficiently using the 2D-FFT, *i.e.*  $(T^\dagger T \mathbf{i})(\vec{r}_p) = (\text{IFFT}(\text{FFT}(Q) \cdot \text{FFT}(I))) (\vec{r}_p)$ .

Now we return to the calculation of  $\mathbf{i}^\dagger T^\dagger T \mathbf{i}$ ,

$$\mathbf{i}^\dagger T^\dagger T \mathbf{i} = \sum_p I^*(\vec{r}_p) \cdot (\text{IFFT}(\text{FFT}(Q) \cdot \text{FFT}(I)))(\vec{r}_p) \quad (4.33)$$

$$= \sum_m (\text{FFT}(I)^*)(\vec{k}_m) \cdot (\text{FFT}(Q) \cdot \text{FFT}(I))(\vec{k}_m) \quad (4.34)$$

$$= \sum_m (\text{FFT}(Q))(\vec{k}_m) \cdot (|\text{FFT}(I)|^2)(\vec{k}_m). \quad (4.35)$$

Eq. 4.34 is derived from Eq. 4.33 by using the unitarity of the FFT or equivalently by using Parseval's relation. This speeds up the calculation of  $\mathbf{i}^\dagger T^\dagger T \mathbf{i}$ , since it removes an FFT ( $\text{FFT}(I)$  is present twice in Eq. 4.34).

Because  $Q(\vec{r}_p)$  is a constant during the minimization of the log-posterior, we can store the 2D-FFT of  $Q(\vec{r}_p)$  instead of  $Q(\vec{r}_p)$  itself. This removes yet another FFT, reducing the number of FFT's required to calculate the log-likelihood to only one. However, we need to FFT a twice as large matrix as in the uniform case. Moreover, if in the uniform case one direction is fully sampled, we can already apply the FFT along that direction reducing all operations to 1D.

The gradient likelihood is written as

$$\nabla \log P(S|I) = \frac{T^\dagger \mathbf{s} - T^\dagger T \mathbf{i}}{\sigma_S^2}. \quad (4.36)$$

Again the term  $T^\dagger \mathbf{s}$  arises. The second term  $(T^\dagger T \mathbf{i})(\vec{r}_p) = (\text{IFFT}(\text{FFT}(Q) \cdot \text{FFT}(I)))(\vec{r}_p)$ . However, in the computation of the log-likelihood we have already calculated  $\text{FFT}(Q) \cdot \text{FFT}(I)$ , see Eq. 4.35. This means that we can compute the gradient with only one IFFT. It is advantageous to have this IFFT in the gradient and not in the computation of the log-likelihood, because as we will see the log-likelihood is computed three times as often as its gradient.

#### 4.4.6 Choice of the hyperparameters [1] $\sigma_S$ and $a$

$\sigma_S$  can be estimated by assuming that at the edge of  $k$ -space the signal has decayed and only noise remains. This is usually the case since otherwise there would be a discontinuity at the edge of  $k$ -space giving Gibbs artefacts in the image.

The  $a$  parameter of the Lorentzian prior is linked to  $\sigma_I$  in the following way. In flat parts of the image ( $\xi \approx 0$  and  $\eta \approx 0$ ) the intensity changes are dominated by noise. The source of this image noise is the noise in  $k$ -space. Since the noise in  $k$ -space is Gaussian distributed, so will the noise in the image, because the Fourier transform is a linear operator. Therefore, in flat parts of the image the distribution of  $\xi$  and  $\eta$  should approximate a Gaussian with zero mean and twice the variance of the image noise. This can be accomplished by setting the second derivative of the Lorentzian equal to the second derivative of the Gaussian, both at the peak. The value for  $a$  found in this way is equal to

$$a = 2\sigma_I \quad (4.37)$$

Finally, I approximate  $\sigma_I$  by  $N\sigma_S$ , which can be found by applying Eqs. 3.45 - 3.47 to a Cartesian sample distribution.

#### 4.4.7 Posterior optimization

Substituting Eq. 4.25 and Eq. 4.28 in Eq. 4.12, while ignoring any constant terms, one gets the following optimization problem:

$$\min_I \left\{ \frac{|\underline{\mathbf{s}} - T\underline{\mathbf{i}}|^2}{2\sigma_S^2} + \frac{3}{2} \sum_{n_x, n_y} \log \left\{ a^2 + (\xi_{n_x, n_y})^2 + (\eta_{n_x, n_y})^2 \right\} \right\} \quad (4.38)$$

The image  $I_{\text{MAP}}$  which results from Eq. 4.38 is the reconstructed image. The minimization is done with the conjugate gradient method.

We could start directly with Eq. 4.38, and associate the prior with a regularization term, leading to a constrained maximum likelihood method. The preceding sections was an attempt to derive Eq. 4.38 using the Bayesian approach and showing why we use this type of regularization instead of just trying one or several [88].

#### 4.4.8 Optimization method

The function to be minimized is the minus log-posterior probability,  $-\log P(I|S)$ . The independent variables in this function are the  $N_x \times N_y$  pixel values in the image  $I$ . The number of independent variables is huge, so I cannot use an optimization method based on inverting the Hessian, like Gauss-Newton or Broyden-Fletcher-Goldfarb-Shanno (BFGS [43]), since these require storage on the order of  $(N_x N_y)^2$ . This leaves simple gradient search algorithms of which I chose the conjugate gradient method. This will be explained briefly.

At each iteration  $n$ , the image is updated as

$$I_{n+1} = I_n + \lambda_n \delta I_n, \quad (4.39)$$

in which  $\delta I_n$  is the search direction at iteration step  $n$ . The conjugate gradient algorithm chooses the search direction as follows

$$\delta I_n = -\nabla f(I_n) + \frac{|\nabla f(I_n)|^2}{|\nabla f(I_{n-1})|^2} \delta I_{n-1} \quad (4.40)$$

$$\delta I_0 = 0, \quad (4.41)$$

in which  $\nabla f(I_n) = \frac{\partial f(I_n)}{\partial I_n}$ .

The iteration starts with image  $I_1$  ( $n = 1$ ). Since we have no knowledge of how the image should look like we set  $I_1 = 0$ .

Having the search direction,  $\delta I_n$ , the scalar  $\lambda_n$  is chosen in such a way that  $-\log P(I_n + \lambda_n \delta I_n | S)$  is a minimum. This one-dimensional problem is called line-search. Eq. 4.39 is repeated until convergence is achieved.

Finding the minimum of  $-\log P(I_n + \lambda_n \delta I_n | S)$  is performed in two steps. First the minimum is bracketed. The second step is to actually find the location of the minimum. These steps will now be explained more thoroughly.

#### 4.4.8.1 Bracketing the minimum

The bracketing procedure ensures that we know that there is a minimum in the first place. It also gives us an interval where to look for the minimum.

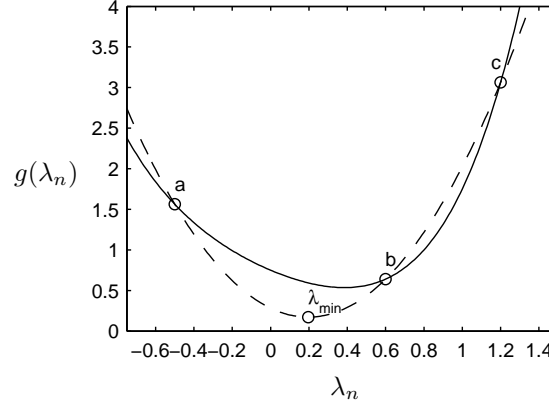


Figure 4.14: Bracketing the minimum.  $a$ ,  $b$  and  $c$  as in the figure ensures that  $g(\lambda_n) = -\log P(I_n + \lambda_n \delta I_n | S)$  has a minimum between  $\lambda = a$  and  $\lambda = c$ .

The bracketing procedure finds  $\lambda_n^a$ ,  $\lambda_n^b$  and  $\lambda_n^c$  in such a way that  $-\log P(I_n + \lambda_n^b \delta I_n | S) < -\log P(I_n + \lambda_n^a \delta I_n | S)$  and  $-\log P(I_n + \lambda_n^b \delta I_n | S) < -\log P(I_n + \lambda_n^c \delta I_n | S)$ . This ensures that there is a minimum of  $-\log P(I_n + \lambda_n \delta I_n | S)$  between  $\lambda_n = \lambda_n^a$  and  $\lambda_n = \lambda_n^c$ . I choose  $a = 0$ . This saves one function evaluation, because  $-\log P(I_n | S)$  is the function evaluated at the minimum of the previous iteration.

Assuming that the size of the search steps (*i.e.*  $\lambda_n$ ) do not change a lot from iteration to iteration, choosing  $\lambda_n^b = \lambda_{n-1}$  and  $c = 2\lambda_{n-1}$ , will ensure fast bracketing. If  $-\log P(I_n + \lambda_n^a \delta I_n | S)$ ,  $-\log P(I_n + \lambda_n^b \delta I_n | S)$  and  $-\log P(I_n + \lambda_n^c \delta I_n | S)$  are not as required, I set  $\lambda_n^b = 2\lambda_{n-1}$  and  $c = 4\lambda_{n-1}$ , etc. This requires only one extra function evaluation per increase of  $\lambda^b$  and  $\lambda^c$ . This procedure stops when bracketing is achieved.

#### 4.4.8.2 Isolating the minimum along the search direction

Finding the location of the exact minimum is difficult when the function is non-linear. However, it is not necessary to find the exact minimum because this is a minimization along this search direction and the next search direction will be in a different direction.

Therefore, the exact minimum is approximated by the minimum of the parabola passing through  $-\log P(I_n + \lambda_n^a \delta I_n | S)$ ,  $-\log P(I_n + \lambda_n^b \delta I_n | S)$  and  $-\log P(I_n + \lambda_n^c \delta I_n | S)$ . This procedure does not require any extra function evaluations, which makes it fast. The value for  $\lambda_n$  thus found is used for  $\lambda_n$ .

Finally, we need to calculate the new value of the criterion function  $-\log P(I_n + \lambda_n \delta I_n | S)$ .

#### 4.4.8.3 Number of function evaluations

Using the above procedure for the line-search requires the following number of operations per iteration

- The bracketing requires at least 2 function evaluations. When the bracketing does not succeed immediately, each successive try requires one extra function evaluation. With the above procedure bracketing usually succeeds with 2 function evaluations.
- The evaluation of the criterion function at the ‘minimum’ along the search direction requires one function evaluation.
- Finding the next search direction requires one gradient evaluation.

We see that it is beneficial to have a faster function than gradient evaluation, simply because the number of function evaluations is at least three times the number of gradient evaluations per iteration. This is the reason of calculating the likelihood and its gradient in the way explained in section 4.4.5.

## 4.5 Summary and conclusions

This chapter treated image reconstruction from incomplete data. According to my definition ‘incomplete’ or ‘sparse’ pertains to undersampling. Undersampling brings about irretrievable loss of necessary information. Hence, inference of the true image from incomplete data is impossible. This predicament can be alleviated by invoking prior knowledge (physical approach) or by regularization (mathematical approach). Undersampling can have at least two causes: 1) Omission of trajectories. This occurs when the scan time is to be reduced. 2) Correction for patient motion, directly in  $k$ -space.

When trajectories have been omitted, an obvious procedure is to estimate the samples on these trajectories and then proceed as if one deals with complete data; this was called type 1 reconstruction. However, when originally complete data are corrected for motion, (irregular) changes of most sample positions occur. As a result, the notion of missing data does no longer apply. In this case, I reconstruct an image that satisfies the available (locally undersampled) data, without estimating any other data; this was called type 2 reconstruction.

The material treated in this chapter allows the following conclusions.

- Reconstruction from incomplete data requires prior knowledge or regularization.
- Type 1 image reconstruction can be done with edge-preserving filtering.
- Edge-preserving filtering requires gridding, which in turn requires the density compensation. When the sampling is irregular, the density compensation is error-prone.
- Type 2 image reconstruction can be done with Bayesian estimation. Properties of Bayesian estimation are
  - it obviates the density compensation

- it can handle arbitrary sample positions
- it can be applied to data with omissions as well as motion-corrected data.
- The algorithm is iterative and therefore computationally intensive. To alleviate this aspect, the computation of the ‘likelihood term’ has been streamlined considerably.

## Chapter 5

# Evaluation of image reconstruction from incomplete data

### 5.1 Introduction

This chapter concerns performance evaluation of image reconstruction from incomplete  $k$ -space data. The word incomplete pertains to local violation of the Nyquist sampling criterion, also alluded to as local under-sampling or sparse sampling. In this thesis there are two sources of local under-sampling. One is intentional omission of  $k$ -space trajectories aimed at reduction of scan time, the other is unintentional changes of the angles between successive  $k$ -space trajectories due to patient motion, called intra-scan rotations.

Ultimately, only the judgment of persons applying a reconstruction algorithm counts. Regular contacts with clinicians are therefore desirable. However, developers of reconstruction algorithms cannot efficiently progress under solely such a regime. Consequently, various performance estimators, producing a single number that represents the results for all pixels, have been devised. An important aspect in this is the choice of the reference image. One may refer to the image estimated from complete (no Nyquist violation)  $k$ -space data, or to the image obtained when omitted samples are zero-filled or to a combination of the two. In the case of intra-scan rotations, one refers to the case of no rotation. In general, a complication arises due to the fact that even fully sampled data are necessarily truncated at the edge of  $k$ -space which causes Gibbs artefacts (ringing). The next section treats several alternative performance estimators.

## 5.2 Criteria

### 5.2.1 Image performance measures

For easy evaluation of the performance of the image reconstruction from the sparse data-set, this performance has to be specified by a number. Ideally, we would like a number given only the reconstructed image. However, practical performance measures compare the reconstructed image to a reference image. Whenever we only have the sparse data-set, the choice of reference is not a trivial matter. However, I will assume that the sparse data-set was constructed from a full data-set from which one can reconstruct a reference image or that a separate reference image (Cartesian measurement) is available. Note that the reference image is in no way ideal since it may have Gibbs-artefacts due to truncation of  $k$ -space or other artefacts related to system imperfections. A reason for using simulations is to avoid these problems since we then know the true image. However, we have to make sure that the simulations are representative of the type of data-sets/images we are likely to find in real-world situations.

To compare the reconstructed image with the reference image, it is convenient to introduce the residue. The residue is defined as the difference between the reference image,  $I_{\text{ref}}$ , and the reconstructed image,  $I_{\text{recon}}$ . These images may have been scaled differently and may have different DC-offsets. Therefore I define the residue as follows

$$\mathbf{i}_{\text{residu}} \stackrel{\text{def}}{=} \mathbf{i}_{\text{ref}} - \alpha \mathbf{i}_{\text{recon}} + \beta, \quad (5.1)$$

in which the factor  $\alpha$  compensates difference in scaling and  $\beta$  is a difference in DC-offset. In this and subsequent equations, the images are written as vectors by stacking their columns on top of each other. Thus,  $\mathbf{i}_{\text{residu}}$  corresponds to  $I_{\text{residu}}$ , etc.

The first three performance measures are all based on the energy of the residue  $|\mathbf{i}_{\text{residu}}|^2 = \mathbf{i}_{\text{residu}}^\dagger \mathbf{i}_{\text{residu}}$ . They differ in whether or not they include the scaling and/or DC-offset in the calculation of the residue.

The first performance measure ignores any scaling and DC-offset difference ( $\alpha = 1$  and  $\beta = 0$ ). This results in the widely used mean squared error [1]

$$\text{MSE}(I_{\text{ref}}, I_{\text{recon}}) = \frac{|\mathbf{i}_{\text{ref}} - \mathbf{i}_{\text{recon}}|^2}{N_x N_y}, \quad (5.2)$$

where the size of both images is  $N_x \times N_y$ .

The MSE depends on the scaling of  $I_{\text{ref}}$  and  $I_{\text{recon}}$ . For instance, if we multiply both images by 10, the MSE increases by a factor of 100. To eliminate this effect we calculate the MSE relative to the energy of the reference image (actually  $\sqrt{|\mathbf{i}_{\text{ref}}|^2 |\mathbf{i}_{\text{recon}}|^2}$ , to keep the measure symmetric in  $\mathbf{i}_{\text{ref}}$  and  $\mathbf{i}_{\text{recon}}$ ), which gives the relative mean squared error

$$\text{rMSE}(I_{\text{ref}}, I_{\text{recon}}) = \frac{|\mathbf{i}_{\text{ref}} - \mathbf{i}_{\text{recon}}|^2}{\sqrt{|\mathbf{i}_{\text{ref}}|^2 |\mathbf{i}_{\text{recon}}|^2}} \quad (5.3)$$

Using the rMSE, the first performance number is defined as

$$\text{perf}_1 \stackrel{\text{def}}{=} -10 \log_{10}(\text{rMSE}) = -10 \log_{10} \left( \frac{|\mathbf{i}_{\text{ref}} - \mathbf{i}_{\text{recon}}|^2}{\sqrt{|\mathbf{i}_{\text{ref}}|^2 |\mathbf{i}_{\text{recon}}|^2}} \right) \quad (5.4)$$



The measure  $\text{perf}_1$  only gives usable results whenever  $I_{\text{ref}}$  and  $I_{\text{recon}}$  are scaled equally. The performance should not be influenced by a change of scaling because a change of scaling does not really change the image. The second performance measure eliminates the necessity of equal scaling by adapting  $\alpha$  in the residue. The value of  $\alpha$  is found by a least squares fit ( $\beta$  is still fixed at 0). This minimizes the energy of the residue, which maximizes the resulting performance measure. Using this value for  $\alpha$  and ensuring symmetry in  $\mathbf{i}_{\text{ref}}$  and  $\mathbf{i}_{\text{recon}}$  of the resulting performance measure, one gets

$$\text{perf}_2 = -10 \log_{10} \left( 1 - \frac{(\mathbf{i}_{\text{recon}}^\dagger \mathbf{i}_{\text{ref}})(\mathbf{i}_{\text{ref}}^\dagger \mathbf{i}_{\text{recon}})}{|\mathbf{i}_{\text{ref}}|^2 |\mathbf{i}_{\text{recon}}|^2} \right). \quad (5.5)$$

Finally, the third performance measure compensates also for a possible DC-offset difference by adapting  $\beta$ . Although I will not use this performance measure, I include it here for completeness. Again minimizing the energy of the residue in least squares sense, results in the following performance number

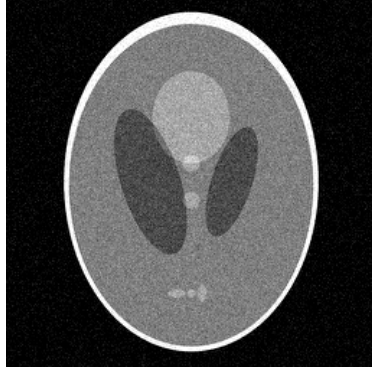
$$\text{perf}_3 = -10 \log_{10} (1 - |r|^2), \quad (5.6)$$

where  $r$  is the correlation coefficient of  $I_{\text{ref}}$  and  $I_{\text{recon}}$ . Note that Eq. 5.6 follows easily from Eq. 5.5 by subtracting the mean value from both  $I_{\text{ref}}$  and  $I_{\text{recon}}$ . M.R. Smith has devised a rather more elaborate performance criterion [92].

