Exploring subspace-constrained approaches to low-rank fMRI acceleration

Harry T. Mason

Wadham College

University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Hilary 2020

Abstract

Functional magnetic resonance imaging (fMRI) is a medical imaging technique that measures brain activity non-invasively. One of the fundamental quandaries in fMRI is the balance that must be struck between spatial fidelity and temporal resolution. An increase in sampling efficiency could improve either or both of these metrics, allowing images to be created from fewer data points than would otherwise be required. This process is referred to as acceleration.

Some degree of acceleration is already standard in fMRI scans. The most common acceleration methods are parallel imaging methods, which utilise the spatial sensitivity profiles of the receiver coils in order to separate the aliased artefacts that result from using fewer data to create each image. However, there are other mathematical properties which can also be incorporated into the reconstruction process in order to allow a higher degree of acceleration.

One such property is the inherently low-rank nature of fMRI data, which was introduced by Chiew et al. in 2015 as the k-t FASTER method (fMRI Accelerated in Space-Time via Truncation of Effective Rank). The authors also demonstrated in 2016 that the low-rank information could be combined with the coil sensitivity profiles to achieve a higher acceleration factor than either low-rank information or coil profiles could achieve alone.

In this thesis, the k-t FASTER approach is expanded upon by incorporating additional, subspace-specific constraints into the reconstruction process. First, k-t FASTER will be reformulated as an alternating minimisation problem in order to more easily allow subspace-specific regularisation terms. Then, a variety of constraints will be explored in an artificial framework. The constraints being tested are: Tikhonov constraints (which encourage the subspaces to take more minimal energy forms), low-resolution priors (which more greatly weight the oversampled central k-space in radial sampling), and a temporal subspace smoothing constraint (which minimises the variation between frames). These constraints will be applied to real fMRI data acquired with a TURBINE trajectory (Trajectory Using Radially Batched Internal Navigator Echoes), a hybrid radial-Cartesian 3D trajectory with Golden Ratio angle increments in the radial orientation.

The aforementioned subspace-constrained approaches could be seen to achieve better classification of the underlying functional activation over a k-t FASTER reconstruction in real data at R=16, or TR_{vol}=0.5s. Ultimately, Tikhonov constraints are found to provide consistently high-quality reconstructions at a range of acceleration factors and SNR values, but in real data with a slow-paradigm task fMRI experiment at high acceleration (R=26, TR_{vol}=0.3s), the temporal subspace smoothing constraints can outperform Tikhonov constraints.

Acknowledgements

Is there anything more individual than a university degree?
The singular demonstration of one's intellectual capacity.
A recognition of new knowledge you've uncovered on your own.
The real demonstration of what you can achieve alone.

Well, actually...
I couldn't have done this without my parents, who I know as Mum and Dad.
They worked so hard to support me with the upbringing that I had.
They helped me grow and develop, they looked after me when I was sick.
They shined their lights for me while wearing themselves down their wick.

Or without the rest of my family too. I have three siblings who are all close to me.
All older and some wiser and so if you're reading this, thank you Robin, Bernard, and Suzie.
During this degree I became a proud uncle. I lost my dear grandma. Life is wonderfully and tragically dynamic.
But still thank you all for the good times, both the calm and the manic.

And, while I'm thinking of people, my school teachers (like Mr Candlin) certainly guided me,
I'm thankful for all the knowledge and encouragement that they provided to me.
They somehow got me to Oxford, where other professors helped me through,
So to Professors Martin, Grau, Clifton, and Chappell, specific thanks to you too.

But life isn't just academic. And so, to that end,
Let me just say I've been supported by some very kind friends,
And while I know them all from utterly different places,
My heart still lights up each time I see any one of their faces.

To Joe and Beth - thank you for letting me be your best man.
To Will, Fiona, Tim, Farai, and others - let's keep meeting up when we can.
To Matt, Ann, Juliane, and Rachel - keep writing things with such creative weight.
To Tom. Oh Tom. My long long suffering housemate.

It would be dishonest of me not to also give praise,
To He who has guided me through my toughest days.
But also to my friends and pastors at St. Aldates then,
I can say nothing more than thank you, and Amen.

What is life without sport? I'd say I've played my fair share,
And I've played higher than this asthmatic kid ever dreamed to dare.
There are so many I could thank, and it may seem kitsch,
But there are too many people that I'm glad I shared a pitch.

Specifically thanks to those who shared the dark blue kit with me at varsity,
But also to Josh, Chris, Ali, Nic, Jack, Steve, and Matt, and Duncan and the rest at S & C,
To those teammates who helped me take on the world in Frankfurt, Venice, Wrocław, and Ohio.
Thanks to all those who helped me write, commentate, and grow.

I simply do not have space to give thanks to you all.
So many wonderful people to fit in a space so small.
However I know you, from growing up in York to growing more in Wadham MCR,
Thank you. Thank you for helping me get this far.

But while I'm grateful to all these people, specifically for my degree:
I have to give thanks to Professor Jezzard, and all of my cohort in ONBI.
Thank you all of FMRIB physics group for tolerating my eccentricity,
For helping me feel less lost, thanks to my wonderfully smart reconstruction buddy Sophie.

But the biggest thanks of all has to go to Professor Miller and Dr. Chiew.
I truly, truly, would not be submitting this without you.
I have been in awe of all you've known, amazed at what you've done.
You've shown so much patience and guidance helping me be what I've become.
Through some tough love and kindness, the determination to finish what we started,
Beyond this degree, I will be forever grateful for the wisdom you've imparted.

Wisdom I've hoped to capture, in the pages before you.
So, to the dear soul reading this, I have thanks for you too.

Figure 1: People who mean a lot to me.

Contents

List of Figures

Abbreviations

AUC Area Under the Curve

BOLD Blood Oxygenation Level-Dependent

CC Canonical Correlation

CCS Canonical Correlation Score

CF Cost Function

CNR Contrast-to-Noise Ratio

CS Compressed Sensing

DoF Degrees of Freedom

EPI Echo Planar Imaging

fMRI functional Magnetic Resonance Imaging

FWHM Full Width Half Maximum

GRAPPA Generalized Autocalibrating Partially Parallel Acquisitions

HRF Haemodynamic Response Function

IHT Iterative Hard Thresholding

k-t FASTER FMRI Accelerated in Space-time via Truncation of Effective Rank (in k-t space)

k-t PSF Partially Separable Functions

LRP Low Resolution Priors

MC Matrix Completion

MRI Magnetic Resonance Imaging

NUFFT Non-Uniform Fast Fourier Transform

PCA Principal Component Analysis

PF Partial Fourier

PI Parallel Imaging

PSF Point Spread Function

R The Acceleration Factor

RMSE Root Mean Square Error

ROC Receiver Operating Characteristic

SENSE SENSitivity Encoding

SMS Simultaneous Multi-slice

SNR Signal-to-Noise Ratio

SVD Singular Value Decomposition

T Tesla (magnetic field strength)

T1 Longitudinal relaxation time

T2 Transverse relaxation time

T2* Transverse relaxation time including inhomogeneities

TE Echo Time

tSNR Temporal SNR

TV Total Variation

TR Repetition Time

TURBINE Trajectory Using Radially Batched Internal Navigator Echoes

1 | Introduction

Contents

1.1 Motivation

Magnetic Resonance Imaging (MRI) is a non-invasive imaging technique that can produce detailed internal images of an object, most notably human anatomy. These images can be explicitly weighted by biological function. Dynamic MRI is a broad term that refers to MRI techniques that take multiple images in order to observe the changes in anatomy or biological function over time. fMRI (functional MRI) is a non-invasive, dynamic functional imaging technique that is used to spatially and temporally localise brain activity.

fMRI must balance the temporal resolution and the spatial resolution when reconstructing the dynamic image dataset from the acquired scanner data. Acceleration aims to increase the temporal resolution through higher spatial sampling efficiency in conjunction with advanced image reconstruction techniques that leverage additional information and/or constraints, without harming the image quality at the chosen spatial resolution. An increased sampling efficiency will allow a smaller set of samples to reconstruct the same underlying information with less loss in image fidelity than would be expected compared to conventional reconstruction techniques. By providing increased temporal degrees of freedom in a given scan duration, acceleration can: improve sensitivity to temporal features; reduce physiological noise aliasing; and improve statistical power. Depending on the application, the increased sampling efficiency garnered from acceleration could also be used to reduce scan times, or to increase the spatial resolution (by capturing higher spatial frequencies using the same number of samples per frame, rather than by reducing the number of samples per frame to increase the temporal resolution).

Various acceleration techniques have been widely adopted in fMRI data acquisition. Parallel imaging methods rely on the spatial variation of sensitivity profiles of multi-channel receiver coils, which provide additional spatial information in image reconstruction. This can occur in the image domain (e.g. SENSE [1]) or in the sampling domain (e.g. GRAPPA [2]). These original methods accelerated within a given slice while reducing the impact of the resulting image distortions and aliasing artefacts. Simultaneous Multi-Slice imaging extends these techniques [3, 4] to accelerate across slices, thereby increasing the achievable temporal resolution.

The aforementioned coil sensitivity methods are timepoint-by-timepoint approaches to reconstruction that do not leverage any temporal information for increased acceleration. Methods that do jointly consider k-space and time (k-t methods) can take many forms [5, 6, 7, 8, 9, 10, 11]. By focusing on redundancies and/or structural information in k-t space, k-t methods have the potential for much greater degrees of acceleration than time-independent methods.

One of the exploitable structures in k-t space is the inherent low-rank nature of fMRI [12], which can be understood as the combination of a few spatially coherent temporal processes (i.e. activation maps that identify voxels with a common time-series). fMRI often applies a dimensionality reduction [13] which explicitly enforces a low-rank representation of the system prior to resting-state analysis methods such as independent component analysis (ICA) [14]. It should be noted that various noise sources (e.g. thermal noise, physiological noise, etc.), motion and image artefacts make the system only approximately low rank, although some of the temporally-coherent confounds can be estimated as low-rank processes. This dimensionality reduction gives the final dataset a naturally low mathematical rank, and direct reconstruction of this low-rank representation of the dataset will require far fewer degrees of freedom, representing a high potential for acceleration.

Low-rank methods often split the dataset into a spatial subspace and a temporal subspace, and try to estimate one and then the other. An early such approach in fMRI is Partially Separable Functions (k-t PSF) [15, 16], a data-driven method which rigidly defines the temporal subspace with a hard constraint (taken from a low spatial resolution and high temporal resolution acquisition), before reconstructing the spatial subspace using a simultaneously acquired trajectory with high spatial resolution and low temporal resolution. An alternative approach to combining fMRI reconstruction with a rank-constrained approach was k-t FASTER (fMRI Accelerated in Space-Time via Truncation of Effective Rank [12, 17]), which did not constrain either subspace in order to allow the reconstruction to map the subspaces which best describes the entirety of the acquired data. The rank constraint alone was enough to achieve notable acceleration over other k-t methods [12], but rank-constrained methods may also be combined with coil sensitivity information [17] for increased acceleration.

In addition to the rank and coil sensitivity constraints, other information may also be incorporated into the reconstruction. Tikhonov regularisation enacts a basic overfitting restriction on the temporal and spatial domains, and serves as a way to regularise the energy content of the reconstruction. Regularisation applied to more greatly weight central k-space will utilise the uneven distribution of information in radially acquired k-t datasets. Radial trajectories have a higher spatial sampling density in central k-space than the outer k-space, and so a stronger weighting on low-resolution k-space will allow the reconstruction to be more strongly constrained in the densely sampled region of k-space. The importance of central k-space more generally in MRI reconstruction has previously been used in approaches such as keyhole [10], k-t SPARSE [9], and k-t PCA [7]. The use of central k-space specifically as a soft subspace constraint with radial trajectories in fMRI is novel, as is the use of Tikhonov subspace constraints in fMRI reconstruction (although Tikhonov weighting would not require radial sampling). Temporal smoothing regularisation of some form has previously been incorporated into fMRI reconstruction [18, 19], as temporal smoothing can help remove some of the higher temporal frequency components. If Golden Ratio angle radial sampling is used (such as in TURBINE [20]), the spokes will never overlap and the resulting sampling point spread function will vary between frames. This should produce artefacts which will be especially susceptible to removal by temporal smoothing.

In this thesis, k-t FASTER is reformulated to fit an alternating minimisation framework with incorporated coil sensitivity maps. The alternating minimisation framework allows for straightforward incorporation of subspace constraints without the need for SVDs (Singular Value Decompositions). Tikhonov constraints, low-resolution priors, and temporal subspace smoothing constraints all are explored within that framework as soft constraints. These constrained-subspace low-rank fMRI reconstructions are compared to unconstrained-subspace k-t FASTER and rigidly-constrained k-t PSF reconstructions of retrospectively and prospectively sampled datasets. In particular, these methods are evaluated with regards to accurate representations of the spatial and temporal components, and the accuracy of the recovered statistical parameter maps. The ultimate goal of this thesis is to demonstrate the feasibility of high fidelity reconstructions using the subspace-constrained k-t FASTER reconstruction at a high undersampling factor, rather than to consider the extent of the statistical benefit gained by such a highly accelerated reconstruction.

1.2 Thesis Outline

The work presented in the thesis aims to explore and develop the idea of subspace-constrained low-rank reconstruction of undersampled fMRI data in an alternating minimisation framework. The body of this thesis contains two background chapters which will lay out the motivation behind this approach by examining the current state of the field, and three research chapters which will expand on current approaches (notably k-t FASTER) in order to examine the limits of acceleration achievable in a constrained low-rank framework. Introduction and conclusion chapters will introduce and conclude the thesis respectively, with the conclusion chapter containing any notes regarding future work.

Background Chapters

Chapter 2: MRI will introduce the reader to the basic concepts of MRI signal formation, acquisition, reconstruction, and within-volume acceleration. Then, the core concepts behind how fMRI data is acquired, processed, and accelerated will also be discussed (with more advanced acceleration techniques saved for the next chapter).

Chapter 3: Advanced Image Reconstruction Techniques provides a mathematical grounding to the mathematics of regularised optimisation, and explains how mathematical structures (sparsity and low-rank) can be incorporated into fMRI reconstruction for acceleration purposes.

Research Chapters

In **Chapter 4: Updating the Formulation of k-t FASTER**, the concepts behind the low-rank fMRI approach k-t FASTER are adapted for use in an alternating minimisation framework. Various parameters are explored to ensure an optimal solution is reached while limiting the computation time required. The alternating minimisation framework is better suited for incorporating constraints in the reconstruction, and is tested to ensure that it performs equivalently to the original k-t FASTER approach.

Chapter 5: Exploring Constrained Low-Rank Reconstruction takes the alternating minimisation framework of the previous chapter and applies additional constraints to the reconstruction of an artificial dataset. Tikhonov regularisation, low-resolution priors, and temporal subspace smoothing are implemented, tested, and analysed for the effect they impose on the reconstructed data. This includes a variety of Tikhonov and low-resolution prior implementations that seek to gain further insight into the impact of these constraints on the reconstruction process.

Chapter 6: Implementing Constrained Low-Rank Reconstruction, the subspace-constraints of the previous chapter are tested in realistic fMRI datasets to evaluate the potential use in actual reconstruction. The nature of optimal weightings for the different constraints is also investigated.

Bibliography

[1] Klaas P. Pruessmann, Markus Weiger, Markus B. Scheidegger, and Peter Boesiger. SENSE: Sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, 42(5):952–962, 1999.

[2] Mark A. Griswold, Peter M. Jakob, Robin M. Heidemann, Mathias Nittka, Vladimir Jellus, Jianmin Wang, Berthold Kiefer, and Axel Haase. Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA). *Magnetic Resonance in Medicine*, 47(6):1202–1210, 2002.

[3] Kawin Setsompop, Borjan A. Gagoski, Jonathan R. Polimeni, Thomas Witzel, Van J. Wedeen, and Lawrence L. Wald. Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magnetic Resonance in Medicine*, 67(5):1210–1224, 2012.

[4] Markus Barth, Felix A. Breuer, Peter J. Koopmans, David G. Norris, and Benedikt A. Poser. Simultaneous multislice (SMS) imaging techniques. *Magnetic Resonance in Medicine*, 75(1):63–81, 2016.

[5] Bruno Madore, Gary H. Glover, and Norbert J. Pelc. Unaliasing by Fourier-encoding the overlaps using the temporal dimension (UNFOLD), applied to cardiac imaging and fMRI. *Magnetic Resonance in Medicine*, 42(5):813–828, nov 1999.

[6] Jeffrey Tsao, Peter Boesiger, and Klaas P. Pruessmann. k-t BLAST and k-t SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations. *Magnetic Resonance in Medicine*, 50(5):1031–1042, nov 2003.

[7] Henrik Pedersen, Sebastian Kozerke, Steffen Ringgaard, Kay Nehrke, and Yong Kim Won. k-t PCA: Temporally constrained k-t BLAST reconstruction using principal component analysis. *Magnetic Resonance in Medicine*, 62(3):706–716, 2009.

[8] Feng Huang, James Akao, Sathya Vijayakumar, George R. Duensing, and Mark Limkeman. K-t GRAPPA: A k-space implementation for dynamic MRI with high reduction factor. *Magnetic Resonance in Medicine*, 54(5):1172–1184, 2005.

[9] Michael Lustig, David L. Donoho, and John M. Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007.

[10] Joop J. Van Vaals, Marijn E. Brummer, W. Thomas Dixon, Hans H. Tuithof, Hans Engels, Rendon C. Nelson, Brigid M. Gerety, Judith L. Chezmar, and Jacques A. Den Boer. “Keyhole” method for accelerating imaging of contrast agent uptake. *Journal of Magnetic Resonance Imaging*, 3(4):671–675, jul 1993.

[11] Cristina Chavarrías, J. F. P. J. Abascal, P. Montesinos, and M. Desco. Exploitation of temporal redundancy in compressed sensing reconstruction of fMRI studies with a prior-based algorithm (PICCS). *Medical Physics*, 42(7):3814–3821, 2015.

[12] Mark Chiew, Stephen M. Smith, Peter J. Koopmans, Nadine N. Graedel, Thomas Blumensath, and Karla L. Miller. k-t FASTER: Acceleration of functional MRI data acquisition using low rank constraints. *Magnetic Resonance in Medicine*, 74(2):353–364, 2015.

[13] Martin J. McKeown, Tzyy Ping Jung, Scott Makeig, Greg G. Brown, Sandra S. Kindermann, Te Won Lee, and Terrence J. Sejnowski. Spatially independent activity patterns in functional MRI data during the Stroop color-naming task. *Proceedings of the National Academy of Sciences of the United States of America*, 95(3):803–810, 1998.

[14] Aapo Hyvärinen. Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. *IEEE Transactions on Neural Networks*, 10:626–634, 1999.

[15] Zhi-Pei Liang. Spatiotemporal Imaging with Partially Separable Functions. *IEEE International Symposium on Biomedical Imaging*, 2:988–991, 2007.

[16] Fan Lam, Bo Zhao, Yinan Liu, Zhi-Pei Liang, Michael Weiner, and Norbert Schuff. Accelerated fMRI using Low-Rank Model and Sparsity Constraints. In *Proceedings of the International Society for Magnetic Resonance in Medicine 21*, volume 21, page 2620, 2013.

[17] Mark Chiew, Nadine N. Graedel, Jennifer A. Mcnab, Stephen M. Smith, and Karla L. Miller. Accelerating functional MRI using fixed-rank approximations and radial-cartesian sampling. *Magnetic Resonance in Medicine*, 00:1–12, 2016.

[18] Xuesong Li, Xiaodong Ma, Lyu Li, Zhe Zhang, Xue Zhang, Yan Tong, Lihong Wang, Sen Song, and Hua Guo. Dual-TRACER: High resolution fMRI with constrained evolution reconstruction. *NeuroImage*, 164(February 2017):172–182, 2018.

[19] Mark Chiew and Karla L. Miller. Improved statistical efficiency of simultaneous multi-slice fMRI by reconstruction with spatially adaptive temporal smoothing. *NeuroImage*, 203(August):116165, 2019.

[20] Nadine N. Graedel, Mark Chiew, Jennifer A. McNab, and Karla L. Miller. FMRI using a 3D radial-Cartesian trajectory: spatio-temporal tunability and artefact correction. In *Proceedings of the International Society for Magnetic Resonance in Medicine 23*, volume 23, page 1, 2015.

2 | MRI

Contents

This chapter aims to introduce the reader to the field of MRI by covering the key concepts of MRI physics, MRI reconstruction, and fMRI-specific reconstruction and analysis. This will form the groundwork for the later chapters in this thesis.

2.1 Data Acquisition and Signal Formation

2.1.1 Origin of the Signal

Nuclei with an odd number of neutrons and/or protons possess nuclear spin angular momentum, a property more commonly known as spin. The most abundant source of spin signal in human biology is ^{1}Hydrogen, a magnetically-sensitive nucleus which is most commonly found in water as H_2O. Spin (S) is linearly proportional to the magnetic dipole moment μ. This proportional relation is linked by the gyromagnetic ratio γ, a fundamental nuclei-specific constant which relates the ratio between the magnetic moment and the angular momentum of a system. For hydrogen, $\gamma_H = 267.5$ Mrad s^{-1} T^{-1}.

$$\mu = \gamma S \tag{2.1}$$

In the absence of an external magnetic field, spins are randomly orientated and the net magnetic moment is zero. If a sample containing hydrogen is placed in an external magnetic field, a net magnetic moment is formed in that sample in the direction of that field. In MRI, the static external field (B_0) is generated by a large external magnet; typical B_0 field strengths are 1.5T or 3T. The direction of the B_0 field defines the longitudinal axis (or z-axis), which the nuclei rotate around in an act known as precession. The rate of change of angular momentum ($\frac{dS}{dt}$) is defined as the product between the magnetic moment and the combined magnetic influence of all external fields (B):

$$\frac{dS}{dt} = \mu \times B \tag{2.2}$$

Physically, equation 2.2 states that the axis of precession will never change unless an external field is applied in any direction that is neither parallel nor anti-parallel to μ. Multiplying both sides of equation 2.2 by the gyromagnetic ratio γ and summing over a unit volume produces a relationship between M (the net magnetic moment) and the rate of change of M. This new equation is a dynamic macroscopic description of the relationship between the external magnetic field B and the net magnetic moment.

$$\frac{dM}{dt} = M \times \gamma B \tag{2.3}$$

The solution produced by equation 2.3 defines the characteristic frequency known as the Larmor frequency (ω). The Larmor frequency is the natural precession frequency for nuclear spins in a given magnetic field, with the proportionality again characterised by the gyromagnetic ratio.

$$\omega = \gamma B \tag{2.4}$$

Equation 2.4 is known as the Larmor equation, and was derived in 1897 [1] to explain the splitting of optical spectral lines due to electron resonance (rather than Nuclear Magnetic Resonance, which would not be described until 1938 by Rabi et al. [2]). Regardless, Larmor demonstrated that electrons should precess in the direction of a magnetic field and while laying out the mathematical basis that later MRI research would build on. Appreciating how varying the magnetic field can modulate the characteristic frequency ω is vital to understanding how magnetisation can be manipulated to produce an MR image.

2.1.2 RF Pulses

Radio Frequency (RF) energy is typically defined by a frequency range of 10^4 to 10^{11} rad s^{-1}, and a magnetic pulse which oscillates within the RF frequency range is called an RF pulse. RF pulses at the Larmor frequency can be used to "excite" the magnetisation, provided that the applied pulse has a non-zero magnitude in the transverse direction to the current axis of spin, causing the spin to be "tipped" away from the current precession angle (see equation 2.3). The pulse can be thought of as a torque being applied to the system. The angle at which the net magnetic moment is tipped is referred to as the flip angle. The temporal spacing between RF pulses is referred to as the repetition time (TR).

After the RF pulse is applied, the system will recover slowly into an equilibrium state of realignment with the B_0 field. The longitudinal magnetisation recovery $M_z(t)$ and the transverse magnetisation decay $M_{xy}(t)$ are respectively governed by the T1 and T2 time constants, or equivalently by the R1 and R2 rate constants. The equilibrium magnetisation conditions for a fully relaxed system are $M_z = M_0$ and $M_{xy} = 0$. A 90° RF pulse would result in the field being fully tipped: $M_z = 0$ and $M_{xy} = M_0$. Other flip angles and initial magnetisation states result in intermediate magnetisation conditions for M_{xy} and M_z.

The longitudinal magnetisation recovery $M_z(t)$ is defined by the rate at which the longitudinal component of the magnetisation M_z can recover to the equilibrium state M_0 from the post-tip magnetisation $M_z(0)$. The T1 time lengthens as field strength increases.

$$M_z(t) = M_0 + (M_z(0) - M_0)e^{-t/T_1} \tag{2.5}$$

The transverse magnetisation decay describes the degeneration of the transverse component of the tipped magnetisation. This decay occurs due to the loss of phase coherence (dephasing), and is weakly dependent on the field strength.

$$M_{xy}(t) = M_0 e^{-t/T_2} \tag{2.6}$$

It can be assumed for all tissues that T2 $\leq$ T1, i.e. that transverse decay occurs quicker than longitudinal recovery. Typical values at 3T define T1 on the order of 1000ms, and T2 on the order of 100ms for grey and white matter (with the precise value dependent on tissue type and many other factors) [3]. These dynamics can be combined to describe the nuclear magnetisation of the full system with the Bloch equation, observed phenomenologically by Bloch in 1946 [4] ($\mathbf{i}$ and $\mathbf{j}$ are transverse axes, $\mathbf{k}$ is the longitudinal axis).

$$\frac{dM}{dt} = M \times \gamma B - \frac{M_x(t)\mathbf{i} + M_y(t)\mathbf{j}}{T_2} - \frac{(M_z(t) - M_0)\mathbf{k}}{T_1} \tag{2.7}$$

The Bloch equation is sometimes referred to as the equation of motion for nuclear magnetisation. In the absence of relaxation, equation 2.7 reduces to equation 2.3, which defines the dynamic properties of M through simple precession in the presence of an external magnetic field. For most situations, the Bloch equation accurately describes the system, although extensions exist to capture additional dynamic magnetisation properties where required (e.g. Bloch-McConnell [5] which encodes chemical exchange).

2.1.3 Detecting the Signal

A receiver RF coil tuned to the appropriate frequency will measure locally-generated magnetic signal as an induced electrical current via Faraday's law. There are typically many receiver coils, with each coil receiving a slightly different signal through receptive fields that are distributed over the whole object in unique but overlapping spatial sensitivity profiles. These coil spatial sensitivity distributions can be exploited in parallel imaging (section 2.2.2).

Nuclei which are perturbed by a weak oscillating magnetic pulse produce an electromagnetic signal response at a characteristic frequency, a phenomenon known as Nuclear Magnetic Resonance (NMR). The signal is generated purely through the naturally decaying precession in the transverse plane (equation 2.6), and is known as the Free Induction Decay (FID, figure 2.1). FID is the most basic signal in MRI and was first described by Erwin Hahn in 1949 [6].

Additional gradient fields and RF pulses can be used to generate signals with a peak after the

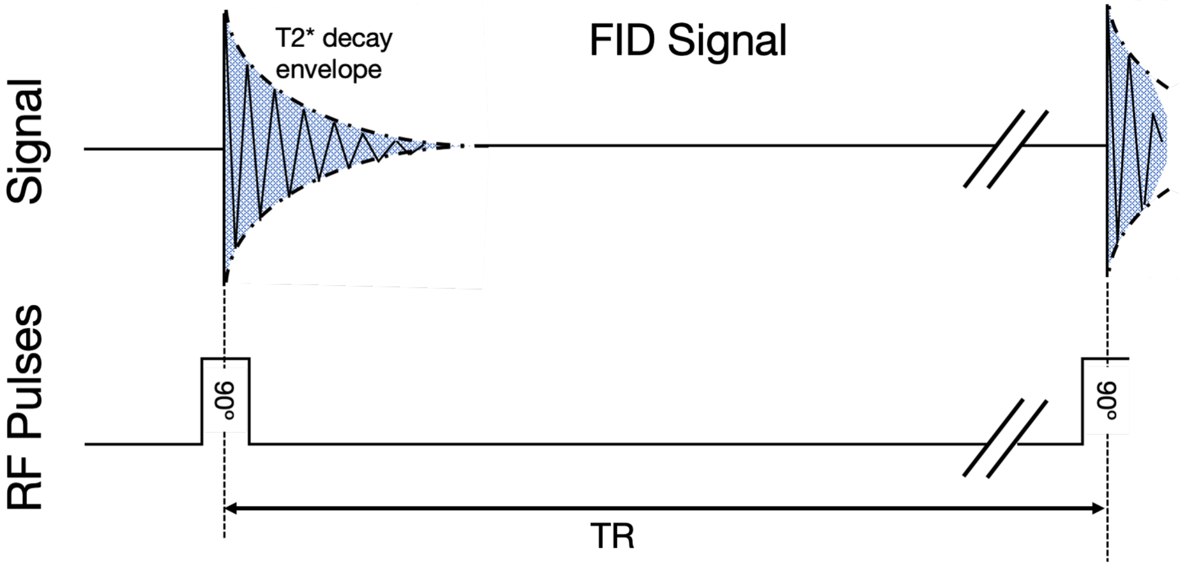


Figure 2.1: The FID pulse sequence and signal. An initial 90° RF pulse tips the magnetisation to form a signal in the transverse domain, which decays according to a T2* constant.

initial excitation (an echo). The time between the initial excitation and the peak is called the Echo Time (TE). The precise timings of RF and gradient pulses are collectively referred to as a pulse sequence.