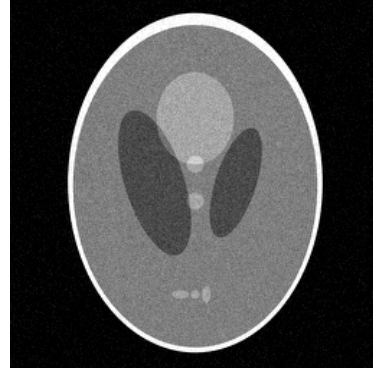
Table 5.1: Performances of IFFT reconstruction for full Cartesian scanning, using alternative noiseless reference images,  $I_{\text{cart}}$  and  $I_{\text{true}}$ . The central column represents the effect of noise, the right-hand column represents the combined effects of noise plus Gibbs-artefacts.

SNR	$\text{perf}_2$ (in dB) $I_{\text{ref}} = I_{\text{cart}}$	$\text{perf}_2$ (in dB) $I_{\text{ref}} = I_{\text{true}}$
10	15.0	14.0
15	18.4	16.6
20.7	21.2	18.2
30	24.4	19.5
50	28.8	20.5
$\infty$	$\infty$	21.2

Fig. 5.1 shows IFFT reconstructions from four fully sampled Cartesian Shepp-Logan simulations, generated with increasing noise. These reconstructions serve to provide some insight in the relation between a performance number and the visual judgment of image quality. At SNR = 10 and 15, noise visibly perturbs the image; the performance is below 20 dB. At SNR = 30, noise barely perturbs the image, while the performance approaches 25 dB. Gibbs artefacts can be detected clearly only after zooming.



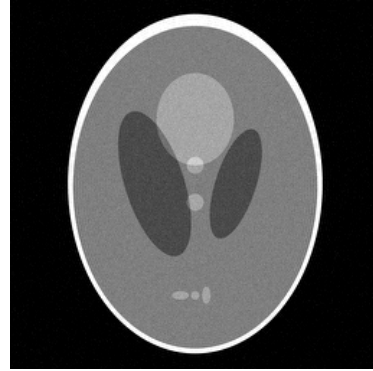
SNR = 10,  $\text{perf}_2 = 15.0$  dB



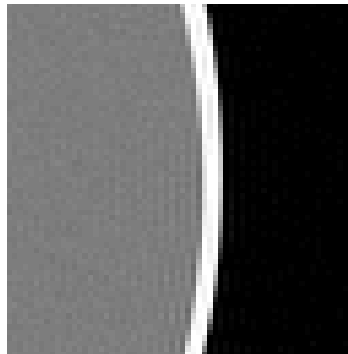
SNR = 15,  $\text{perf}_2 = 18.4$  dB



SNR = 20.7,  $\text{perf}_2 = 21.2$  dB



SNR = 30,  $\text{perf}_2 = 24.4$  dB



Zoomed, SNR = 50

Figure 5.1: Reconstructions from a full Cartesian-sampled Shepp-Logan simulation.

Table 5.1 shows the performance of IFFT reconstruction from a full Cartesian scan, at various noise levels and using alternative noiseless reference images. In Cartesian scanning, one can use as reference image either that obtained from IFFT of the full data set –  $I_{\text{cart}}$  – or the true image –  $I_{\text{true}}$  (if known). When using  $I_{\text{cart}}$  as reference, the reconstructed image and the reference image possess the same Gibbs artefacts; hence, the performance number is affected only by noise. As expected then, at  $\text{SNR} = \infty$ , the performance is  $\infty$  too. When using  $I_{\text{true}}$  as reference, the performance is of course lower: at  $\text{SNR} = \infty$ , it is only 21.2 dB. The latter value is attained already at  $\text{SNR} = 20.7$  when using  $I_{\text{cart}}$  as reference.

### 5.2.2 Special performance measure for scan-time reduction

The three performance measures (Eqs. 5.4, 5.5 and 5.6) discussed so far each compare a reconstructed image with a reference image, therefore using two images. The performance measure I discuss in this section uses three images, one reference image and two reconstructed images. It was especially developed for Cartesian scan-time reduction [71, 74]. The main idea of this performance measure is to compare the quality (actually the MSE's) of two reconstructed images.

The first reconstructed image,  $I_{\text{trunc}}$ , is obtained from the full Cartesian data-set with the simplest way of scan-time reduction, *i.e.* omission of the outer trajectories resulting in a truncation of  $k$ -space, see Fig. 4.1a. The reason for omitting the outer trajectories is that they contain the least signal energy, resulting in lowest reduction of the MSE.  $I_{\text{trunc}}$  is reconstructed from the truncated  $k$ -space in the simplest way, *i.e.* by zero-filling the missing trajectories and subsequently applying IFFT.

The second reconstructed image,  $I_{\text{recon}}$ , is obtained from the *same* full data-set by omitting trajectories in a sophisticated irregular way, see Fig. 4.1b. The image should be reconstructed using a special algorithm capable of handling such data-sets, for example, one of the reconstruction methods explained in chapter 4; see [71] for the dedicated Cartesian case. Using zero-filling and IFFT results in a worse image than  $I_{\text{trunc}}$ , *i.e.* lower MSE. The reason is that with irregular omission, using the same amount of scan-time reduction, we omit trajectories closer to  $k = 0$  than in the case of truncation. The samples on these trajectories have higher signal energy. Without a special algorithm, this approach would result in a lower quality (*i.e.* lower MSE).

The definition of this performance criterion is such that it provides a measure of whether or not the sophisticated way of scan-time reduction was worthwhile with respect to simple truncation

$$\text{perf}_4 = 1 - \frac{\text{MSE}(I_{\text{ref}}, I_{\text{recon}})}{\text{MSE}(I_{\text{ref}}, I_{\text{trunc}})} = 1 - \frac{\|\mathbf{i}_{\text{ref}} - \mathbf{i}_{\text{recon}}\|^2}{\|\mathbf{i}_{\text{ref}} - \mathbf{i}_{\text{trunc}}\|^2}, \quad (5.7)$$

where  $I_{\text{ref}}$  is the reference image reconstructed (by 2D-IFT) from the full Cartesian data-set.

The value of  $\text{perf}_4$  is smaller than, or equal to, 1.  $\text{perf}'_4 \leq 0$  occurs when  $\text{MSE}(I_{\text{ref}}, I_{\text{recon}})$  is larger than, or equal to,  $\text{MSE}(I_{\text{ref}}, I_{\text{trunc}})$ . In that case, the image obtained by scan-time reduction through irregular omission of trajectories is worse than the image obtained by simple truncation of  $k$ -space. This indicates a failure of the scan-time reduction

through irregular omission. A  $\text{perf}_4$  between 0 and 1 indicates that  $I_{\text{recon}}$  is ‘closer’ to  $I_{\text{ref}}$  than  $I_{\text{trunc}}$  is. This indicates a successful scan-time reduction through irregular omission of the trajectories. The maximum value of  $\text{perf}_4$  is 1. This occurs when  $I_{\text{recon}}$  equals  $I_{\text{ref}}$ , indicating perfect estimation of the omitted data/trajectories. Typical values of  $\text{perf}_4$  are around 0.65 for 30% scan-time reduction [71].

#### 5.2.2.1 Problems with applying $\text{perf}_4$ to non-Cartesian data-sets

For  $\text{perf}_4$  one needs two sparse data-sets. In the Cartesian case one can construct both of them from the full data-set simply by omitting trajectories. Hence, it is not necessary to measure separately. This is important because such measurements are usually unavailable. In the following discussion, I will assume that we have only the complete non-Cartesian data-set. Any sparse data-set must be constructed from the complete data-set by omitting trajectories.

First of all, there is no problem with  $I_{\text{ref}}$  or  $I_{\text{recon}}$ .  $I_{\text{ref}}$  is reconstructed using an appropriate algorithm suitable for full non-Cartesian data-sets; see chapter 3. For  $I_{\text{recon}}$ , one of the algorithms described in chapter 4 should be used. Which trajectories to omit is discussed in the next chapter. This corresponds exactly to what was done in the Cartesian case.

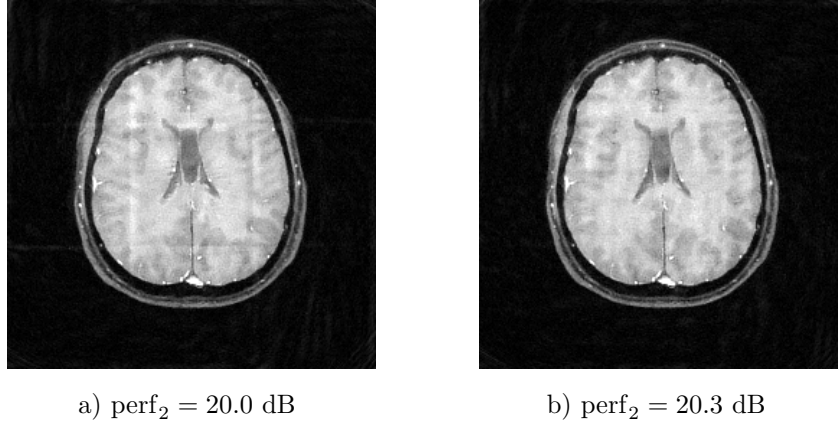


Figure 5.2: Bayesian reconstruction from sparse spiral data for two alternative approaches to omitting trajectories. a) Every third trajectory of the full set is omitted. The remaining ones stay put. b) Trajectories are omitted at random from the full set, the remaining ones staying put.

The first problem in applying  $\text{perf}_4$  lies in the fact that in the non-Cartesian case there is no simple way of omitting trajectories. Remember that in the Cartesian case simple scan-time reduction was accomplished by omitting the outer trajectories corresponding to a truncation of  $k$ -space. The reason for omitting these trajectories, as already explained above, is that we know that the outer trajectories have lowest signal energy, and

omitting them results in the lowest reduction of the MSE. Moreover, since we omit trajectories with small signal energy we can reconstruct  $I_{\text{trunc}}$  with zero-filling and applying a reconstruction algorithm suitable for complete data-sets, *i.e.* IFFT in the Cartesian case.

In the non-Cartesian case there are not (necessarily) any outer trajectories. For instance, in spiral scanning all trajectories start at the centre of  $k$ -space. Moreover, all spiral trajectories have more or less equal signal energy, so we do not have simple selection of which trajectories to omit based on their energy content.

Therefore we must have a different simple selection rule; for example, for 25% scan-time reduction ‘omit every fourth trajectory’. However, we then arrive at the second problem. Since we are still omitting trajectories with high signal energy content, we expect image artefacts if we reconstruct using a ‘simple’ reconstruction algorithm, *i.e.* zero-filling and application of one of the algorithms described in chapter 3. This would correspond to what was done in the Cartesian case. However, using zero-filling and IFFT in the Cartesian case still resulted in an acceptable  $I_{\text{trunc}}$ , because the omitted trajectories contained little signal energy.

One could ignore the fact that  $I_{\text{trunc}}$  contains artefacts, but this will cause  $\text{perf}_4$  to be nearly equal to 1. The reason for this is that  $\text{MSE}(I_{\text{ref}}, I_{\text{trunc}})$  is so large, that any reasonable  $I_{\text{recon}}$  will have a smaller MSE. However, this does not mean that  $\text{MSE}(I_{\text{ref}}, I_{\text{recon}})$  is small; therefore  $I_{\text{recon}}$  can still be pretty bad. Values of 0.9 for  $\text{perf}_4$  are easily attainable, but still represent a bad  $I_{\text{recon}}$ . In the Cartesian case  $\text{perf}_4 = 0.9$  is considered very good.

Fig. 5.2 shows that even though two images may have close performance measures, their subjective qualities can be completely different: The image on the left has a severe square-like artefact.

### 5.3 Summary and conclusions

This chapter concerned quality evaluation of image reconstruction from the raw  $k$ -space data. In my opinion, evaluation of a reconstructed image should ultimately be left to the clinical users. However, continual involvement of clinical users during the development of a reconstruction algorithm is impractical. Hence, the need for evaluation algorithms that provide some measure of image quality in the form of a single number. My evaluation algorithms are based on the relative energy of the difference between the reconstructed image and a reference image. The choice of the reference image is not trivial.

For sparse scans, I made a distinction between Cartesian and non-Cartesian sampling. With Cartesian sampling, one can restrict omission of trajectories to the ones with a relatively weak signal. This allows use of two reference images, one obtained from the complete scan and the other from the truncated scan. With radial and spiral scanning, on the other hand, all trajectories have roughly equal signal intensity. The latter implies that the notion of truncation ceases to exist, and with it the corresponding reference image. This lowered the level of sophistication.

Finally, an application of my quality evaluation algorithm to reconstructions from two alternative sparse spiral scans yielded a high quality for a visually bad result. This

result suggests that automatic image-quality ranking without human interaction, may pose problems.

## Chapter 6

# Optimal sparse sample distributions

### 6.1 Introduction

A sample distribution is a collection of sample points in  $k$ -space. For instance, spiral scanning gives us a spiral sample distribution. Usually the distribution is chosen in such a way that the Nyquist criterion is fulfilled everywhere in  $k$ -space. Image reconstruction from a data-set sampled according to such a complete sample distribution can be accomplished by one of the methods described in chapter 3.

To accomplish scan-time reduction I omit certain trajectories from the complete sample distribution. This sparse sample distribution is bound to violate the Nyquist criterion at certain locations in  $k$ -space and I will need to apply one of the methods described in chapter 4, or accept any image artefacts due to the  $k$ -space under-sampling.

This chapter concerns optimal sparse sample distributions. Such a distribution allows reconstruction of images with minimal artefacts. In other words, which trajectories do we have to omit from the complete sample distribution in order to have optimal image reconstruction and how to choose them.

I shall restrict myself to radial and spiral scanning. I shall distinguish two types of sparse sample distributions. The distinction pertains to the angle between successive radial or spiral trajectories. In one category this angle is constant during a scan, in the other category it is varied during a scan. In the sequel, I shall speak of regular and irregular sparse scanning, respectively.

For sparse Cartesian scanning, I refer the reader to Refs. [71, 93].

### 6.2 Regular sparse spiral scanning

This section pertains to sparse radial or spiral sampling with a constant angle between successive trajectories, *i.e.* the trajectories are distributed regularly over  $k$ -space. I begin by recalling the ideal relation between a continuous image and corresponding continuous

$k$ -space signal

$$S(\vec{k}) = \iiint_{\text{FOV}} I(\vec{r}) e^{-2\pi i \vec{k} \cdot \vec{r}} dV, \quad (6.1)$$

and vice versa. If  $I(\vec{r})$  is real-valued, it follows immediately from Eq. 6.1 that

$$S(-\vec{k}) = S^*(\vec{k}), \quad (6.2)$$

where  $*$  denotes complex conjugation. In words, Eq. 6.2 implies that reflection of the signal with respect to  $\vec{k} = \vec{0}$  results in complex conjugation, if the image is real-valued. This in turn, suggests that sampling at both  $\vec{k}$  and its origin-reflected value  $-\vec{k}$  yields redundant information.

As derived in chapter 3, the reconstructed image can be written as the convolution of the point-spread function and ‘true’ image obtained from the continuous Fourier transform

$$I_{\text{recon}}(\vec{r}) = \text{PSF}(\vec{r}) * I_{\text{true}}(\vec{r}) = [\text{Re PSF}(\vec{r}) + i \text{Im PSF}(\vec{r})] * I_{\text{true}}(\vec{r}) \quad (6.3)$$

in which  $*$  stands for convolution. See Eq. 3.6 for the definition of the point-spread function.

Assuming  $I_{\text{true}}(\vec{r})$  to be real-valued – a simple matter in the case of simulations – Eq. 6.3 says that in general a reconstructed image is complex-valued. However,  $I_{\text{recon}}(\vec{r})$  becomes real-valued when the point-spread function  $\text{PSF}(\vec{r})$  is real-valued. According to Eq. 3.6,  $\text{PSF}(\vec{r})$  in turn becomes real-valued if the distribution of sampling points in  $k$ -space is symmetric with respect to  $\vec{k} = \vec{0}$ . This symmetry condition can be satisfied by taking an even number of radial or spiral trajectories, with equal angles between successive trajectories.

If the number of radial and spiral trajectories is uneven, and the angles between all successive trajectories are equal, one can consider two different realisations related to each other by the operation  $\vec{k} \rightarrow -\vec{k}$ . I shall distinguish between the two realizations and corresponding results by adding a prime to the symbols involved, *i.e.* I write  $\text{PSF}(\vec{r})$ ,  $\text{PSF}'(\vec{r})$ , and  $I_{\text{recon}}(\vec{r})$ ,  $I'_{\text{recon}}(\vec{r})$ . From Eq. 3.6, it follows for an uneven number of radial or spiral trajectories

$$\text{Re PSF}'(\vec{r}) = \text{Re PSF}(\vec{r}), \quad (6.4)$$

$$\text{Im PSF}'(\vec{r}) = -\text{Im PSF}(\vec{r}), \quad (6.5)$$

and subsequently,

$$\text{Re } I'_{\text{recon}}(\vec{r}) = \text{Re } I_{\text{recon}}(\vec{r}), \quad (6.6)$$

$$\text{Im } I'_{\text{recon}}(\vec{r}) = -\text{Im } I_{\text{recon}}(\vec{r}), \quad (6.7)$$



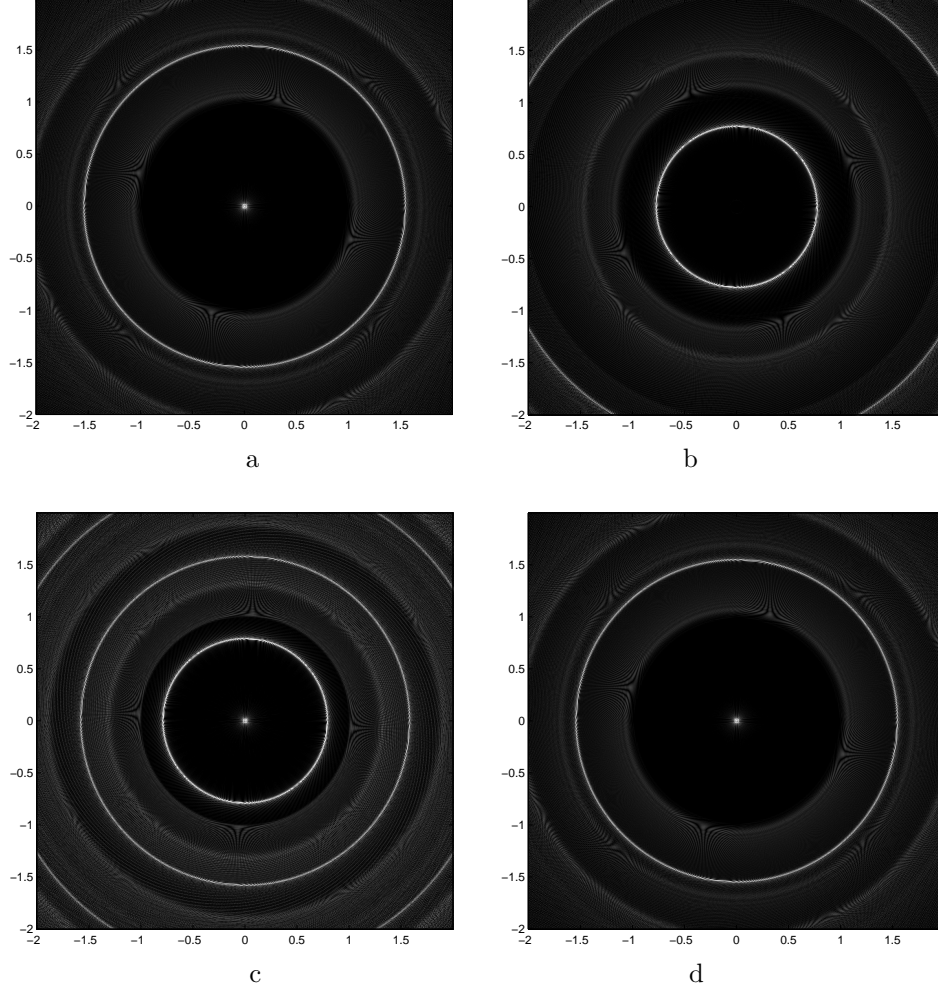


Figure 6.1: Examples of point-spread functions of spiral scans. The region shown is four times the ROI as defined by the gridding procedure described in chapter 3. a) 49 spirals, real part, b) 49 spirals, imaginary part, c) 50 spirals, real part, d) 98 spirals, real part. The intensity scales in a), b), c), and d) are equal to each other. In a), b), and d), the peak at  $\vec{r} = \vec{0}$  – ideally, a ' $\delta$ -function' – has been truncated. Slices of a), b), c), and d), along the  $x$ -axis, are displayed in Fig. 6.2. The latter indicate clearly that the intensity of the peak at  $\vec{r} = \vec{0}$  is much higher than that of the circular features which are sampling artefacts.

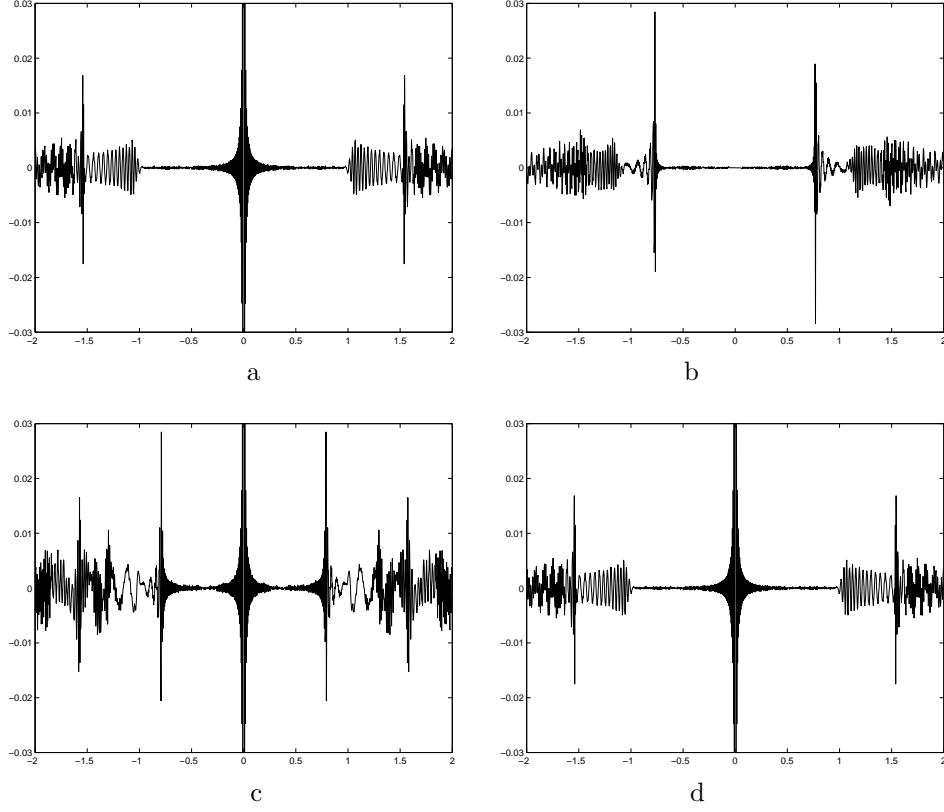


Figure 6.2: Slices through  $\vec{r} = \vec{0}$ , parallel to the  $x$ -axis, of the point-spread functions shown in Fig. 6.1. The region shown,  $-2 \leq x \leq 2$ , is four times the FOV as defined by the gridding procedure described in chapter 3. a) 49 spirals, real part, b) 49 spirals, imaginary part, c) 50 spirals, real part, d) 98 spirals, real part. The height of the peak at  $x = 0$  – which is actually a sinc function – in a), c), d) has been normalized to unity. It exceeds the height of the sampling artefacts seen for  $|x| > 0.5$  by about two orders of magnitude.

so that

$$\text{Re } I_{\text{recon}}(\vec{r}) = \frac{1}{2} (I_{\text{recon}}(\vec{r}) + I'_{\text{recon}}(\vec{r})) . \quad (6.8)$$

Eqs. 6.4 through 6.8 are useful for investigating optimal sparse sampling. For the same purpose, I provide Figures of  $\text{Re PSF}(\vec{r})$  and  $\text{Im PSF}(\vec{r})$ .

Fig. 6.1 shows real and imaginary parts of 2D PSF's for 50 and 49 spiral trajectories. To visualize the structure at lower intensity, the central peaks of  $\text{Re PSF}(\vec{r})$  at  $\vec{k} = \vec{0}$  have been truncated, while the remainders are depicted with equal grey scale. An important

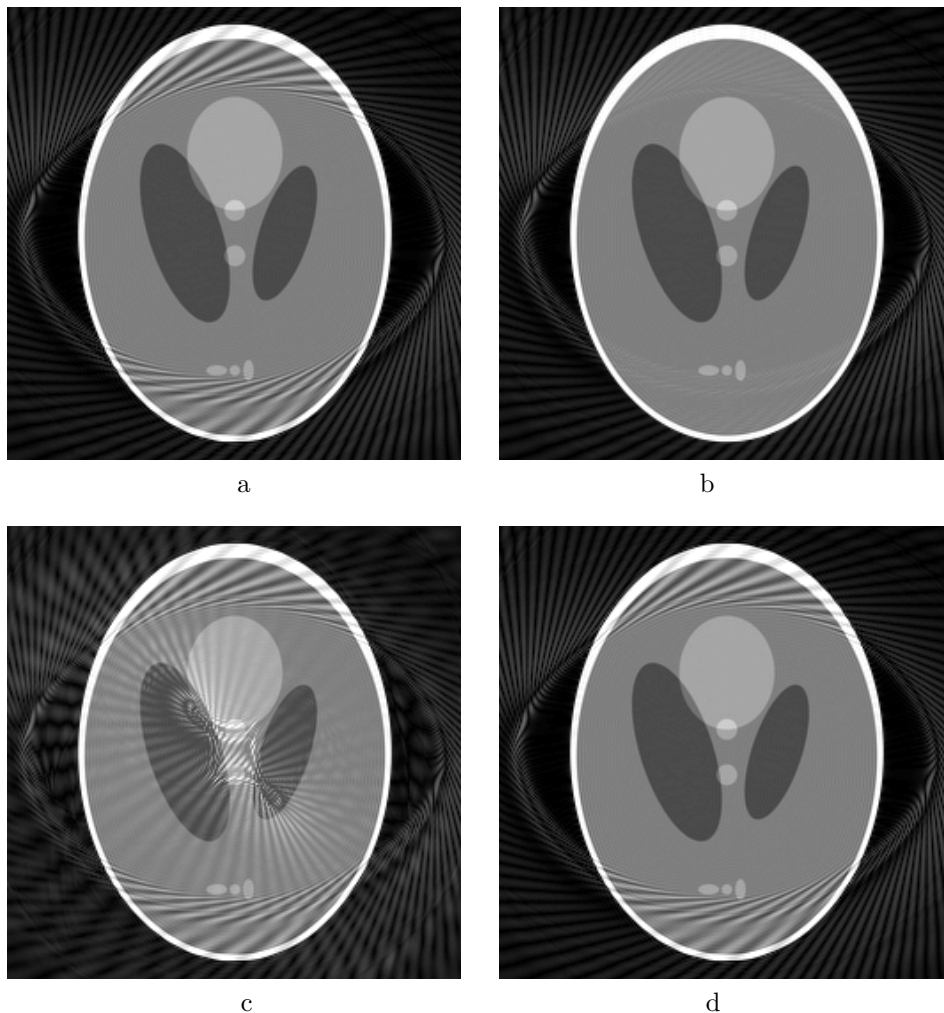


Figure 6.3: Reconstruction from sparse spiral scans, using gridding; the angles between successive spirals are all equal. a) 50 spirals, absolute-value image, b) 49 spirals, absolute-value image, c) 25 spirals, absolute-value image, d) 25 spirals, real part of image.

result is that Re PSF for 49 spirals (b) has significantly fewer artefacts than Re PSF for 50 spirals (a). This will be explained later on in this section. The dominance of the central peaks is apparent from the 1D slices of  $\text{PSF}(\vec{r})$  plotted in Figs. 6.2. Fig. 6.2b shows that the imaginary part of a PSF for an uneven number of spiral trajectories is indeed non-zero; however it is much less intense than the peak of the real part of the same PSF. This phenomenon grows stronger for increasing number of spirals.

I shall now show results of a Shepp-Logan simulation of spiral scanning using 50 and 49 spiral trajectories. I chose the number of 50 (and 49, for that matter) such that the Shepp-Logan object is larger than the spiral FOV. The reconstruction method was gridding. In conformity with the prevailing convention, I display the absolute value of an image,  $|I_{\text{recon}}(\vec{r})|$ . Fig. 6.3 shows the results.

Fig. 6.3a shows  $|I_{\text{recon}}(\vec{r})|$  for 50 spiral trajectories. Two things are important here: 1) Satisfying the Nyquist sampling criterion would require about 60 spirals. 2) 50 is an even number. It can be seen that the quality is strongly degraded by streaks. Next, I decrease the number of spirals to 49 which is an uneven number. The resulting image is shown in Fig. 6.3b. Although the number of spirals has been lowered with respect to Fig. 6.3a, the quality of the image is rather better. This phenomenon can be explained via Eq. 6.8 combined with the fact that the imaginary part of  $I_{\text{recon}}(\vec{r})$  is rather less intense than the real part. Eq. 6.8 implies that  $\text{Re } I_{\text{recon}}(\vec{r})$  equals the sum of two images, one obtained from 49 spirals and the other from 49 spirals obtained by reflection with respect to  $\vec{k} = \vec{0}$ . Note, that the reflected spirals do not form the same set as the original one because their number is uneven. In fact, the sum of the two sets comprises 98 unique spirals with equal angles between neighbours. It follows that  $\text{Re } I_{\text{recon}}^{49}(\vec{r})$  equals  $I_{\text{recon}}^{98}$ ; the superscripts 49 and 98 indicate the number of spirals used in the scan. An image from 98 spirals is sufficiently densely sampled to be artefact-free. Because I display the absolute value of  $I_{\text{recon}}^{49}(\vec{r})$ , one can see some perturbation by the non-zero imaginary part.

Next, I consider the case of  $50/2=25$  spirals. Fig. 6.3c shows the corresponding image  $|I_{\text{recon}}^{25}(\vec{r})|$ ; it is heavily degraded by undersampling and the influence of the imaginary part. In agreement with Eq. 6.8,  $\text{Re } I_{\text{recon}}^{25}(\vec{r})$ , shown in Fig. 6.3d, equals  $|I_{\text{recon}}^{50}(\vec{r})|$  which is shown in Fig. 6.3a.

The above results indicate that although a number of 49 spiral trajectories formally amounts to undersampling, appropriate distribution of these trajectories over  $k$ -space still enables one to retrieve all information required to estimate the image. As a consequence, I recommend that regular sparse spiral scanning (and radial scanning as well) be done with an uneven number of trajectories.

## 6.3 Irregular sparse spiral scanning

### 6.3.1 Introduction

The rationale for varying the angle between successive radial or spiral trajectories stems from sparse Cartesian scanning practice. Marseille found that images resulting from irregular omission of Cartesian trajectories are superior to those resulting from regular omissions [71]. It is natural to investigate whether the same tendency exists for radial or spiral scanning. Following the Cartesian ‘recipe’, one can omit radial or spiral trajectories at irregular positions, leaving the remaining trajectories on their original positions. Evidently, this recipe leads to irregular distributions of trajectories. Hence the name irregular sparse spiral scanning.

The reference distribution in this investigation is a full scan comprising an even num-

ber of trajectories with equal angle between successive trajectories. Next, I omit trajectories pair-wise. The members of a pair are related by point-symmetry with respect to  $\vec{k} = \vec{0}$ . An example is presented in Fig. 6.4. The pair-wise omission avoids entanglement with the point-symmetry effect investigated above. This is necessary to distinguish between the effectiveness of the two approaches.

Below, two feasible methods of selecting an optimal irregular sparse spiral sample distribution are described. The first is to select the best distribution from an appreciable number – *e.g.* thousand – of randomly chosen distributions. The second is a brute-force search over all possible distributions, subject to certain restrictions. I have applied the first method to a real-world spiral scan.

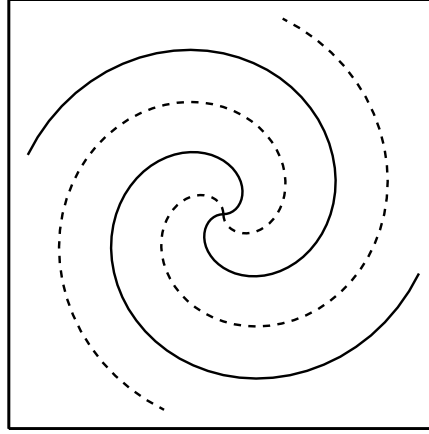


Figure 6.4: Two pairs of spiral trajectories in  $k$ -space, indicated by a drawn line and a dashed line. The members of a pair are related point-symmetry with respect to  $\vec{k} = \vec{0}$ .

The selection criterion is based on the performance measure described in chapter 5. The sparse sample distributions in question being irregular, Bayesian reconstruction is needed.

### 6.3.2 Best of random distributions

This method selects the best sample distribution from a certain number of randomly chosen sample distributions. In this approach one counts on covering a representative subset from the total number of combinations. The result is that we find very bad and good distributions, with all sorts of intermediate cases. This enables one to extract desirable properties a sample distribution must have in order to produce artefact-free reconstructions. The method was applied to an *in vivo* measurement of a human skull. The full scan consisted of 30 pairs of regularly distributed spiral trajectories. From 30 of these pairs, I omitted 10, in 1000 randomly chosen ways. For reconstruction, I used the

Bayesian algorithm described in section 4.4. The distributions were sorted according to the performance defined by Eq. 5.5. Fig. 6.5 shows the result. The performance varies between 17.2 dB and 19.2 dB.

Also shown is the case of omitting every third pair, resulting in a rather ordered distribution. The latter yields a performance of 19 dB, which ranks amongst the best. However, note that the attendant image suffers from the same square artefact as in Fig. 5.2a. (The performance in Fig. 5.2a is 1 dB higher than in Fig. 6.5, because the former underwent an extra treatment after Bayes reconstruction.