2.1.3.1 Gradient Echo

The Gradient Recalled Echo (GRE) sequence (first explicitly described by Abragam in 1961 [7]) applies a linearly varying field known as a dephasing gradient after the RF pulse. The gradient systematically dephases and rephases the spins to allow an image to form. After the signal is received the remaining magnetisation can be either RF-spoiled or gradient-spoiled to destroy any remaining transverse magnetisation and reduce the overall TR.

A small flip angle is typically used in GRE imaging, allowing for short TRs since the longitudinal magnetisation is mostly preserved and so requires a shorter recovery time. The optimal flip angle (in terms of Signal-to-Noise Ratio) for a given TR and T1 is defined by the Ernst angle [8]:

$$\theta_E = acos(e^{-TR/T1}) \tag{2.8}$$

A GRE sequence is susceptible to field inhomogeneities giving rise to an effective T2 value known as T2*; sequences that are more susceptible to field inhomogeneities will have a T2* that diverges further from T2 value. The influence of the inhomogeneities on the transverse decay can be equivalently referred to as T2' or $1/\gamma\Delta B_0$ (the difference in local field strength). Inhomogeneities can be caused by both macroscopic and microscopic sources.

$$\frac{1}{\mathrm{T2}^*} = \frac{1}{\mathrm{T2}} + \frac{1}{\mathrm{T2'}} \tag{2.9}$$

2.1.3.2 Spin Echo

The Spin Echo sequence (SE) uses a 180° RF pulse to flip the magnetisation. Up until the TE/2 mark where the pulse is applied, field inhomogeneity will cause the magnetisation to dephase (figure 2.2). The RF pulse flips the spins, which causes the spins to rephase at the same rate they were previously dephasing, as all local inhomogeneity effects are reversed. This causes the spins to form a coherent echo at TE seconds after the RF pulse. The coherent echoes make SE sequences a true measure of T2-weighted images. SE was discovered by Hahn in 1950 [9]. Subsequent refocusing pulses can also be used to form secondary echoes (still working on the tipped excitation from the original 90° RF pulse).

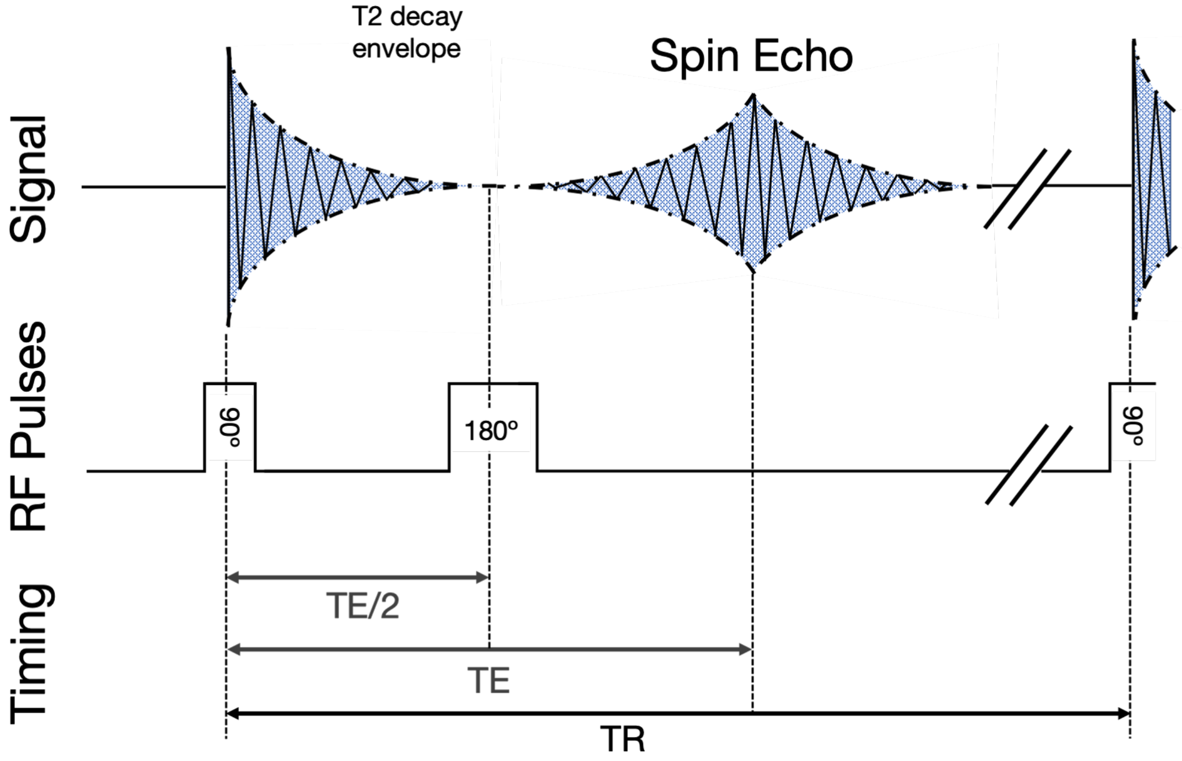


Figure 2.2: The SE pulse sequence and signal. An initial 90° tips the magnetisation to form a signal, which slowly decays. A 180° RF pulse after TE/2 seconds reverses the spins, which form a coherent echo and then decay again. The time between 90° pulses is the TR. The reduction in the echo peak compared to the initial signal magnitude is caused by T2 decay.

2.1.4 Spatial Encoding

The previous sections discussed how to create a non-selective signal from an object in a magnetic field. A non-selective sequence excites all the tissue, and so all the tissue is generating signal. This section will deal with creating spatial localisation within the tissue, with the spatial localisation information then being used to form an image.

Gradient coils are loops of wire within an MR scanner that are capable of producing spatially varying magnetic fields, resulting in a spatially varying precession frequency. This pattern of spin variation can be used to encode physical location into the MR signal detected by the receiver coils. Linear gradient coils are assumed here.

The expression of the Larmor equation (equation 2.4) in the presence of gradient fields is:

$$\omega = \gamma B_0 + \gamma G.r = \gamma B_0 + \gamma(G_x.x + G_y.y + \gamma G_z.z) \tag{2.10}$$

G refers to the gradient field and is often recorded in units of mT m^{-1}. G is a three-element vector made up of the orthogonal elements G_x, G_y, and G_z, with G_z orientated in the direction of B_0. r is a three-element vector of location (x, y, and z).

2.1.4.1 Slice Selection

In non-selective excitation, a homogenous B will produce identical precession in all locations, as all locations are excited by the excitation pulse. Once tipped, all locations will produce a signal at the same frequency. By applying a gradient along a given direction while transmitting an RF pulse at a specific frequency, only a slice of the object (which is perpendicular to the applied gradient) will still be precessing near enough to the RF frequency that it will be excited. The thickness of the slice is proportional to the RF pulse bandwidth and inversely proportional to the applied gradient strength. The rest of the tissue will be precessing at either higher or lower frequencies, and so the spins will not be tipped. This general process of exciting only a specific area is called “selective excitation”, with the term “slice selection” referring specifically to selective excitation used to excite a thin plane.

By convention, the slice direction is often designated as the z-axis, but any axis of excitation can be achieved through the correct combination of the three gradients. After the excitation process described above, the entire excited plane is producing the same signal. This makes it impossible to distinguish contrast within this plane. Using gradients after the excitation can help by producing a set of spatial frequency coefficients which collectively describe the excited plane. In MRI, the spatial frequency domain is referred to as k-space.

2.1.4.2 k-Space

The concepts of time and frequency are mathematically linked through a Fourier Transform (FT). The frequency content of any time-based signal can be uncovered with a 1D FT.

$$G(f) = F(g(t)) = \int_{-\infty}^{\infty} g(t)e^{-i.2\pi.t.f}dt \tag{2.11}$$

F represents the FT, $g(t)$ is an arbitrary time-based signal, and $G(f)$ is the frequency spectra which characterises $g(t)$. The FT is reversible, meaning the inverse FT can yield a time signal from a frequency-based signal.

$$g(t) = F^{-1}(G(f)) = \int_{-\infty}^{\infty} G(f)e^{+i.2\pi.t.f}df \tag{2.12}$$

The FT is a continuous linear transform with the following properties worth noting among many interesting features of the transform:

- $F(g(t-t_0)) = G(f)e^{-i.2\pi.f.t_0}$: A linear shift in one domain is a phase shift in the other
- $F(g(t)*h(t)) = G(f) \times H(f)$: A convolution operation between two functions in one domain is a multiplication operation in the other
- $F(G(t)) = g(-f)$: The duality property

The mapping of the 1D relationship between time and temporal frequency represents a common use of Fourier analyses, although the FT can be used to analyse any parameter and the equivalent reciprocal domain. The FT can also be applied across more than one dimension, e.g. a 2D FT applied to a two-dimensional spatial domain produces a representation of 2D k-space (spatial frequencies). Equation 2.13 shows a continuous 2D FT and equation 2.14 shows a continuous 2D inverse FT.

$$G(k_x, k_y) = F(g(x,y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)e^{-i.2\pi.(x.k_x+y.k_y)}dxdy \tag{2.13}$$

$$g(x,y) = F^{-1}(G(k_x, k_y)) = \int_{-\infty}^{\infty} G(k_x, k_y)e^{+i.2\pi.(x.k_x+y.k_y)}dk_xdk_y \tag{2.14}$$

The common use of the letter k in the algebra led to the representation of spatial frequencies being called "k-space". k_x and k_y represent orthogonal spatial frequencies in the x and y directions respectively. The coefficients of $G(k_x,k_y)$ corresponds to the degree that a given spatial frequency (k_x,k_y) is represented in the spatial signal $g(x, y)$. This is equivalent to the coefficients of $g(x, y)$ reflecting the signal energy at location (x,y). All of the 1D properties of the Fourier transform hold in two or more dimensions.

A discrete representation of the Fourier transform is required for MRI. However, the discretisation of a continuous signal will typically lead to an inherent loss. The spatial resolution is a sampling property which defines the distance at which two peaks in the continuous signal can be distinguished. The image resolution $(\frac{1}{\Delta x})$ is the resolution at which the sampled signal is displayed, and is defined by the range of k-space sampled $(2k_{max})$.

$$2k_{max} = \frac{1}{\Delta x} \tag{2.15}$$

Relatedly, the resolution in k-space (Δk) will define the field of view (FOV) of the final image. The chosen k-space resolution Δk must be small enough to encompass the desired FOV (e.g. that the FOV is wide enough to capture the whole brain if one is attempting to image the whole head).

$$\frac{1}{\Delta k} = FOV \tag{2.16}$$

The sampling process acquires a finite set of data that attempts to represent the underlying continuous truth. The continuous k-space signal $s(k_x, k_y)$ is sampled by a 2D Dirac comb ${}^{2D}\mathrm{III}(\frac{k_x}{\Delta k_x}, \frac{k_y}{\Delta k_y})$, an infinite string of equally spaced delta functions in 2D k-space separated by repeated distances of Δk_x and Δk_y. A finite sampling extent is represented by multiplication with a 2D box function ${}^{2D}\sqcap(\frac{k_x}{W_{kx}}, \frac{k_y}{W_{ky}})$ which is designed to extend half a sampling period beyond the highest k-space frequencies in each dimension (e.g. the box width in x is $W_{kx} = 2(k_{x\ max} + \frac{\Delta k_x}{2})$).

$$S_{sampled}(k_x, k_y) = s(k_x, k_y) \left(\frac{1}{\Delta k_x \Delta k_y}\right)^2 {}^{2D}\mathrm{III}\left(\frac{k_x}{\Delta k_x}, \frac{k_y}{\Delta k_y}\right) {}^{2D}\sqcap\left(\frac{k_x}{W_{kx}}, \frac{k_y}{W_{ky}}\right) \tag{2.17}$$

In fully-sampled Cartesian MRI, k-space is sampled using N_{fe} lines in the x-direction and N_{pe} lines in the y-direction. The fully-sampled k-space can then be transformed into an image through an inverse Discrete Fourier Transform (inverse DFT).

$$s_{x,y} = \sum_{k_x=-N_{fe}/2+1}^{N_{fe}/2} \sum_{k_y=-N_{pe}/2+1}^{N_{pe}/2} S_{sampled}(k_x, k_y) e^{-i.2\pi.(k_x.x/N_{fe}+k_y.y/N_{pe})} \tag{2.18}$$

The sampled k-space signal can also be formulated as the inverse problem, where the signal sampled in k-space (S) is formed by a DFT (E) applied to some underlying unknown true image I. Given that this equation deals with discrete variables, this linear equation can be presented more neatly in matrix form.

$$S = EI \tag{2.19}$$

In a 1D example for a vector of length N, the DFT-encoding E for fully-sampled k-space is an N by N matrix, whose entries are defined as $E_{j,k} = e^{-i2\pi\frac{jk}{N}}, -N/2+1 \geq j,k \geq N/2$. E can then be represented as the following matrix:

$$E = \frac{1}{\sqrt{N}} \begin{pmatrix} e^{-i2\pi\frac{(N/2-1)(N/2-1)}{N}} & \cdots & e^{-i2\pi\frac{N/2-1}{N}} & 1 & e^{+i2\pi\frac{N/2-1}{N}} & e^{+i2\pi\frac{(2)(N/2-1)}{N}} & \cdots & e^{+i2\pi\frac{(N/2)(N/2-1)}{N}} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ e^{-i2\pi\frac{N/2-1}{N}} & \cdots & e^{-i2\pi\frac{1}{N}} & 1 & e^{+i2\pi\frac{1}{N}} & e^{+i2\pi\frac{2}{N}} & \cdots & e^{+i2\pi\frac{N/2}{N}} \\ 1 & \cdots & 1 & 1 & 1 & 1 & \cdots & 1 \\ e^{+i2\pi\frac{N/2-1}{N}} & \cdots & e^{+i2\pi\frac{1}{N}} & 1 & e^{-i2\pi\frac{1}{N}} & e^{-i2\pi\frac{2}{N}} & \cdots & e^{-i2\pi\frac{N/2}{N}} \\ e^{+i2\pi\frac{(N/2-1)(2)}{N}} & \cdots & e^{+i2\pi\frac{2}{N}} & 1 & e^{-i2\pi\frac{2}{N}} & e^{-i2\pi\frac{(2)(2)}{N}} & \cdots & e^{+i2\pi\frac{(2)(N/2)}{N}} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ e^{+i2\pi\frac{(N/2-1)(N/2)}{N}} & \cdots & e^{+i2\pi\frac{N/2}{N}} & 1 & e^{-i2\pi\frac{N/2}{N}} & e^{-i2\pi\frac{(N/2)(2)}{N}} & \cdots & e^{-i2\pi\frac{(N/2)(N/2)}{N}} \end{pmatrix} \tag{2.20}$$

I and S in equation 2.19 are typically vectorised versions of 2D matrices (of size $xy \times 1$ and $k_x k_y \times 1$). E then requires some rearranging, but can still represent a 2D DFT in matrix form (of size $k_x k_y \times xy$). Practically, the most efficient way of implementing a discrete Fourier transform is the Fast Fourier transform (FFT), which factorises the encoding matrix E into a set of factors, many of which are zero (this representation is called sparse). This sparse representation simplifies the DFT complexity from $\mathrm{O}(N^2)$ to $\mathrm{O}(N \log N)$.

2.1.4.3 Aliasing

Aliasing is the mismapping of multiple signals onto the same image space location due to missing information in the k-space domain. The minimum sampling rate required is defined by the Nyquist-Shannon criterion, which states that sampling must occur at twice the rate of the maximum frequency in the underlying truth that the sampler is hoping to represent. That is, from equation 2.16, Nyquist is satisfied along the x-axis if $\Delta k_x \geq \frac{1}{N_{fe}\Delta x}$ for an image with N_{fe} lines.

In general, aliasing in k-space can be represented as the convolution of the image with the point spread function (PSF) of a given sampling pattern. For many patterns, such as those depicted in figure 2.3, the PSF can be readily derived using an FT of the sampling pattern. However, PSFs are an entirely generalisable way of representing the resulting artefacts from a broad range of sampling patterns. Artefacts are any feature which appear in an image but are not present in the original object.

Figure 2.3 shows some characteristic sampling patterns, their corresponding PSFs, and a phantom image sampled with the equivalent sampling patterns. Undersampling leads to replications of the underlying image (Nyquist ghosts), which overlap with the main image (leading aliasing to also be known as wrap-around). Aliasing can be represented as repeated peak intensities in the corresponding point spread function, which define the centre for each aliased copy. The very central point in k-space represents the mean intensity of an image.

Undersampling will not necessarily produce the aliasing artefacts of figure 2.3, a fact demonstrated in figure 2.4. The centre of k-space defines the low-frequency information of an image and, if isolated, will produce a blurred version of that image. The outer portion of k-space defines the high-frequency information (the detail) of an image and represents rapid changes such as edges within the image.

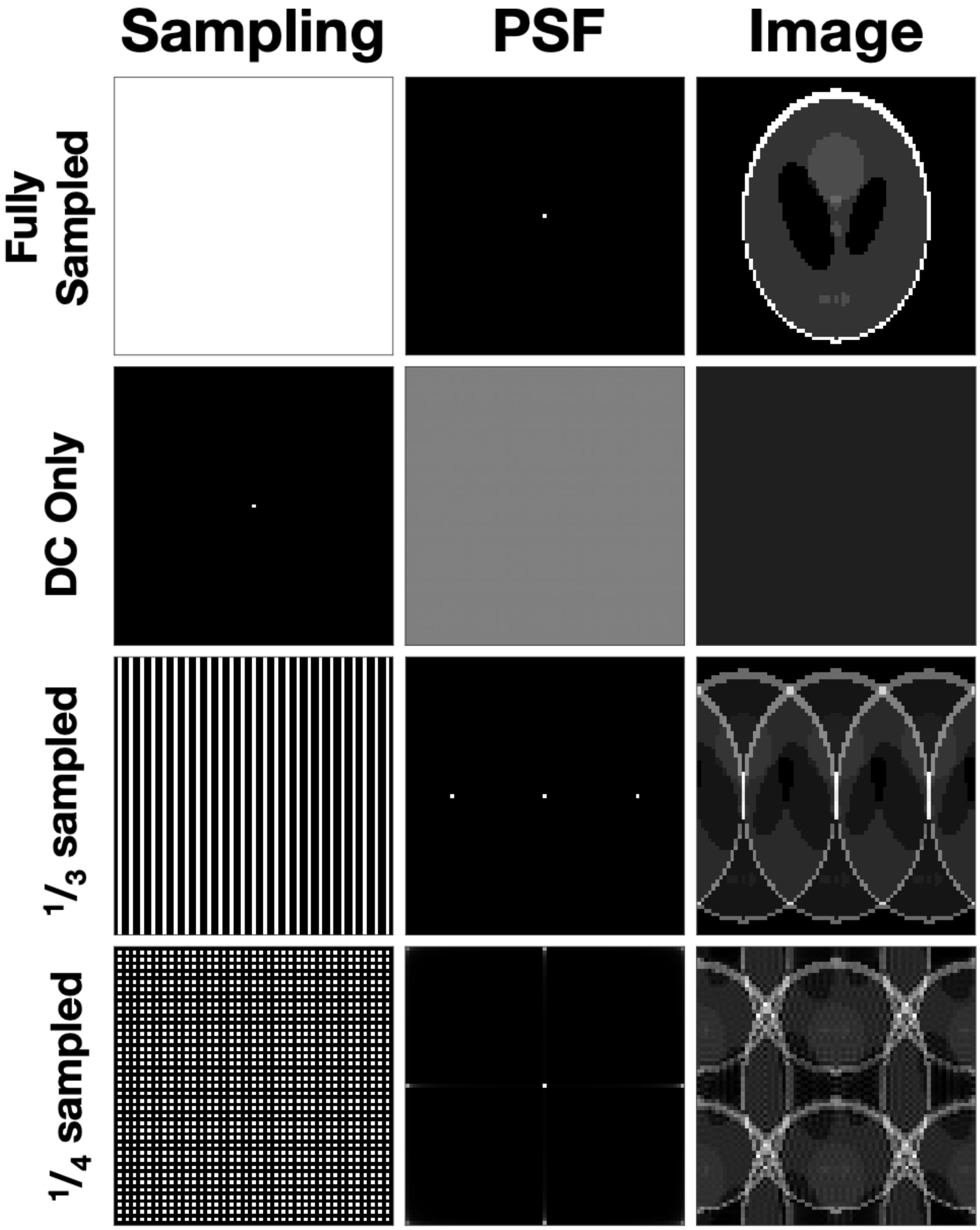


Figure 2.3: A representation of common Cartesian aliasing patterns. The first column shows sampling patterns, where a white pixel represents a sample, and a black pixel represents an unsampled k-space location. The second column displays the corresponding PSF in image space. The final column shows a visual representation of a phantom image being sampled by the associated sampling pattern in k-space. Top row: A fully sampled representation of k-space. 2nd row: Only the central frequency in k-space is sampled. 3rd row: Only 1/3 of the required points are sampled (an undersampling factor of 3 in k_x). Bottom row: Only 1/4 of the required points are sampled (an undersampling factor of 2 in both k_x and k_y)

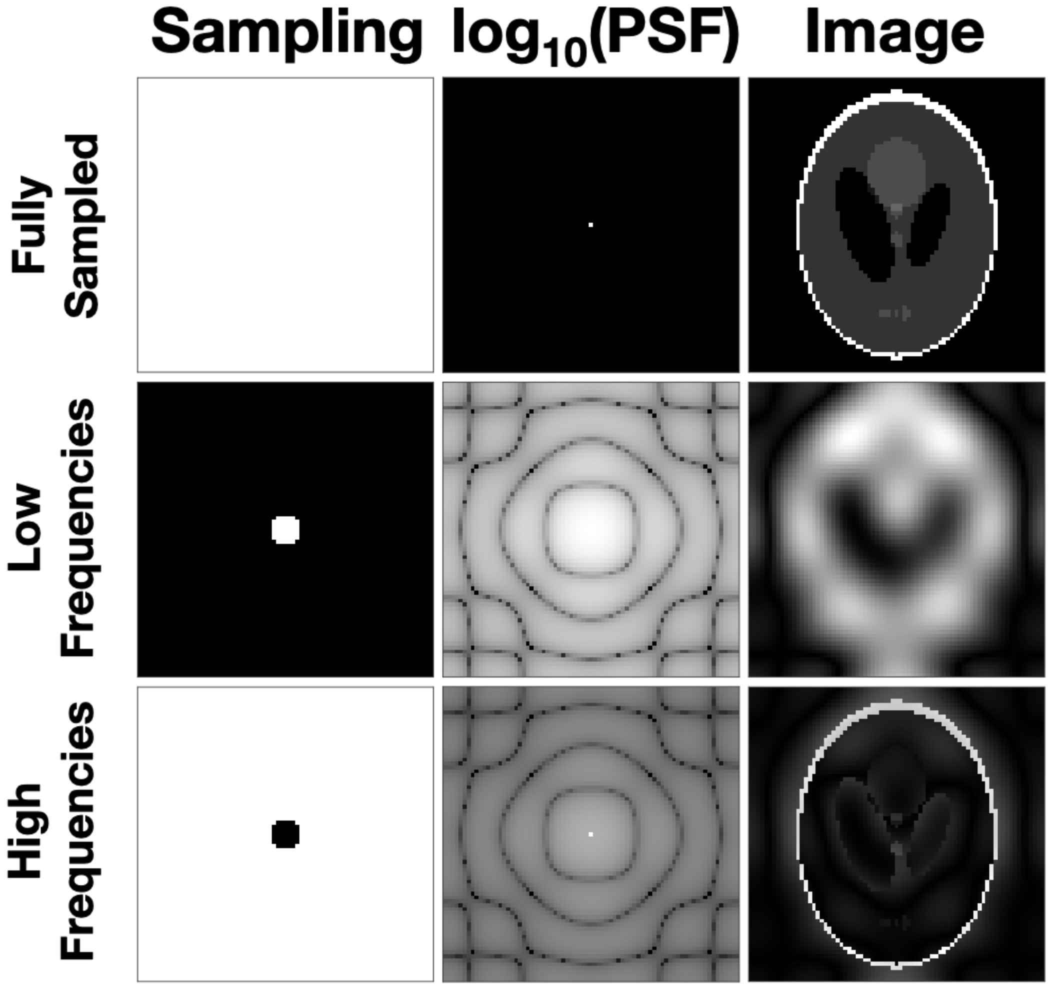


Figure 2.4: A demonstration of the separate contributions of high and low spatial frequencies within an image. The first column shows sampling patterns, where a white pixel represents a sample, and a black pixel represents an unsampled k-space location. The second column displays the $\log_{10}$ transform of the PSF in image space. The final column shows a visual representation of a phantom image being sampled by the associated sampling pattern in k-space. Top row: A fully sampled representation of k-space. Middle row: Only the central low-resolution frequencies are sampled. Bottom row: Only the outer high-resolution frequencies are sampled.

2.1.4.4 Spatial Phase Encoding

At the start of section 2.1.4, it was demonstrated that a linear gradient applied during the RF excitation pulse can ensure only spins within a desired plane are excited. In this section, we describe how linear gradients can be applied after the RF excitation pulse to produce k-space information.

Once the spins are excited, they will all start precessing at the same frequency. If a linear gradient is applied, then a linearly-varying frequency pattern is generated in the subject via equation 2.4. This means that nuclei on the higher end of the gradient field will spin at a faster frequency than nuclei located in a position to experience a lower field strength. The phase (Φ) accumulates over time, but the amount of phase-accumulation is frequency-dependent.

$$\Phi(t) = \int_0^t \omega(\tau) d\tau \tag{2.21}$$

A greater frequency will then result in a greater amount of phase accumulation over a well-defined spatial pattern, and this phase accumulation is then used to encode additional spatial information in the signal. The measured signal in the coil S(k_x,k_y,t) is phase-dependent.

$$S(k_x, k_y, t) = \int_x \int_y m(x, y) e^{-t/T2} e^{-i.2\pi.\Phi(x,y,t)} dxdy \tag{2.22}$$

The phase at a particular point in time is expressed as the integral of the applied gradients over time by noting the relationship between ω at a given location (x,y) and the applied transverse gradient field.

$$\Phi(x, y, t) = \int_0^t \Big(\omega(x, y, \tau)\Big) d\tau = \gamma \int_0^t \Big(xG_x(\tau) + yG_y(\tau)\Big) d\tau \tag{2.23}$$

The following substitutions can be used to represent the signal in k-space:

$$\begin{aligned} k_x(t) &= \frac{\gamma}{2\pi} \int_0^t G_x(\tau) d\tau \\ k_y(t) &= \frac{\gamma}{2\pi} \int_0^t G_y(\tau) d\tau \end{aligned} \tag{2.24}$$

Which transforms equation 2.22 into the form of a forward Fourier transform:

$$S(k_x, k_y, t) = \int_x \int_y m(x,y) e^{-t/T2} e^{-i.2\pi.(k_x(t).x + k_y(t).y)} dxdy \tag{2.25}$$

The substitution equations 2.24 are important because they allow the integral of G_x or G_y to define a point in k-space, meaning a non-zero gradient will change the k-space co-ordinate which is being represented. The signal in the receiver coil at that particular point in time then represents the k-space coefficient across the excited tissue. A simplification of the signal equation (equation 2.25) can present the signal as the inner product of the object magnetisation and the encoding by the gradients.

The key outcome of equation 2.25 is that varying gradients can create a spatially encoded signal. A simple demonstration of this is shown in figures 2.5 and 2.6 to demonstrate GRE and SE imaging respectively. A dephasing gradient is initially needed so that the k-space coefficient locations covered are symmetric in the readout direction.

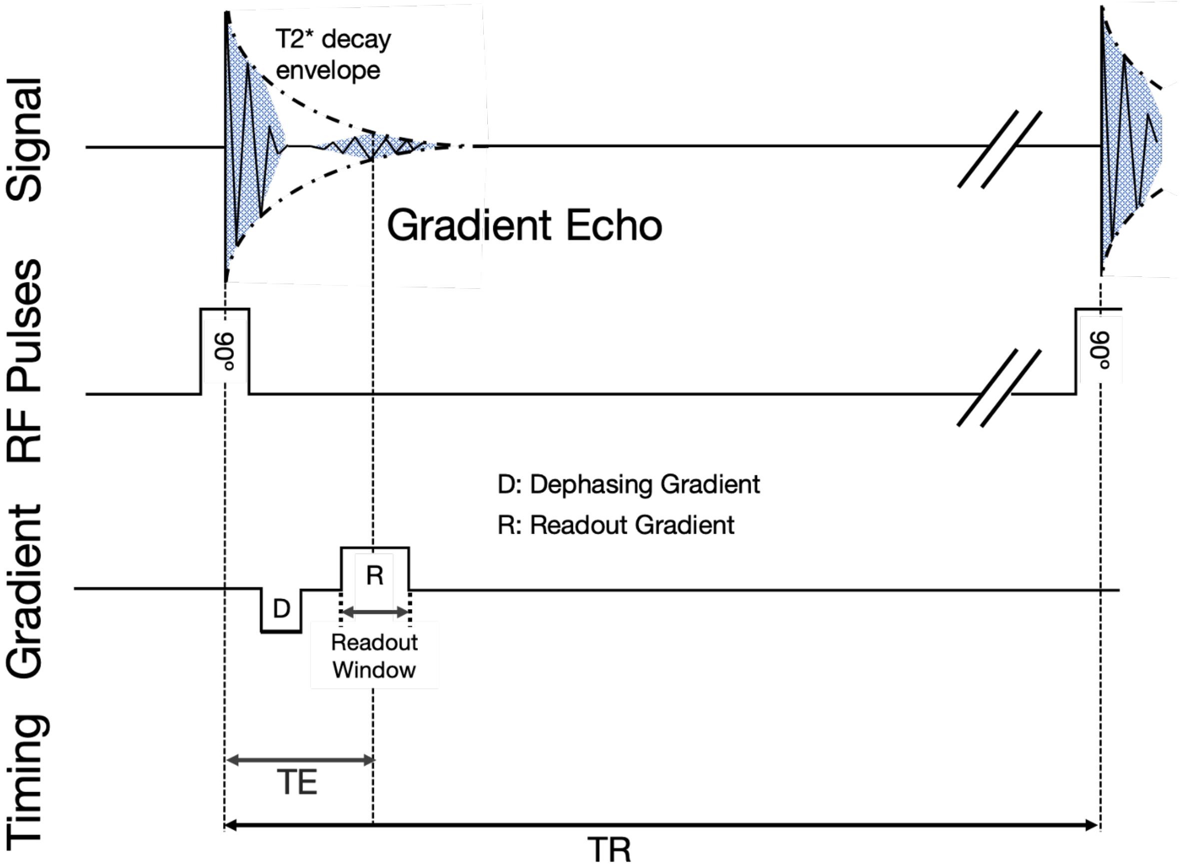


Figure 2.5: The GRE pulse sequence and signal. An initial 90° tips the magnetisation to form a signal, with which an echo is rapidly formed through a linear gradient, which could allow for a rapid TR (especially compared with figure 2.6.

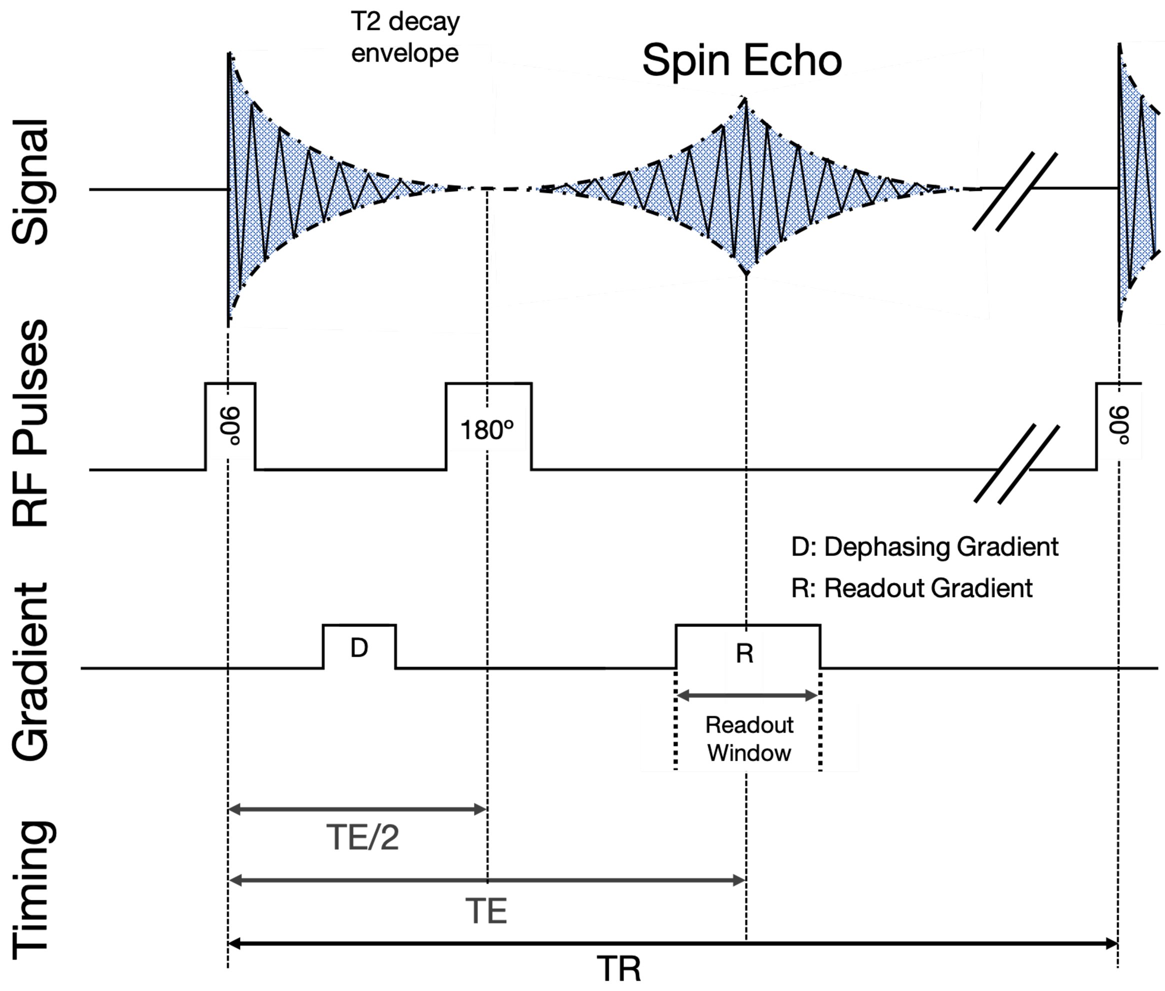


Figure 2.6: The SE pulse sequence and signal, with readout gradients included. An extension of figure 2.2.

2.1.5 k-Space Encoding

k-Space coverage can be broken down into two major types of trajectory – Cartesian and non-Cartesian. The key features a user has to consider in selecting a trajectory are: a) the simplicity of the reconstruction, b) the types of artefacts produced, c) the uniformity of k-space coverage, and d) the time required to acquire the data necessary to produce an image. There is variation within Cartesian and non-Cartesian trajectories for these choices, but in general non-Cartesian reconstructions use a more complicated reconstruction in exchange for more incoherent artefacts.

2.1.5.1 Cartesian

Cartesian sampling collects samples on a Cartesian grid, with the grid axes split into a frequency/readout-encoding axis and a phase-encoding axis (see figure 2.7). The frequency-encode direction is defined by the straight lines along which data is sequentially acquired in k-space. All non-frequency-encoding axes are called phase-encode axes. The readout is achieved with a constant gradient that is applied along this direction while signal is collected. For example, if G_x is constant and G_y is zero, equation 2.24 states that there will be a linear increase in k_x over time. Before a frequency-encoding line is collected, gradient/s in the phase-encode direction/s are applied to move the readout location from the centre of k-space to the starting point of a frequency-encode line. A grid of k-space coefficients is then collected line by line.

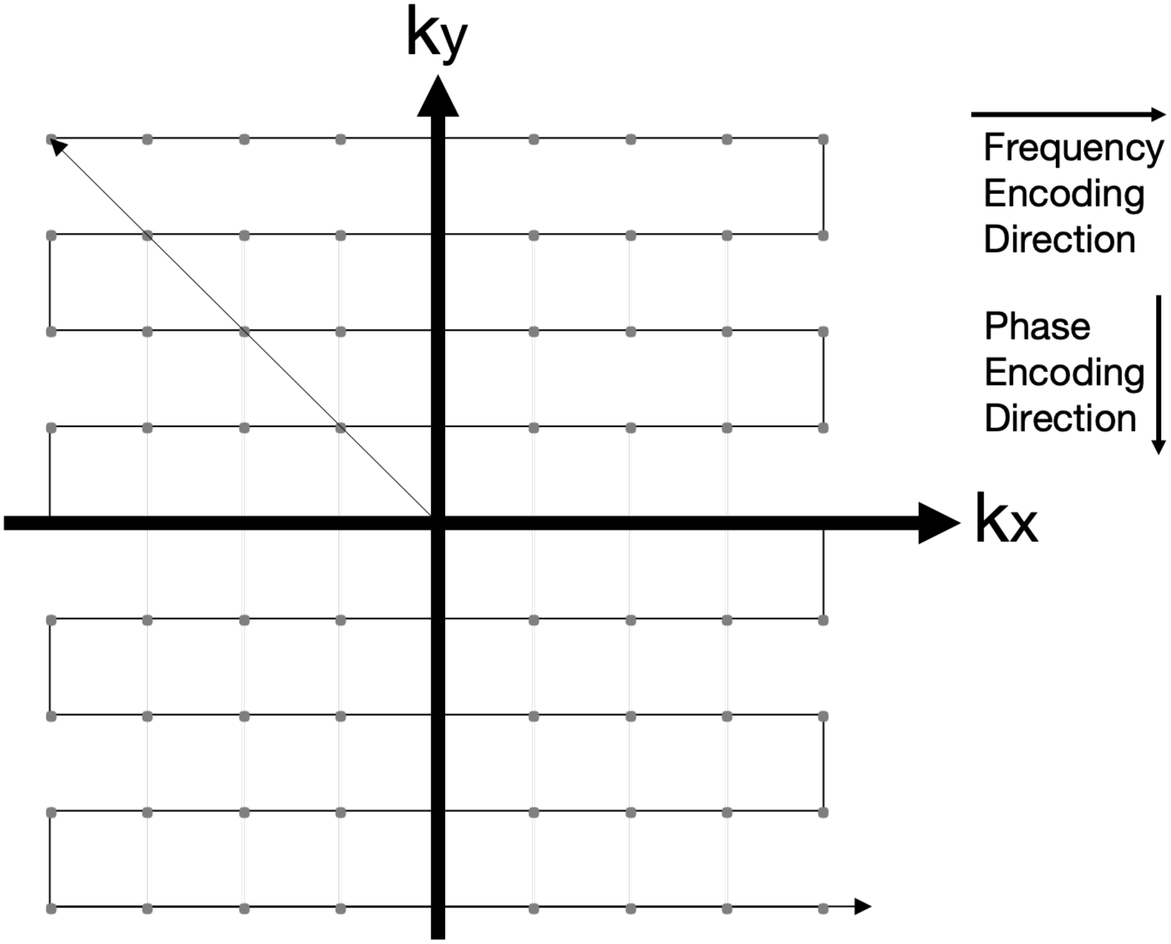


Figure 2.7: A Cartesian single-shot EPI trajectory. The x-axis is known as the frequency-encoding axis, and the y-axis is known as the phase-encoding axis. The trajectory starts in the centre of k-space, traverses to the upper-left corner, and collects data along the frequency-encode direction. The dots represent the collected k-space coefficients.

EPI

An Echo-Planar Imaging (EPI) trajectory is a Cartesian k-space approach that collects multiple parallel lines in a single RF excitation pulse/shot. An EPI sequence which acquires the entire plane of k-space that the user intends to collect after a single RF excitation is known as “single-shot EPI” (see figure 2.7). If more than one excitation is needed, the sequence is labelled as “multi-shot EPI”. Multi-shot EPI has reduced artefacts, but an increased overall acquisition time [10]. In EPI, the frequency-encoding gradient is turned on while the others are turned off (to collect one line), before a phase-encoding gradient is turned on briefly to “blip” the trajectory to a different k-space location, at which point the frequency-encoding gradient is turned on to collect a new line.

EPI was first proposed by Sir Peter Mansfield in 1977 [11], and is now common in many imaging applications [10]. EPI is highly favoured among Cartesian trajectories for its short acquisition times (40-50 ms for a full frame), but is susceptible to off-resonance effects due to the long trajectory. Ghosting can occur in the phase-encode direction due to phase shifts between the odd and even echos, and is caused by many sources (e.g. timing errors, eddy currents, gradient imperfections [12]). An odd/even phase error of zeroth-order (a constant shift) will cause discrete ghosts, and a first-order phase error (a linear shift) will lead to more non-uniform ghosts. It is fairly simple to undersample in EPI (simply by skipping lines), and desirable to do so to ensure a reduced signal decay during acquisition. The undersampling (or acceleration) factor is defined in EPI as the ratio of sampling lines in fully sample k-space (N_{full}) to the number of sampling lines acquired ($N_{sampled}$):

$$R = \frac{N_{full}}{N_{sampled}} \tag{2.26}$$

Echo Volumar Imaging (EVI) is a 3D formulation of EPI (EPI here will refer to 2D EPI). The main advantages of EVI over EPI are the theoretical reduction in thermal noise and the ability to undersample in multiple spatial dimensions [13]. EVI is particularly favoured for high spatial resolution + high field applications. Single-shot EVI is very sensitive to distortion and blurring [14]. Segmented EVI improves these issues [15], but is still very sensitive to physiological noise and motion [16].

2.1.5.2 Non-Cartesian

Non-Cartesian trajectories come in many forms. Typically they aim to do one or more of the following: oversample central k-space (as most MRI typically has a greater density of information in the central frequencies); collect a large amount of k-space information with each shot; sample central k-space with a very short (or sometimes very long) echo time; and/or create more incoherent artefacts if undersampled. More incoherent artefacts are less likely to be confused for signal in an accelerated reconstruction, and so are more desirable than coherent artefacts, see figure 2.8. Non-Cartesian trajectories may also have secondary goals such as reducing the trajectory duration or being more resistant to motion. The trade-off is a more complicated trajectory and reconstruction (see section 2.2.1 for details). Radial and spiral trajectories are common, but more complicated trajectories like rosettes [17, 18] have been attempted too.

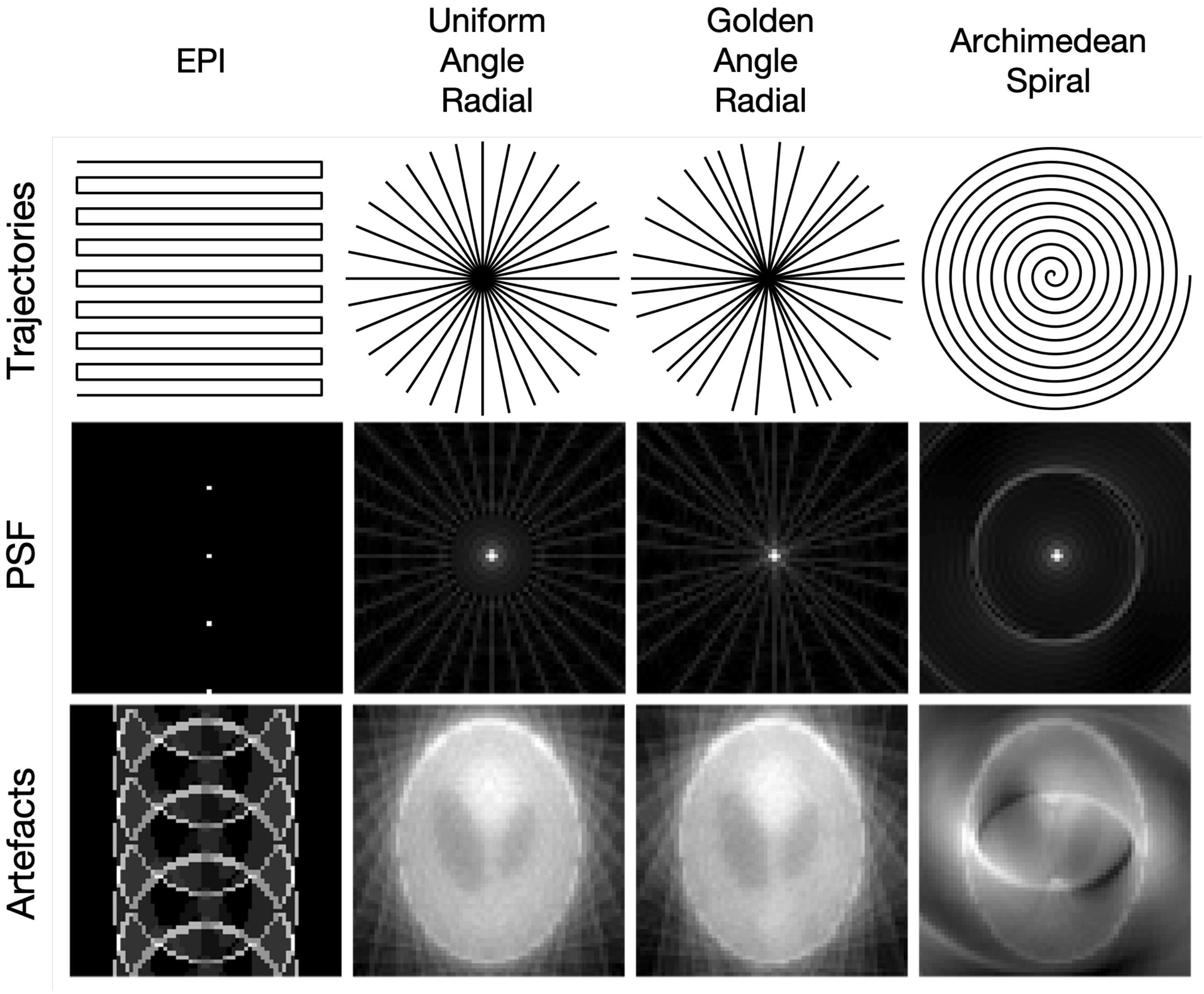


Figure 2.8: A demonstration of different trajectory artefacts at equivalent levels of undersampling. EPI artefacts have the Cartesian ghosting artefacts demonstrated in figure 2.3. The two radial trajectories both show characteristic radial streaking with only minor differences. The spiral undersampling shows a characteristic smearing pattern.

Radial

Radial trajectories consist of straight lines which all pass through central k-space. Rather than sampling over a Cartesian grid, the radial trajectories will instead fill a central "disc" of k-space, not covering the outer corners covered in a Cartesian grid. The incoherence properties of radial sampling (see figure 2.8) are very good for undersampling [19, 20]. Due to repeatedly sampling the centre of k-space, radial trajectories also tend to be more resistant to motion as the redundancy can be used to detect object movement.

Typically each radial spoke satisfies Nyquist along the line of the acquisition [21]. However, because the gaps in k-space between spokes are not uniform, a greater number of lines are needed in radial sampling compared to Cartesian sampling. Specifically, $\frac{\pi}{2}$ more lines are needed to define the acceleration factor R through N_{full} in equation 2.26 [22] (e.g. 157 radial lines are needed for a 100×100 image).

The most basic approach to radial trajectories is a uniform angle approach that uses an equal angle between all lines for a predetermined number of lines per image. An alternative approach spaces the radial projections at constant azimuthal increments of $\frac{180^\circ}{\Phi} \approx 111.25^\circ$ [23], known as a Golden Ratio angle profile. This scheme ensures near-uniform coverage of k-space from any arbitrary retrospective combination of consecutive projections, allowing for a flexible approach to acceleration [24] which trades the reduction of image artefacts (favouring lower acceleration) against an increased temporal resolution (favouring higher acceleration). Adding small perturbations to the spoke angle can also promote the incoherence [25].

2D Golden Ratio angle radial imaging can be used to gather a 3D cylinder of k-space through a Golden Ratio angle stack-of-stars approach [26, 27], or the same theory can be used to gather a 3D Golden Ratio angle sphere [28].

Spiral

Spiral sampling [29, 30] is the other common non-Cartesian trajectory. Unlike radial sampling, spiral sampling does not require the sampling to be straight lines which span k-space – instead, spirals either traverse from the centre to the edge of k-space (spiral-out) or from the edge to the centre (spiral-in) in a curved trajectory. Spiral trajectories reduce the downtime that the signal is not measured when traversing k-space. This leads to a shorter overall scan time, despite requiring a longer time per shot than radial acquisitions as fewer shots are required to cover k-space. This scan time reduction is ideal for real-time imaging [30]. Like EPI, single-shot and multi-shot variants of spiral imaging exist [31]. Spiral sampling artefacts are incoherent and typically akin to blurring; adjacent points in k-space are acquired with a smaller temporal gap which helps reduce artefact strength (see figure 2.8).

A common example of spiral sampling is the Archimedes spiral, which uses a radius which is linearly proportional to the angle (with angle $\theta(t)$ defined over the range $0^\circ \rightarrow 360^\circ \times N_{turns}$). The trajectory to cover all of k-space is then the characterised purely by the number of interleaving spiral spokes N_{int}, the FOV, and the angle over time $\theta(t)$:

$$k(t) = k_x(t) + ik_y(t) = \frac{N_{int}\theta(t)}{\mathrm{FOV}} e^{i\theta(t)} \tag{2.27}$$

A variable density spiral [32] can follow a more unconventional curve to ensure greater oversampling of central k-space. Approximate analytic formulas for designing spiral trajectories are readily available [33]. Spiral 3D sampling is typically a stack-of-spirals scan, which scans in-plane spirals and stacks the results vertically.

2.1.5.3 Hybrid Trajectories

Hybrid trajectories collect samples in a manner that combines Cartesian and non-Cartesian trajectories. PROPELLER (Periodically Rotated Overlapping ParalEL Lines with Enhanced Reconstruction) is a 2D hybrid trajectory [34] which takes a mini EPI plane covering central k-space, and then rotates that EPI trajectory within the plane to cover a k-space disc. Sampling central k-space in each acquisition allows for basic position/rotation correction, recognition and rejection of acquisitions acquired in heavy motion, and an averaging of low-spatial frequencies to further decrease motion artefacts.

TURBINE (Trajectory Using Radially Batched Internal Navigator Echoes) [35] is a hybrid 3D radial-Cartesian k-space cylinder which acquires an EPI blade that is rotated around the phase encoding axis at constant azimuthal increments. This captures a 3D k-space cylinder. The TURBINE sequence is beneficial for many reasons. The hybrid of radial and Cartesian sampling combines the benefits of both methods by ensuring constant oversampling of k-space to enable motion correction, while maintaining the encoding efficiency achieved by EPI. The radial aspect ensures increased incoherence compared with 3D EPI imaging alone.

As with radial sampling, a Golden Ratio angle approach can be used in TURBINE as an alternative to uniform angles [36] (see figure 2.9). The flexibility of acceleration that the Golden Ratio angle allows for the final reconstruction is one benefit (see section 2.1.5.2). Additionally, the blades can be batched to form a high temporal resolution navigator time series which can be used to estimate rigid body parameters for motion correction [37]. A navigator is a secondary signal which is used to derive some sort of image property (typically motion) for the purpose of correction. In TURBINE, the navigator can be derived without the need for any additional RF pulses.

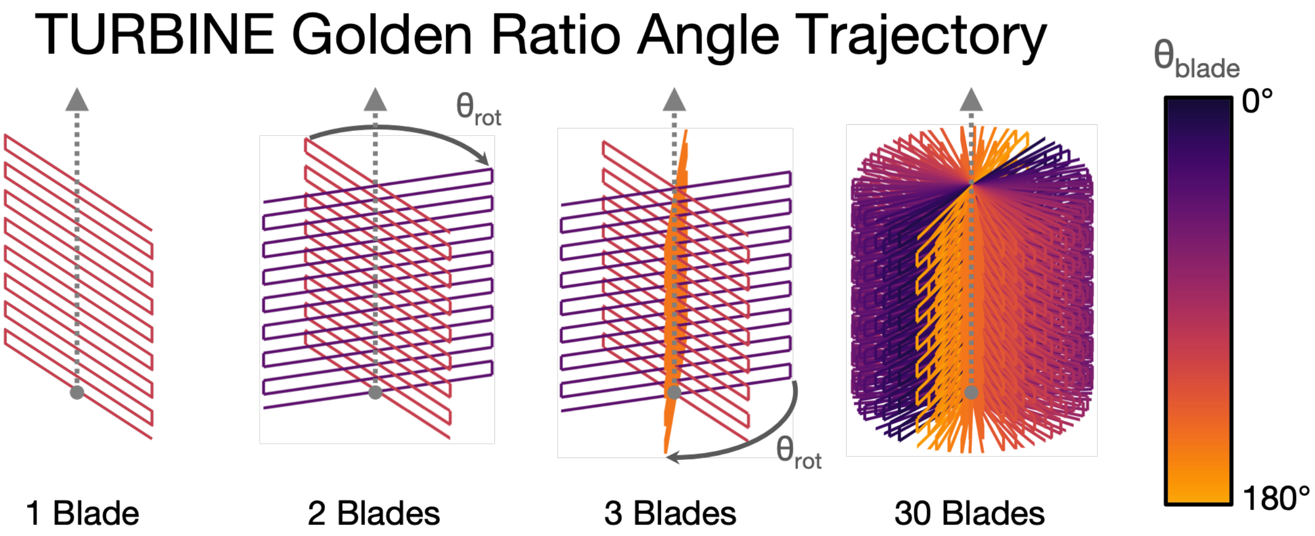


Figure 2.9: A demonstration of the rotation of EPI blades using the Golden Ratio angle in a TURBINE trajectory. Each subsequent blade is rotated by $\theta_{rot} \approx 111.25°$ around the phase-encode axis to ensure the next blade lies in the biggest remaining gap. The blade colour encodes θ_{blade}, the absolute angle of each blade. The blades will never overlap, and collectively fill out a cylinder in k-space.

2.1.6 SNR

The Signal-to-Noise Ratio (SNR) is a very important concept in MRI, as it defines how reliable the useful information (signal) is for a given acquisition. SNR is defined in many ways [38] with the most straight-forward definition being the mean of the signal divided by the standard deviation of the noise :

$$SNR = \frac{\mu_{signal}}{\sigma_{Noise}} \tag{2.28}$$

Many factors influence the underlying SNR [39]. In terms of acquisition parameters, the SNR is directly proportional to the square root of total measurement time. The resulting time relation encompasses an increase of the signal averaging factor N_{ave} (acquiring the same k-space region more than once and averaging the results), and a reduction of the temporal sampling rate $\frac{1}{\Delta t}$ (this reduces the signal bandwidth, and so the reduces the noise variance per sample, increasing the SNR). If these factors double the signal amplitude, they will also double the noise variance and therefore increase the noise standard deviation by $\sqrt{2}$.

SNR is directly proportional to the voxel volume ($\Delta x\ \Delta y\ \Delta z$), as the amount of signal in a voxel is proportional to the volume, while the noise variance present in different voxel sizes is equal if all other factors are held constant. Combining voxels together would only yield a square root increase in SNR, and so acquiring at high resolution and combining voxels would be less efficient than an acquisition at that combined voxel size.

The relationship of SNR to all of these parameters can be combined as follows:

$$SNR \propto \Delta x \Delta y \Delta z \sqrt{\frac{N_{ave}}{\Delta t}} \tag{2.29}$$

SNR is also roughly proportional to the main magnetic field B_0. The signal is directly proportional to B_0^2, but the subject-based noise variance also grows roughly proportionally to B_0^2 [40]. Very high field strengths begin to show a greater than linear relation. It should be noted that at very low field strengths (<0.05T) the coil-based noise variance begins to dominate and grows proportionally to $\sqrt{B_0}$ [41]. The SNR alone doesn't capture the other factors which are also affected by increasing field strength, such as increasing T1 (leading to less recovery per TR and reduced contrast) or increased field inhomogeneities [40].