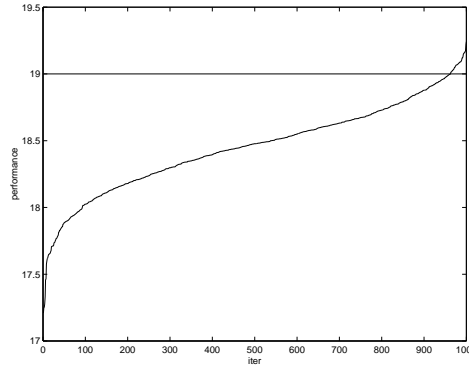


Figure 6.5: Performances obtained with 1000 randomly selected sample distributions, sorted according to increasing value. The extra horizontal line indicates the performance value obtained with ‘equidistant’ omission of every third pair.

The ten worst and the ten best distributions are shown in Fig. 6.6a and b respectively, black and white representing omitted and measured respectively. It can be seen that the best distributions 1) are mostly irregular, but 2) avoid clustering. As an exception, distribution 995 has a regular appearance. The ten worst distributions appear irregular too, but clustering of omissions is significant.

Further investigation awaits availability of more spiral scans of a variety of objects. In addition, it is to be ascertained that the square artefact – mentioned before, and to be discussed in the next chapter – be absent in future scan-time reduction experiments.

### 6.3.3 Brute force search, by minimax

A brute-force search tries *all* possible combinations of omissions of trajectory pairs. This approach guarantees finding the optimal combination. However, the total number of combinations is usually prohibitively large. For example, with a total number of 30 pairs and omitting 10 – resulting in a 30% scan time reduction – the total number of combinations is  $\binom{30}{10} \approx 3 \times 10^7$ . Even if a reconstruction would last only one second, the total calculation time would be nearly a whole year. However, this example is rather optimistic: The number of pairs may be as high as 128 and the reconstruction may take

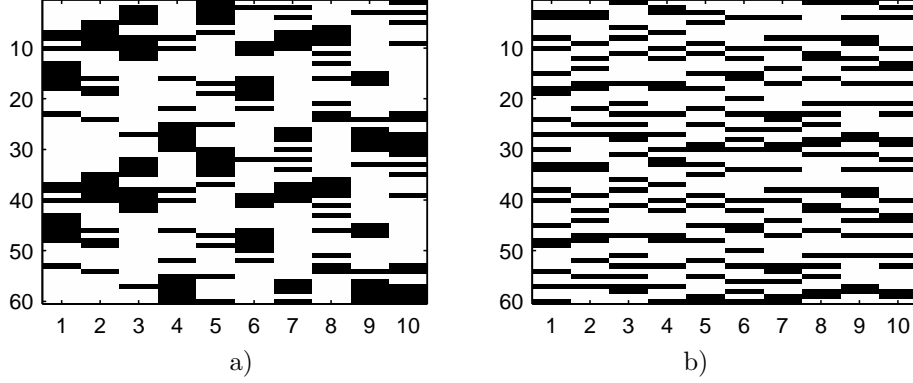


Figure 6.6: Twenty radial sample distributions obtained from a random generator. The black regions indicate omissions. From the set of 1000, this Figure shows those for the a) 10 worst performances, b) 10 best performances.

several minutes, resulting in an unrealistic search method. Under the circumstances, one must resort to the minimax method [71].

The minimax search strategy starts by considering all measured trajectory pairs. From these, it omits the pair that contributes least to the performance. Using the numbers above, this first step requires 30 reconstructions. From the remaining pairs, one then again omits only the one that contributes least to the performance, etc. Note that omitted pairs are not reconsidered. The procedure is to be pursued till the number of omitted pairs reaches 10. The total number of reconstructions required amounts to  $30 + 29 + 28 + \dots + 21 = 255$ . Since omitted pairs are rejected for good, this search strategy can only be suboptimal.

Application of the minimax method awaits the availability of a set of spiral scans of different objects.

## 6.4 Summary and conclusions

This chapter concerned mainly sparse spiral scanning. I have made a distinction between regular and irregular distributions of trajectories, defined as follows.

In a regular distribution, all angles between successive trajectories are equal.

An irregular distribution is derived from a full – Nyquist satisfied everywhere – regular distribution by irregular omission of trajectories and fixing the remaining trajectories to their original positions.

For sparse regular spiral distributions, I conclude

- $2N - 1$  trajectories yield better image quality than do  $2N$  trajectories. ( $N$  is integer.)

- It is advisable to routinely scan with an uneven number of trajectories, irrespective of exploitation of point-symmetry with respect to  $\vec{k} = \vec{0}$ .
- Image reconstruction can be done with gridding.

For sparse irregular spiral distributions, I conclude

- Irregular omission yields better image quality than does regular omission.
- Omissions should not be strongly clustered.
- Image reconstruction needs Bayesian estimation.



## Chapter 7

# Applications

### 7.1 Introduction

In this chapter, I apply the 2D Bayesian reconstruction program described in section 4.4.

My predecessor G.J. Marseille has successfully devised a Bayesian reconstruction algorithm for sparse 2D Cartesian scans [71]. As explained in chapter 2 of the present thesis, Cartesian sparsity pertains to a single dimension. My task was to expand the Bayesian methodology to sparse 2D *spiral* scanning which requires a 2D approach; see chapter 2.

Aiming at more general applications than just spiral scanning, I have adapted the algorithm to arbitrary 2D sample positions. Thus, the algorithm starts by reading a list of complex-valued samples and corresponding  $\vec{k}$ 's; the values of these vectors can be arbitrary. The algorithm checks whether the reconstructed image agrees with the available samples, wherever their positions. Moreover, the procedure obviates the notion of sampling density. In the case of undersampling, the adopted 2D prior should warrant a correct image. These properties need testing.

I distinguish two test phases. In phase 1, one tests image reconstruction with simulated  $k$ -space data. A simulation can be made realistic by ‘reverse gridding’ of a real-world Cartesian full scan, as pointed out in chapter 3. An advantage of simulations is that the true image is available for comparison, preventing doubts about the quality of the result. In phase 2, one tests image reconstruction with raw data coming straight from a clinical scanner. In the case of non-Cartesian scanning, *e.g.* radial or spiral, the algorithm has to cope with deficiencies of the signal at the starting point  $\vec{k} = \vec{0}$  of the trajectories. The latter test may be very hard.

Before proceeding, I mention use of the algorithm by others – in the context of TMR/Networks cooperation. Bourgeois has applied 2D Bayesian reconstruction to motion-corrected  $k$ -space data [40, 94–97], and Reczko did the same for sparse spiral data [98, 99]. Martinez has extended the 2D algorithm to 3D and has applied it to arbitrarily sampled data and real-world animal radial scans [100, 101]. Lethmate has applied both the 2D and 3D versions to real-world radial scans of various objects [102, 103].

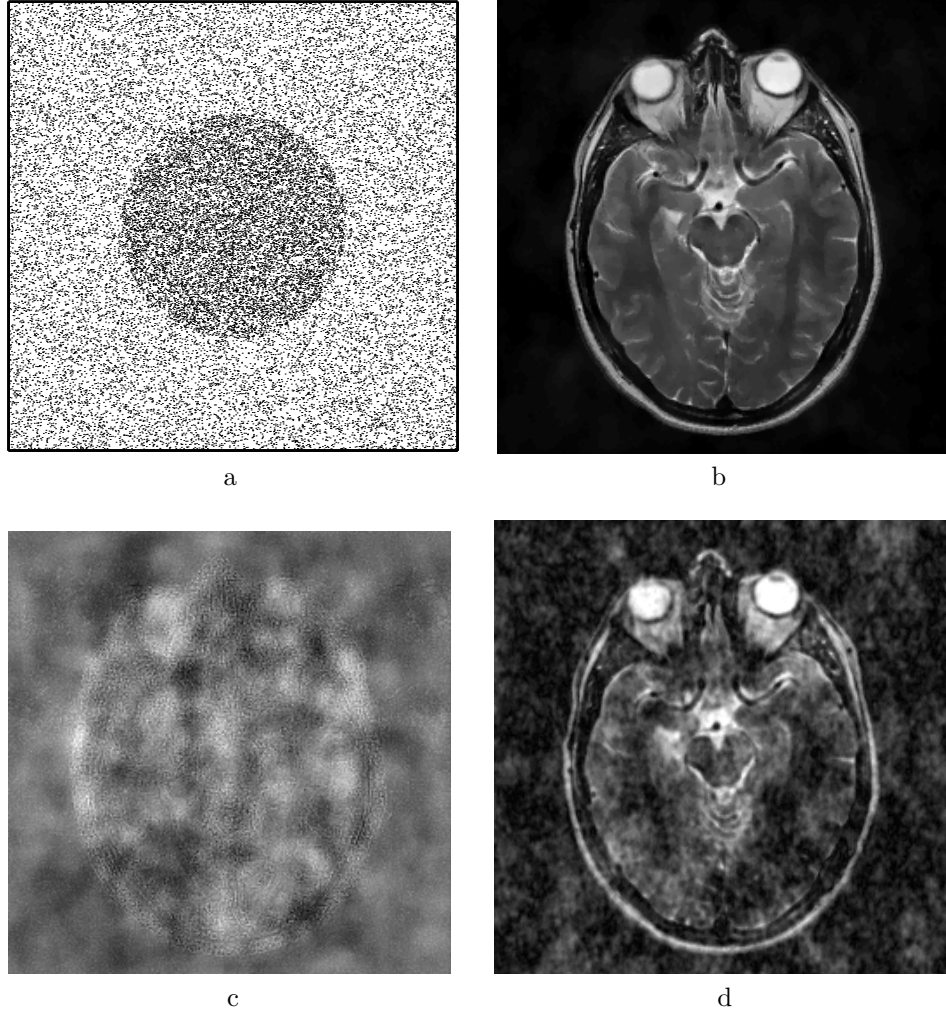


Figure 7.1: Reconstruction from sparse randomly positioned samples, generated by reverse gridding of a Cartesian brain scan. a) Random sample positions. b) Bayesian reconstruction. The hyperparameters  $\sigma_S$  and  $a$  were set according to section 4.4.6. c) Difference between b) and the inverse FFT of the original Cartesian data. d) Gridding reconstruction.

## 7.2 Simulations

### 7.2.1 Random scan

In order to investigate whether the algorithm can cope with both undersampling and arbitrary sample positions, I generated randomly sampled scan data by reverse gridding of Cartesian human brain scan data. Fig. 7.1a displays the sample positions. In a circular central region of  $k$ -space,  $|k| \leq 0.25$ , the average density of sample positions satisfies the Nyquist criterion, but beyond  $|k| = 0.25$  I reduced the average density to 40% of this value. Results obtained with this sample distribution are displayed in Fig. 7.1b,c,d, and described below. I assume that Bayesian reconstruction can handle sparse Cartesian, radial, and spiral sample distributions equally well.

Fig. 7.1b shows the successful Bayesian reconstruction [104]. For reference, Fig. 7.1c shows the difference with respect to the FFT of the original Cartesian data. The residue is seen to be small. This indicates that the Bayesian algorithm can cope with random sample positions that do not satisfy – not even on average – the Nyquist sampling rule.

Fig. 7.1d shows the result of gridding. Visual inspection confirms immediately what was to be expected: Gridding cannot cope well with this case, because

1. The density compensation for sparse random sample positions is problematic.
2. Prior knowledge cannot be used, while undersampling makes it indispensable.

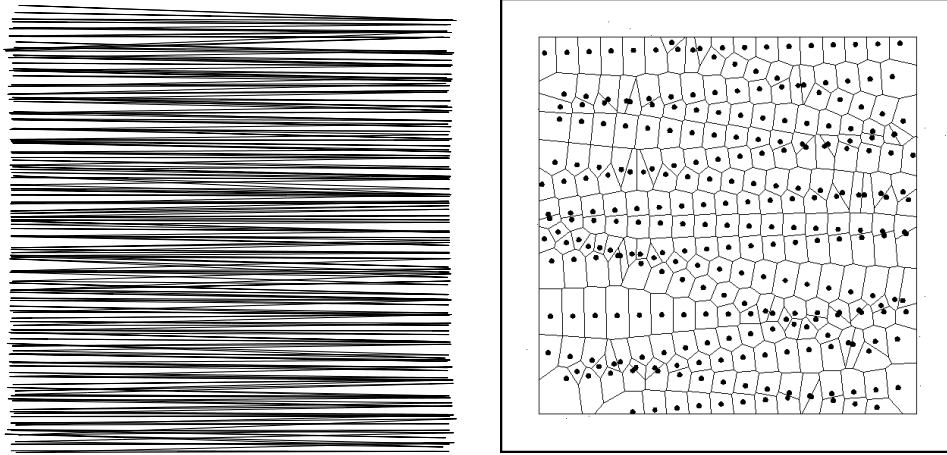


Figure 7.2: Effect of intrascan patient-rotation on a Cartesian scan. a) Full 2D scan. b) Details of the perturbed sample positions, with corresponding Voronoi polygons. [96]

### 7.2.2 Intra-scan rotation

The performance of the Bayesian algorithm for arbitrary sparse sample distributions proves very useful for handling motion-corrected data in functional MRI, abbreviated as fMRI [40, 42, 94–97]. In fMRI, one wishes to detect *small differences* in image intensity induced by some function. Motion artefacts hamper detection of the fMRI effect. Hence the need for motion-correction.

The present application pertains to motion that takes place between measuring successive trajectories. The measurement time of a trajectory is assumed sufficiently small to assume patient-immobility during its course. Translation of the object (image) results in a phase-change in  $k$ -space while trajectories retain their directions. A rotation of the object (or image), however, amounts to a rotation of  $k$ -space and therefore the trajectories will rotate along. Fig. 7.2 shows the effect of intrascan-rotation on a Cartesian scan. Each trajectory is rotated by a ‘random’ patient rotation, which in turn causes under-sampled regions of  $k$ -space. If ignored prior to applying the reconstruction algorithm, these rotations inevitably result in image artefacts.

Bourgeois has developed the estimation of true sample positions from the raw data; see [96, 97] for details. As an illustration of the role of Bayesian reconstruction, I reproduce briefly his simulation of intrascan motion-correction with Shepp-Logan data; see Fig. 7.3. Fig. 7.3a shows the image obtained by 2D-FFT, ignoring the simulated intrascan rotations applied to the dataset. Attendant artefacts are clearly visible. Fig. 7.3c shows the signal after estimation of the intrascan rotations. Subsequent gridding applied to these data did not produce a satisfactory result. Next, the Bayes algorithm was applied; see Fig. 7.3b. The quality of this image proves that both the motion-correction and the subsequent Bayes reconstruction were successful. Finally, Fig. 7.3d is the reconstruction of the original motion-free  $k$ -space signal.

## 7.3 Real-world spiral scans

### 7.3.1 Preamble

The previous section dealt with image reconstruction from simulated signals. Although simulations are indispensable for testing reconstruction algorithms, it is difficult to include *in vivo* conditions or scanner imperfections. This section deals with Bayes reconstructions from real-world spiral scans. First, I report experimental conditions that gave rise to unacceptable artefacts in Bayes reconstructions from sparse scans.

### 7.3.2 Problems at $\vec{k} = \vec{0}$

As mentioned in chapter 2, spiral trajectories start at  $\vec{k} = \vec{0}$  where the signal is strongest. In addition, the sampling density for small  $|\vec{k}|$  is high. As a result, spiral scanning is less sensitive to object motion than is Cartesian scanning. A disadvantage of spiral scanning

is uncertainty about the initial part of the trajectory; this has to do with the finite build-up time of the field gradient. Since the signal is strongest at  $\vec{k} = \vec{0}$ , uncertainty about the exact sample positions around and at  $\vec{k} = \vec{0}$  can have serious consequences for image reconstruction.

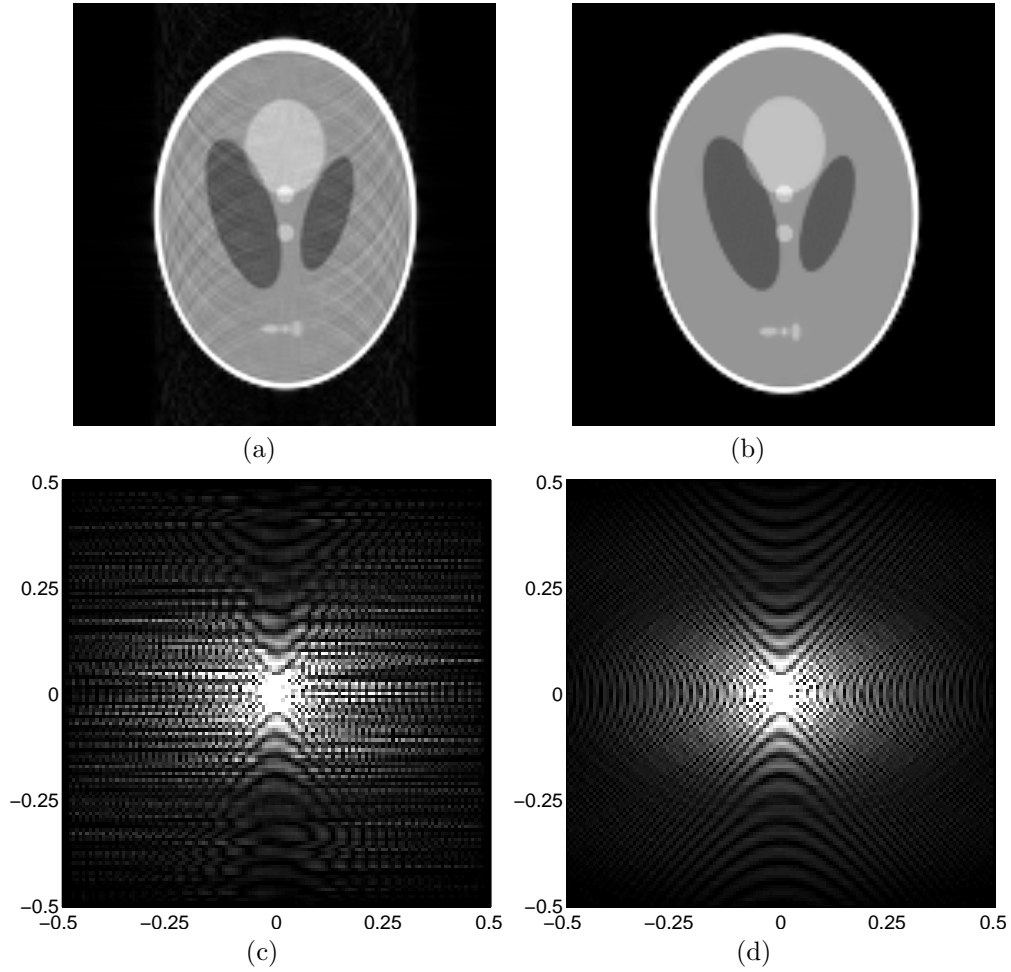


Figure 7.3: Correction for intrascan rotations using a Shepp-Logan simulation [42,96,97]. Trajectories were rotated over random angles having a Gaussian distribution ( $\mu = 0.0, \sigma = 1.0$  degree): (a) Distorted image reconstructed using 2D IFFT and perturbed by motion artefacts. (b) Bayesian reconstruction of the image. (c) The  $k$ -space data with rotation artefacts. (d) The  $k$ -space data after Bayesian reconstruction.