Special care must be taken when calculating equation 2.28. Typically the mean is taken only over the desired signal region (ignoring any pixels containing only close-to-zero background noise), and the variance calculated from an area that does not contain the active subject. The SNR is multiplied by 0.65 to account for the underestimation of the noise standard deviation in magnitude images. This SNR calculation holds as long as the noise is spatially homogeneous and follows a Rayleigh distribution [42]. The Contrast-to-Noise Ratio (CNR) is also commonly used. The difference between the two signals of interest Δ_{Signal} used in place of the signal mean. Δ_{Signal} represents how distinct two signals are and could be calculated between the mean signal of two separate tissues, or between a peak signal and signal baseline.

$$CNR = \frac{\Delta_{Signal}}{\sigma_{Noise}} \tag{2.30}$$

An additional SNR definition is temporal SNR (tSNR), which accounts for varying noise over time in dynamic MRI (e.g. from cardiac and respiratory sources in the subject). tSNR is calculated by dividing the mean signal intensity by the standard deviation over time, and is used to evaluate the quality of the time series (whereas SNR evaluates the quality of a single image/volume). The physiological noise is more impactful in large voxel sizes and at high field strength [43].

2.2 Image Reconstruction

An acquired signal must be converted from k-space into image space. Solving equation 2.19 given a known S (a sampled signal vector) will provide I, the desired image vector. The simplest way to do this is by fully sampling over one coil on a Cartesian grid and simply reconstructing the image with an inverse 2D FT applied to the collected k-space. The mathematics behind more complicated reconstructions will be covered in chapter 3.

In reality, MRI often undersamples k-space in order to reduce the scan time required to produce each image. Acceleration can often reduce the SNR (signal-to-noise ratio) in exchange for faster sampling. As mentioned in subsection 2.1.5.2, non-Cartesian sampling produces more incoherent aliasing artefacts in acceleration but requires a more complicated reconstruction approach. If the signal is received in multiple coils with overlapping sensitivity profiles, these can be used to de-alias accelerated signal. This coil information can be incorporated in image space (section 2.2.2.1) or k-space (section 2.2.2.2). Partial Fourier is also included briefly as a final common in-plane image reconstruction method in this section.

2.2.1 Non-Cartesian Reconstruction

Non-Cartesian sampling methods (such as radial) will produce discrete samples that do not lie on a Cartesian grid. A Fourier transform can be directly calculated from the acquired sample points, or a process known as gridding can be applied which interpolates the data onto a Cartesian grid so that the computationally efficient Fast Fourier Transform (FFT) algorithm can be used. Other methods (e.g. nearest neighbour) have been attempted, but are not common. Whichever approach is taken is incorporated in equation 2.19 into the encoding operator E.

2.2.1.1 Gridding

Gridding [44, 45] is the interpolation of arbitrarily located data onto a uniform Cartesian grid, after density pre-compensation is applied to account for sampling density being non-uniform in k-space [46]. This interpolation is done by convolution with a gridding kernel, with many options of kernel available [21]. The optimal kernel is an infinite sinc function, but the Kaiser-Bessel kernel is an appropriately finite kernel equivalent with a reasonable computation time [21]. Once the gridding is complete, an inverse FFT can convert the gridded k-space into an image.

The inverse gridding process must compute the non-uniform k-space data from the Cartesian sampled image. This inverse process is very important for iterative reconstruction algorithms, which must alternate between image and k-space.

2.2.1.2 NUFFT

The NUFFT is a generalised version of the gridding algorithm with a fast implementation that better suits iterative algorithms [47, 48], and a common toolbox is provided by Fessler et al. [49]. Different approaches to speeding up the NUFFT are still being explored [50, 51], but the most significant addition may be Toeplitz Embedding [52, 53]. A practical walkthrough of Toeplitz Embedding is provided in appendix 2.A. Briefly, Toeplitz Embedding exploits the diagonalisation properties of Toeplitz matrices in order to provide a more computationally efficient implementation, and the NUFFT Gram matrix $E'_{NUFFT}E_{NUFFT}$ contains a block Toeplitz structure which encourages a Toeplitz Embedding approach.

2.2.2 Parallel Imaging

Parallel Imaging (PI) aims to only acquire a subset of k-space data, and then estimate the unacquired samples based on the spatial profiles of multi-channel receive coils. Equation 2.19 can be updated to include C, a full set of coil sensitivities which multiplexes the vector I into c copies. c is the number of coil channels and each image copy is weighted by the individual coil profile. Coil sensitivity maps are smooth and non-uniform (coils are more sensitive to closer magnetisation). E now operates on all c coil images in order to produce a set of k-space samples at the same locations for each coil [54]. I and S are both vectors. Figure 2.10 shows a basic coil multiplication applied in image space, where the effect of the individual coil weighting can be clearly seen.

$$S = ECI \tag{2.31}$$

The coil sensitivity maps are not orthogonal. This gives them some compressibility, and in fact C is often reduced to the few most dominant components [55, 56]. The compression does not fundamentally change equation 2.31, merely reduces the dimensionality of C (and of the combination operator in E). This, in turn, reduces the computation power required to find a solution. The differing coil profiles allow for recovery of the ’true’ signal through de-aliasing of the acquired signal. This recovery can occur in both the image domain and the k-space sampling domain. Additionally, the coil sensitivity information can be used to separate multiple slices that were collected in the same RF pulse. A generalised overview shown in figure 2.11, and will be described more fully in the upcoming sections.

The g-factor (geometry factor) is a PI-specific cost to acceleration. The g-factor is a spatially-varying noise amplification term which encompasses the number, size, and orientation of surface coils, but also reflects the corresponding image geometry (location of the imaging plane, FOV,

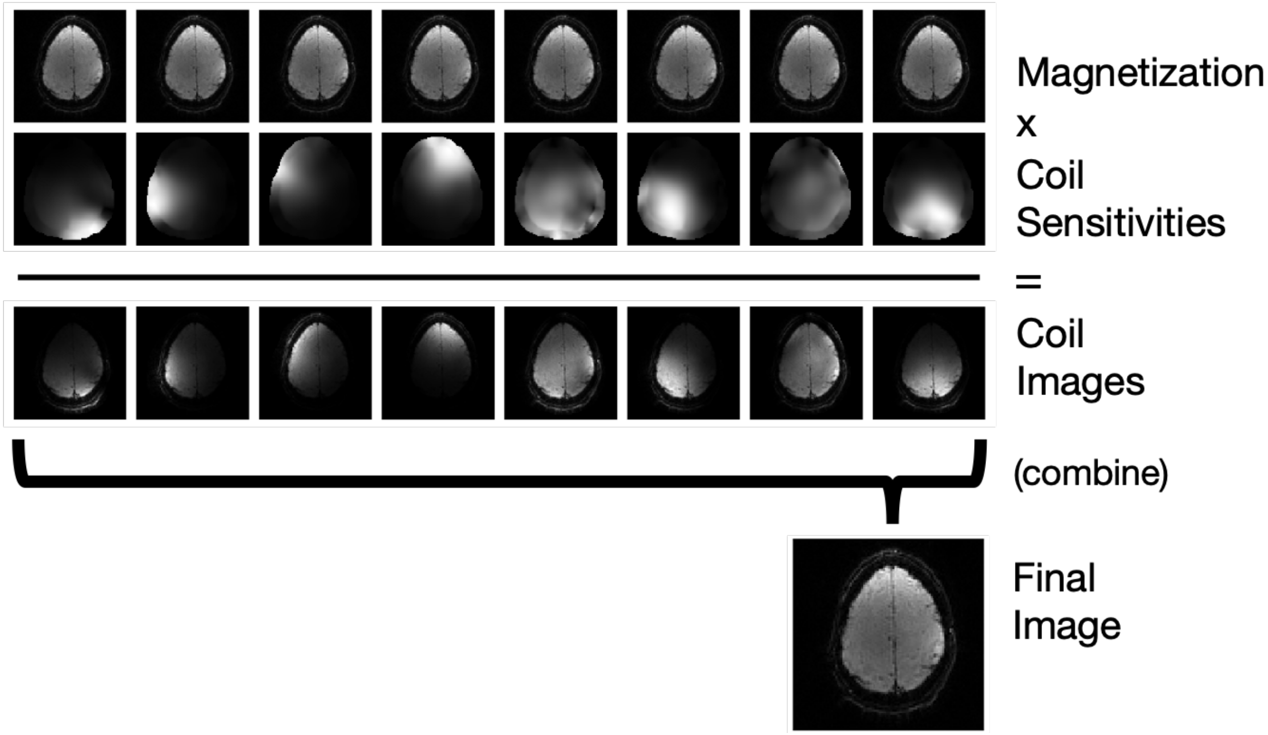


Figure 2.10: A demonstration of how the coil sensitivities (represented with C in equation 2.31) weight the underlying magnetisation through a unique spatial profile, which then each generates a different coil image. These coil images can then be combined into a full image using a least squares operator.

etc.) and the k-space sampling pattern (with the distance between k-space samples affecting noise distribution). A high g-factor term arises from an ill-conditioned matrix with coil sensitivities that are too similar. The g-factor is combined with the square root of the acceleration factor R to represent the reduction in SNR which arises from PI.

$$\mathrm{SNR}_{acc}(x,y) = \frac{\mathrm{SNR}_{full}(x,y)}{g(x,y)\sqrt{R}} \tag{2.32}$$

2.2.2.1 Applying Coil Sensitivities in Image Space

The most common approach to applying coil sensitivities in image space is SENSE (SENSitvity Encoding) [57] which seeks to unfold overlapping pixels in the image domain. As long as the acceleration factor does not exceed the number of linearly-independent coils, the image is theoretically recoverable (although typically R factors of 2-4 are used, as the coil spatial sensitivity maps are not orthogonal and the achievable acceleration is much lower).

Again, c will represent the number of coils. A pixel at the same sampled image location for all of the coils (a, a c-element vector) will contain an aliased (replicated) signal from R points of the

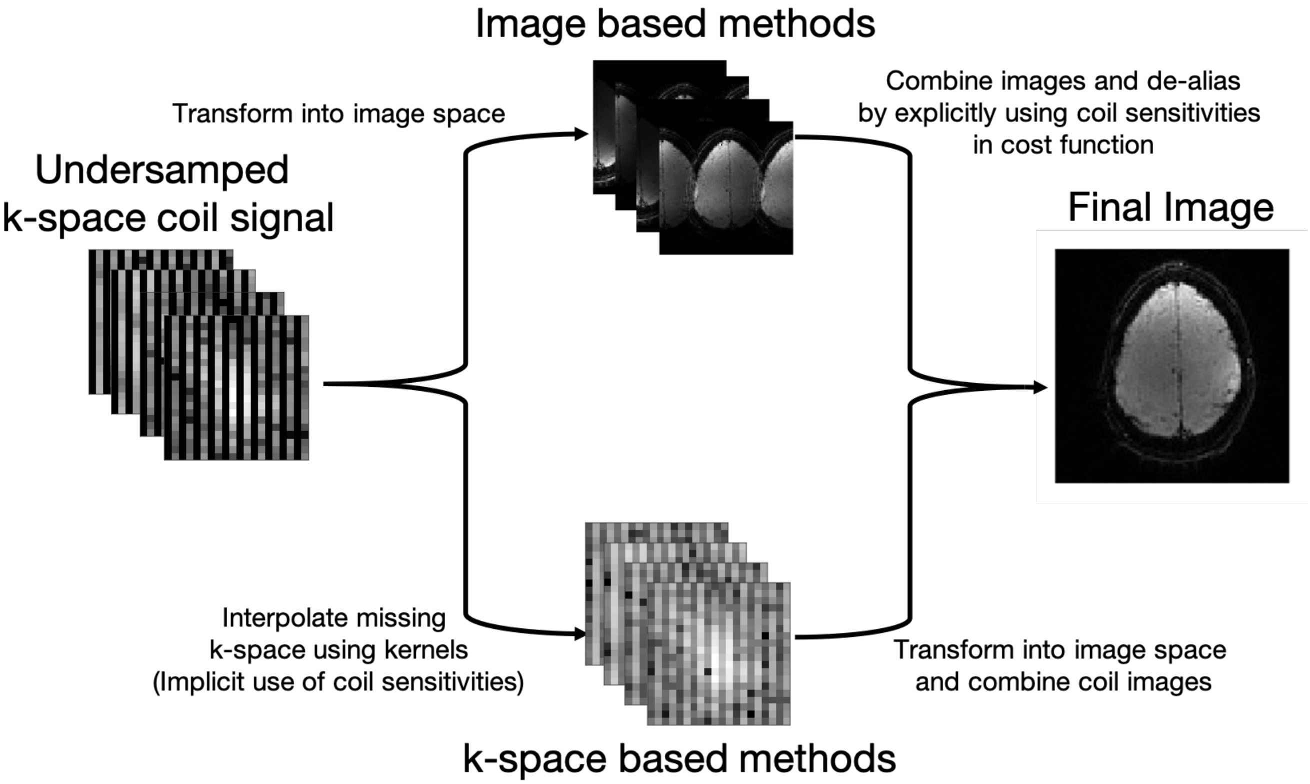


Figure 2.11: A demonstration of the different paths of parallel imaging. Both approaches generate an undersampled k-space pattern for each coil (here the undersampling is shown as zero-filled for clarity, but this is not always the case). Image-based methods will transform the undersampled k-space into image space, and use aliasing location knowledge to recover the image. Methods that are k-space based will attempt to recover the missing k-space directly (often with a kernel calculated from a fully sampled centre of k-space, not shown here). This recovered k-space is then transformed into image space.

true underlying image (b, an R-element vector). These points are multiplied by the R sampled points for each coil sensitivity map (C_{vox}, a $c \times R$ matrix representing the coil sensitivity matrix for a single voxel).

$$a = C_{vox} b \tag{2.33}$$

a is detected output, and C_{vox} must be pre-calculated. b is the underlying unknown truth. Equation 2.33 can be solved through standard least-squares techniques (see section 3.2). Additional regularisation in an iterative reconstruction can also improve the reconstruction power of SENSE.

2.2.2.2 Applying Coil Sensitivities in k-Space

PI can also be applied in k-space (see figure 2.11). GRAPPA (GeneRalized Partially Parallel Acquisitions) [58] is the most common approach in the field, although alternate approaches such as SMASH [59] are also used. GRAPPA synthesises the missing k-space data directly for each coil. In image space, the sampled magnetisation is multiplied by each of the coil profiles (figure 2.10). This means that the samples and coil profiles are convolved together in k-space, and a sampled k-space point will contain information about the surrounding k-space.

The amount of spread occurring for each sample is calculated from a set of autocalibration signal lines which fully sample a central band of the underlying k-space [60]. The value at a chosen point (a target) is estimated as the weighted combination of the point's sampled surroundings (the sources). The combination of spatially-arranged weights required to match the sources to the target is the GRAPPA kernel. This process is repeated for all known target points in which the kernel is calculated from a set of fully known source points. The same kernel is then used throughout all of k-space for each coil. Typically a 2D kernel is favoured (even in 3D MRI) [61].

The kernel weights w for a target coil c_t are calculated as the optimal fit for a given set of N_c source coils, and k-space spacings $(\Delta k_x, \Delta k_y)$, given a known signal at a specific k-space location s_{target} and a known source signal s_{source} .

$$s_{target}(c_t, k_x, k_y) = \sum_{c_s=1}^{N_c} \sum_{k_x} \sum_{k_y} w(c_t, c_s, \Delta k_x, \Delta k_y) s_{source}(c_s, k_x + \Delta k_x, k_y + \Delta k_y) \tag{2.34}$$

The above equation can also be formulated as a matrix of GRAPPA weights W. The same equation can be used once W is formed to calculate unknown s_{target} locations.

$$s_{target} = W s_{source} \tag{2.35}$$

The GRAPPA reconstruction can be visualised as a sliding kernel convolution with all the unacquired k-space locations to fill out a k-space matrix for every coil. These can be transformed into individual coil images, which are then combined using a sum-of-squares approach. As GRAPPA does not require explicit knowledge of coil sensitivities, it is more robust than SENSE in the case of slight inaccuracies in the coil sensitivity maps.

2.2.2.3 Simultaneous Multi-Slice

Simultaneous Multi-Slice (SMS) [62] methods seek to spatially accelerate between planes. SMS imaging excites multiple slices concurrently with multiband RF excitation pulses. A linearly varying slice selection gradient is applied, and a composite multiband pulse composed of a standard pulse and slice dependent phase modulations will tip the magnetisation in all the excited slices. The slices will then be acquired simultaneously with a single k-space trajectory, and are then separated in image reconstruction using parallel imaging techniques. The number of slices excited by this pulse is known as the multiband factor (MB). A $\frac{1}{\sqrt{MB}}$ improvement in SNR efficiency is found compared with single slice acquisition, although a g-factor penalty will still apply. Many techniques for de-aliasing the slices exist [63, 64], including both SENSE and GRAPPA-based approaches.

CAIPIRINHA (Controlled Aliasing In Parallel Imaging Results In Higher Acceleration), more commonly known as CAIPI, is a modified trajectory that is very common in 2D multi-slice approaches [65]. CAIPI shifts the sampling lines to better separate the aliased copies without requiring any additional sampling. By shifting the aliased copies, CAIPI achieves better separation between images (and a more unique projection onto each coil profile), which reduces error in reconstruction. An example is shown in figure 2.12.

2.2.3 Partial Fourier

Coil-independent in-plane acceleration is also possible. Partial Fourier (PF) uses knowledge of k-space possessing conjugate (Hermitian) symmetry - rather than the coil sensitivity information of PI - to recover the missing k-space signal. If the data was acquired with no phase errors, diagonally opposed points from the centre of k-space will be the complex conjugate of one another (i.e. they will share the same amplitude but opposite phase). Theoretically, this means only half of k-space needs to be acquired, and the rest of k-space can be estimated by conjugation [66]. In reality, a certain amount of central k-space beyond the halfway line must be sampled in order to account for field inhomogeneities and phase shifts [67]. Note that PF produces blurring artefacts, rather than the aliasing artefacts of section 2.2.2. Various types of Partial Fourier reconstruction exist [68], including conjugate synthesis, Homodyne [69], Margosian [70], and Projection Onto Convex Sets [71].

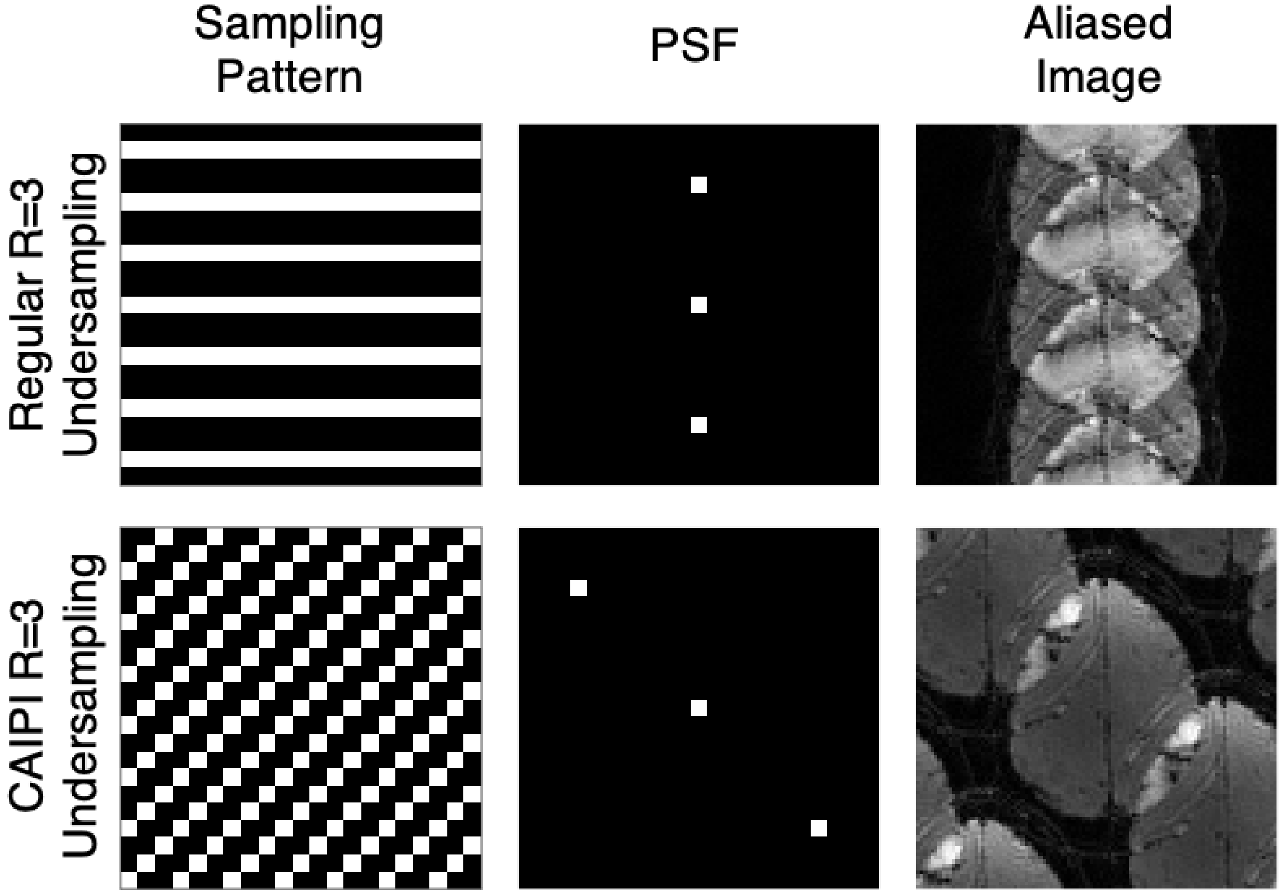


Figure 2.12: A demonstration of CAIPI in 2D. Each white dot in the sampling pattern represents a single sample in 2D, or a sampling line which is orthogonal to the plane in 3D. The PSF shows an increased spread of the ghosting artefact for CAIPI sampling, a fact that can be visualised in the example aliased image. The k-space of the aliased image was also sampled with similar R=3 undersampling patterns, but is shown at a higher resolution for visual clarity.

2.2.4 Conclusion

Reconstruction of a single clean MR image (or volume) from an undersampled signal can be achieved through the use of coil sensitivity maps or k-space knowledge. In dynamic MRI, multiple images are collected and acceleration can also be applied through a variety of methods to account for redundancies in the temporal information too. This next section will introduce functional MRI as a dynamic imaging method, explain how the data is commonly processed to provide functional information, and then address the advantages of acceleration in fMRI.

2.3 fMRI

The term MRI captures a wide range of approaches that can be used to non-invasively image soft tissue anatomy. Different imaging parameters and magnetisation preparation can yield different biological information of a subject. Dynamic MRI takes multiple MRI images/volumes over a short period of time in order to observe changes in the subject, creating a time-series. Functional MRI (fMRI) is a form of dynamic MRI which detects local changes in biological parameters (such as blood flow) which correspond with brain function [72, 73, 74, 75].

The BOLD (Blood Oxygen Level Dependent) signal [76, 77] is the most common form of fMRI and measures varying oxygenation levels of the blood as a proxy for the functional activity. Other fMRI methods such as VASO (measuring Cerebral Blood Volume: CBV) [78] and ASL (measuring Cerebral Blood Flow: CBF) [79] are not uncommon and provide alternative insights into the workings of the brain. These methods have lower SNR than BOLD fMRI ($\mathrm{SNR_{BOLD}} > \mathrm{SNR_{VASO}} > \mathrm{SNR_{ASL}}$), and have a lower temporal resolution [80]. However, ASL/VASO give more specific detail regarding biological functions (ASL is fully quantitative), and the signals are better localised to the neuronal activity [81, 82].

2.3.1 The BOLD Signal

The main link between oxygen and fMRI relates to the generation of the main energy molecule, adenosine triphosphate (ATP). ATP provides the energy needed to drive many active processes in living cells, including neuron signaling [83]. When oxygen is present, glucose can be broken down into ≈34 ATP molecules via aerobic metabolism. When oxygen is absent, the body must resort to anaerobic metabolism, which produces just 2 ATP molecules from a glucose molecule. O_2 molecules are transported by the haemoglobin molecule in blood to the parts of the body which require energy.

Haemoglobin is an assembly of four globular protein subunits which each contain a haem group – an iron ion in a ring which can each reversibly bind to one of the oxygen atoms in an O_2 molecule (the other atom in O_2 is unbound). Oxyhaemoglobin is a weakly diamagnetic molecule [84] (repulsive in magnetic fields) that is formed by oxygen molecules binding to the haem groups of haemoglobin during respiration. The oxygen binding curve of haemoglobin is sigmoidal – once one O_2 molecule is bound, the binding affinities of subsequent heme groups are increased by the first binding [85]. Oxyhaemoglobin molecules are carried in the blood through the cardiovascular system; they are able to release their oxygen to create deoxyhaemoglobin, which is a paramagnetic molecule (attractive to magnetic fields). In the brain, astrocytes (a star-shaped non-neuronal cell) are able to transport the O_2 molecule across the blood-brain barrier through passive diffusion and are strongly linked to blood flow [86, 87].

In 1990, BOLD contrast was observed in an accentuated manner under conditions that were induced by inspired gases [76]. This led to studies using the BOLD signal to detect brain activity under more natural conditions [72, 73, 74, 75]. While the precise origins of the signal are still being studied [88], the general pathway is known with reasonable certainty (figure 2.13). An increase in neuronal activity causes a change in the oxygenation of the blood and blood flow [89], in a process known as neuronal coupling [90]. The main cause of these physiological changes is driven by neurotransmitter signalling which is mediated by astrocytes [91, 92]. The relative decrease in deoxyhaemoglobin concentration also reduces the susceptibility difference between the vessel and the surrounding tissue. Through equation 2.9, a reduction in local inhomogeneities will also reduce the amount of dephasing, and increases the local T2* values. This effect can be visualised with T2* imaging (n.b. small vessels like capillaries can also be visualised with T2 imaging [93]). Setting the TE equal to T2* will capture the maximum BOLD signal change between the baseline and activated states.

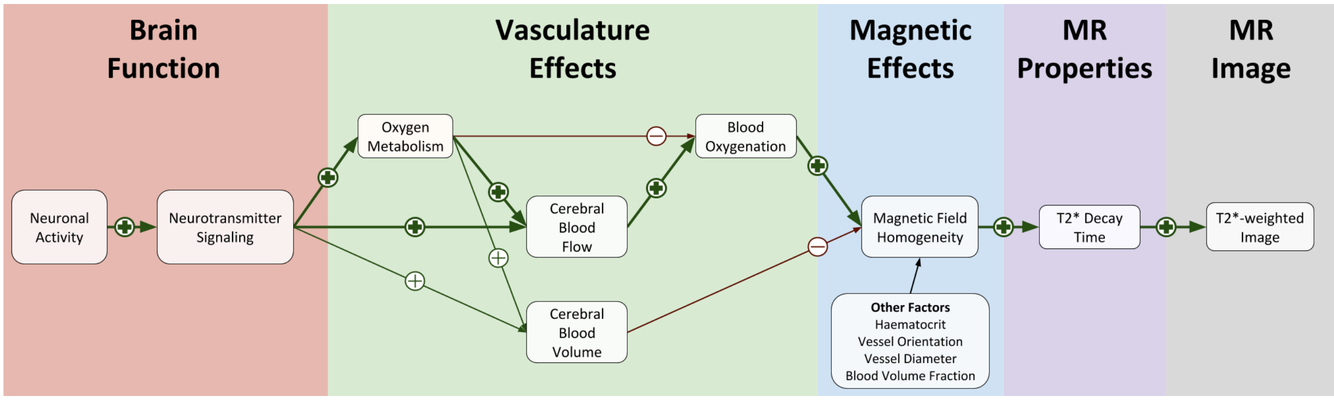


Figure 2.13: A schematic of the interactions involved in generating the BOLD signal. Positive/green and negative/red arrows indicate a corresponding proportional change. The bold arrows signify the most dominant pathway. Adapted from figures by Springer [94] and Noll [95], and updated with information from Logothetis [91].

This increase in blood oxygenation occurs on a slight delay relative to the onset of neural activity, due to the time required for chemical signalling to instantiate the physiological effect changes. The curve characterising this rise and fall is known as the Haemodynamic Response Function (HRF, figure 2.14), which maps the signal that occurs during a single instance of neuronal activity. The combination of hundreds of thousands of neurons lying within each voxel [96] and the broad spatial PSF of the BOLD signal [97] mean that fMRI can only capture mass neuronal activity. The dynamic expression of the BOLD signal in T2*-weighted images can give fundamental insight into deducing the spatially distributed, functionally active regions of the brain (within the limitations of resolution for fMRI).

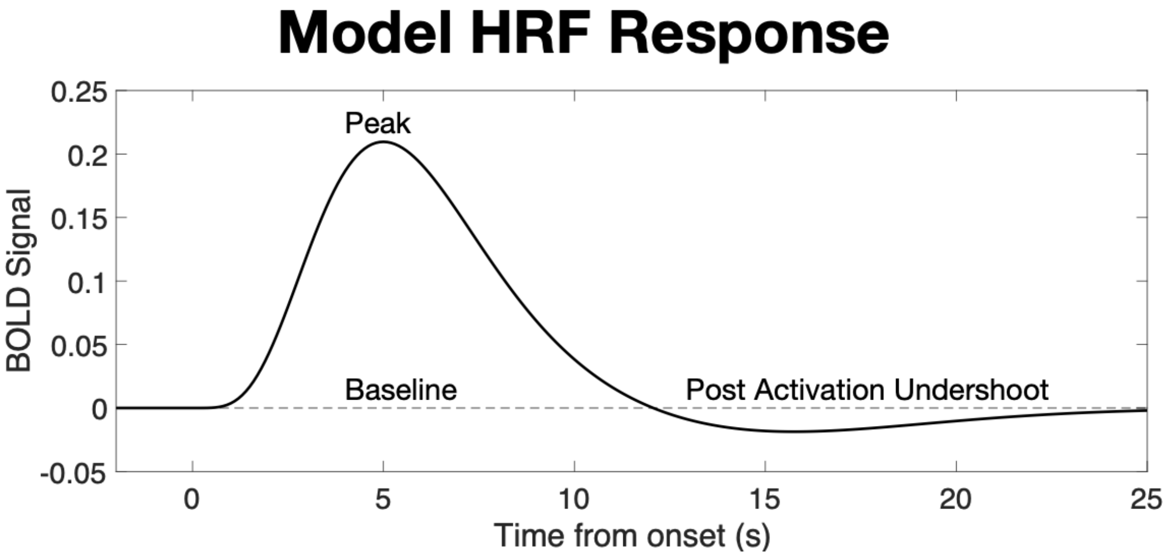


Figure 2.14: The Haemodynamic Response Function. The signal produced in response to neuronal activity. The curve was created through adapting code from the VBA Toolbox [98].

2.3.2 Types of BOLD fMRI

fMRI enables neuroscientists to detect brain activity, which can be used to delineate brain regions and identify temporally correlated responses. This helps elucidate the computational mechanisms underpinning cognition. fMRI provides benefits in many areas of research, such as the examination of connectivity diseases [99] like Alzheimer's disease, dementia, schizophrenia, and multiple sclerosis; or in studies of altered brain circulation such as traumatic brain injury [100].

2.3.2.1 Task fMRI

Task fMRI (tfMRI) presents a subject with a stimulus or cognitive task in order to study the resulting cognitive process. In a simple block design approach, the stimulus is presented continuously for a set period and then taken away for a set period. This stimulus on/off pattern is then repeated, in order to provide a contrast between background levels of neuronal activation and neuronal activation in response to a stimulus. More complicated block design approaches can use multiple tasks (e.g. a "task 1 on/all tasks off/task 2 on/all tasks off" approach). Event-based approaches measure the response to discrete and short-duration events and have increased flexibility in terms of randomisation of stimuli.

A General Linear Model (GLM) is used to statistically analyse tfMRI approaches. The fMRI signal vector for each voxel Y is formulated by the multiplication of a vector of parameter estimates β and a set of linearly independent regressors X (a.k.a. explanatory variables). These regressors are a combination of regressors of interest and nuisance regressors. Regressors of interest represent convolutions of the expected on/off response to certain stimuli with the HRF. The nuisance regressors are additional measures that may confound the analysis such as: white matter signal, cerebrospinal fluid signal, temporal derivatives of those signals, and signal from the edge of the brain. A separate noise term ϵ can be used to represent additional factors not modelled in the GLM.

$$Y = X\beta + \epsilon \tag{2.36}$$

An estimate of the weighting vector $\hat{\beta}$ can then be generated through a simple least squares model of the known regressors and voxel-specific signal [101]:

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{2.37}$$

A set of t-statistics can be calculated to test the null hypothesis. A contrast vector b is used to select the relevant columns of β. In the simplest block case $b = [1, 0]$ and X is a matrix with two columns. The first column measures the expected response to the task if activation is occurring and the second measures the expected response with no reaction to the task. The t-statistic distribution can then be calculated as follows.

$$t_i = \frac{b\beta}{\sqrt{\sigma^2 b(X'X)^{-1}b'}} \tag{2.38}$$

The denominator in equation 2.38 represents the standard error, and σ represents the mean squared residual for each voxel $(Y - \beta X)'(Y - \beta X)/N_{DoF}$. N_{DoF} represents the degrees of freedom, which are the number of time points minus the number of parameter estimates for a voxel. The t-statistic scores for each voxel can be arranged into a statistical parameter map as a visual measure of significance. z-statistic maps are often used to display the same underlying p-scores as the t-map, but with zero mean and unit variance. Each voxel in a z-score statistical parameter map states how many standard deviations away from the mean a particular observation is, indicating likely areas of activation in the brain [101]. A voxel with an appropriately high z-score to reject the null hypothesis is then deemed to show BOLD activation (e.g. $z>1.7$ is $p<0.05$, or $z>2.3$ is $p<0.01$).

The abundance of voxels in a dataset can lead a $p<0.05$ threshold to still select a large number of false positive voxels (known as family-wise errors), and so some multiple-comparison correction is needed. Many approaches have been tried to account for these errors including: a Bonferroni correction, which divides the p-threshold by the number of voxels; a false discovery rate approach which controls for the expected number of false positives [102]); or a clusterwise approach which accounts for activation typically occurring in groups of voxels [103]. The GLM is a univariate approach that assumes all voxel signals are independent. Another assumption is that the residuals are independent and random, with a Gaussian distribution and a mean of zero [104]. Preprocessing can increase the validity of these assumptions, if not satisfy them entirely.

2.3.2.2 Resting State fMRI

tfMRI is typically used to investigate an explicit cognitive process, whereas resting-state fMRI (rsfMRI) [99] is an important tool in calculating Functional Connectivity (FC). FC is the temporal covariance of neuronal activation patterns from anatomically separated brain regions [105]. rsfMRI measures spontaneous neuronal co-activation between brain regions, in the absence of any external stimuli (i.e. in a subject at rest). One of the most common connectivity patterns found is in the default mode network, a region that is highly active until the subject engages in an explicit task, at which point the activity is reduced [106]. The default mode network is just one of the various networks which are active at rest; these various networks are collectively termed the resting state network [107]. It should not be inferred from that name that those same networks can not also be active during tasks [108].

The methods most often used to identify connectivity in rsfMRI data are an independent component analysis (ICA) approach and seed-based methods [109]. Seed-based methods identify time-series correlation between a region of interest and the entire brain, and thresholds the correlation score to identify connected regions. ICA is a multivariate signal analysis tool the separates a signal into non-Gaussian sources of maximal statistically independent patterns of activity [110, 111], and requires an assumption of linear constant mixing of signals in fMRI [112]. All components are equally weighted in ICA, and are not necessarily orthogonal. Uninteresting sources may then be discarded in ICA as a way of cleaning the data. ICA applied to rsfMRI divides the entire brain into statistically independent spatial components to identify regions with similar temporal patterns, and can also be used for denoising (n.b. this is specifically spatial ICA; temporal ICA splits the brain into statistically independent temporal patterns and is less common in fMRI [113]).

Temporal resolution is a limiting factor in rsfMRI, using acceleration to increase temporal resolution also improves the statistical degrees of freedom for both methods. Acceleration is particularly useful in ICA analysis which is more likely to form a mathematically underdetermined equation. rsfMRI requires pre-processing to clean the signal before analysis, just like tfMRI.

2.3.2.3 Common Processing Steps

As mentioned in both the tfMRI and rsfMRI analysis, the BOLD signal is susceptible to corruption by a number of instrumental, physiological, or subject-specific factors [114]. Confounding factors can appear signal-like and lead to uncertainty in interpreting results. There are many actions that can be taken to clean an fMRI signal [115]. These actions are referred to as preprocessing. Preprocessing occurs after acquisition and before the signal analysis.

Autocorrelation is the self-similarity of a signal, and typically indicates that elements within the signal are changing at a slower pace than the sampling rate. Autocorrelation can occur from many sources, and will bias the error estimations (and in fact, are a significant contributor to the error term ϵ in equation 2.36 [116]). Noise autocorrelation can be removed either by "pre-colouring" the signal (trying to model the unknown covariance structure within the model) or "pre-whitening" (modelling and removing the signal prior to analysis) [117]. The modelling of autocorrelation is done through an autoregressive model (AR). Autocorrelation bias can be diminished by null-correction through mixture models (part of the functionality of tools like MELODIC [118]). High-pass temporal filtering can remove low-frequency artefacts to limit autocorrelation.

Motion is a major source of artefacts in MRI [119], and specifically in fMRI [120]. This can be addressed by rigid linear correction between frames in cases of small motion (using robust affine registration tools like MCFLIRT [121]), by using navigators in methods like TURBINE which oversample central k-space and produce a low spatial resolution image for registration [34, 37], and/or by estimating the motion within the scan using a motion tracking system (e.g. cameras [122]). Motion correction methods will either aim to detect motion, then either: undo the effects of motion, remove the artefacts created by those effects, or discard corrupted images.

Many image artefacts (including some artefacts from motion) can be removed via an ICA. ICA is the principal functionality of tools like MELODIC [123]. These ICA components are hand classified as artefact or signal [124], or classified automatically by tools like FIX [125]. A wide variety of artefacts can be removed in this fashion, including hardware artefacts, physiological noise, and spin history effects. Principal Component Analysis (PCA) is first used to reduce the number of components required to be extracted (as another denoising method). ICA can be done on a subject-level or group-level [126].

Of course, there are many extra pre-processing techniques that are also used to clean up a dataset. RETROICOR [127] fits a low order temporal Fourier series to the data in order to correct for respiratory and cardiac effects. Extracting the brain (with a tool like BET) removes any confounding signal that will occur in an anatomically impossible location (outside the brain). Slice timing correction is an operation which shifts each slice to halfway through a volume TR using interpolation. Volume-wise spatial smoothing can remove noise and improve Gaussianity. Low-pass filtering can help smooth the signal but may remove frequencies of interest. These and more are all functionality of the FEAT [128] tool.

2.3.2.4 Acceleration

fMRI data requires both good image fidelity (conferring spatial specificity), and high temporal resolution (conferring statistical robustness). Without acceleration, fMRI scans can localise activity to within millimetres in the spatial domain and no better than a TR of a few seconds on a temporal scale [96]. A GRE-EPI scan will generate a T2* image (capturing the BOLD signal) with high temporal efficiency and good resistance to motion [129].

The Nyquist-Shannon criterion would seem to be a fundamental limit on the minimal limit of acquired data that can be used to reconstruct an fMRI scan. However, the benefits of acceleration within an image volume by parallel imaging [130, 131] (section 2.2.2) and partial Fourier (section 2.2.3) have already been noted, with both approaches able to produce high fidelity images using fewer data samples than Nyquist-Shannon dictates as necessary. More generally, the field of acceleration research aims to maximise the level of undersampling achievable while using the remaining data in an optimal manner to avoid undue sacrifice of SNR and CNR [132]. Increasing sampling efficiency in such a fashion could decrease scan times, improve the signal robustness, and increase the information density (see figure 2.15). This could help distinguish the slowly varying BOLD signal [133, 134, 135, 136, 137], improve statistical power in the GLM [138], and allow more defined separation of functional networks [139]. Acceleration can also encode additional information within the signal [140].

Acceleration in fMRI is fundamentally limited by more than Nyquist-Shannon. The temporal resolution of fMRI is fundamentally limited by the HRF [141], which occurs on the scale of seconds (see figure 2.14). More rapid sampling could better characterise the HRF, but will be unable to represent the brain activity occurring at a shorter timescale (which other methods such as EEG are able to do). Shorter TR can also lead to more complicated autocorrelation, and a higher-order AR model or increased white noise modelling is needed to adjust [138]. This increases the model complexity, which increases the chance of overfitting.

However, temporal acceleration could also help produce a cleaner signal by reducing the amount of T2* blurring in an EPI signal (depending on how the signal is undersampled), or by better characterising other features such as physiological noise. These physiological processes can be clearly resolved (i.e. they no longer alias temporally) with a TR<0.5s [138]. An increased temporal resolution leads to higher statistical scores in task fMRI [142] (particularly with a faster task

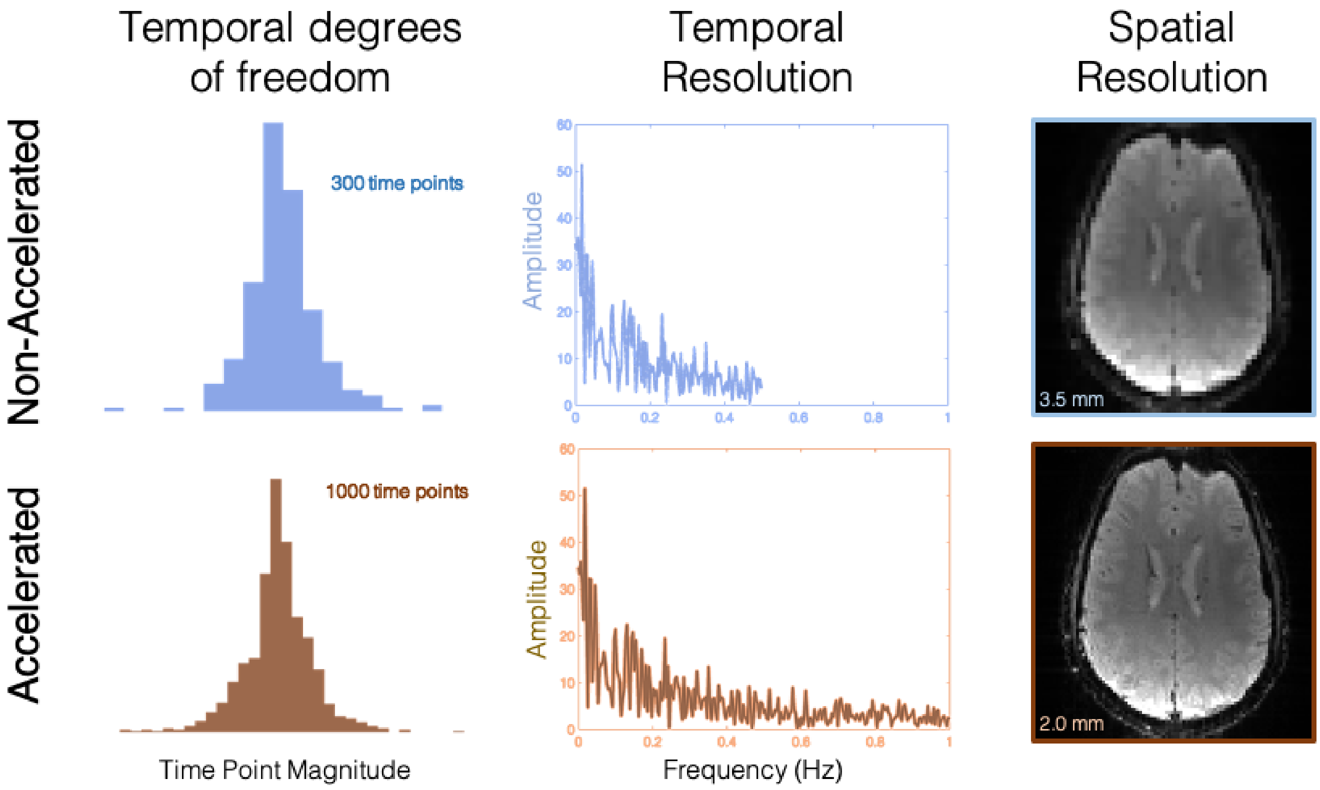


Figure 2.15: An example of some of the benefits of acceleration for dynamic MRI. Collecting fewer points per image could allow increased temporal degrees of freedom (collecting more images over the same time period) and/or temporal resolution. By exchanging fewer lines in k-space for covering a greater range in k-space (and estimating the missing points), the spatial resolution can also be improved. Figure adapted with permission from an original Mark Chiew image.

paradigm), and is good for ICA analysis of rsfMRI (showing enhanced network detection and peak z-scores), although acceleration is less beneficial for seed-based regression [130, 143].

The spatial resolution is limited by a number of factors. In 3T with GRE-EPI, the PSF of BOLD has a Full-Width Half Maximum (FWHM) of 3.9mm [97]. The functional FWHM can be reduced (and better localised) by using: higher field strength, SE imaging (which favours capillaries), or other fMRI approaches like ASL or VASO [93]. Increasing the spatial resolution will also allow better imaging of the underlying structure, especially if the image can resolve functional subunits of the brain which have been previously identified by invasive methods [144]. High field scans particular favour increased spatial resolution due to increased SNR, despite the related distortion increase from T2* decay in high fields (though this distortion is helped by in-plane acceleration) [145].

Small levels of acceleration are now typical in fMRI scans of all types [146, 147]. The most common type of acceleration is PI (see section 2.2.2). However, there are also approaches to acceleration which exploit the dynamic information in a dataset. These will be covered in section 3.5.

2.4 Conclusion

Creating an MRI image is a time-consuming process, fundamentally constrained by many factors. Producing a sequence of images for an fMRI scan will then be restricted by those same factors, limiting the achievable temporal resolution. By using acceleration techniques, less data needs to be collected for each scan. While there is a technical limit to the fMRI resolution benefits achievable with acceleration (defined by the spatial PSF of the BOLD signal and the temporal resolution of the HRF), the increased statistical benefits and improved signal/noise characterisation can be used to better inform the subsequent fMRI analysis.

2.A Toeplitz Embedding

Toeplitz embedding seeks to replace the matrix multiplications typically required for a Gram matrix $A'A$ with more computationally efficient calculations (FFTs and a diagonal matrix multiplication). Assuming A was not already diagonal, Toeplitz embedding dramatically decreases computation time needed to calculate $A'A$, particularly for large matrices.

2.A.1 Theory

A Toeplitz matrix is defined as any matrix containing only constant coefficient diagonals (descending from left to right). The matrix does not have to be square, but must have constant diagonals. An $n \times n$ Toeplitz matrix takes the form:

$$T = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} \\ a_{-1} & a_0 & a_1 & \ddots & & a_{n-2} \\ a_{-2} & a_{-1} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & a_1 & a_2 \\ a_{2-n)} & & \ddots & a_{-1} & a_0 & a_1 \\ a_{1-n} & a_{2-n} & \cdots & a_{-2} & a_{-1} & a_0 \end{pmatrix} \tag{2.39}$$

An $n \times n$ Toeplitz matrix (such as matrix 2.39) has $2n$ - 1 degrees of freedom, rather than the n^2 degrees of freedom that would be assumed in a non-Toeplitz matrix of equivalent size. A square Toeplitz matrix can be fully defined by the first and last column (coefficients $a_{1-n} \rightarrow a_{n-1}$). A Toeplitz system is defined as follows. T is a Toeplitz matrix, and y and x are placeholder vectors.

$$y = Tx \tag{2.40}$$

A circulant matrix is a special case of the Toeplitz matrix. It is a matrix which not only has constant diagonals, but the diagonals can be considered to wrap around once they reach the edge of the matrix. An $n \times n$ circulant matrix takes the form:

$$C = \begin{pmatrix} c_0 & c_{n-1} & c_{n-2} & \cdots & c_2 & c_1 \\ c_1 & c_0 & c_{n-1} & \ddots & & c_2 \\ c_2 & c_1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & c_{n-1} & c_{n-2} \\ c_{n-2} & & \ddots & c_1 & c_0 & c_{n-1} \\ c_{n-1} & c_{n-2} & \cdots & c_2 & c_1 & c_0 \end{pmatrix} \tag{2.41}$$

An n x n circulant matrix only has n degrees of freedom. Circulant matrices often appear as the integration kernel for a convolution operation. Any circulant matrix can be fully described by its first column c, after being diagonalised by the DFT matrix F (see equation 2.20).

$$C = F^{-1} \Lambda F \tag{2.42}$$

Where Λ is the following diagonal matrix:

$$\Lambda = diag(Fc) \tag{2.43}$$

A square Toeplitz matrix of size $n \times n$ can be embedded into a circulant matrix of size $2n \times 2n$, with some simple manipulation.

$$A = \left(\begin{array}{cccccc|cccccc}
a_0 & a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} & 0 & a_{-n+1} & a_{-n+2} & \cdots & a_{-2} & a_{-1} \\
a_{-1} & a_0 & a_1 & \ddots & & a_{n-2} & a_{n-1} & 0 & a_{-n+1} & \ddots & & a_{-2} \\
a_{-2} & a_{-1} & \ddots & \ddots & \ddots & \vdots & a_{n-2} & a_{n-1} & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & a_1 & a_2 & \vdots & \ddots & \ddots & \ddots & a_{n+1} & a_{-n+2} \\
a_{-n+2} & & \ddots & a_{-1} & a_0 & a_1 & a_2 & & \ddots & a_{n-1} & 0 & a_{-n+1} \\
a_{-n+1} & a_{-n+2} & \cdots & a_{-2} & a_{-1} & a_0 & a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} & 0 \\
\hline
0 & a_{-n+1} & a_{-n+2} & \cdots & a_{-2} & a_{-1} & a_0 & a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} \\
a_{n-1} & 0 & a_{-n+1} & \ddots & & a_{-2} & a_{-1} & a_0 & a_1 & \ddots & & a_{n-2} \\
a_{n-2} & a_{n-1} & \ddots & \ddots & \ddots & \vdots & a_{-2} & a_{-1} & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & a_{n+1} & a_{-n+2} & \vdots & \ddots & \ddots & \ddots & a_1 & a_2 \\
a_2 & & \ddots & a_{n-1} & 0 & a_{-n+1} & a_{-n+2} & & \ddots & a_{-1} & a_0 & a_1 \\
a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} & 0 & a_{-n+1} & a_{-n+2} & \cdots & a_{-2} & a_{-1} & a_0
\end{array}\right) \tag{2.44}$$

The same matrix can be more simply portrayed in block matrix form:

$$A = \left(\begin{array}{c|c} T & Q \\ \hline Q & T \end{array}\right) \tag{2.45}$$

T is a Toeplitz matrix (see equation 2.39), and Q is an embedding matrix, which can be generated using the first and last column of T. Any number can be used on the diagonal of Q, with 0 a common choice for the sake of simplicity. This circulant matrix is also described as a block-Toeplitz matrix (although not all block Toeplitz matrices are circulant - Q could be a different Toeplitz matrix, and A would always be a block-Toeplitz matrix, while not necessarily being circulant). A matrix-vector product can be formed using the following relation:

$$\begin{pmatrix} Tx \\ Qx \end{pmatrix} = A \begin{pmatrix} x \\ 0_n \end{pmatrix} \tag{2.46}$$

Where x represents an arbitrary vector in a typical $y = Tx$ system, and A is the circulant matrix. Qx represents an irrelevant noisy vector. 0_n in the right-hand vector indicates that x is zero-padded. Combining equations 2.40, 2.42, 2.43, and 2.46 gives the following equation.

$$\begin{pmatrix} y \\ Qx \end{pmatrix} = \begin{pmatrix} Tx \\ Qx \end{pmatrix} = F^{-1} diag(Fa) F \begin{pmatrix} x \\ 0_n \end{pmatrix} \tag{2.47}$$

a is the first column of the block Toeplitz matrix in equation 2.44. F and F^{-1} can be replaced by an FFT and inverse FFT respectively, and may be point-multiplied (shown with $.*$) to more efficiently calculate the diagonal matrix. The Qx is not needed, so truncation is used for efficiency. This leads to the final equation for an efficient Toeplitz system calculation.

$$y = Tx = truncate\Bigg(ifft\Bigg(fft(a). * fft \begin{pmatrix} x \\ 0_n \end{pmatrix} \Bigg)\Bigg) \tag{2.48}$$

2.A.2 Applying Toeplitz Embedding to MRI

Let's assume a simple MRI scan with Cartesian k-space samples (see section 2.1.5.1). The discretised k-space grid can be transformed into a discretised image space grid through a DFT (equation 2.49). E is the 2D DFT encoding matrix, y is a generic k-space signal input defined on Cartesian coordinates, and x is the output image (also defined on Cartesian coordinates). x and y are vectors, and E has been subsequently reshaped to still perform a 2D DFT.

$$y = Ex \tag{2.49}$$

The problem of solving equation 2.49 for a known x and unknown y can be approached in many ways (see section 3.2). Solving an iterative reconstruction method with a method such as conjugate gradients will lead to the gradient of the squared L2 norm of equation 2.49 being used in the update step in every cycle of the optimisation function.

$$\nabla_{x_i}(\|y - Ex_i\|_2^2) = 2E'y - 2E'Ex_i \tag{2.50}$$

$E'y$ is consistent for every optimisation cycle, so needs only to be calculated once. An important property of the DFT matrix is that its Gram matrix ($E'E$) contains a block Toeplitz structure.

$$E'Ex_i = T_E x_i \tag{2.51}$$

The block Toeplitz within the DFT can be combined with the formulation in equation 2.48.

$$T_E x_i = truncate\Bigg(ifft\Bigg(fft(c_E). * fft \begin{pmatrix} x_i \\ 0_n \end{pmatrix} \Bigg)\Bigg) \tag{2.52}$$

c_E can be formed straight from the first and last columns of the Toeplitz matrix $T_E = E'E$, and so at no point does the circulant matrix actually need to be formed. Given E is the DFT matrix, c_E represents the Point Spread Function (PSF) of the transform.

$$T_E x_i = truncate\Bigg(ifft\Bigg(fft(PSF). * fft \begin{pmatrix} x_i \\ 0_n \end{pmatrix} \Bigg)\Bigg) \tag{2.53}$$

Using c to represent the PSF helps expand the formulation to non-Cartesian k-space sampling. The underlying image x remains Cartesian, and so the diagonalisation of the circulant matrix by FFTs also remains constant. The main difference now is that the PSF cannot be created by simply taking the first column of a Toeplitz matrix. Instead, the PSF is created from the NUFFT operator (section 2.2.1.2) to characterise the sampling pattern. Once the PSF is acquired, a straightforward FFT can be applied to the PSF. It should be noted that using a NUFFT (rather than FFTs) will proportionally increase the overall time taken to optimise by a constant, related to the interpolation time required to map the non-Cartesian sampling onto a Cartesian grid. Equation 2.53 can then be combined with equation 2.50 to form a calculation of the gradient which is far more computationally efficient than the equivalent calculation of $E'E$ using two NUFFT matrices.

$$\nabla_{x_i}(\|y - Ex_i\|_2^2) = E'y - \mathit{truncate}\left(\mathit{ifft}\left(\mathit{fft}(PSF). * \mathit{fft}\begin{pmatrix} x_i \\ 0_n \end{pmatrix}\right)\right) \tag{2.54}$$

Bibliography

[1] Joseph Larmor. On the theory of the magnetic influence on spectra; and on the radiation from moving ions. *Philosophical Magazine*, 44(271):503–512, 1897.

[2] Isidor I. Rabi, Jerrold R. Zacharias, Sidney Millman, and Polykarp Kusch. A new method of measuring nuclear magnetic moment. *Physical Review*, 53(4):318, 1938.

[3] Clotilde Roy, Alisson Slimani, Christophe De Meester, Mihaela Amzulescu, Agnès Pasquet, David Vancraeynest, Jean Louis Vanoverschelde, Anne Catherine Pouleur, and Bernhard L. Gerber. Age and sex corrected normal reference values of T1, T2, T2* and ECV in healthy subjects at 3T CMR. *Journal of Cardiovascular Magnetic Resonance*, 19(1):1–12, 2017.

[4] Felix Bloch. Nuclear induction. *Physical Review*, 70(7-8):460–474, 1946.

[5] Harden M. Mcconnell and Richard E. Robertson. Isotropic nuclear resonance shifts. *The Journal of Chemical Physics*, 29(6):1361–1365, 1958.