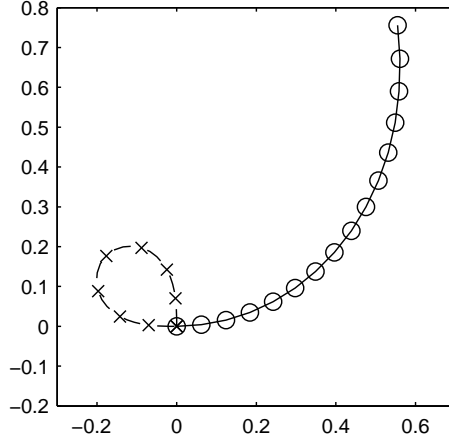


Figure 7.4: Complete spiral trajectory, starting with a small loop – dashed line. Sampling starts when the trajectory is at  $\vec{k} = \vec{0}$  for the second time.

In order to circumvent deficiencies at the start of sampling, the trajectories first make a small loop and return to  $\vec{k} = \vec{0}$  before taking off on their true course; see Fig. 7.4. During this initial loop, no samples are taken. In fact, sampling starts when the trajectory supposedly is at  $\vec{k} = \vec{0}$  for the second time. Note that at  $\vec{k} = \vec{0}$ , the phase of the signal is zero by definition. To verify this property in practice, I display the phases of the first fifteen samples on all trajectories in Fig. 7.5a. It can be seen that the desired effect of the loop does not take place: At index = 0, the phases have not nearly coincided with each other. Back-extrapolation of the lines reveals that this occurs at index  $\approx -4$ . Fig. 7.5b displays the absolute values of the first fifteen samples of all trajectories. Again, the curves do not coincide at the assumed  $\vec{k} = 0$ .

### 7.3.3 Bayes reconstruction from sparse scans

This section pertains to reduction of the scan time by means of omission of trajectories [72]. The number of spiral trajectories required to satisfy the Nyquist criterion depends on the curvature of the trajectories. Going from  $\vec{k} = \vec{0}$  to the edge of  $k$ -space, the direction of our trajectories changes by  $4\pi$ . At this rate of change, one needs about 60 trajectories to avoid undersampling throughout  $k$ -space. This already constitutes a substantial reduction of measurement time with respect to Cartesian and radial scanning. Our aim is to reduce the measurement time further, to  $2/3$ , by omission of 20 trajectories. The omission was done pairwise so as to avoid intermixing of possible point-symmetry effects – see chapter 6.

I chose the following three omission approaches.

- a. The remaining 40 trajectories are evenly spread over  $k$ -space.

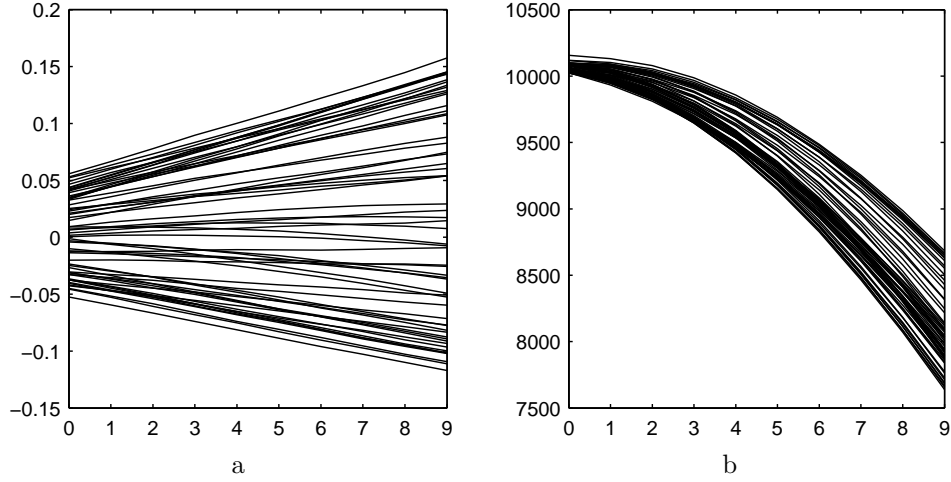


Figure 7.5: Observed properties of the first fifteen samples of all trajectories. Horizontal: index of sample. a) Phases, in units of radians. At index = 0, all phases should be zero. They are not. Extrapolation revealed coincidence at index  $\approx -4$ . b) Amplitudes, in arbitrary units.

- b. Every third trajectory of the full set of 60 is omitted. The remaining 40 retain their positions.
- c. Trajectories are omitted at random from the full set of 60. The remaining 40 retain their positions. This was repeated with 1000 different random omissions; see chapter 6.

For testing, I used a full spiral scan of a human brain made up of 60 trajectories. Fig. 7.6 displays Bayesian reconstructions obtained with the three approaches. As explained in chapter 6, approach (a) causes a reduction of the FOV and hence suffers from the same artefacts as shown in Fig. 6.3a; see also [105]. Apparently, the prior knowledge described in section 4.4.3.4 does not suffice to suppress artefacts of this type.

Fig. 7.6b displays the Bayes reconstruction for approach (b). A square artefact is clearly visible. The cause of this artefact was ascribed to imperfections of the measurements at  $\vec{k} = \vec{0}$ , described in section 7.3.2. The size of the square increases with the number of regularly omitted spirals. With simulations, no such artefact appeared. It follows that approach (b) will be applicable if experimental conditions could be improved. Given the uncertainty around spiral imaging at the point of time of this work, further investigation of the properties of the artefact was abandoned.

Fig. 7.6c displays the Bayesian reconstruction for random omission as described in section 6.3. The best-performing random omission, displayed in Fig. 7.6d, has been chosen. Random omission is optimal for Cartesian scanning [71]. Note that the square

artefact has disappeared. This indicates that the effect of the experimental imperfections at  $\vec{k} = \vec{0}$  is averaged out when spreading the omissions randomly. Further tests await the availability of new spiral scans.

## 7.4 Summary and conclusions

This chapter concerned applications of the 2D Bayes reconstruction algorithm developed in chapter 4. The desired property of the algorithm is that it be capable of handling sparse arbitrary sampling. This includes obviating the sampling density compensation. If sparse arbitrary sampling can be handled, the same capability should be expected for sparse spiral sampling. I tested the 2D Bayes reconstruction algorithm with sparse simulated data and sparse real-world spiral scans.

### Simulations

#### Bayes reconstruction

- The performance for sparse randomly sampled data is high. This indicates that an undersampled scan, made up of any currently accepted trajectory shape, can be accommodated.
- The density compensation is indeed obviated.
- The computation time is long compared to gridding.

#### Gridding

- Gridding performs badly with sparse randomly sampled data. This remains true when using the Pipe-Menon density compensation. As a result, I do not recommend gridding for sparse spiral scans.

### Real-world spiral scans

- The 2D Bayes reconstruction algorithm is sensitive to a deficiency of our real-world spiral scan for small  $|\vec{k}|$ . This leads to square artefacts when omitting trajectories in regular fashion. The nature of the deficiency is not understood, nor is the appearance of the artefact.
- Further investigations await new measurements.



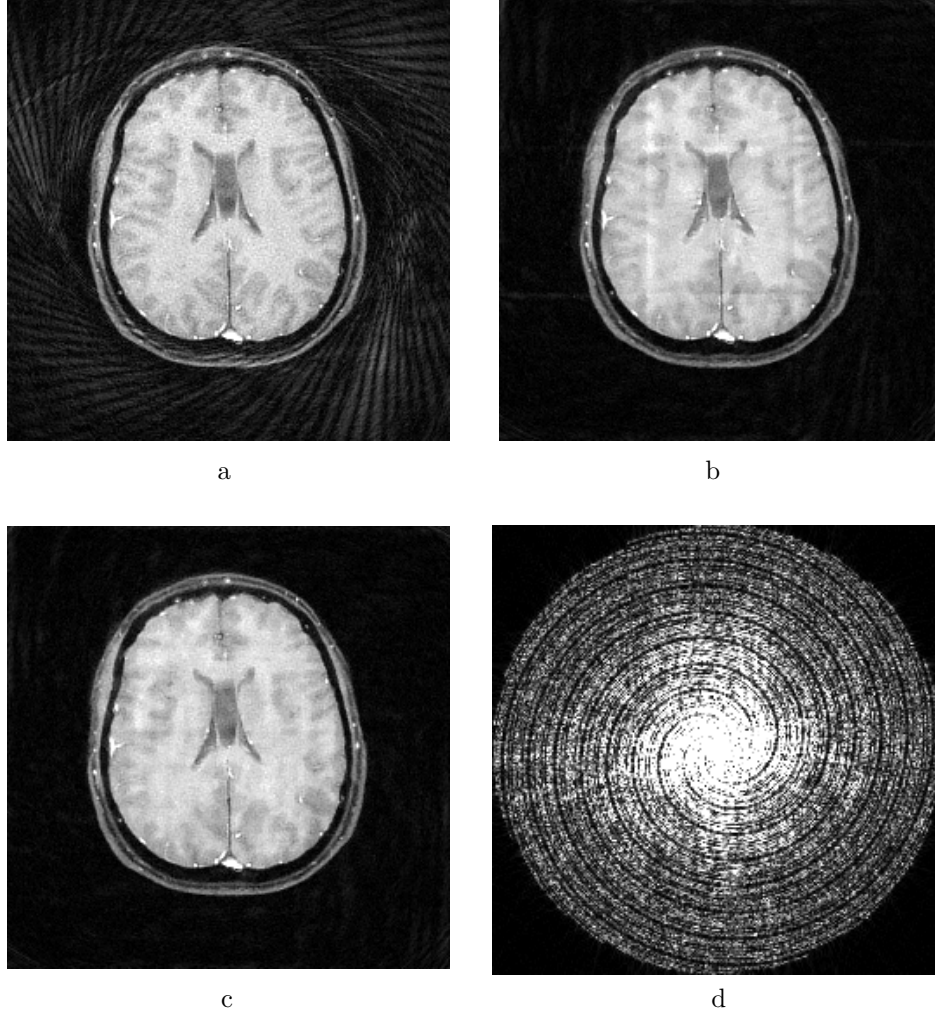


Figure 7.6: Bayesian reconstruction from sparse spiral data for three alternative approaches to omitting trajectories that correspond to 33% reduction of measurement time. (a) The remaining trajectories are evenly spread over  $k$ -space. (b) Every third trajectory of the full set is omitted. The remaining ones remain in place. (c) Trajectories are omitted at random from the full set, the remaining ones remaining in place. The best-performing sparse set of Fig. 6.5 was chosen. (d) Best-performing sparse data set pertaining to (c).



## Chapter 8

# Conclusions and Future Work

### 8.1 Preamble

At the outset, the theme of this thesis was image reconstruction from raw spiral scan data. In particular, it was to be investigated whether scan time reduction through smart omission of spiral trajectories is feasible. To that end, I concentrated on two algorithms: 1) Gridding, 2) Bayesian reconstruction of images from 2D sparse spiral  $k$ -space data.

Initially, the gridding algorithm merely served to produce an image from complete, non-Cartesian  $k$ -space data. The resulting image, in turn, served as reference image for reconstructions from sparse spiral  $k$ -space data. The challenge of the Bayesian reconstruction algorithm was to expand the existing 1D version [71] to 2D and beyond. In addition, Gaussian edge-preserving filtering was investigated.

As time progressed, our industrial partner abandoned activities on spiral scanning. As a result my emphasis on spiral scanning diminished. Instead, I turned to analysis and improvement of gridding, and on alternative non-Cartesian sampling schemes such as radial and random. Radial scanning is applied when the decay of the signal is fast. Random scanning is unintentional, for example, it can be caused by correcting Cartesian sample positions for random patient motion.

First, I summarize and conclude first in general terms. Subsequently, gridding and Bayesian reconstruction are reviewed in some more detail. Finally, I suggest ideas for future research.

### 8.2 General Overview

The first two chapters serve to inform signal-processing readers about some aspects of MRI scanning. The main message is that sampling is done along trajectories in  $k$ -space. The duration of a trajectory is not influenced by the number of samples taken during its course. Between successive trajectories, the scanner waits for (some degree of) thermal re-equilibrium of perturbed magnetic moments. The scan time is determined by the number of trajectories needed to sufficiently cover  $k$ -space.

When the number of samples and their positions in  $k$ -space enable complete reconstruction of the image without use of prior knowledge, I speak of *complete* sampling. Chapter 3 is devoted to image reconstruction – especially gridding – for this case.

Image reconstruction from *incomplete* sampling is impossible unless one adds alternative information (prior knowledge). An algorithm that is capable of taking alternative information into account is the above-mentioned Bayesian reconstruction. This is treated in chapter 4.

Criteria for judging the quality of an image reconstruction constitute a source concern. So long as testing is done with simulations, such as the Shepp-Logan case, a non-medical person should be able to judge the result visually by him/herself. But already, replacement of visual judgment by an algorithm that automatically produces a single number indicative of quality, poses problems. In any case, judgment of reconstruction of a *real-world* image remains very difficult for a non-medical person. It is not uncommon that visual judgment does not agree with the single number-judgment produced by computer. These aspects are treated in chapter 5.

Whereas chapter 4 concerned reconstruction from sparse arbitrarily sampled  $k$ -space data, chapter 6 concerned design criteria for sparse sampling schemes. The treatment is based on the so-called point-spread function. I conclude that for spiral scanning the number of trajectories be *uneven*.

Finally, chapter 7 is devoted to application of the 2D Bayesian reconstruction algorithm described in chapter 4. The main conclusion is that the algorithm can handle arbitrary sparse sampling distributions while obviating the sampling density compensation. At the same time, the algorithm is perturbed by imperfections of real-world spiral scans at the origin of  $k$ -space.

## 8.3 Gridding

This section summarizes gridding, as treated in chapter 3.

- The main algorithm for image reconstruction from complete non-uniform scans is gridding. Algorithms dedicated to special non-uniform sampling schemes – *e.g.* chirp- $z$  for linogram scans – seem no longer required.
- Gridding comprises three steps:
  1. Sampling density compensation
  2. Convolution-interpolation
  3. IFFT
  4. Roll-off correction
- In the case of irregular sampling, the sampling density compensation should preferably be estimated with the method of Pipe and Menon [48]. This method forces the point-spread function to approach a delta function. Yet, as the sample positions become more random, the sampling density compensation becomes the dominant source of error of gridding.

- The accuracy of convolution-interpolation can be set according to needs, using a Kaiser-Bessel kernel. The following three parameters are important.
  1. The width  $L$  of the Kaiser-Bessel kernel.
  2. The overgridding (resampling) factor  $f_{\text{og}}$ . Given  $L$  and  $f_{\text{og}}$ , the accuracy and relative computation time can easily be derived with the theory of chapter 3.
  3. The shape parameter  $B$  of the Kaiser-Bessel kernel. I derived an expression for the optimal value of  $B$ , given  $L$  and  $f_{\text{og}}$ .

## 8.4 Bayesian Reconstruction

This section summarizes 2D Bayesian reconstruction, as treated in chapter 4.

- In the case of incomplete sampling, prior knowledge and or regularization are indispensable. The Bayesian approach facilitates accommodation of prior knowledge. The function to be minimized – the posterior – is the sum of the likelihood term and the prior.
- The Bayesian approach obviates the density compensation. In other words, the likelihood term is not weighted. This implies that arbitrary sample positions are handled without additional effort, which in turn constitutes a decisive practical advantage. Inclusion of weights may accelerate convergence.
- The way of computing the likelihood term saves considerable time. The same principle can be applied to image reconstruction for SENSE, as described in Appendix A.
- The prior is a 2D Lorentzian distribution of nearest-neighbour pixel-intensity differences. For the sake of simplicity, the distribution is isotropic in  $x$  and  $y$ . The width of the distribution is derived from the noise level.

## 8.5 Suggestions for further research

The project enabled me to make some contributions to reconstruction of images from raw, non-uniformly sampled  $k$ -space data. Meanwhile, MRI scanners and their applications evolved and perennially posed new challenges. Below, I mention some subjects that presently need attention.

### Complete sampling

Gridding can now handle all sample positions, provided the latter are sufficiently densely distributed. Points of concern are:

- When the sample positions are irregular, the sampling density compensation – which is part of gridding – becomes inaccurate. This aspect, and the attendant computational load, need further investigation.

- The positions of initial samples on trajectories that start at  $\vec{k} = \vec{0}$  may be inaccurate. Methods to correct these positions ought to be devised.

#### Incomplete sampling

In spite of its long computation time, Bayesian reconstruction seems to be the method of choice for sparse irregular sampling, because it obviates the sampling density compensation. The 2D version is described in this thesis. A 3D version is in progress [100]. I recommend the following continuation of activities on Bayesian reconstruction.

- Presently, the prior is an isotropic multidimensional Lorentzian distribution of nearest neighbour intensity differences. The need for anisotropic distributions should be investigated.
- In my implementation, the chosen width of the distribution has no bearing on the actual width following from a complete scan. Investigation is desired into
  - Automatic estimation of an optimum width for accuracy of reconstruction.
  - Influence of the width on the convergence speed.
- Weighting of the likelihood term may accelerate the convergence. Possible weights are the coefficients of the density compensation. In any case, this would yield a near-gridding result at the first iteration.
- Signal imperfections at the origin of  $k$ -space have a detrimental effect on image quality. The origin of this effect ought to be established, once new spiral scans become available.

As for sparse spiral imaging, I recommend investigation of using an uneven number of trajectories.

# Appendix A

## Major Speedup of Reconstruction for Sensitivity Encoding with Arbitrary Trajectories

Proceedings of the ISMRM, p. 767, 2001

F.T.A.W. Wajer, K.P. Pruessmann\*

*Spin Imaging Group, Faculty of Applied Physics, Delft University of Technology, Delft,  
The Netherlands*

*\*Institute of Biomedical Engineering, University and ETH Zürich, Zürich, Switzerland*

### A.1 Introduction

Sensitivity encoding (SENSE) with receiver coil arrays was recently proposed for enhancing imaging speed in MRI [106]. While relatively easy with Cartesian sampling, SENSE has been difficult with general  $k$ -space trajectories due to enhanced complexity of the reconstruction. For efficient SENSE reconstruction from arbitrary  $k$ -space trajectories, several iterative reconstruction algorithms have recently been proposed [107–109]. As a common feature, in each iteration step these methods require matrix-vector multiplications, which largely determine the reconstruction speed. In this work we propose a novel implementation of the specific matrix-vector multiplications that occur in SENSE reconstruction, enabling substantial further increases in reconstruction efficiency.

### A.2 Method

SENSE reconstruction from arbitrary  $k$ -space trajectories was formally described in [106]. By neglecting noise statistics or decorrelating receiver noise as proposed in [109], the reconstruction problem reduces to

$$\mathbf{E}^H \mathbf{E} \mathbf{I} = \mathbf{E}^H \mathbf{m}, \quad (\text{A.1})$$

where  $E_{(\gamma, \kappa), \rho} = s_{\gamma}(\mathbf{r}_{\rho}) e^{i \mathbf{k}_{\kappa} \cdot \mathbf{r}_{\rho}}$ , with  $s_{\gamma}(\mathbf{r}_{\rho})$  the sensitivity of coil  $\gamma$  at position  $\mathbf{r}_{\rho}$ .  $\rho$  indexes the voxels and  $\kappa$  the  $k$ -space sampling positions, respectively.  $\mathbf{I}$  is the image to be reconstructed from the measured  $k$ -space data  $\mathbf{m}$ .

In iterative solving of Eq. A.1, the key procedures are the calculations of  $E^H \mathbf{m}$ , which is constant during iteration, and  $E^H E \mathbf{I}^{(n)}$ , with  $\mathbf{I}^{(n)}$  the reconstructed image at iteration  $n$ .

As pointed out in [107, 109],  $E^H \mathbf{m}$  can be written as

$$[E^H \mathbf{m}]_\rho = \sum_{\gamma} s_{\gamma}^*(\mathbf{r}_{\rho}) \left( \sum_{\kappa} e^{-i\mathbf{k}_{\kappa} \cdot \mathbf{r}_{\rho}} m_{\gamma, \kappa} \right), \quad (\text{A.2})$$

where the last sum can be evaluated efficiently using gridding.

We will now turn attention to the evaluation of  $E^H E \mathbf{I}^{(n)}$ . In [109]  $E^H E \mathbf{I}^{(n)}$  is evaluated in two steps as  $E^H (E \mathbf{I}^{(n)})$ , which can be accomplished by two gridding steps (per coil). We will now show how to remove the two gridding interpolation steps, and since the calculation time is heavily determined by these two steps we get a major speedup.

Instead of using  $E^H E \mathbf{I}^{(n)} = E^H (E \mathbf{I}^{(n)})$  we examine the components of  $E^H E$ ,

$$[E^H E]_{\rho, \rho'} = \sum_{\gamma} s_{\gamma}^*(\mathbf{r}_{\rho}) \left( \sum_{\kappa} e^{-i\mathbf{k}_{\kappa} \cdot (\mathbf{r}_{\rho} - \mathbf{r}_{\rho'})} \right) s_{\gamma}(\mathbf{r}_{\rho'}) \quad (\text{A.3})$$

$$= \sum_{\gamma} s_{\gamma}^*(\mathbf{r}_{\rho}) Q(\mathbf{r}_{\rho} - \mathbf{r}_{\rho'}) s_{\gamma}(\mathbf{r}_{\rho'}), \quad (\text{A.4})$$

with

$$Q(\mathbf{r}) = \sum_{\kappa} e^{-i\mathbf{k}_{\kappa} \cdot \mathbf{r}}. \quad (\text{A.5})$$

Note the similarity between the summation in Eq. A.5 and the term in parentheses in Eq. A.2. They can be evaluated at the same time prior to starting the iterations using gridding. The only thing we need to take care about is that the size of  $Q$  is twice the size of  $\mathbf{I}^{(n)}$ , since  $\mathbf{r}_{\rho} - \mathbf{r}_{\rho'}$  has twice the range of  $\mathbf{r}_{\rho}$ .

The elements of  $E^H E \mathbf{I}^{(n)}$  are now evaluated as

$$[E^H E \mathbf{I}^{(n)}]_{\rho} = \sum_{\gamma} s_{\gamma}^*(\mathbf{r}_{\rho}) \left( \sum_{\rho'} Q(\mathbf{r}_{\rho} - \mathbf{r}_{\rho'}) s_{\gamma}(\mathbf{r}_{\rho'}) I_{\rho'}^{(n)} \right) \quad (\text{A.6})$$

$$= \sum_{\gamma} s_{\gamma}^*(\mathbf{r}_{\rho}) \left[ Q * (s_{\gamma}(\mathbf{r}_{\rho'}) I_{\rho'}^{(n)}) \right]_{\rho}, \quad (\text{A.7})$$

where the  $*$  in Eq. A.7 stands for a discrete convolution [72]. This convolution can be efficiently evaluated using the FFT [49], *i.e.*  $Q * (s_{\gamma}(\mathbf{r}_{\rho'}) I_{\rho'}^{(n)}) = \text{IFFT}(\text{FFT}(Q) \cdot \text{FFT}(s_{\gamma}(\mathbf{r}_{\rho'}) I_{\rho'}^{(n)}))$ . Instead of storing  $Q$ ,  $\text{FFT}(Q)$  is stored, which leads to  $Q * (s_{\gamma}(\mathbf{r}_{\rho'}) I_{\rho'}^{(n)})$  being evaluated using only two FFT's. The size of these FFT's is equal to the size of  $Q$ , requiring  $s_{\gamma}(\mathbf{r}_{\rho'}) I_{\rho'}^{(n)}$  to be zero-filled before applying the FFT. This size is the same as used in [109], so there is no increase in calculation time of the FFT's. Therefore, the gain is that we have replaced two gridding interpolations (per coil) by a coil-wise multiplication in  $k$ -space by  $\text{FFT}(Q)$ , which can be evaluated much faster.

Finally, including a diagonal preconditioning matrix  $D$  into the algorithm, as suggested in [109], *i.e.*  $E^H D E \mathbf{I} = E^H D \mathbf{m}$ , presents no problem and does not increase



computation time. In Eq. A.2 we include an extra term  $D_{\kappa,\kappa}$  and the definition of  $Q$  changes to

$$Q(\mathbf{r}_\rho) = \sum_{\kappa} D_{\kappa,\kappa} e^{-i\mathbf{k}_\kappa \cdot \mathbf{r}_\rho}. \quad (\text{A.8})$$

### A.3 Results

The proposed convolution approach was implemented in the iteration scheme described in [109]. Its performance was compared with that of the original method using explicit gridding. Single-shot spiral acquisition with radial density reduced to 40% of the Nyquist limit ( $R=2.5$ ) was simulated using an array of six circular receiver coils surrounding a cylindrical phantom. All calculations were performed on a 400 MHz DEC Alpha workstation using PV Wave (Visual Numerics Inc., Houston, Texas).

The speed benefit of the new approach was assessed for typical image sizes. The results are summarized in Tab. A.1, proving a substantial speed advantage of the convolution method. As expected, except for negligible numerical error, the novel implementation yielded exactly the same image progressions as did the original method.

Table A.1: Comparison of CPU times per iteration.

Image size	$64 \times 64$	$128 \times 128$	$256 \times 256$
Convolution	1.1 s	2.7 s	12 s
Gridding	4.5 s	17 s	68 s

### A.4 Conclusions

The novel method is really fast (gain of  $3 - 6\times$ ) and makes SENSE in combination with spiral even more practical. The actual speed benefit is greatest for low reduction factors, because gridding is most costly with dense  $k$ -space sampling.

### Acknowledgements

This work is supported by the Dutch Technology Foundation (STW).



## Appendix B

# On the use of Gaussian prior knowledge

This appendix is a short explanation of why Gaussian-type prior knowledge is not able to estimate missing  $k$ -space datapoints in case the  $k$ -space measurement noise is uncorrelated and Gaussian-distributed. The derivation is for the 1D Cartesian case. Extension to the two- or more-dimensional case is straightforward. The non-Cartesian case is included by taking the Cartesian grid in the following explanation fine enough.

I start by setting up a model of the measurement for the 1D Cartesian sampling case. Since  $k$ -space is the Fourier transform of the image, I write

$$\underline{\mathbf{s}} = M F \underline{\mathbf{i}} + \underline{\mathbf{n}}, \quad (\text{B.1})$$

where  $\underline{\mathbf{i}}$  is the image,  $\underline{\mathbf{s}}$  the  $k$ -space data at measured locations,  $\underline{\mathbf{n}}$  is the noise contribution.  $F$  is the unitary FFT matrix and  $M$  is a matrix selecting the appropriate columns of  $F \underline{\mathbf{i}}$ , where  $\underline{\mathbf{s}}$  is sampled. Each row of  $M$  contains only zeros, except for one element which is 1.

Assuming Gaussian-type prior knowledge, image reconstruction amounts to a minimization of

$$\frac{1}{\sigma_S^2} |\underline{\mathbf{s}} - M F \underline{\mathbf{i}}|^2 + \frac{\lambda}{\sigma_S^2} |D \underline{\mathbf{i}}|^2, \quad (\text{B.2})$$

in which  $D$  is a circulant matrix.  $D$  is circulant, because all pixels in  $\underline{\mathbf{i}}$  are assumed to obey the same prior. A well-known example for  $D$  is the unit matrix, for minimum norm estimation. Another example is  $D$  having ones on the diagonal and minus ones on the first subdiagonal. In this case the prior knowledge amounts to the gradient of the image having a Gaussian distribution.

The reconstructed image can now be found by setting the gradient of Eq. B.2 to zero, *i.e.*

$$-\frac{F^\dagger M^\dagger}{\sigma_S^2} (\underline{\mathbf{s}} - M F \underline{\mathbf{i}}_{\text{recon}}) + \frac{\lambda}{\sigma_S^2} D^\dagger D \underline{\mathbf{i}}_{\text{recon}} = 0 \quad (\text{B.3})$$

The resulting reconstructed image can be written as

$$\mathbf{i}_{\text{recon}} = (F^\dagger M^\dagger M F + \lambda D^\dagger D)^{-1} F^\dagger M^\dagger \underline{\mathbf{s}} \quad (\text{B.4})$$

Taking the FFT of the reconstructed image gives

$$\underline{\mathbf{s}}_{\text{recon}} = F \mathbf{i}_{\text{recon}} = F(F^\dagger M^\dagger M F + \lambda D^\dagger D)^{-1} F^\dagger M^\dagger \underline{\mathbf{s}} \quad (\text{B.5})$$

$$= (F[F^\dagger M^\dagger M F + \lambda D^\dagger D]F^\dagger)^{-1} M^\dagger \underline{\mathbf{s}} \quad (\text{B.6})$$

$$= (M^\dagger M + \lambda F D^\dagger F^\dagger F D F^\dagger)^{-1} M^\dagger \underline{\mathbf{s}} \quad (\text{B.7})$$

$$= (M^\dagger M + \lambda Y^\dagger Y)^{-1} M^\dagger \underline{\mathbf{s}} \quad (\text{B.8})$$

The circulant matrix  $D$  is diagonalised by the FFT matrix  $F$ , which implies that  $Y$  is a diagonal matrix.  $Y^\dagger Y$  is therefore also diagonal, as is  $M^\dagger M$ . Since  $M^\dagger \underline{\mathbf{s}}$  contains zeros at the missing  $k$ -space positions, so does  $\underline{\mathbf{s}}_{\text{recon}}$ .

# Summary

## Non-Cartesian MRI Scan-Time Reduction through Sparse Sampling

Magnetic resonance imaging (MRI) signals are measured in the Fourier domain, also called  $k$ -space. Samples of the MRI signal can not be taken at will, but lie along  $k$ -space trajectories determined by the magnetic field gradients.

MRI measurements are usually Cartesian, where the trajectories are parallel and equidistant and sampling along the trajectories is also equidistant. This allows fast reconstruction using the inverse Fast Fourier transform (IFFT). However, this thesis focuses on non-Cartesian MRI. Typical trajectories in this case are radial and spiral, but there exists a multitude of other possibilities. Chapter 2 gives a basic introduction on MRI relevant for this thesis.

Image reconstruction in the non-Cartesian case can not be accomplished by IFFT. In certain cases, however, dedicated reconstruction algorithms are available. For example, for radial scanning there exists the Filtered Back Projection algorithm. Another possibility, aiming at maintaining the IFFT for the transformation to the image domain, is the gridding algorithm. This algorithm, which is capable of image reconstruction from a  $k$ -space sampled along arbitrary trajectories, is given extensive attention in this thesis.

A major complication in non-Cartesian sampling is the compensation for the non-uniformity of the sampling density. In case the trajectories are rather regular, an analytical expression for the density may be derived or Voronoi triangulation can be applied. Another more recent approach is the Pipe-Menon algorithm. However, all these approaches fail in more irregular cases.

The above-mentioned image reconstruction algorithms are described in chapter 3. They are based on the inverse Fourier transform and therefore require  $k$ -space sampling to obey the Nyquist criterion. These algorithms cannot cope with  $k$ -space undersampling. However, in certain cases there may not be enough time to fully sample  $k$ -space; or scan-time is deliberately reduced by omitting trajectories, the total scan-time being proportional to the number of trajectories. Under these circumstances we still want to be able to reconstruct an image.

Chapter 4 presents two algorithms that are able to cope with undersampled  $k$ -space data, and still reconstruct artefact free images.

The first, based on work by G.J. Marseille, who worked on Cartesian scans, aims on estimating values for the missing data. In this approach the missing data are estimated iteratively by shuttling back and forth between image and  $k$ -space, while smoothing the image with an edge-preserving filter and resetting the measured data to their original values. This algorithm has a major drawback in that it requires density compensation. This means that this algorithm is only applicable if the trajectories are regular. Moreover, the algorithm requires user input on which  $k$ -space points are missing. In certain cases, especially when sampling is irregular, this may be impossible or not desirable. Note that this difficulty is absent in undersampled Cartesian scans.

The second algorithm, also based on work by G.J. Marseille, directly estimates the image from the available  $k$ -space data. It is based on Bayes theorem, which allows both incorporation of consistency with the measured data, as in maximum likelihood estimation, *and* prior knowledge about the image. If  $k$ -space is undersampled, the  $k$ -space data alone give insufficient information to reconstruct satisfactory images. The prior knowledge gives the required additional information.

In the Cartesian case, one dimension is always completely measured. Along this dimension IFFT can already be applied. This effectively reduces the image reconstruction problem to one dimension, meaning that different columns in the image can be treated separately. Consequently, the used prior knowledge, *i.e.* the Lorentzian edge distribution model, is a function taking only into account edges in one direction.

In the non-Cartesian case, no dimension is completely sampled. Therefore the image can not be treated column-wise. Moreover, the prior has to take into account edges in both directions, since there is no preferred direction. In this thesis the prior used in Cartesian work is extended to include edges in more than one direction.

In contrast to the first algorithm, the Bayesian approach allows one to obviate density compensation. Therefore, this algorithm can handle any type of sampling, to whatever degree of irregularity of sampling. In addition, since the image is directly estimated there is no need for the user to input which  $k$ -space points are missing.

Chapter 5 discusses image quality measures. The measures are necessary for evaluation of the developed reconstruction algorithms.

The performance of the mentioned reconstruction algorithms, in case of deliberate omission of trajectories to reduce the scan-time, depends on which trajectories are omitted. Chapter 6 is devoted to the matter of how to omit trajectories. Ideally, the raw data satisfy Hermitian symmetry. This property can be exploited when omitting trajectories. This chapter also describes two ways of finding optimal omission of trajectories from a full measurement. Finally, one of these methods is applied to an in vivo spiral scan. Optimal distributions seem to omit trajectories irregularly, without clustering too many omitted trajectories together so as to keep local undersampling to a minimum.

Finally, chapter 7 deals with applications. The methods alluded to above are tested on simulations and real-world data. The Bayesian estimator appears particularly suited for pseudo-random sample positions.

Frank Wajer, Delft University of Technology

# Samenvatting

## Niet-Cartesische MRI Scan-Tijd Reductie door Schaars Bemonsteren

Magnetische resonantie (MRI) signalen worden gemeten in de Fourier ruimte, in MRI ook wel  $k$ -space genoemd. Bemonsteren van het MRI signaal kan niet op een willekeurige manier, maar langs  $k$ -space trajecten, die bepaald worden door de aangelegde magnetische veldgradienten.