[6] Erwin Hahn. An accurate Nuclear Magnetic Resonance Method for Measuring Spin-Lattice Relaxation Times. *Physical Review*, pages 145–146, 1949.

[7] Anatole Abragam. *The Principles of Nuclear Magnetism.* Oxford : Clarendon Press, 1961.

[8] Richard R. Ernst and Weston A. Anderson. Application of Fourier Transform Spectroscopy to Magnetic Resonance. *Review of Scientific Instruments*, 37(1):93–102, 1966.

[9] Erwin Hahn. Spin Echos. *Physical Review*, 80(4):580–594, 1950.

[10] Mehdi Poustchi-Amin, Scott A. Mirowitz, Jeffrey J. Brown, Robert C. McKinstry, and Tao Li. Principles and Applications of Echo-planar Imaging : A Review for the General Radiologist. *Radiographics : a review publication of the Radiological Society of North America, Inc*, 21:767–779, 2001.

[11] Peter Mansfield. Multi-planar image formation using NMR spin echoes. *Journal of Physics C: Solid State Physics*, 10(3):L55–L58, 1977.

[12] Peter Jezzard and Stuart Clare. Sources of distortion in functional MRI data. *Human Brain Mapping*, 8(2-3):80–85, 1999.

[13] Benedikt A. Poser, Peter J. Koopmans, Thomas Witzel, Lawrence L. Wald, and Markus Barth. Three dimensional echo-planar imaging at 7 Tesla. *NeuroImage*, 51(1):261–266, 2010.

[14] Wietske Van Der Zwaag, Susan Francis, and Richard Bowtell. Improved Echo Volumar Imaging (EVI) for functional MRI. *Magnetic Resonance in Medicine*, 56(6):1320–1327, 2006.

[15] W. van der Zwaag, T. Kober, J.P. Marques, Gary Glover, R. Gruetter, and Gunnar Krueger. Comparison of single-shot 2D EPI and segmented 3D EVI acquisition for fMRI at 7T a 2D d. *Proc Intl Soc Mag Reson Med*, 17:1550, 2009.

[16] Nadine N. Graedel. *Three-Dimensional Hybrid Radial Cartesian Echo Planar Imaging for Functional MRI*. PhD thesis, 2016.

[17] Douglas C. Noll, Scott J. Peltier, and Fernando E. Boada. Simultaneous Multislice Acquisition using Rosette Trajectories (SMART): A New Imaging Method for Functional MRI. *Magnetic Resonance in Medicine*, page 1, 1997.

[18] Douglas C. Noll. Multishot rosette trajectories for spectrally selective MR imaging. *IEEE transactions on medical imaging*, 16(4):372–377, 1997.

[19] Klaus Scheffler and Jürgen Hennig. Reduced Circular Field-of-View Imaging. *Magnetic Resonance in Medicine*, 40(3):474–480, 1998.

[20] Kai Tobias Block, Martin Uecker, and Jens Frahm. Undersampled Radial MRI with Multiple Coils . Iterative Image Reconstruction Using a Total Variation Constraint. *Magnetic Resonance in Medicine*, 57:1086–1098, 2007.

[21] John I. Jackson, Craig H. Meyer, Dwight G. Nishimura, and Albert Macovski. Selection of a Convolution Function for Fourier Inversion Using Gridding. *IEEE Transactions on Medical Imaging*, 10(3):473–478, 1991.

[22] Matt A. Bernstein, Kevin F. King, and Xiaohong Joe Zhou. Common Image Reconstruction Techniques. In *Handbook of MRI Pulse Sequences*, chapter 13, pages 491–571. Elsevier Sceince and Technology, 2004.

[23] Stefanie Winkelmann, Tobias Schaeffter, Thomas Koehler, Holger Eggers, and Olaf Doessel. An optimal radial profile order based on the golden ratio for time-resolved MRI. *IEEE Transactions on Medical Imaging*, 26(1):68–76, 2007.

[24] Yoon-Chul Kim, Shrikanth S. Narayanan, and Krishna S. Nayak. Flexible Retrospective Selection of Temporal Resolution in Real- time Speech MRI Using a Golden-Ratio Spiral View Order. *Magnetic Resonance in Medicine*, 48(Suppl 2):1–6, 2011.

[25] Mark Chiew, Nadine N. Graedel, and Karla L. Miller. Promoting incoherence of radial x-f point spread functions using randomly perturbed golden angles. In *Proceedings of the 24th Annual Meeting of ISMRM*, number 2, pages 8–10, Singapore, 2016.

[26] Li Feng, Robert Grimm, Kai Tobias Block, Hersh Chandarana, Sungheon Kim, Jian Xu, Leon Axel, Daniel K. Sodickson, and Ricardo Otazo. Golden-angle radial sparse parallel MRI: combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI. *Magnetic resonance in medicine*, 72(3):707–717, 2014.

[27] Ziwu Zhou, Fei Han, Lirong Yan, Danny J.J. Wang, and Peng Hu. Golden-ratio rotated stack-of-stars acquisition for improved volumetric MRI. *Magnetic Resonance in Medicine*, 78(6):2290–2298, 2017.

[28] Rachel W. Chan, Elizabeth A. Ramsay, Charles H. Cunningham, and Donald B. Plewes. Temporal stability of adaptive 3D radial MRI using multidimensional golden means. *Magnetic Resonance in Medicine*, 61(2):354–363, 2009.

[29] Chang Beom Ahn, Jong Hoon Kim, and Zang Hee Cho. High-Speed Spiral-Scan Echo Planar NMR Imaging-I. *IEEE Transactions on Medical Imaging*, 5(1):2–7, 1986.

[30] Craig H. Meyer, Bob S. Hu, Dwight G. Nishimura, and Albert Macovski. Fast Spiral Coronary Artery Imaging. *Magnetic Resonance in Medicine*, 28(2):202–213, dec 1992.

[31] Gary H. Glover and Song Lai. Self-navigated spiral fMRI: Interleaved versus single-shot. *Magnetic Resonance in Medicine*, 39(3):361–368, 1998.

[32] Daniel M. Spielman, John M. Pauly, and Craig H. Meyer. Magnetic resonance fluoroscopy using spirals with variable sampling densities. *Magnetic Resonance in Medicine*, 34(3):388–394, 1995.

[33] James G. Pipe and Nicholas R. Zwart. Spiral trajectory design: A flexible numerical algorithm and base analytical equations. *Magnetic Resonance in Medicine*, 71(1):278–285, 2014.

[34] James G. Pipe. Motion correction with PROPELLER MRI: Application to head motion and free-breathing cardiac imaging. *Magnetic Resonance in Medicine*, 42(5):963–969, 1999.

[35] Jennifer A. McNab, Daniel Gallichan, and Karla L. Miller. 3D steady-state diffusion-weighted imaging with trajectory using radially batched internal navigator echoes (TURBINE). *Magnetic Resonance in Medicine*, 63(1):235–242, 2009.

[36] Nadine N. Graedel, Mark Chiew, Jennifer A. McNab, and Karla L. Miller. FMRI using a 3D radial-Cartesian trajectory: spatio-temporal tunability and artefact correction. In *Proceedings of the International Society for Magnetic Resonance in Medicine 23*, volume 23, page 1, 2015.

[37] Nadine N. Graedel, Jennifer A. McNab, Mark Chiew, and Karla L. Miller. Motion Correction for Functional MRI With Three-Dimensional Hybrid Radial-Cartesian EPI. *Magnetic Resonance in Medicine*, 78:527–540, 2017.

[38] Marijke Welvaert and Yves Rosseel. On the definition of signal-to-noise ratio and contrast-to-noise ratio for fMRI data. *PLoS ONE*, 8(11):1010, 2013.

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

[40] Rolf Pohmann, Oliver Speck, and Klaus Scheffler. Signal-to-noise ratio and MR tissue parameters in human brain imaging at 3, 7, and 9.4 tesla using current receive coil arrays. *Magnetic Resonance in Medicine*, 75(2):801–809, 2016.

[41] T. W. Redpath. Signal-to-noise ratio in MRI. *British Journal of Radiology*, 71(JULY):704–707, 1998.

[42] Olaf Dietrich, José G. Raya, Scott B. Reeder, Maximilian F. Reiser, and Stefan O. Schoenberg. Measurement of signal-to-noise ratios in MR images: Influence of multichannel coils, parallel imaging, and reconstruction filters. *Journal of Magnetic Resonance Imaging*, 26(2):375–385, 2007.

[43] Christina Triantafyllou, Richard D. Hoge, G. Krueger, C. J. Wiggins, A. Potthast, G. C. Wiggins, and Lawrence L. Wald. Comparison of physiological noise at 1.5 T, 3 T and 7 T and optimization of fMRI acquisition parameters. *NeuroImage*, 26(1):243–250, may 2005.

[44] John D. O'Sullivan. Fast Sinc Function Gridding Algorithm for Fourier Inversion in Computer Tomography. *IEEE Transactions on Medical Imaging*, MI-4(4):200–207, 1985.

[45] Volker Rasche, Roland Proksa, Ralph Sinkus, Peter Börnert, and Holger Eggers. Resampling of data between arbitrary grids using convolution interpolation. *IEEE Transactions on Medical Imaging*, 18(5):385–392, 1999.

[46] James G. Pipe and Padmanabhan Menon. Sampling density compensation in MRI: Rationale and an iterative numerical solution. *Magnetic Resonance in Medicine*, 41(1):179–186, 1999.

[47] Bradley P. Sutton, Douglas C. Noll, and Jeffrey A. Fessler. Fast, iterative image reconstruction for MRI in the presence of field inhomogeneities. *IEEE Transactions on Medical Imaging*, 22(2):178–188, 2003.

[48] Jeffrey A. Fessler, Sangwoo Lee, Valur T. Olafsson, Hugo R. Shi, and Douglas C. Noll. Toeplitz-based iterative image reconstruction for MRI with correction for magnetic field inhomogeneity. *IEEE Transactions on Signal Processing*, 53(9):3393–3402, sep 2005.

[49] Jeffrey A. Fessler. MIRT toolbox.

[50] June-Yub Lee and Leslie Greengard. The type 3 nonuniform FFT and its applications. *Journal of Computational Physics*, 206(1):1–5, 2005.

[51] Leslie Greengard and June-Yub Lee. Accelerating the Nonuniform Fast Fourier Transform. *SIAM Review*, 46(3):443–454, 2004.

[52] Benjamin E. Barrowes, Fernando L. Teixeira, and Jin A. Kong. Fast Algorithm for Matrix-Vector Multiply of Asymmetric Multilevel Block-Toeplitz Matrices. *IEEE*, pages 630–633, 2001.

[53] Rizwan Ahmad, C. D. Austin, and Lee C. Potter. Toeplitz embedding for fast iterative regularized imaging. In *Proceedings of the SPIE*, volume 8051, pages 1–10, 2011.

[54] Shih Yi Lee, Chien Liang Wu, Li Kuo Kuo, Chuan Tsai Lai, Ching Ping Hsu, Hwung Ying Hwung, Yu Wei Chen, Yueh Hsiu Ho, Hui Chuan Hsu, Fang Ju Sun, Ping Fang Yuan, Ivy Pan, and Wen Han Chang. The NMR Phased Array. *Magnetic Resonance in Medicine*, 16:192–225, 1990.

[55] Martin Buehrer, Klaas P. Pruessmann, Peter Boesiger, and Sebastian Kozerke. Array compression for MRI with large coil arrays. *Magnetic Resonance in Medicine*, 57(6):1131–1139, 2007.

[56] Tao Zhang, John M. Pauly, Shreyas S. Vasanawala, and Michael Lustig. Coil compression for accelerated imaging with Cartesian sampling. *Magnetic Resonance in Medicine*, 69(2):571–582, 2013.

[57] Klaas P. Pruessmann, Markus Weiger, Markus B. Scheidegger, and Peter Boesiger. SENSE: Sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, 42(5):952–962, 1999.

[58] Mark A. Griswold, Peter M. Jakob, Robin M. Heidemann, Mathias Nittka, Vladimir Jellus, Jianmin Wang, Berthold Kiefer, and Axel Haase. Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA). *Magnetic Resonance in Medicine*, 47(6):1202–1210, 2002.

[59] Daniel K. Sodickson and Warren J. Manning. Simultaneous acquisition of spatial harmonics (SMASH): Fast imaging with radiofrequency coil arrays. *Magnetic Resonance in Medicine*, 38(4):591–603, 1997.

[60] Peter M. Jakob, Mark A. Griswold, Robert R. Edelman, and Daniel K. Sodickson. AUTO-SMASH: A self-calibrating technique for SMASH imaging. *Magma: Magnetic Resonance Materials in Physics, Biology, and Medicine*, 7(1):42–54, 1998.

[61] Martin Blaimer, Felix A. Breuer, Matthias F. Mueller, Nicole Seiberlich, Dmitry Ebel, Robin M. Heidemann, Mark A. Griswold, and Peter M. Jakob. 2D-GRAPPA-operator for faster 3D parallel MRI. *Magnetic Resonance in Medicine*, 56(6):1359–1364, 2006.

[62] David J. Larkman, Joseph V. Hajnal, Amy H. Herlihy, Glyn A. Coutts, Ian R. Young, and Gsta Ehnholm. Use of multicoil arrays for separation of signal from multiple slices simultaneously excited. *Journal of Magnetic Resonance Imaging*, 13(2):313–317, 2001.

[63] Markus Barth, Felix A. Breuer, Peter J. Koopmans, David G. Norris, and Benedikt A. Poser. Simultaneous multislice (SMS) imaging techniques. *Magnetic Resonance in Medicine*, 75(1):63–81, 2016.

[64] David A. Feinberg and Kawin Setsompop. Ultra-Fast MRI of the Human Brain with Simultaneous Multi-Slice Imaging. *Journal of Magnetic Resonance Imaging*, 144(5):90–100, 2013.

[65] Felix A. Breuer, Martin Blaimer, Robin M. Heidemann, Matthias F. Mueller, Mark A. Griswold, and Peter M. Jakob. Controlled aliasing in parallel imaging results in higher acceleration (CAIPIRINHA) for multi-slice imaging. *Magnetic Resonance in Medicine*, 53(3):684–691, 2005.

[66] David A. Feinberg, James D. Hale, Jeffery C. Watts, Leon Kaufman, and Alexander Mark. Halving MR imaging time by conjugation: Demonstration at 3.5 kG. *Radiology*, 161(2):527–531, 1986.

[67] James R. MacFall, Norbert J. Pelc, and Robert M. Vavrek. Correction of Spatially Dependent Phase Shifts for Partial Fourier Imaging. *Magnetic Resonance Imaging*, 6:143–155, 1988.

[68] G. McGibney, Michael R. Smith, S. T. Nichols, and A. Crawley. Quantitative evaluation of several partial fourier reconstruction algorithms used in mri. *Magnetic Resonance in Medicine*, 30(1):51–59, 1993.

[69] Douglas C. Noll, Dwight G. Nishimura, and Albert Macovski. Homodyne Detection in Magnetic Resonance Imaging. *IEEE Transactions on Medical Imaging*, 10(2):154–163, 1991.

[70] Paul M. Margosian and Franz Schmitt. Faster MR Imaging Methods. *Medical Image Processing*, 0593:6, 1986.

[71] Ewart M. Haacke, E. D. Lindskog, and Weili Lin. A fast, iterative, partial-fourier technique capable of local phase recovery. *Journal of Magnetic Resonance (1969)*, 92(1):126–145, 1991.

[72] Peter A. Bandettini, Eric C. Wong, R. Scott Hinks, Ronald S. Tikofsky, and James S. Hyde. Time course EPI during task activation. *Magnetic Resonance in Medicine*, 25:390–397, 1992.

[73] Kenneth K. Kwong, John W. Belliveau, David A. Chesler, Inna E. Goldberg, Robert M. Weisskoff, Brigitte P. Poncelet, David N. Kennedy, Bernice E. Hoppel, Mark S. Cohen, Robert Turner, Hong-Ming Cheng, Thomas J. Brady, and Bruce R. Rosen. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences*, 89(12):5675–5679, jun 1992.

[74] Seiji Ogawa, David W. Tank, Ravi S. Menon, Jutta M. Ellermann, Seong-Gi Kim, Hellmut Merkle, and Kâmil Uğurbil. Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences of the United States of America*, 89(13):5951–5955, 1992.

[75] Seiji Ogawa, Ravi S. Menon, David W. Tank, Seong-Gi Kim, Hellmut Merkle, Jutta M. Ellermann, and Kâmil Uğurbil. Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. *Biophysical Journal*, 64(3):803–812, mar 1993.

[76] Seiji Ogawa, T. M. Lee, A. R. Kay, and David W. Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, 87(24):9868–9872, dec 1990.

[77] Keith R. Thulborn, John C. Waterton, Paul M. Matthews, and George K. Radda. Oxygenation dependence of the transverse relaxation time of water protons in whole blood at high field. *Biochimica et Biophysica Acta*, 714(2):265–270, 1982.

[78] Hanzhang Lu, Xavier Golay, James J. Pekar, and Peter C. M. Van Zijl. Functional magnetic resonance imaging based on changes in vascular space occupancy. *Magnetic Resonance in Medicine*, 50(2):263–274, 2003.

[79] John A. Detre, John S. Leigh, Donald S. Williams, and Alan P. Koretsky. Perfusion imaging. *Magnetic Resonance in Medicine*, 23(1):37–45, 1992.

[80] Ke Zhang, Dengfeng Huang, and N. Jon Shah. Comparison of resting-state brain activation detected by bold, blood volume and blood flow. *Frontiers in Human Neuroscience*, 12(November):1–10, 2018.

[81] T. Q. Duong, D. S. Kim, K. Uğurbil, and S. G. Kim. Localized cerebral blood flow response at submillimeter columnar resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 98(19):10904–10909, 2001.

[82] Tao Jin and Seong Gi Kim. Cortical layer-dependent dynamic blood oxygenation, cerebral blood flow and cerebral blood volume responses during visual stimulation. *NeuroImage*, 43(1):1–9, 2008.

[83] R. Douglas Fields and Beth Stevens. ATP: An extracellular signaling molecule between neurons and glia. *Trends in Neurosciences*, 23(12):625–633, 2000.

[84] Linus Pauling and Charles D. Coryell. The Magnetic Properties and Structure of Hemoglobin, Oxyhemoglobin and Carbonmonoxyhemoglobin. *Proceedings of the National Academy of Sciences*, 22(4):210–216, 1936.

[85] Max F. Perutz. Stereochemistry of cooperative effects in haemoglobin: Haem-Haem interaction and the problem of allostery. *Nature*, 228(5273):726–734, 1970.

[86] Arthur M. Butt. ATP: A ubiquitous gliotransmitter integrating neuron-glial networks. *Seminars in Cell and Developmental Biology*, 22(2):205–213, 2011.

[87] Cameron Rink and Savita Khanna. Significance of brain tissue oxygenation and the arachidonic acid cascade in stroke. *Antioxidants and Redox Signaling*, 14(10):1889–1903, 2011.

[88] Seong-Gi Kim and Seiji Ogawa. Biophysical and physiological origins of blood oxygenation level-dependent fMRI signals. *Journal of Cerebral Blood Flow and Metabolism*, 32(7):1188–1206, 2012.

[89] Richard D. Hoge, Jeff Atkinson, Brad Gill, Gérard R. Crelier, G. Bruce Pike, and Sean Marrett. Linear coupling between cerebral blood flow and oxygen consumption in activated human cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 96:9403–9408, 1999.

[90] Nikos K. Logothetis. Neurovascular uncoupling: Much ado about nothing. *Frontiers in Neuroenergetics*, 2(June):1–4, 2010.

[91] Nikos K. Logothetis and Josef Pfeuffer. On the nature of the BOLD fMRI contrast mechanism. *Magnetic Resonance Imaging*, 22(10 SPEC. ISS.):1517–1531, 2004.

[92] Dinesh G. Nair. About being BOLD. *Brain Research Reviews*, 50(2):229–243, 2005.

[93] Peter J. Koopmans and Essa Yacoub. Strategies and prospects for cortical depth dependent T2 and T2* weighted BOLD fMRI studies. *NeuroImage*, 197(February 2018):668–676, 2019.

[94] Charles S. Springer Jr., Clifford S. Patlak, Ildiko Palyka, and Wei Huang. Principles of Susceptibility Contrast-Based Functional MRI: The Sign of the Functional MRI Response. In Chrit T.W. Moonen and P. A. Bandetttini, editors, *Functional MRI*, chapter 9, pages 91–102. Springer-Verlag, 1999.

[95] Douglas C. Noll. A primer on MRI and functional MRI, 2001.

[96] Peter A. Bandettini. Seven Topics In Functional Magnetic Resonance Imaging. *Journal of Integrative Neuroscience*, 8(3), 2009.

[97] Laura M. Parkes, Jens V. Schwarzbach, Annemieke A. Bouts, Roel H.R. Deckers, Pim Pullens, Christian M. Kerskens, and David G. Norris. Quantifying the spatial resolution of the gradient echo and spin echo BOLD response at 3 Tesla. *Magnetic Resonance in Medicine*, 54(6):1465–1472, 2005.

[98] Jean Daunizeau. MBB-team/VBA-toolbox, 2018.

[99] Martijn P. van den Heuvel and Hilleke E. Hulshoff Pol. Exploring the brain network: A review on resting-state fMRI functional connectivity. *European Neuropsychopharmacology*, 20(8):519–534, 2010.

[100] Max Wintermark, Pina C. Sanelli, Yoshimi Anzai, Apostolos J. Tsiouris, and Christopher T. Whitlow. Imaging Evidence and Recommendations for Traumatic Brain Injury: Advanced Neuro- and Neurovascular Imaging Techniques. *American Journal of Neuroradiology*, 36:1–11, 2014.

[101] Karl J. Friston, Andrew P. Holmes, Keith J. Worsley, Jean-Baptiste P. Poline, Chris D. Frith, and Richard S. J. Frackowiak. Statistical Parametric Maps in Functional Imaging: A General Linear Approach. *Human Brain Mapping*, 2:189–210, 1995.

[102] Christopher R. Genovese, Nicole A. Lazar, and Thomas Nichols. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, 15(4):870–878, 2002.

[103] K. J. Worsley, A. C. Evans, S. Marrett, and P. Neelin. A three-dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism*, 12(6):900–918, 1992.

[104] Martin M. Monti. Statistical analysis of fMRI time-series: A critical review of the GLM approach. *Frontiers in Human Neuroscience*, 5(MARCH):1–13, 2011.

[105] Bharat Biswal, F. Zerrin Yetkin, Victor M. Haughton, and James S. Hyde. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic Resonance in Medicine*, 34(4):537–541, 1995.

[106] Michael D. Greicius, Ben Krasnow, Allan L. Reiss, and Vinod Menon. Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 100(1):253–258, 2003.

[107] Marcus E Raichle. The restless brain. *Brain Connectivity*, 1(1):3–12, 2011.

[108] Stephen M. Smith, Peter T. Fox, Karla L. Miller, David C. Glahn, P. Mickle Fox, Clare E. Mackay, Nicola Filippini, Kate E. Watkins, Roberto Toro, Angela R. Laird, and Christian F. Beckmann. Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences of the United States of America*, 106(31):13040–13045, 2009.

[109] Neha and Tapan K. Gandhi. Resting state fMRI analysis using Seed based and ICA methods. *IEEE*, 16:2551–2554, 2016.

[110] Aapo Hyvärinen. Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. *IEEE Transactions on Neural Networks*, 10:626–634, 1999.

[111] Aapo Hyvärinen and Erkki Oja. Independent Component Analysis : Algorithms and Applications. *Neural Networks*, 13(4-5):411–430, 2000.

[112] Martin J. McKeown and Terrence J. Sejnowski. Independent component analysis of fMRI data: examining the assumptions. *Human Brain Mapping*, 6(5-6):368–372, 1998.

[113] Xin Gao, Tao Zhang, and Jinhu Xiong. Comparison between spatial and temporal independent component analysis for blind source separation in fMRI data. *Proceedings - 2011 4th International Conference on Biomedical Engineering and Informatics, BMEI 2011*, 2:690–692, 2011.

[114] César Caballero-Gaudes and Richard C. Reynolds. Methods for cleaning the BOLD fMRI signal. *NeuroImage*, 154(December 2016):128–149, 2017.

[115] Keith J. Worsley. Statistical Analysis of Activation Images. In Peter Jezzard, P. M. Matthews, and Stephen M. Smith, editors, *Functional MRI: An introduction to methods*, chapter 14, pages 251–270. Oxford University Press, Oxford, 2001.

[116] Edward Bullmore, Michael Brammer, Steve C.R. Williams, Sophia Rabe-Hesketh, Nicolas Janot, Anthony David, John Mellers, Robert Howard, and Pak Sham. Statistical methods of estimation and inference for functional MR image analysis. *Magnetic Resonance in Medicine*, 35(2):261–277, 1996.

[117] Mark W. Woolrich, Brian D. Ripley, J. Michael Brady, and Stephen M. Smith. Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage*, 14(6):1370–1386, 2001.

[118] Christian F. Beckmann and Stephen M. Smith. Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging. *IEEE Transactions on Medical Imaging*, 23(2):137–152, 2004.

[119] Julian Maclaren, Michael Herbst, Oliver Speck, and Maxim Zaitsev. Prospective motion correction in brain imaging: A review. *Magnetic Resonance in Medicine*, 69(3):621–636, 2013.

[120] Karl J. Friston, Steven Williams, Robert Howard, Richard S. J. Frackowiak, and Robert Turner. Movement-related effects in fMRI time-series. *Magnetic Resonance in Medicine*, 35(3):346–355, 1996.

[121] Mark Jenkinson, Peter R. Bannister, J. Michael Brady, and Stephen M. Smith. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage*, 17(2):825–841, 2002.

[122] Lei Qin, Peter van Gelderen, John Andrew Derbyshire, Fenghua Jin, Jongho Lee, Jacco A. de Zwart, Yang Tao, and Jeff H. Duyn. Prospective head-movement correction for high-resolution MRI using an in-bore optical tracking system. *Magnetic Resonance in Medicine*, 62(4):924–934, oct 2009.

[123] Thomas P Minka. Automatic choice of dimensionality for PCA, 2000.

[124] Ludovica Griffanti, Gwenaëlle Douaud, Janine Bijsterbosch, Stefania Evangelisti, Fidel Alfaro-Almagro, Matthew F. Glasser, Eugene P. Duff, Sean Fitzgibbon, Robert Westphal, Davide Carone, Christian F. Beckmann, and Stephen M. Smith. Hand classification of fMRI ICA noise components. *NeuroImage*, 2017.

[125] Gholamreza Salimi-khorshidi, Gwenaëlle Douaud, Christian F. Beckmann, Matthew F. Glasser, Ludovica Griffanti, and Stephen M. Smith. Automatic Denoising of Functional MRI

Data: Combining Independent Component Analysis and Hierarchical Fusion of Classifiers. *NeuroImage*, 44(0):449–468, apr 2015.

[126] Christian F. Beckmann, Clare E. Mackay, Nicola Filippini, and Stephen M. Smith. Group comparison of resting-state FMRI data using multi-subject ICA and dual regression. *NeuroImage*, 47(Suppl 1):S148, 2009.

[127] Gary H. Glover, Tie-qiang Li, and David B. Ress. Image-Based Method for Retrospective Correction of Physiological Motion Effects in fMRI : RETROICOR. *Magnetic Resonance in Medicine*, 44:162–167, 2000.

[128] Stephen M. Smith, Mark Jenkinson, Mark W. Woolrich, Christian F. Beckmann, Timothy E. J. Behrens, Heidi Johansen-Berg, Peter R. Bannister, Marilena De Luca, Ivana Drobnjak, David E. Flitney, Rami K. Niazy, James Saunders, John Vickers, Yongyue Zhang, Nicola De Stefano, J. Michael Brady, and Paul M. Matthews. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23(SUPPL. 1):208–219, 2004.

[129] Jacco A. De Zwart, Peter Van Gelderen, Peter Kellman, and Jeff H. Duyn. Application of sensitivity-encoded echo-planar imaging for blood oxygen level-dependent functional brain imaging. *Magnetic Resonance in Medicine*, 48(6):1011–1020, 2002.

[130] Lysia Demetriou, Oliwia S. Kowalczyk, Gabriella Tyson, Thomas Bello, Rexford D. Newbould, and Matthew B. Wall. A comprehensive evaluation of increasing temporal resolution with multiband-accelerated protocols and effects on statistical outcome measures in fMRI. *NeuroImage*, 176(May):404–416, 2018.