MRI metingen zijn meestal Cartesisch, waarmee bedoeld wordt dat de trajecten parallel en even ver van elkaar liggen en dat het bemonsteren langs de trajecten regelmatig is. Dit staat toe de zogenaamde inverse Fast Fourier transformatie (IFFT) te gebruiken voor beeldreconstructie. Dit proefschrift richt zich op niet-Cartesische MRI. Typische trajecten in dit geval zijn radiaal en spiraal, maar er zijn nog verscheidene andere mogelijkheden. Hoofdstuk 2 geeft een basis uitleg over MRI die voor het lezen van dit proefschrift relevant is.

Beeldreconstructie in het niet-Cartesische geval kan niet met behulp van de IFFT. Echter, in sommige gevallen bestaan er specifieke algoritmes. Bijvoorbeeld, voor radiaal meten bestaat er het zogenaamde Filtered Back Projection algoritme. Een andere mogelijkheid, die probeert het voordeel van de IFFT te behouden, is het gridding algoritme. Aan dit algoritme, dat beelden kan reconstrueren uit een willekeurig bemonsterde  $k$ -space, wordt uitvoerig aandacht geschonken in dit proefschrift.

Een grote moeilijkheid bij niet-Cartesisch bemonsteren is het compenseren van de bemonsterdichtheid. In het geval de trajecten redelijk regelmatig zijn, kan een analytische uitdrukking voor de dichtheid worden gevonden of kan het Voronoi algoritme uitkomst bieden. Recent is hier het Pipe-Menon algoritme bij gekomen. Helaas bieden al deze aanpakken geen uitkomst in de onregelmatige gevallen.

De bovengenoemde beeldreconstructie methoden worden beschreven in hoofdstuk 3. Zij zijn allen gebaseerd op de inverse Fourier transformatie en daarom moet de bemonstering voldoen aan het Nyquist criterium. Deze algoritmen kunnen niet overweg met een onderbemonsterd  $k$ -space. Echter, in sommige gevallen kan het gebeuren dat het niet mogelijk is een volledig bemonsterd  $k$ -space te meten of er is, om de meettijd te verkorten, met opzet voor gekozen om bepaalde trajecten weg te laten; de meettijd is evenredig met het aantal trajecten. Ook in deze gevallen willen we nog steeds een beeld reconstrueren.

Hoofdstuk 4 behandelt twee algoritmen die overweg kunnen met onderbemonsterde  $k$ -space data.

De eerste, afgeleid van werk van G.J. Marseille, die gewerkt heeft aan Cartesische metingen, probeert weggelaten meetpunten te schatten. De missende data wordt iteratief geschat door heen en weer te pendelen tussen beelddomein en  $k$ -space, terwijl het beeld gefilterd wordt met een edge-preserving filter en de gemeten data teruggezet worden op hun gemeten waarde. Dit algoritme heeft als grote nadeel dat het de dichtheidcompensatie benodigt. Dit maakt het algoritme alleen toepasbaar in gevallen met regelmatige trajecten. Bovendien benodigt dit algoritme invoer van de gebruiker welke datapunten missen. Vooral in gevallen waarbij  $k$ -space zeer onregelmatig bemonsterd is, is dit ofwel niet mogelijk ofwel niet wenselijk. Deze problematiek doet zich in het Cartesische geval niet voor.

Het tweede algoritme, ook afgeleid van werk van G.J. Marseille, schat het beeld direct gegeven de beschikbare  $k$ -space data. Hierbij wordt gebruik gemaakt van het theorema van Bayes. Dit schept de mogelijkheid om naast het inbouwen van consistentie met de beschikbare data, zoals met maximum likelihood methoden, ook voorkennis over het beeld in te bouwen. De gemeten data geeft te weinig informatie om geschikte beelden te reconstrueren indien  $k$ -space onderbemonsterd is. De voorkennis levert in deze gevallen additionele informatie.

Bij Cartesisch meten is altijd één richting volledig bemonsterd. In deze richting kan dus de IFFT al toegepast worden. Hiermee wordt beeldreconstructie gereduceerd tot een één dimensionaal probleem; elke kolom in het beeld kan apart behandeld worden. Dientengevolge is de gebruikte voorkennis, nl. het Lorentziaans zijn van de distributie van de randen in het beeld, slechts een functie die alleen randen in één richting meeneemt.

In het niet-Cartesische geval is geen richting volledig bemonsterd. Daarom is een kolomsgewijze aanpak niet mogelijk. Bovendien is het nu noodzakelijk om randen in beide richtingen mee te nemen in de voorkennis, want er is nu geen voorkeursrichting. In dit proefschrift wordt de voorkennis die gebruikt werd in het Cartesische geval uitgebreid om ook randen in meer dan één richting mee te nemen.

In tegenstelling tot het eerste algoritme heeft de Bayesiaanse aanpak het voordeel dat de dichtheidscompensatie overbodig is. Dit betekent dat dit algoritme elk bemonsterd  $k$ -space aankan, hoe willekeurig ook. Bovendien vervalt de noodzaak voor invoer van de gebruiker welke datapunten missen.

Hoofdstuk 5 behandelt verscheidene beeldkwaliteit maten. Deze zijn nodig voor evaluatie van de behandelde reconstructie methoden.

De kwaliteit van de genoemde reconstructiemethoden, in het geval van opzettelijk weglaten van trajecten om de meettijd te verkorten, hangt af van welke trajecten weggelaten worden. Hoofdstuk 6 is gewijd aan het op een optimale manier weglaten van trajecten. In het ideale geval voldoet de data aan Hermitische symmetrie. Deze eigenschap kan gebruikt worden als er trajecten worden weggelaten. Dit hoofdstuk behandelt ook twee manieren om op een optimale manier trajecten weg te laten uit een volledige meting. Een van deze methoden wordt toegepast op een in vivo spiraal meting. Optimale



distributies blijken trajecten op een willekeurige manier weg te laten, waarbij er zoveel mogelijk wordt voorkomen dat er trajecten naast elkaar worden weggelaten. Dit houdt lokale onderbemonstering tot een minimum.

Als laatste, behandelt hoofdstuk 7 toepassingen. The genoemde methoden worden getest op simulaties en real-world data. De Bayes schatter lijkt vooral geschikt in gevallen waarbij de bemonsterpunten zeer willekeurig verspreid liggen.

Frank Wajer, Technische Universiteit Delft



# Bibliography

- [1] E. Walter and L. Pronzato, *Identification of Parametric Models*, Springer, 1997.
- [2] M.T. Vlaardingerbroek and J.A. den Boer, *Magnetic Resonance Imaging: Theory and Practice*, Springer, Berlin, 2nd edition, 1999.
- [3] G.A. Wright, “Magnetic resonance imaging,” *IEEE Signal Processing Magazine*, vol. 14, pp. 56–66, 1997.
- [4] W.R. Nitz, “MR imaging: acronyms and clinical applications,” *European Radiology*, vol. 9, pp. 979–997, 1999.
- [5] A. Abragam, *The Principles of Nuclear Magnetism*, Clarendon, 1961.
- [6] M.K. Stehling, R. Turner, and P. Mansfield, “Echo-Planar Imaging: Magnetic Resonance Imaging in a fraction of a second,” *Science*, vol. 254, pp. 43–50, 4 October 1991.
- [7] K. Scheffler and J. Hennig, “Frequency-resolved single-shot MR imaging using stochastic  $k$ -space trajectories,” *Magnetic Resonance in Medicine*, vol. 35, pp. 569–576, 1996.
- [8] G.H. Glover and S. Lai, “Self-navigated spiral fMRI - Interleaved versus single-shot,” *Magnetic Resonance in Medicine*, vol. 39, pp. 361–368, 1998.
- [9] D.B. Twieg, “The  $k$ -trajectory formulation of the NMR imaging process with application and synthesis of imaging methods,” *Medical Physics*, vol. 10, pp. 610–621, 1983.
- [10] J. Hennig, “ $k$ -Space sampling strategies,” *European Radiology*, vol. 9, no. 6, pp. 1020–1031, 1999.
- [11] G.H.L.A. Stijnman, “Alternative radial scan technologies,” M.S. thesis, Delft University of Technology, June 1998.
- [12] A.L. Alexander, J. Roberts, E. Kholmovski, and D.L. Parker, “MR imaging with 3d fast spin echo using projection reconstruction: technique description and applications,” in *Conference Record. Nuclear Science Symposium and Medical Imaging Conference, 24-30 Oct 1999*, 1999, vol. 2, pp. 1087–1090.

- [13] P. Edholm and G.T. Herman, "Linograms in image reconstruction from projections," *IEEE Trans. Medical Imaging*, vol. MI-6, no. 4, pp. 301–307, 1987.
- [14] P. Edholm, G.T. Herman, and D.A. Roberts, "Image reconstruction from linograms: Implementation and evaluation," *IEEE Trans. Medical Imaging*, vol. 7, no. 3, pp. 239–246, 1988.
- [15] M. Magnusson, *Linogram and other direct Fourier methods for tomographic reconstruction*, Ph.D. thesis, Linköping University, S-581 83, Sweden, 1993.
- [16] C. Jacobson, *Fourier Methods in 3D-Reconstruction from Cone-Beam Data*, Ph.D. thesis, Linköpings Universitet, EE Dept, Linköping University, S-581 83, Sweden, 1996.
- [17] P. Malm and J. Carlsson, "Implementation and evaluation of 2D Linogram technique in magnetic resonance imaging," Tech. Rep. LiTH-ISY-EX-1724, Linköpings Universitet, EE Dept, Linköping University, S-581 83, Sweden, May 1997.
- [18] J.M. Wagner, "Reconstruction of three-dimensional images by "linogram" techniques," *Bulletin Scientifique de l'Association des Ingenieurs Electriciens sortis de l'Institut-Montefiore*, vol. 111, no. 1, pp. 3–9, april 1998.
- [19] T. Yokoyama, K. Nagai, K. Mizutani, and A. Hasegawa, "Image reconstruction of diffraction tomography using linograms," in *IEEE Ultrasonics Symposium. Proceedings.*, 1998, vol. 2, pp. 1655–1658.
- [20] N. Gai and L. Axel, "Characterization of and correction for artifacts in linogram MRI," *Magnetic Resonance in Medicine*, vol. 37, pp. 275–284, 1997.
- [21] N. Gai and L. Axel, "A dual approach to linogram imaging for MRI," *Magnetic Resonance in Medicine*, vol. 38, pp. 337–341, 1997.
- [22] D. Graveron-Demilly, Y. Crémillieux, S. Cavassila, and C. Mauger, "A 3-D Fourier imaging method based on radial scan," in *Proc. Int. Society of Magnetic Resonance in Medicine, Third Scientific Meeting*, Nice, France, August 1995, p. 671.
- [23] D. Graveron-Demilly, G.J. Marseille, Y. Crémillieux, S. Cavassila, and D. van Ormondt, "SRS-FT: A Fourier imaging method based on sparse radial scanning and Bayesian estimation," *Journal of Magnetic Resonance B*, vol. 112, pp. 119–123, 1996.
- [24] G.H.L.A. Stijnman, D. Graveron-Demilly, F.T.A.W. Wajer, and D. van Ormondt, "MR image estimation from sparsely sampled radial scans," in *ProRISC IEEE Workshop on Circuits, Systems and Signal Processing*, Mierlo (The Netherlands), 1997, pp. 603–610.
- [25] E. Yudilevich and H. Stark, "Spiral sampling: theory and an application to magnetic resonance imaging," *J. Optical Society of America*, vol. 5, pp. 542–553, 1988.

- [26] C.H. Meyer, B.S. Hu, D.G. Nishimura, and A. Macovski, "Fast spiral coronary artery imaging," *Magnetic Resonance in Medicine*, vol. 28, pp. 202–213, 1992.
- [27] D.C. Noll, "Methodologic considerations for spiral  $k$ -space functional MRI," *International Journal of Imaging Systems and Technology*, vol. 6, pp. 375–383, 1995.
- [28] W.A. Kaiser, "Dynamic spiral MR mammography," *Radiology*, vol. 215, no. 3, pp. 919–920, June 2000.
- [29] C. Salustri, Y. Yang, J.H. Duyn, and G.H. Glover, "Comparison between analytical and numerical solutions for spiral trajectory in 2d  $k$ -space," in *Proc. Int. Society of Magnetic Resonance in Medicine, Fifth Scientific Meeting*, Vancouver, B.C., Canada, 12-18 April 1997, p. 1813.
- [30] O. Heid, "Archimedian spirals with euclidean gradient limits," in *Proc. Int. Society of Magnetic Resonance in Medicine, Fourth Scientific Meeting*, New York, USA, 27 April-3 May 1996, p. 114.
- [31] F.T.A.W. Wajer, R. de Beer, M. Fuderer, A.F. Mehlkopf, and D. van Ormondt, "Spiral MRI scan-time reduction through omission of interleaves and iterative gaussian filtering," in *Proc. Int. Soc. Magn. Reson. Med, Fifth Sci. Meeting*, Vancouver, Canada, april 1997, p. 134.
- [32] F.T.A.W. Wajer, R. de Beer, M. Fuderer, A.F. Mehlkopf, and D. van Ormondt, "Spiral MRI scan-time reduction by undersampling and iterative gaussian filtering," in *Proc. 3rd annual conference Advanced School for Computing and Imaging*, Heyen, NL, 2-4 June 1997, pp. 55–60.
- [33] F.T.A.W. Wajer, R.de Beer, M. Fuderer, A.F. Mehlkopf, and D. van Ormondt, "Spiral mri scan-time reduction through omission of interleaves and Bayesian image reconstruction," in *Proc. Int. Society of Magnetic Resonance in Medicine, Sixth Scientific Meeting*, Sydney, Australia, April 1998, p. 1960.
- [34] X.H. Zhou, Z.P. Liang, S.L. Gewalt, G.P. Cofer, P.C. Lauterbur, and G.A. Johnson, "Fast spin-echo technique with circular sampling," *Magnetic Resonance in Medicine*, vol. 39, pp. 23–27, 1998.
- [35] N.E. Myridis and C. Chamzas, "Sampling on concentric circles," *IEEE Transactions on Medical Imaging*, vol. 17, pp. 294–299, 1998.
- [36] N.E. Myridis and C. Chamzas, "Rhombus hankel transform reconstruction (RHTR) and rosette trajectories," in *Proc. Int. Society of Magnetic Resonance in Medicine, Sixth Scientific Meeting*, Sydney, Australia, April 1998, p. 677.
- [37] D.C. Noll, "Multishot rosette trajectories for spectrally selective MR imaging," *IEEE Transactions on Medical Imaging*, vol. 16, pp. 372–377, 1998.
- [38] D.C. Noll, S.J. Peltier, and F.E. Boada, "Simultaneous multislice acquisition using rosette trajectories (SMART) - A new imaging method for functional MRI," *Magnetic Resonance in Medicine*, vol. 39, pp. 709–716, 1998.

- [39] R. Proksa V. Rasche, “Reconstruction of MR images from data sampled along arbitrary  $k$ -space trajectories,” in *Proc. Int. Society of Magnetic Resonance in Medicine, Sixth Scientific Meeting*, Sydney, Australia, April 1998, p. 668.
- [40] M. Bourgeois, F. Wajer, G. Stijnman, Y. Crémillieux, D. van Ormondt, A. Briguet, and D. Graveron-Demilly, “Reduction of random in-plane intrascan rotation artefacts,” in *Proc. Int. Soc. Magn. Reson. Med, Sixth Sci. Meeting*, Sydney, Australia, april 1998, p. 1963.
- [41] F.T.A.W. Wajer, M. Bourgeois, D. Graveron-Demilly, and D. van Ormondt, “Nonuniform Sampling in Magnetic Resonance Imaging,” in *Proc ICASSP*, Istanbul, Turkey, 5-9 June 2000, vol. VI, pp. 3846–3849, Invited Talk for Special Session on ‘Theory and Practice of Nonuniform Sampling’, organized by F. Marvasti and P.J.S.G. Ferreira.
- [42] F.T.A.W. Wajer, G.H.L.A. Stijnman, M. Bourgeois, D. Graveron-Demilly, and D. van Ormondt, *Magnetic resonance Image Reconstruction from Nonuniformly sample  $k$ -Space Data*, chapter 10, Plenum, 2001, pp. 439-478, Chapter in Sampling Theory and Practice, F. Marvasti Ed.
- [43] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C*, Cambridge UP, 2nd edition, 1997, reprinted version 2.08.
- [44] A. Okabe, B. Boots, and K. Sugihara, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, Wiley, New York, 1992.
- [45] A. Webster, “Callan’s canyons and Voronoi’s cells,” *Nature*, vol. 391, pp. 430, 1998.
- [46] E. Woudenberg, “Optimization of Gridding for Non-Uniform MR Image Reconstruction ,” M.S. thesis, Delft University of Technology, June 1999.
- [47] R.A.J. de Jong, “Optimization of Radial Scan Technology for MRI,” M.S. thesis, Delft University of Technology, February 2000.
- [48] J.G. Pipe and P. Menon, “Sampling Density Compensation in MRI: Rationale and an Iterative Numerical Solution,” *Magnetic Resonance in Medicine*, vol. 41, pp. 179–186, 1999.
- [49] A.V. Oppenheim and A.S. Willsky, *Signals and Systems*, Prentice hall, 1983.
- [50] P. C. Lauterbur, “Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance,” *Nature*, vol. 242, pp. 190, 1973.
- [51] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, Prentice Hall, Englewood Cliffs (New Jersey), 1989.
- [52] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms and Applications*, Prentice Hall, New York, 1992.