[131] Benjamin B. Risk, Mary C. Kociuba, and Daniel B. Rowe. Impacts of simultaneous multislice acquisition on sensitivity and specificity in fMRI. *NeuroImage*, 172(January):538–553, 2018.

[132] Todd B. Parrish, Darren R. Gitelman, Kevin S. LaBar, and M. Marsel Mesulam. Impact of signal-to-noise on functional MRI. *Magnetic Resonance in Medicine*, 44(6):925–932, 2000.

[133] Bixente Dilharreguy, Richard A. Jones, and Chrit T.W. Moonen. Influence of fMRI data sampling on the temporal characterization of the hemodynamic response. *NeuroImage*, 19(4):1820–1828, 2003.

[134] Jozien Goense, Hellmut Merkle, and Nikos K. Logothetis. High-Resolution fMRI Reveals Laminar Differences in Neurovascular Coupling between Positive and Negative BOLD Responses. *Neuron*, 76(3):629–639, 2012.

[135] Catie Chang, Moriah E. Thomason, and Gary H. Glover. Mapping and correction of vascular hemodynamic latency in the BOLD signal. *NeuroImage*, 43(1):90–102, 2008.

[136] Richard B. Buxton, Kâmil Uludağ, David J. Dubowitz, and Thomas T. Liu. Modeling the hemodynamic response to brain activation. *NeuroImage*, 23(SUPPL. 1):220–233, 2004.

[137] Laura D. Lewis, Kawin Setsompop, Bruce R. Rosen, and Jonathan R. Polimeni. Fast fMRI can detect oscillatory neural activity in humans. *Proceedings of the National Academy of Sciences*, 113(43):E6679–E6685, oct 2016.

[138] Jingyuan E. Chen, Jonathan R. Polimeni, Saskia Bollmann, and Gary H. Glover. On the analysis of rapidly sampled fMRI data. *NeuroImage*, 188(January):807–820, 2019.

[139] Stephen M. Smith, Karla L. Miller, Steen Moeller, Junqian Xu, Edward J. Auerbach, Mark W. Woolrich, Christian F. Beckmann, Mark Jenkinson, Jesper L. Andersson, Matthew F. Glasser, David C. Van Essen, David A. Feinberg, Essa Yacoub, and Kâmil Uğurbil. Temporally-independent functional modes of spontaneous brain activity. *Proceedings of the National Academy of Sciences*, 109(8):3131–3136, 2012.

[140] S. Sophie Schauman, Mark Chiew, and Thomas W. Okell. Highly accelerated vessel-selective arterial spin labeling angiography using sparsity and smoothness constraints. *Magnetic Resonance in Medicine*, 83(June):892–905, mar 2019.

[141] Seong Gi Kim, Wolfgang Richter, and Kâmil Uğurbil. Limitations of temporal resolution in functional MRI. *Magnetic Resonance in Medicine*, 37(4):631–636, 1997.

[142] L. Chen, A. T. Vu, J. Xu, S. Moeller, K. Ugurbil, E. Yacoub, and D. A. Feinberg. Evaluation of highly accelerated simultaneous multi-slice EPI for fMRI. *NeuroImage*, 104:452–459, 2015.

[143] David A. Feinberg, Steen Moeller, Stephen M. Smith, Edward J. Auerbach, Sudhir Ramanna, Matthew F. Glasser, Karla L. Miller, Kâmil Uğurbil, and Essa Yacoub. Multiplexed Echo Planar Imaging for Sub-Second Whole Brain FMRI and Fast Diffusion Imaging. *PLoS ONE*, 5(12), 2010.

[144] Nikos K. Logothetis. What we can do and what we cannot do with fMRI: Supplementary Material. *Nature*, 453(7197):869–878, jun 2008.

[145] Kawin Setsompop, David A. Feinberg, and Jonathan R. Polimeni. Rapid brain MRI acquisition techniques at ultra-high fields. *NMR in Biomedicine*, 29(9):1198–1221, 2016.

[146] Stephen M. Smith, Fidel Alfaro-Almagro, and Karla L. Miller. UK Biobank Brain Imaging Documentation. Technical Report May, UK Biobank, 2019.

[147] Kâmil Uğurbil, Junqian Xu, Edward J. Auerbach, An T. Vu, Pierre Francois Van de Moortele, John Strupp, Guillermo Sapiro, Federico De Martino, Dingxin Wang, Noam Harel, Michael Garwood, Liyong Chen, David A. Feinberg, Stephen M. Smith, Karla L. Miller,

Stamatios N. Sotiropoulos, Saad Jbabdi, Jesper L. Andersson, Timothy E. J. Behrens, Matthew F. Glasser, David C. Van Essen, and Essa Yacoub. Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. *NeuroImage*, pages 80–104, 2013.

3 | Advanced Image Reconstruction Techniques

Contents

3.1 Introduction

In 1929, surrealist painter Renè Magritte painted a very ordinary-looking pipe, and captioned his painting "Ceci n'est pas une pipe", which translates as: "This is not a pipe". He titled the piece "The

Treachery of Images" with his point being thus - that all images are merely a flawed representation of the object they are hoping to represent, rather than a direct replication of the object. The act of representing an object as an image will undoubtedly cause some corruption in how the object looks and behaves. In Magritte's case, he regretted the two-dimensionality of the representation which meant he was unable to stuff his representation of a pipe with tobacco [1].

The act of reconstructing a representation of the human brain from an MRI signal (particularly the functional representation of a BOLD fMRI signal) is, similar to Magritte's pipe, filled with challenges.

The first challenge is that sampling occurs in a domain (k-space) which is different from the domain which will be used to represent the data (an image space representation). A transform is required between the two domains in order to visualise the image. This first challenge is surmountable with a perfect transform, but certain issues (e.g. field inhomogeneity, T2* readout decay, motion) mean that some distortion in the representation of the underlying object will likely occur during the act of sampling.

The second challenge is that the discrete voxels used to represent the data will never be able to precisely represent the continuous nature of the real world. The sampling point spread function in MRI is unable to resolve down to a neuronal level, meaning each voxel contains a mixture of signals from various indistinguishable sources.

Additionally, fMRI seeks to represent a functional image of the brain using the BOLD signal as a proxy for local brain function. While the BOLD signal is a good proxy for mass neuronal activity, it can be influenced by other factors while also introducing an inherent time lag [2]. Various noise factors and confounding physiological artefacts provide extra challenges that may muddy both the signal and the subsequent image reconstruction further.

The act of reconstructing an image representation of the brain is a difficult task, even in the best possible scenario. Further to this, section 2.3.2.4 highlighted the desire to undersample in k-space while still producing an accurate representation of the functional information in the brain. In image reconstruction, the undersampling limit set by Nyquist-Shannon (see section 2.1.4.3) states that the sampling rate in each direction (i.e. Δk_x in the x-direction) must not be less than the number of rows of the image N_x multiplied by the distance between each row Δx:

$$\Delta k_x \geq \frac{1}{\Delta x N_x} \tag{3.1}$$

Reconstruction of a signal which has violated the Nyquist-Shannon sampling requirement will lead to aliasing unless additional information is incorporated into the reconstruction. This additional information can take many forms, e.g. spatial variation in coil sensitivity maps, or mathematical dataset structures such as low rank and sparsity. The act of incorporating the mathematical dataset structure information into image reconstruction in order to sample below the Nyquist-Shannon limit (without leading to aliasing) will be the focus of this chapter. The physics behind image reconstruction were covered in sections 2.1.4, 2.1.5, and 2.2. This chapter will instead focus on the mathematics behind image reconstruction in an undersampled domain, before finishing with a discussion on popular dynamic undersampling approaches.

3.2 Optimisation

Optimisation is the process of selecting the best solution from a set of possible solutions. The best solution is defined with respect to some criteria, which can be captured inside the objective function. The objective function maps a set of input parameters onto a single real value, with that value defining how well that particular set of parameters fits the outlined criteria. If the objective function is to be minimised by the optimisation process it is called a cost function.

Optimisation can be split into two approaches: direct optimisation methods, and iterative optimisation methods. If a solution is computationally expensive or mathematically impossible to find through the direct approach, then an iterative approach is used. The precise dividing line between direct and iterative methods is not well defined. Here, direct methods are defined as methods with a fixed number of steps, whereas iterative methods contain a convergence criterion. The following section will work through the motivation behind optimisation, and outline a typical formulation to solve an optimisation problem through iteration.

For this section, x will refer to a ground truth vector (e.g. the image of a brain in MRI), A is an operator matrix (e.g. the Fourier/Coil encoding in MRI) that acts upon x, and y is a vectorised set of the set of measurements (e.g. the k-space signal). A' is the adjoint of A. e refers to the noise, which is assumed here to have a zero-mean random Gaussian distribution. They are related as follows:

$$y = Ax + e \tag{3.2}$$

3.2.1 Forming the Optimisation Problem

There are two sets of approaches to solving the cost function - direct approaches and iterative approaches. For now, only one input variable into the cost function will be considered.

3.2.1.1 Direct Methods

Direct methods take an analytic approach in order to find an optimal solution to a defined problem. The forward problem is a simple example of a direct method, as it solves for an unknown signal ($y_{unknown}$) by multiplying two known variables (the transform A and a known MRI image x_{known}).

$$y_{unknown} = Ax_{known} \tag{3.3}$$

The inverse problem is more complicated to solve. Here, we solve for $x_{unknown}$ (the image of the brain) using y_{known} (the sampled signal) and A (the transform):

$$y_{known} = Ax_{unknown} \tag{3.4}$$

The most straightforward approach to solving the inverse problem is to pre-multiply each side of equation 3.4 by the transform inverse A^{-1}, forming the solution $x = A^{-1}y$. If A is either not square or has a zero determinant, then it is non-invertible and a different approach must be taken.

If A is non-square then a solution can be found by using a Moore-Penrose pseudoinverse matrix [3, 4], which forms $x = (A'A)^{-1}A'y$ as a solution. If A has more rows than columns, the system is overdetermined and many solutions exist. If A has more columns than rows, the system is underdetermined and no true solution exists. The solution provided by the pseudoinverse is the least-squares solution, which minimises the Euclidean distance ϵ (a.k.a. the error) between the transformed estimate Ax and the observation y. This distance can be alternatively stated as the L2 norm of the difference between the two terms (for more on norms see section 3.A).

$$\|y - Ax\|_2 = \epsilon \tag{3.5}$$

Squaring both sides of equation 3.5 removes the need for a square root calculation in the L2 norm, and also allows the function to be represented through simple matrix algebra: $\|y - Ax\|_2^2 = (y - Ax)'(y - Ax)$. This formulation is also known as the sum of squares, and can form a cost function $\|y - Ax\|_2^2$ inside an optimisation formulation:

$$\hat{x} = argmin_x \Bigg(\|y - Ax\|_2^2 \Bigg) \tag{3.6}$$

$argmin(\ )$ states that the x parameter chosen as the output $\hat{x}$ provides the smallest possible output of the cost function $\|y - Ax\|_2^2$. If the min( ) function is used, the output is the minimum value of the cost function, rather than the parameter which provides that solution. The use of the L2 norm ensures a single stable solution which minimises the overall energy in the error term. The cost function in equation 3.6 is typically referred to as a “data consistency” term, and is often used in conjunction with other terms. In order to solve the optimisation problem in equation 3.6, we will start by finding the gradient.

$$\begin{aligned}\nabla_x(\|Ax-y\|_2^2) &= \nabla_x\Big((Ax-y)'(Ax-y)\Big)\\ &= \nabla_x\Big((x'A'-y')(Ax-y)\Big)\\ &= \nabla_x\Big((y'y-x'A'y-y'Ax+x'A'Ax)\Big)\\ &= (-2A'y+2A'Ax)\end{aligned} \tag{3.7}$$

If the gradient is zero then the optimal solution is $\hat{x} = (A'A)^{-1}A'y$, which is the pseudoinverse matrix. The Hessian matrix (the second differential of the cost function with respect to x) is a positive definite matrix ($\nabla_x\nabla_x(\|Ax-y\|_2^2) = A'A$), provided that A has full column rank. If the Hessian is positive definite, the solution $\hat{x}$ at $\nabla_x(\|Ax-y\|_2^2) = 0$ is a minimum solution, and therefore the optimum solution to the cost function. The derivation in equation 3.7 uses the following definitions:

$$\|Ax\|_2^2 = x'A'Ax \tag{3.8}$$

$$\nabla_x(Ax) = A \tag{3.9}$$

$$\nabla_x(x'A') = A' \tag{3.10}$$

$$\nabla_x(x'A'Ax) = 2A'Ax \tag{3.11}$$

The inversion approach $\hat{x} = A^{-1}y$ will find the exact solution, and the pseudoinverse will provide the minimum distance solution defined by equation 3.6. However, both approaches contain a matrix inversion of some kind, which can be extremely costly from a computational perspective. Other approaches (such as matrix decomposition) can also be used as a way to more efficiently implement matrix algorithms, including matrix inversion.

3.2.1.2 Iterative Methods

A major issue with inversion approaches in problems with a large truth and/or a large number of measurements is that inverting a large matrix takes a huge amount of calculations. This requires a lot of time and computing power, even with the more efficient inversion techniques.

Instead, iterative approaches are often a more attractive option for solving optimisation problems. Unlike direct approaches, iterative reconstructions contain an initially unknown number of steps. Iterative methods vary the unknown parameter x to try to find a minimum solution to the cost function. The parameter must be initialised, and then an update to that parameter is required (equation 3.12). This update is then applied multiple times in order to improve the cost function compared to the initial estimate (not necessarily monotonically).

$$x_{i+1} \leftarrow f(x_i) \tag{3.12}$$

x_i is the initial guess, and x_{i+1} is the updated guess. The precise form of the update can take many forms, all of which trade-off between accuracy, robustness, and speed. Many approaches involve finding the gradient of the cost function (e.g. conjugate gradients), although there are also many that don't (e.g Nelder-Mead, Monte Carlo methods). The gradient can be found for a simple data consistency term using the approach in equation 3.7.

From this point, an iterative solution to the inverse problem can be found. Given an observed dataset y, a known transform A, a cost function formulation, an update step, and an initial guess of x; a reasonable approximate value of x can be calculated by iterating between cost function calculation and update steps. Once the value of x is deemed to have converged with respect to some criteria, the iterative problem is solved.

3.2.2 Regularisation

As previously defined, optimisation is the process of finding the best solution to a problem (given a set of constraints). The form of the best solution can be altered by changing the form of the cost function. While the cost function can be just a data consistency term (as outlined in equation 3.6), solving the cost function in such a form can lead to undesirable solutions, particularly where the noise is significant.

Regularisation is the practice of adding additional terms to the cost function and is typically used either where the problem is ill-conditioned or where the problem is susceptible to overfitting. Adding a regularisation term changes the form that x is likely to take as a solution, with the precise nature of restriction being dependent on the type of regularisation.

One common approach is to try to find the minimal energy form of x which still minimises the errors arising from the data consistency term.

$$\text{minimise } \|x\|_2^2 \text{ such that } \|Ax - y\|_2^2 < \epsilon \tag{3.13}$$

These terms can be combined into one cost function, known as Lagrangian form.

$$\hat{x} = argmin_x \bigg(\|y - Ax\|_2^2 + \lambda \|x\|_2^2 \bigg) \tag{3.14}$$

The relative weightings within a cost function are controlled with λ parameters (a.k.a. Lagrangian Multipliers). Separate norm terms within a cost function tend to be controlled by a separate λ weighting parameters. The relative λ size between different terms will control the influence of those terms. The first term usually has an implicit weighting of 1 or an explicit weighting of $\frac{1}{2}$.

The solution of regularised optimisation is always a saddle point of the Lagrangian [5]. Iterative reconstructions can be quite slow, therefore careful choices must be made regarding the form of the update step and the cost function if an iterative approach is taken (although a direct approach can be taken to solve equation 3.14 if appropriate). Forming the regularised optimisation problem to use the L2 norm of a parameter (as in equation 3.14) is referred to as Tikhonov regularisation (or ridge regression), and is often used to prevent overfitting. The $\lambda\|x\|_2$ term is referred to as a Tikhonov constraint. Other $\|x\|_p$ terms in a cost function will guide x towards other characteristic forms (e.g. an L1 term will maximise sparsity).

3.2.3 Optimisation and Image Reconstruction

This section has served as a guide to the basic principle behind optimisation. The building blocks of norms, regularisation, and iterative optimisation methods are vital in understanding how advanced image reconstruction methods can incorporate a variety of mathematical information into the construction of different optimisation problems.

It is good at this point to take a small step back and relate this framework back to the previously stated equations. Equation 2.19 states that $S = EI$, where S is the signal obtained from the scanner, I is an unknown image, and E is an operator which multiplies the images by coil sensitivity maps, and then transform the output from image space into k-space. This directly relates to the $y = Ax$ problem outlined in this section, as S and E are known and can be used to solve for I. An undersampled signal will begin to produce aliasing artefacts in I. While coil sensitivities can help to limit the impact of the artefacts (see section 2.2.2), including additional regularisation in the optimisation problem can help constrain the problem, allowing for additional undersampling. The next sections will walk through how sparsity and low-rank information respectively can be incorporated into an optimisation problem.

3.3 Compressed Sensing

Compressed Sensing (CS) is a way of exploiting structure to reconstruct a signal with less information than is dictated by Nyquist-Shannon. In order to comprehend CS there are two key concepts that need to be fully grasped: compression and sparsity. The general principles behind CS will be discussed first, before applying CS to an MRI framework.

Compressibility

The first concept involved in CS is compressibility. A compressed dataset aims to capture the same information as an uncompressed dataset - but to do so with far fewer bits of data. The fewer bits of data you need, the more compressible a dataset is. Compression can be lossless (all information is preserved) or lossy (information is irreplaceably lost). Lossy compression algorithms aim to lose information that is irrelevant to the user while focusing on preserving the information that defines the usefulness of the dataset.

To introduce a simple example of compressibility, let's form a numberline N of the numbers 1, 3, 5, 7, 9, 11, 13, 15, 17, 19 (figure 3.1).

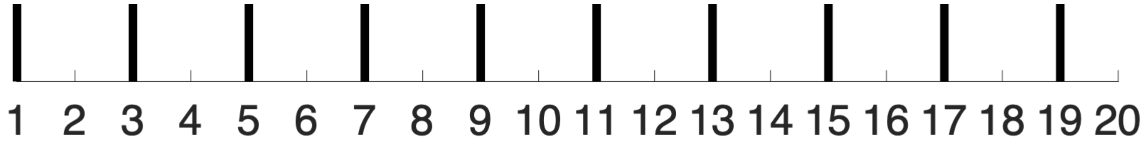


Figure 3.1: A number line of odd numbers between 1 and 20.

Now, while this set of numbers could be referred to by the full set of elements every time, there are other more efficient ways to express this information requiring far fewer coefficients. For example, the following ways to define N are all equivalent in MATLAB:

$$N = [1\ 3\ 5\ 7\ 9\ 11\ 13\ 15\ 17\ 19]$$
$$N = [1 : 10] * 2 - 1$$
$$N = [0 : 9] * 2 + 1$$
$$N = 1 : 2 : 20$$

This final expression ($N = 1 : 2 : 20$) is probably the closest to the way a person would naturally describe N - the odd numbers between 1 and 20. It is far easier to remember N in that form than as each of the 10 individual numbers, as it only requires three coefficients (1, 2, 20) and knowledge of the structure that relates them.

This is an example of lossless compression - a more efficient representation of the same underlying information. Sometimes loss is considered acceptable if it enables significant compression. A numberline of all odd numbers between 1 and 10,000 - but also the number 14.37 - would lose comparatively little information by expressing it only as $N = 1 : 2 : 10,000$. Rather than stating 5001 numbers, instead only 3 numbers are used in combination with a specific operator, and yet only 0.02% of the information is lost in doing so. Outliers in a structure are often noise, but are occasionally the most desirable knowledge in a dataset. In our example, 14.37 is an outlier, not because it lies outside of the range described, but because it does not fit the pattern that we wish to enforce on the data. Knowledge about the nature of non-zero coefficients within a dataset is vital in discerning which type of information is likely to be lost in a particular mode of compression.

For a more visual example, consider the field of compression in images. JPEG is a well-known image compression format first proposed in 1972 [6]. JPEG uses a Discrete Cosine Transform (DCT), which is a discrete Fourier transform that only uses real numbers (i.e. as a weighted sum of cosine functions at different frequencies). As many of the weights on the basis functions in the DCT will be very small, they can be set to zero without much loss in image quality. The locations of the more abundant coefficients in the DCT are efficiently stored through Huffman encoding [7], compressing the image. Natural images are often highly compressible [8].

The effect of compression (and specifically overcompression) can be seen in figure 3.2. At 50% compression there is almost zero loss in image quality which is visible to the naked eye when compared to the uncompressed image. This is because JPEG specifically aims to compress the image in ways that are hard to detect in the human eye (e.g. the human eye struggles to detect differences in brightness of <1%, the eye detects brightness on a logarithmic scale) [9].

At 90% compression, a significant amount of data has been lost due to these assumptions being pushed too far. Smooth backgrounds are transformed into visually discrete blocks of colour with sharp boundaries, and the streaky ray of light in the detailed image has lost its clearly defined edges. However, if the only signal of interest in the patterned image was the zigzag pattern, then the relevant information of the first image would still be recoverable. Compression applied to an already formulated dataset is typically done for data storage reasons.

Compressed sensing is compression that occurs at the point of sensing (gathering the signal), as opposed to compression applied to an already sampled image. CS requires prior knowledge regarding the domain in which compression can occur, which leads to the second key concept in

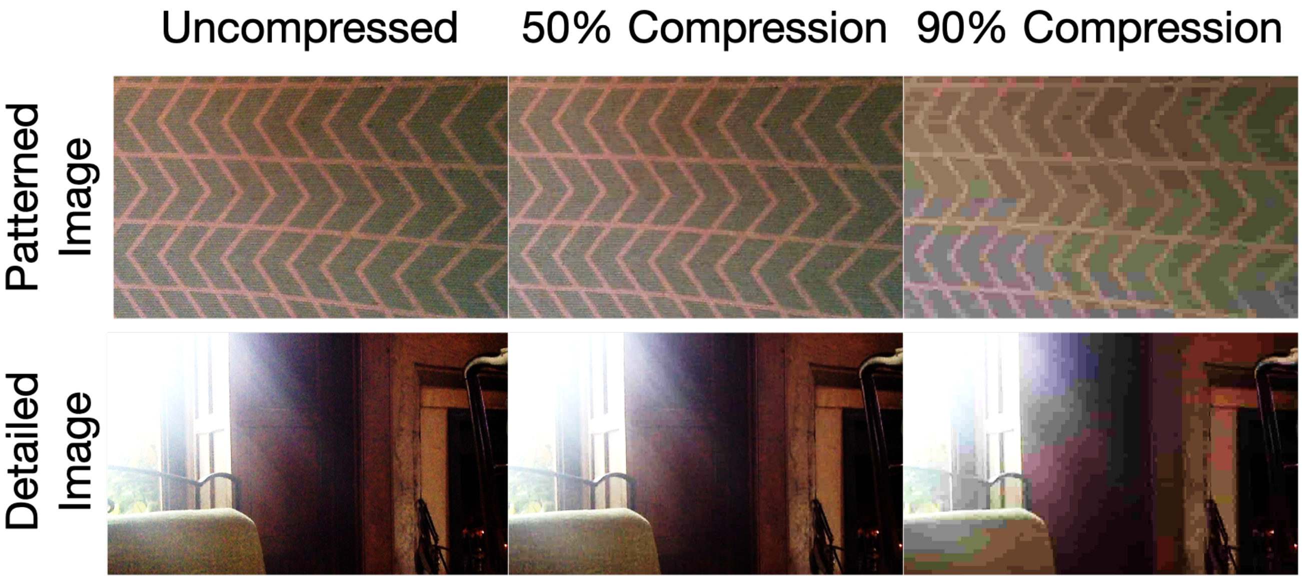


Figure 3.2: A demonstration of JPEG compression of two separate images. Top row: an image with a simple pattern. Bottom row: an image with more detailed features. The images are displayed with zero compression, 50% compression, and 90% compression. Compression was done using the JPEG algorithm on hnet.com.

CS: sparsity.

Sparsity

Sparsity is defined as the number of zero elements in a dataset compared to the overall number of elements. Any matrix with a large ratio of zero elements to non-zero elements is known as a sparse matrix (otherwise it is known as a dense matrix). An alternate definition could state: in a sparse signal, all of the energy is concentrated in a few coefficients. A distinction is also sometimes made between strong sparsity (very few non-zero elements) and weak sparsity (more non-zero coefficients while still being sparse), although these terms do not have a strict definition. Sparsity is a very desirable property in CS, as it defines a domain in which a dataset may be compressible. Once a dataset is compressed, it can be represented by fewer coefficients. Sparsity is commonly maximised, although since absolute maximal sparsity would zero all coefficients, a data consistency term is also needed to ensure that the underlying measurements are captured with some accuracy within the sparse solution.

Sparsity is not necessarily the same in different domains; a dataset may be very sparse and compressible in one domain but be dense in another. Figure 3.3 demonstrates this principle. If only the 10% largest image space coefficients are kept, the image degrades heavily. However, the 2D Fourier transform is dominated by a relatively small number of coefficients (the biggest coefficient is 50,000x bigger than the median). Retaining only the top 10% of the Fourier coefficients and removing the rest still produces a visually identical dataset in image space, despite the fact that the Fourier compression retained the same number of coefficients as the image space compression. Finding a particular sparse domain is then vital in CS, but is not the only condition that must be satisfied when carrying out a CS reconstruction.

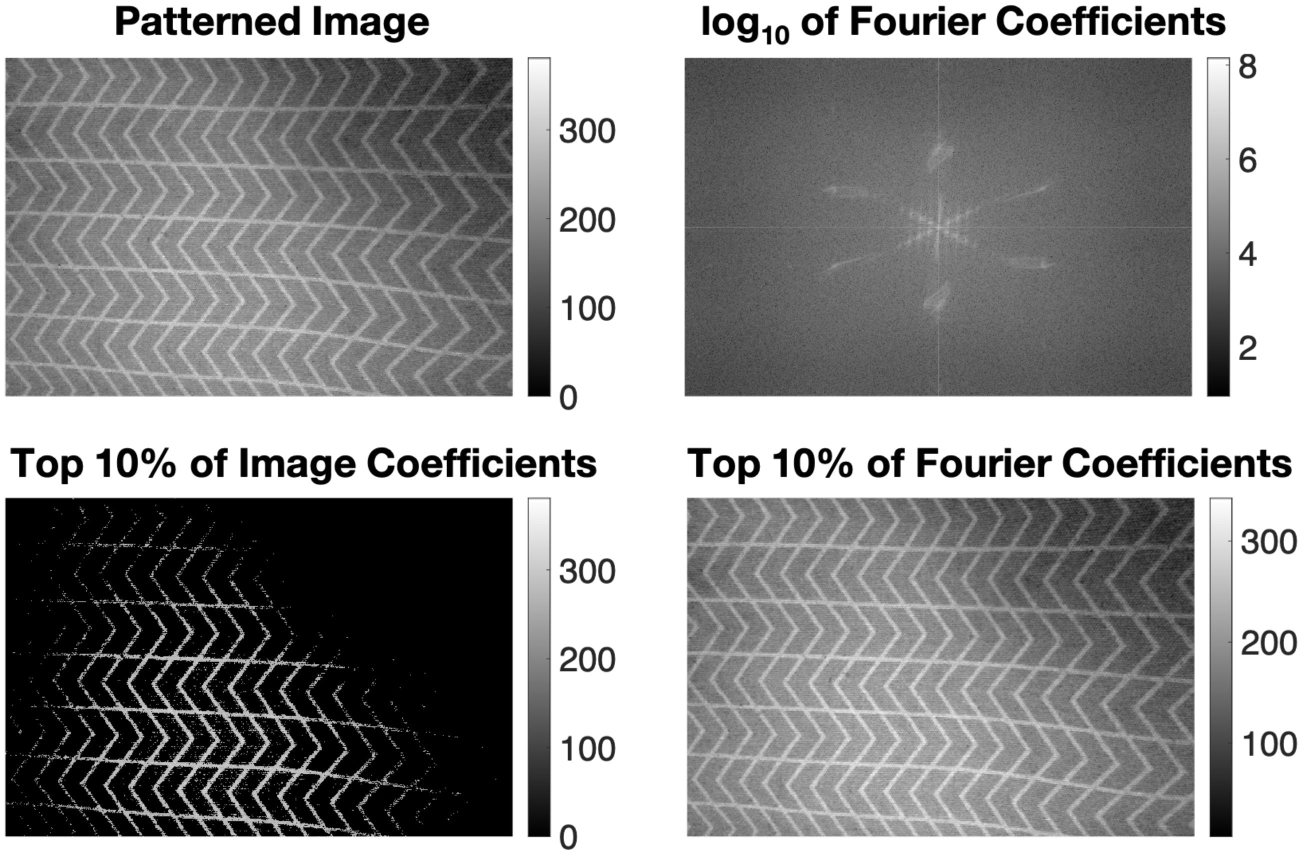


Figure 3.3: A representation of domain-specific sparsity. The patterned image is very dense in the image domain (top left), but very close to sparse in k-space (top right). The bottom images show the effect of enforcing sparsity in image space (bottom left) and k-space (bottom right).