- [53] J.H. Zwaga, “MR image reconstruction from nonuniform samples using convolution gridding,” M.S. thesis, Delft University of Technology, NL, 1997.
- [54] F.T.A.W. Wajer, R. Lethmate, J. van Osch, D. Graveron-Demilly, M. Fuderer, and D. van Ormondt, “Interpolation from arbitrary to Cartesian sample positions: Gridding,” in *ProRISC, IEEE Benelux*, Velthuisen, The Netherlands, November 2000, pp. 571–577.
- [55] R.N. Bracewell, *The Fourier Transform and its Applications*, McGraw-Hill, 1978.
- [56] D. Slepian, “Some comments on Fourier analysis, uncertainty, and modeling,” *SIAM Rev.*, vol. 25, pp. 379–393, 1998.
- [57] J.F. Kaiser, “Digital filters,” in *System Analysis by Digital Computer*, F. Kuo and J.F. Kaiser, Eds., New York, 1966, pp. 218–285, Wiley.
- [58] M.A. Schonewille, *Fourier reconstruction of irregularly sampled seismic data*, Ph.D. thesis, Delft University of Technology, NL, 2000.
- [59] J.I. Jackson, C.H. Meyer, D.G. Nishimura, and A. Macovski, “Selection of a convolution function for Fourier inversion using gridding,” *IEEE Trans. Medical Imaging*, vol. 10, pp. 473–478, 1991.
- [60] H. Schomberg and J. Timmer, “The gridding method for image reconstruction by Fourier transformation,” *IEEE Trans. Medical Imaging*, vol. 14, no. 3, pp. 596–607, 1995.
- [61] F.T.A.W. Wajer, E. Woudenberg, R. de Beer, L.T. Martinez, M. Fuderer, A.F. Mehlkopf, and D. van Ormondt, “Simple Equation for Optimal Gridding Parameters,” in *Proc. Int. Society of Magnetic Resonance in Medicine, Seventh Scientific Meeting*, Philadelphia, USA, 22-28 May 1999, p. 663.
- [62] Y. Bard, *Nonlinear Parameter Estimation*, Academic Press, 1974.
- [63] S.L. Marple, *Digital spectral analysis: with applications*, Prentice-Hall, 1987.
- [64] S.M. Kay, *Fundamentals of statistical signal processing; estimation theory*, Prentice-Hall, 1993.
- [65] F. Godtliessen, “Noise reduction Using Markov Random Fields,” *Journal of Magnetic Resonance*, vol. 92, pp. 102–114, 1991.
- [66] G. Strang, *Linear Algebra and its Applications*, Harcourt Brace Jovanovich, 1988.
- [67] G.H. Golub and C.F. van Loan, *Matrix Computations*, North Oxford Academic, 1983.
- [68] D. Rosenfeld, “An optimal and efficient new gridding algorithm using singular value decomposition,” *Magnetic Resonance in Medicine*, vol. 40, pp. 14–23, 1998.

- [69] H. Sedarat and D. Nishimura, “Generalized gridding reconstruction from nonuniformly sampled data,” in *Proc. SPIE*, 1999, vol. 3808, pp. 447–456.
- [70] C. Mosquera, P. Irarrázabal, and D.G. Nishimura, “Noise behaviour in gridding reconstruction,” in *Proceedings ICASSP*, 1995, vol. 4, pp. 2281–2284.
- [71] G. J. Marseille, *MRI Scan Time Reduction through Nonuniform Sampling*, Ph.D. thesis, Delft University of Technology, NL, 1997.
- [72] F.T.A.W. Wajer, G.H.L.A. Stijnman, M. Fuderer, D. Graveron-Demilly, A.F. Mehlkopf, and D. van Ormondt, “Sparse sampling distributions for MRI,” in *ProRISC, IEEE Benelux*, Mierlo, The Netherlands, November 1998, pp. 603–608.
- [73] G.J. Marseille, M. Fuderer, R. de Beer, A.F. Melkopf, and D. van Ormondt, “Non-uniform phase-encode distributions for MRI scan time reduction,” *Journal of Magnetic Resonance B*, vol. 111, pp. 70–75, 1996.
- [74] M. Fuderer, “Ringing artefact reduction by an efficient likelihood improvement method,” in *Proc. SPIE*, 1989, pp. 84–90.
- [75] A.H. Lettington and Q.H. Hong, “Image restoration using a Lorentzian probability model,” *Journal of Modern Optics*, vol. 42, pp. 1367–1376, 1995.
- [76] B. McNally, “Lorentzian probability model,” Internal Report, pp. 16-17, King’s College, London, U.K., 1996.
- [77] S. Geman and D.E. McClure, “Statistical Methods for Tomographic Image Reconstruction,” *Bulletin of the International Statistical Institute*, vol. 4, pp. 5–21, 1987.
- [78] L. Kaufman and A. Neumaier, “Image Reconstruction through Regularization by Envelope-Guided Conjugate Gradients,” Tech. Rep., AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, NJ 07974-0636, USA, 1991, 31 pp.
- [79] S. Geman and S. Geman, “Stochastic relaxation, gibbs distributions, and the bayesian restoration of images,” *IEEE Trans Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 6, pp. 721–741, November 1984.
- [80] L.H. Yu, “A Simple Statistical Project: Image Reconstruction,” *The American Statistician*, vol. 48, no. 1, pp. 30–34, February 1994.
- [81] A. Tarantola, *Inverse Problem Theory. Methods for Data Fitting and Model Parameter Estimation*, Elsevier, 1987.
- [82] S. Iyanaga and Y. Kawada, Eds., *Encyclopedic Dictionary of Mathematics*, MIT Press, paperback edition, 1980, Mathematical Society of Japan, English translation.
- [83] J.P. Norton, *An Introduction to Identification*, Academic Press, 1986.



- [84] D.S. Sivia, *Data Analysis. A Bayesian Tutorial*, Clarendon Press, Oxford, 1996.
- [85] D. Malakoff, "Bayes offers a 'New' Way to make Sense of Numbers," *Science*, vol. 286, pp. 1460–1464, 19 november 1999.
- [86] M. Dwass, *Probability and Statistics*, W.A. Benjamin, New York, 1970.
- [87] J.R. Blum and J.I. Rosenblatt, *Probability and Statistics*, Saunders, Philadelphia, 1972.
- [88] D.S. Lalush and B.M.W. Tsui, "Simulation Evaluation of Gibbs Prior Distributions for Use in Maximum *A Posteriori* SPECT Reconstructions," *IEEE Trans Medical Imaging*, vol. 11, no. 2, pp. 267–275, 1992.
- [89] G.H.L.A. Stijnman, D. Graveron-Demilly, F.T.A.W. Wajer, and D. van Ormondt, "MR Image estimation from Sparsely Sampled Radial Scans," in *ProRISC, IEEE Benelux*, Mierlo, The Netherlands, November 1997, pp. 603–610.
- [90] F.T.A.W. Wajer, G.H.L.A. Stijnman, M. Fuderer, A.F. Mehlkopf, D. Graveron-Demilly, and D. van Ormondt, "MRI Measurement Time Reduction by Under-sampling and Bayesian Reconstruction using a 2D Lorentzian Prior," in *Proc. 4rd annual conference Advanced School for Computing and Imaging*, Lommel, BE, x-y June 1998, pp. 147–153.
- [91] H. Feichtinger and K. Gröchenig, "Theory and practice of irregular sampling," in *Wavelets: Mathematics and Applications*, J. Benedetto and M. Frazier, Eds. CRC Press, 1993, pp. 305–363.
- [92] T. Mathews and M.R. Smith, "Objective Image Measures for Evaluating Advanced MRI Reconstruction Methods," in *Canadian Conference on Electrical and Computer Engineering, IEEE/CCECE*, 1996, pp. 359–361.
- [93] J.M. Cuppen and A. van Est, "Reducing MR Imaging Time by One-Sided Reconstruction," in *Book of Abstracts, Topical Conference on Fast Imaging*, Cleveland, 15-17 May 1987.
- [94] M. Bourgeois, Y. Crémillieux, C. Delon-Martin, A. Briguet, and D. Graveron-Demilly, "Reduction of Random In-Plane Intrascan Rotation Artefacts," in *Proc. Int. Soc. Magn. Reson. Med, Sixth Sci. Meeting*, Aveiro, Portugal, 16-19 June 1997, pp. 34–36.
- [95] M. Bourgeois, F.T.A.W. Wajer, G.H.L.A. Stijnman, Y. Crémillieux, D. van Ormondt, A. Briguet, and D. Graveron-Demilly, "Réduction des artéfacts de rotation aléatoire intra-image," in *GRAMM*, Toulouse, France, february 1998, p. 28.
- [96] M. Bourgeois, *Corrections des Mouvements Intra-Image en Imagerie d'Activation Cérébrale par Résonance Magnétique*, Ph.D. thesis, Université Claude Bernard-Lyon I, 9 June 1999.

- [97] M. Bourgeois, F.T.A.W. Wajer, D. van Ormondt, and D. Graveron-Demilly, *Reconstruction of MRI Images from Non-uniform sampling: Application to Intrascan Motion Correction in Functional MRI*, chapter 16, Birkhauser, 2000, 25 pages, Chapter in Modern Sampling Theory, Mathematics and Applications, P.J.S.G. Ferreira and J.J. Benedetto, Eds.
- [98] M. Reczko, D.A. Karras, B.G. Mertzios, D. Graveron-Demilly, and D. van Ormondt, "Neural Networks in MR Image Estimation from Sparsely Sampled Scans," in *Proceedings of the First International Workshop on Machine Learning and Data Mining in Pattern Recognition (MLDM'99)*, in *Lecture Notes in Artificial Intelligence*. 16-18 September 1999, number 1715 in Subseries in Computer Science, pp. 75–86, Springer.
- [99] M. Reczko, D.A. Karras, B.G. Mertzios, D. Graveron-Demilly, and D. van Ormondt, "Improved MR image reconstruction from sparsely sampled scans based on neural networks," *Pattern Recognition Letters*, vol. 22, no. 1, pp. 35–46, January 2001.
- [100] L.T. Martinez, F.T.A.W. Wajer, R. Lethmate, D. van Ormondt, and D. Graveron-Demilly, "Bayesian reconstruction of arbitrarily sampled 3-D MR images," in *IX congrès du GRAMM*, Lille, FR, 2-4 February 2000.
- [101] L.T. Martinez, F.T.A.W. Wajer, R. Lethmate, R.A.J. de Jong, G.H.L.A. Stijnman, Y. Crémillieux, D. van Ormondt, and D. Graveron-Demilly, "Sparse 3D Radial Scanning and Reconstruction," in *Proc. Int. Society of Magnetic Resonance in Medicine, Seventh Scientific Meeting*, Denver, USA, 1-7 April 2000, p. 1723.
- [102] F.T.A.W. Wajer, R. Lethmate, R.A.J. de Jong, L.T. Martinez, D. Graveron-Demilly, and D. van Ormondt, "MR Image reconstruction from Sparse and Corrupted  $k$ -space data," in *ProRISC, IEEE Benelux*, J.P. Veen, Ed., Mierlo, The Netherlands, November 1999, STW, pp. 577–584.
- [103] R. Lethmate, J.A.C. van Osch, F. Lamberton, F.T.A.W. Wajer, D. van Ormondt, and D. Graveron-Demilly, "MR Image Reconstruction from 3D Radial Sampling Distributions. A Comparison," in *ProRISC, IEEE Benelux*, J.P. Veen, Ed., Velthuisen, NL, November 2000, STW, pp. 381–387, IEEE-Benelux.
- [104] F.T.A.W. Wajer, R. de Beer, M. Fuderer, A.F. Mehlkopf, and D. van Ormondt, "Bayesian image reconstruction from arbitrarily sampled  $k$ -space without density correction," in *Proc. Int. Society of Magnetic Resonance in Medicine, Sixth Scientific Meeting*, Sydney, Australia, April 1998, p. 667.
- [105] K. Scheffler and J. Hennig, "Reduced circular field of view imaging," *Magnetic Resonance in Medicine*, vol. 40, pp. 474–480, 1998.
- [106] K.P. Pruessmann, M. Weiger, M.B. Scheidegger, and P. Boesiger, "SENSE: Sensitivity Encoding for Fast MRI," *Magnetic Resonance in Medicine*, vol. 42, pp. 952–962, 1999.

- [107] K.P. Pruessmann, M. Weiger, P. Boernert, and P. Boesiger, “A Gridding Approach for Sensitivity Encoding with Arbitrary Trajectories,” in *Proc. Int. Society of Magnetic Resonance in Medicine, Seventh Scientific Meeting*, Denver, USA, 1-7 April 2000, p. 276.
- [108] S.A.R. Kannengiesser, A.R. Brenner, and T.G. Noll, “Accelerated Image Reconstruction for Sensitivity Encoded Imaging with Arbitrary k-Space Trajectories,” in *Proc. Int. Society of Magnetic Resonance in Medicine, Seventh Scientific Meeting*, Denver, USA, 1-7 April 1999, p. 155.
- [109] K.P. Pruessmann, “Advances in Sensitivity Encoding with Arbitrary k-Space Trajectories,” in *Proc. Int. Society of Magnetic Resonance in Medicine, Seventh Scientific Meeting*, Glasgow, UK, 21-27 April 2001, p. 4.



# List of symbols and abbreviations

$a$	Width of the Lorentzian prior
$B$	Shape parameter of Kaiser-Bessel kernel
$B_0$	Main magnetic field strength [T]
$c(x)$	Gridding convolution kernel in image space
$C(k)$	Gridding convolution kernel in $k$ -space
$f_{\text{og}}$	Overgridding factor
FOV	Size of field of view
$G_x, G_y, G_z$	Magnetic field gradients along $x$ , $y$ and $z$ [T/m]
$I$	Image
$\underline{\mathbf{i}}$	Image written as a vector
$\vec{k}$	Position vector in $k$ -space, $(k_x, k_y, k_z)$
$L$	Width of gridding convolution kernel
$\vec{M}$	Macroscopic magnetization
$M_{\perp}$	Transverse magnetization
$\vec{m}$	Magnetisation
$m_{\perp}$	Transverse magnetisation
$N_{\text{rad}}$	Number of radial trajectories
$\mathcal{N}_I$	Noise in image domain
$\mathcal{N}_S$	Noise in $k$ -space
$\underline{\mathbf{n}}$	Measurement noise written as a vector
perf	Performance measure
PSF	Point spread function
$P(I S)$	Posterior
$P(S I)$	Likelihood
$P(I)$	Prior knowledge
$P(S)$	Evidence
ROI	Size of region of interest
$T$	Discrete Fourier transform written as matrix
$T_2^*$	Transverse relaxation time including field inhomogeneities [s]
$\vec{r}$	Position vector in image space, $(x, y, z)$
$S$	$k$ -space data

$\underline{s}$	$k$ -space data written as a vector
$t$	Time [s]
$\delta I$	Gradient search direction
$\Delta k_n$	Sampling density compensation at location $\vec{k}_n$ [ $\text{m}^{-1}$ ]
$\Delta k_{\text{Cart}}$	Cartesian $k$ -space sampling step [ $\text{m}^{-1}$ ]
$\Delta r_{\text{Cart}}$	Step size in image space [m]
$\gamma$	Gyromagnetic ratio of proton [ $\text{rad s}^{-1}\text{T}^{-1}$ ]
$\lambda$	Step-size in gradient search direction
$\xi$	Image intensity difference in $x$ direction
$\eta$	Image intensity difference in $y$ direction
$\sigma_I$	Standard deviation of the noise in image space
$\sigma_S$	Standard deviation of the noise in $k$ -space
$\phi$	Accumulated phase of magnetization [rad]
$\omega$	Resonance frequency [ $\text{rad s}^{-1}$ ]

DFT	Discrete Fourier Transform
FirLS	Finite Impulse Response Least Squares
(I)FFT	(Inverse) Fast Fourier Transform
ML	Maximum Likelihood
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error
PR	Projection Reconstruction
PSF	Point Spread Function
SENSE	Sensitivity Encoding
SNR	Signal to Noise Ratio
SRSFT	Sparse Radial Scanning and Fourier Transform
ROI	Region of Interest
FOV	Field of View

# Dankwoord/Acknowledgement

Dit proefschrift zou nooit tot stand zijn gekomen zonder de hulp en medewerking van vele personen. Op deze plek wil ik iedereen hiervoor heel erg bedanken. In het bijzonder dank ik de volgende personen.

In de eerste plaats wil ik Dik van Ormondt danken voor de dagelijkse begeleiding en onuitputtelijke steun, vooral tijdens het schrijven van mijn proefschrift. Ik had me geen betere begeleider kunnen wensen.

Mijn promotor Toon Mehlkopf wil ik danken voor de algemene leiding, de contacten met Philips Medical Systems en de mogelijkheid om promovendus te zijn in de Spin Imaging groep. Verder dank ik Ron de Beer, Alain Coron, Cesar Couturier, Loic Martinez en Stelian Nastase, voor de prettige samenwerking in de groep.

Mijn (oud)collega-promovendi Tanja Claasen, Ole Karlsen, Erik Konijnenberg, Gert-Jan Marseille en Tom Rozijn dank ik voor de gezellige tijd, de nuttige en onnuttige discussies.

Miha Fuderer en Jan Groen dank ik voor de twee-maandelijke vergaderingen en het leveren van de in-vivo spiral datasets.

I would like to thank Danielle Graveron-Demilly, Marc Bourgeois and Ralf Lethmate for our nice cooperation and for applying my Bayesian algorithm on your fMRI and radial datasets.

De samenwerking met de studenten Jeroen de Jong, Jochen van Osch, D  nis Riedijk, Gijs Stijnman, Martijn Vogten, Evert Woudenberg en Jitse Zwaga heb ik zeer op prijs gesteld.

Mijn ouders wil ik bedanken voor hun onvoorwaardelijke steun tijdens de afgelopen jaren.

Als laatste, Josephine, dank je voor je steun en begrip, vooral tijdens de laatste loodjes.





# Curriculum vitae

Naam	Frank Theo Anthonie Willem Wajer
Geboortedatum	26 juli 1970
Geboorteplaats	Amsterdam

1982 - 1986	VWO, St. Vituscollege te Bussum
1986 - 1988	VWO, Rijnlands Lyceum te Oegstgeest
1988 - 1994	Technische Natuurkunde, Technische Universiteit Delft Afstuderen bij vakgroep Spin Imaging, afstudeeronderwerp: ‘Op SVD gebaseerde signaalverwerking van MRSI-signalen’
1995 - 1999	Onderzoeker in opleiding bij Spin Imaging STW onderzoek: DTN 3509
1999 -	Toegevoegd onderzoeker bij Spin Imaging STW onderzoek: DTN 4967