3.3.1 CS Implementation

3.3.1.1 Conditions

CS has two conditions that must be satisfied to allow for equivalent reconstruction between a fully sampled signal and a signal acquired at a sampling rate below that required by Nyquist-Shannon [10]. A third condition is sometimes also included to explicitly state that sparsity must be promoted in a relevant domain in the reconstruction process, while still ensuring reconstruction consistency with the acquired samples.

The two core conditions are:

1. The dataset is sparse in some domain which is accessible via a mathematical transform
2. The sampling domain and the sparse domain must be incoherent with each other

Condition 1 is satisfied by finding a domain in which the desired dataset is sparse. Common domains include image space, spatial/temporal frequency domains, wavelet domains, or gradient domains, the latter of which has sparsity promoted through total variation norms (TV) [11]. The isotropic spatial TV norm of an image I could be represented as $\|I\|_{TV}$ or $TV_x(I)$, and is defined as follows:

$$TV_x(I) = \|I\|_{TV} = \sum_{i=1}^{N-1}\sum_{j=1}^{M-1}\sqrt{|I_{i+1,j} - I_{i,j}|^2 + |I_{i,j+1} - I_{i,j}|^2} \tag{3.15}$$

Condition 2 is difficult to calculate precisely, but broadly states that a CS reconstruction requires some degree of incoherence between the sparse domain and the domain in which sampling occurs. As a domain cannot be incoherent with itself, it is worth emphasising that the sparse domain cannot be the sampling domain (and so k-space sparsity is not used in MRI).

More strict attempts to define condition two are the Null Space Property (NSP) and the Restricted Isometry Property (RIP). The NSP is a necessary and sufficient fulfilment of condition 2 in the absence of noise [12], and roughly states that two uniquely sparse vectors cannot be subtracted from each other and lie in the operator null space (for more on null spaces, see section 3.4.2). The Restricted Isometry Property (RIP) is an alternate test of incoherence which works with noise and also satisfies the NSP [13]. The RIP states that, for a vector x, encoding matrix G there exists a constant spacing $0 < \delta_s < 1$ where δ_s satisfies the following equation for all x:

$$(1 - \delta_s)\|x\|_2^2 \leq \|Gx\|_2^2 \leq (1 + \delta_s)\|x\|_2^2 \tag{3.16}$$

In MRI, the user has a very limited choice of G, and the RIP is best defined for random Gaussian matrices [14] (which are impossible to construct as a sensing matrix in MRI). As such, rather than satisfy condition 2 with these formal properties, the accepted solution in MRI is simply to make G as incoherent as possible. The coherence is defined as the maximum inner product between any two columns of G:

$$\mathit{Coherence} = \max(\ \langle\ G(:,i), G(:,j)\ \rangle\ )\quad \text{for } i = 1:N, j = 1:N. \tag{3.17}$$

Maximising incoherence (or minimising coherence) is done by minimising the maximum column-wise inner product in the operator. Practically, this is done by measuring the maximum off-diagonal value in a normalised point spread function of the samples. Incoherence should spread energy as uniformly as possible across the image. Non-Cartesian sampling such as radial [15] and variable-density spirals [16, 17] are commonly used due to their favourable incoherence properties. Incoherence is far easier to calculate than the RIP and still produces satisfactory results. The satisfaction of condition 2 is required for the L1 norm to serve as an accurate substitution for the L0 norm when calculating sparsity [18].

3.3.1.2 L1 Norm

In order to incorporate a sparsity-enforcing term into the reconstruction, a particular type of regularisation is required.

The L0 norm is strictly a count of non-zero elements, and is the basic definition of the sparsity of a vector. However, the L0 norm is not by nature convex, which is a desirable property when forming cost functions. Practically, the L1 norm can be used instead of the L0 norm as an alternate measure of sparsity [19], as long as the Restrictive Isometry Property (RIP) is obeyed [20]. The L1 norm is an acceptable substitute because it has a tendency to select solutions with fewer coefficients when used in a cost function. The L1 norm is defined as the sum of the absolute element values of x:

$$\|x\|_1 = \left(\sum_{i=1}^{n} |x_i| \right) \tag{3.18}$$

The L1 norm is less sensitive to outliers than the L2 norm. While the L1 norm is convex, it is not smooth. A small positive perturbation μ ($1 \times 10^{-15} < \mu < 1 \times 10^{-6}$) can be added to the absolute function to form an approximation to the L1 norm which is smooth and differentiable [21]:

$$|x| \approx \sqrt{x'x + \mu} \tag{3.19}$$

The L1 norm is used as a sparsity-enforcing regularisation term by adding a transform into a sparse domain (ψ)

$$\lambda\|\psi x\|_1 \tag{3.20}$$

The precise manner in which the L1 norm is minimised by enforcement of sparsity is implementable in many ways.

3.3.1.3 Solving a CS Optimisation Problem

The implementation of a simple CS cost function can be done most straightforwardly with a conjugate gradient method on the combined equation with an L2 norm and an L1 norm. Here, E represents the transform between k-space and image space.

$$CF_{CS} = \|Ex - y\|_2^2 + \lambda\|\psi x\|_1 \tag{3.21}$$

Using the adapted L1 norm [21] defined in equation 3.19, the gradient can then be defined as:

$$\nabla CF_{CS} = 2E'(Ex - y) + \lambda\psi' W^{-1}\psi x \tag{3.22}$$

Where W is a diagonal weighting matrix with elements $w_{i,i} = \sqrt{(\psi x_i)'(\psi x_i) + \mu}$.

Other approaches to CS have gained traction and overtaken simple convex optimisation as the preferred methodology. One such approach is ISTA (Iterative Shrinkage-Thresholding Algorithm) [22], which applies a shrinking soft threshold to a gradient of the cost function in the iterative step. FISTA (Fast ISTA) [23], seeks to do the same thing as ISTA, but achieve convergence at a much faster rate by applying a threshold to a linear combination of previous guesses. An Iterative Reweighted Least Squares (IRLS) algorithm [24] uses a weighted L2 norm minimisation on x instead of L1, with elementwise weights of $w_i = \sqrt{|x_i|^2 + \gamma}^{-1}$, where γ is a stabilisation parameter. The Split Bregman method [25] decouples the L1 and L2 elements of the optimisation, allowing for the L1 portion to be solved by thresholding and shrinking the number of non-zero elements and for the data consistency L2 portion to be solved analytically.

3.4 Low Rank

Rank, like sparsity; is an exploitable structure in a dataset; there are many parallels that can be drawn between a CS and a low-rank reconstruction [26]. While the fields of CS methods and low-rank reconstruction methods are both relatively recent, the application of low-rank methods to MRI (and specifically fMRI) is still in its infancy, despite the inherently low-rank nature of fMRI. Much of the grounding and principles behind the slightly more developed CS field can then be applied to low-rank approaches.

3.4.1 Rank and Decomposition

The rank of a matrix is defined as the dimension of vector space spanned by matrix's columns, or alternately as the number of linearly independent columns (or linearly independent rows, as the row rank and column rank of a matrix are the same [27]). An $m \times n$ matrix A with $m > n$ has a maximum rank of n. A matrix that has a rank equal to the maximum rank is called a full rank matrix, otherwise it is rank deficient. A matrix that is rank deficient also has reduced Degrees of Freedom (DoF).

Because a low-rank approximation to a dataset contains a dramatic reduction in DoF, it can be expressed in far fewer coefficients than the original dataset. Typical fMRI analysis will apply a matrix decomposition to form a low-rank estimation to the dataset. In a similar manner to how CS methods can use a reduced set of samples to directly approximate a sparse representation of the data, low-rank methods can also use a reduced set of samples to directly reconstruct the low-rank estimation that would have been produced by the analysis process anyway. In addition, matrix decomposition is often used to reduce the dimensionality of the set of coil sensitivity matrices [28, 29], a vital tool in speeding up the optimisation process.

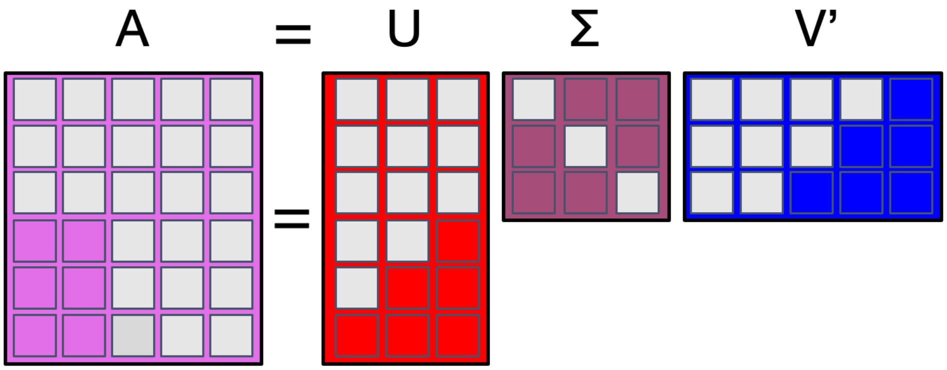
According to the fundamental matrix decomposition theorem [30], any $m \times n$ matrix of rank r can be decomposed as follows by Singular Value Decomposition (SVD).

$$A = U \Sigma V' \tag{3.23}$$

U and V are orthogonal matrices, i.e. $U'U = I$. Σ is a diagonal matrix with r non-zero diagonal elements (σ) which are conventionally ordered to be non-ascending (i.e. $\sigma_i \geq \sigma_{i+1}$). If A is rank deficient then the decomposed matrices (U, Σ, and V) can be shrunk so that Σ is an $r \times r$ matrix. SVD decomposition can be used to visualise the reduction of DoF in a rank-deficient matrix (figure 3.4), defined in equation 3.24.

$$DoF(A_{rank=r}) = r(m + n - r) \tag{3.24}$$

Some matrices with a high rank contain many singular values which are relatively low in magnitude. If these low-magnitude singular values are set to zero, a new low-rank approximation is created. The distribution of singular values is a vital tool in deciding the extent to which the rank can be reduced; the greater the proportion of the curve integral covered by the initial values, the more compressible the data is. Figure 3.5 shows the rank compression of a 1600×1600 natural image. The first 25 components of the image contain 41% of the total energy, compared to 61%



Dimensions	m n		m r		r^2		r n
Degrees of Freedom	$r(m+n-r)$	=	$mr-\frac{r(r+1)}{2}$	+	r	+	$rn-\frac{r(r+1)}{2}$

Figure 3.4: The reduction in DoF in a rank deficient matrix. A full-rank matrix A has $m \times n$ DoF. If matrix A is rank deficient, the reduction in DoF can be most clearly seen through an SVD. The columns of U form an orthonormal set, so one coefficient per column is fixed to make $\|u_{i,1:m}\|_2 = 1$, and an additional coefficient in each column is fixed to ensure an inner product of zero with previous columns $\langle u_{i,1:m}, u_{i-j,1:m} \rangle = 0$. This equates to an $\frac{r(r+1)}{2}$ non-independent DoF in U (and V by the same principles). Matrix Σ only has r independent diagonal elements. DoF$_A$ can be found by summing the DoF across U, Σ, and V. If $r = n$, then DoF$_A = m \times n$, as expected.

of the total energy in the first 100 components. The associated images show the impact of the different compression levels. The cross-like artefact in the overcompressed image belies the nature of rank compression: underlying relationships between the two dimensions are required in order to adequately express the matrix. Although the rank in the 100 component image is reasonably large, there is still an 88% reduction in the DoF compared to the full rank original image.

The following relations are extensions of equation 3.23, and show eigenvalue decomposition:

$$A'A = U\Sigma^2 U' \tag{3.25}$$

$$AA' = V\Sigma^2 V' \tag{3.26}$$

Matrices $A'A$ and AA' are square, which is a necessary condition for eigenvalue decomposition (but not for SVD). Eigenvalue decomposition can be carried out by Principal Component Analysis (PCA), which is a transform that maximises the remaining variance explained in successive eigenvectors. The eigenvectors of $A'A$ are defined by U, and the eigenvectors of AA' are defined by V. The eigenvalues of $A'A$ and AA' are both defined by the matrix Σ^2. The elements of Σ represent both the singular values of A and the square roots of the eigenvalues of $A'A$ and AA'. V can be used to reduce the dimensionality of A through post-multiplication by V.

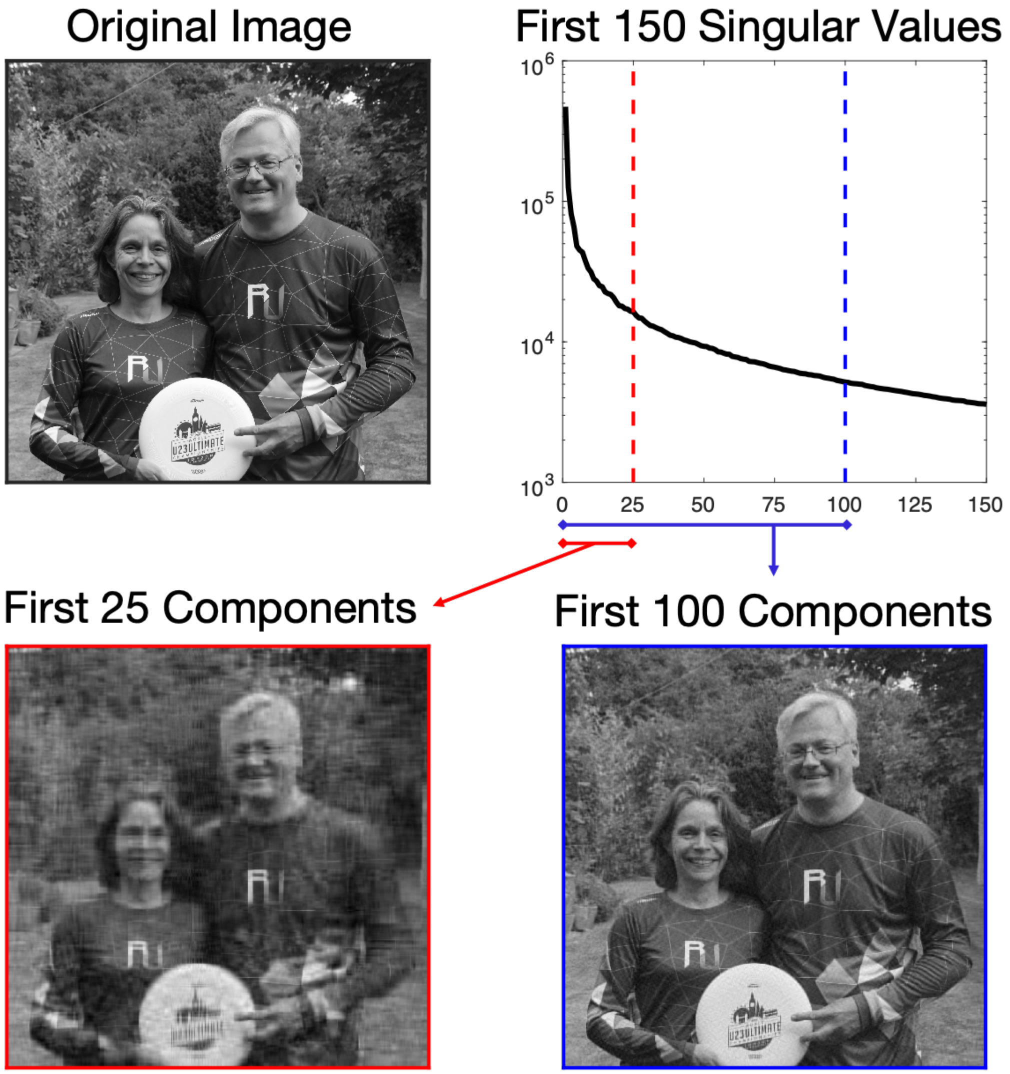


Figure 3.5: A demonstration of overcompression of singular values. The original full rank image and the corresponding SVD curve is shown in the top row. The first 25/100 components were preserved to form the left-hand/right-hand images in the bottom row.

$$A_{compressed} = AV = U\Sigma V'V = U\Sigma \tag{3.27}$$

PCA and SVD both produce a data-driven basis set, which is more adaptable to the data than a specific transform basis set such as the FFT or a wavelet transform. The decomposition approaches define a new orthogonal coordinate system that optimally describes the variance in a single dataset. This means that if there is a low-rank approximation to the data, the PCA/SVD will find it. This stands in contrast to the CS approaches, where prior knowledge regarding the sparse basis set is required.

3.4.2 Subspace Analysis

A subspace is defined by the span of the set of vectors, and each subspace is a subset of a larger vector space. For example, the component vectors in decomposed matrices U and V define the column subspace and row subspace of the full matrix A in equation 3.23. The subspace dimensionality is the rank of A. If A is invertible (e.g. A is full rank and square), row space and column space are equivalent.

3.4.2.1 Canonical Correlation

Image reconstruction quality can be inherently difficult to determine, with conventional methods such as root mean square error (RMSE) or structural similarity index giving great weight to the first component and not considering underlying patterns [31]. As such, when judging the reconstruction quality of subspaces, a separate image quality metric is required.

Canonical correlation (CC) [32] is a multivariate analysis of correlation that is a good way of comparing the similarity of two subspaces, and so will be used extensively in this thesis (particularly when normalised, which will be referred to as the CC score). CC calculates the cosine of the angles between two subspaces [33] (figure 3.6), which is a good measure of subspace similarity [34]. These angles are referred to as principal (or canonical) angles, defined by the range $0° \leq \theta \leq 180°$. The angle between two unit vectors x and y is defined as follows:

$$\cos(\theta) = |x'y| \tag{3.28}$$

The set of principal angles between two subspaces χ and Υ are defined recursively between orthogonal principal components x and y which span the subspace [35]. This process orders the principal components from most aligned to least aligned, or equivalently from smallest angle to biggest. The recursive condition assumes that each subsequent principal component is orthogonal to the components already acquired.

$$\cos(\theta_i) = \max_{x \in \chi} \max_{y \in \Upsilon} |x'y| \tag{3.29}$$

A CC calculation between matrices A and B searches for two vectors (c, d), such that $c'A$ and $d'B$ are maximally correlated. This finds the principal components in the subspace defined by A and B.

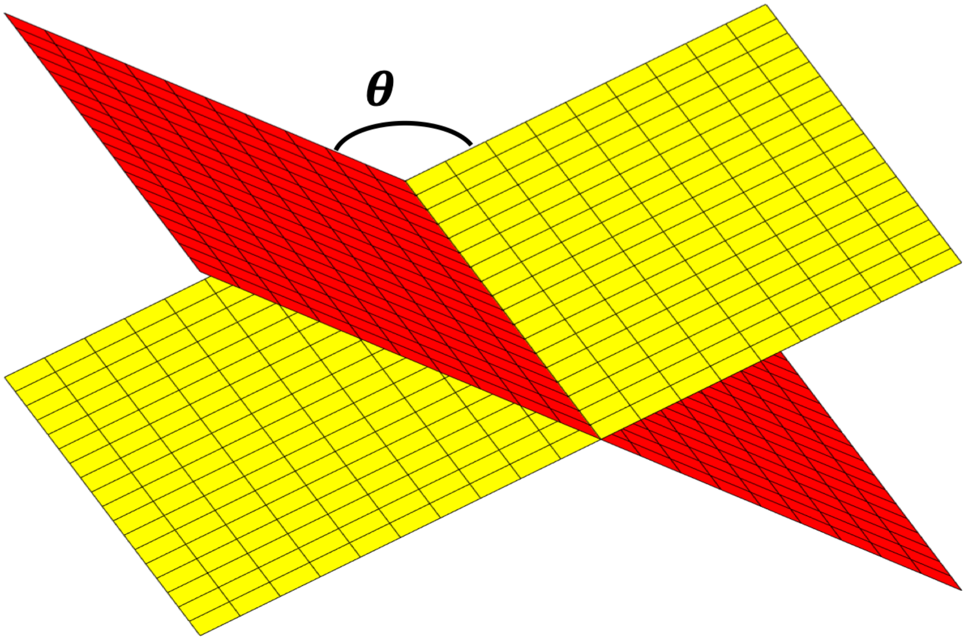


Figure 3.6: A visual demonstration of a principal angle between two simple subspaces.

$$
\begin{aligned}
\hat{c}, \hat{d} &= argmax_{c,d}\big(corr(c'A, d'B)\big) \\
&= argmax_{c,d}\left(\frac{c'\Sigma_{AB}d}{\sqrt{c'\Sigma_{AA}c}\sqrt{d'\Sigma_{BB}d}}\right)
\end{aligned}
\tag{3.30}
$$

$\Sigma_{A,A}$ is the variance matrix of A, and $\Sigma_{A,B}$ is the covariance matrix between A and B. This CC calculation is not affected by the scaling vectors c or d, and so no priority is given to larger components, yielding an absolute score of $0 \rightarrow 1$ for each component. This process is repeated to find a new set of canonical vectors which are completely uncorrelated with the previously found sets, also creating a set of r correlation coefficients sorted by maximal alignment (where $r =$ min(rank(A),rank(B))). In this way, CC analysis will find the optimal orientation between the two datasets (analogous to SVD/PCA finding the optimal orientation of a single dataset).

By summing the angles and dividing by r, a normalised Canonical Correlation Score (CCS) can be defined. CCS provides a single value measure between the alignment of two subspaces, and so provides a good way to either compare reconstruction of an underlying subspace with a known truth or to comparing two separate reconstructions.

3.4.2.2 Null Space

In addition to the row and column subspaces, two other fundamental subspaces can be defined: the null space and the left null space. The null space of matrix A is termed $N(A)$, and can be defined as the set of unit vectors x which satisfy the following equation:

$$N(A) : Ax = 0 \tag{3.31}$$

The row space, column space, null space, and left null space are the four fundamental subspaces of an $m \times n$ matrix, and are linked by orthogonal complements. The orthogonal complement of A is $A^{\perp}$ and is defined as the set of vectors that are orthogonal to every vector in A. An additional property of the orthogonal complement is that $(A^{\perp})^{\perp} = A$. The left null space is the orthogonal complement to row space, and the null space is the orthogonal complement to column space. The orthogonality properties of null space will later be exploited as an alternative method of regularisation.

3.4.3 Different Approaches to Low-Rank Reconstruction

One of the main motivations behind choosing a low-rank solution is the dramatic reduction in the number of unknowns, as shown in equation 3.24. In the same manner that sparsity may be enforced through an L1 constraint which minimises the number of coefficients in a solution, a rank constraint may also be used to minimise the number of singular values (and so the number of DoF). While both methods reduce the dimensionality of the problem, CS methods use a pre-specified basis, and low-rank methods find a basis that is tailored to the data. In the presence of noise or incomplete datasets, direct decomposition methods like PCA can struggle. A low-rank estimation with only a few corrupted signal points formed using PCA will be very different from a low-rank estimation without those points [36]. Iterative methods can be used to ensure the robust reconstruction of these decomposed matrices, whilst still enforcing a low-rank structure [37].

The rank constraint can either be a hard constraint which fixes the rank of the system or a relaxed constraint which allows flexibility of rank in reconstruction. They are known as rank-specification problems or rank-minimisation problems respectively, and there exist numerous approaches for both problems [38].

The Nuclear norm ($\|A\|_*$) is the sum of all n of A's singular values $\sigma(A)$ (equation 3.32). In a similar manner to how the L1-norm - the sum of all coefficients - can be used in CS can be used to find sparse solutions, the Nuclear norm can find a sparse encoding of singular values if the RIP is satisfied [26, 39].

$$\|A\|_* = \sum_{i=1}^{n} \sigma_i(A) \tag{3.32}$$

Relaxed rank approaches typically seek to minimise the nuclear norm [40]. Fixed-rank methods will try to encode the rank in the dimensionality of the matrices, or use an adapted version of thresholding in combination with the nuclear norm.

3.4.3.1 Matrix Completion

Matrix completion takes an undersampled depiction of a low-rank system, and reconstructs a low-rank representation of the undersampled dataset using only the limited data available. Explicitly, matrix completion fills in the missing gaps in the sampled matrix.

Matrix completion has applications in many fields, with the most famous being the Netflix Problem [41, 42] in the domain of collaborative filtering. The problem was this: given a sparsely sampled movie-ratings matrix of users and movies, can an algorithm be found to estimate what rating a user would give a movie they haven't seen yet? The matrix of users and movies is heavily sparse (due to a large number of users, and the existence of far more movies than any person can/should watch). The matrix is also low rank, as users are likely to consistently enjoy movies that are defined by certain underlying categories, and movies are likely enjoyed by a similar set of people. Filling out this sparse matrix using rank information to find new recommendations is a matrix completion problem, and the principle of predicted movie ratings can be extended to reconstruct any matrix defined by relatively few underlying components.

Conditions

In order to apply matrix completion algorithms, the following conditions are often defined [43]:

1. Observed entries are sampled uniformly at random
2. Sampling must exceed a lower bound on the number of observed entries
3. Decomposed matrices must not be overly sparse

Condition 1: In order to predict the missing information in a row/column, some information is needed regarding all underlying patterns. As such, the samples are usually spread between the columns and the rows. This is a commonly-stated sufficient (but not necessary) condition and does not cover the application of matrix completion to MRI (which typically has non-uniform sampling), despite the successful application of matrix completion within MRI [44]. Instead, an isometric condition [45] can be used which combines properties of rank and incoherence.

Condition 2: A low-rank matrix of size $m \times n$ will have $r(m+n-r)$ DoF (equation 3.24). With a sampling level below this DoF limit, it is highly improbable that a single complete matrix can be inferred from the partial set of samples. This condition contains the assumption that the underlying matrix is not full rank, otherwise the lower bound for sampling would be every element in the matrix.

Condition 3 can alternatively be stated that the singular vectors must be incoherent with respect to the standard basis set (the unit vectors pointing in the direction of Cartesian axes), and it is here that low-rank methods really differ from CS. Most sparse matrices are low rank, but are unsuitable for matrix completion as a randomly-sampled sparse matrix is likely to produce a matrix of zeros from which it is hard to infer relations. As such, the coherence μ of the projection of decomposed matrix ($\mathbf{P}_U$) onto the standard basis set $\mathbf{e}_i$ must be less than some positive value μ_0.

$$\mu(U) = \frac{n}{r} \max_{1 \leq i \leq n} (\|\mathbf{P}_U \mathbf{e}_i\|_2^2) \tag{3.33}$$

Just as CS was searching for incoherence with respect to coefficients (because it was searching for sparsity), low-rank methods are seeking to maximise incoherence with the basis functions.

Solving Matrix Completion Problems

Matrix completion can be calculated with a relaxed rank constraint or a fixed rank constraint. Section 4.2 walks through the mathematics behind two common approaches.

Relaxed rank constraints usually involve or approximate the Nuclear norm (often requiring an SVD to calculate/truncate the singular values). A soft shrinkage operator on the singular values [46] is a way to slowly reduce matrix rank (analogous to the CS shrinking operator [22] that applies shrinkage on the number of non-zero coefficients). An IRLS minimisation method [47] can be used to approximate the Nuclear norm with a reweighted version of the Frobenius norm ($W = \sqrt{AA' + \gamma I}^{-1}$), analogous to the approximation of L1 norms using L2 norms in CS [24].

Some fixed-rank methods apply hard singular value thresholds while also using a nuclear norm [48, 49], although other methods can also apply a fixed norm while using a Frobenius norm minimisation approach. Alternating minimisation is a category of fixed rank methods that compute a partial SVD of an undersampled matrix $A = UV'$, and solve for U and V:

$$argmin_{U,V}\left(\frac{1}{2}\|A - UV'\|_F^2\right) \tag{3.34}$$

PowerFactorization [50] is one such alternating method which alternately updates U and V with conjugate gradients. This introduced nonconvexity into the problem, but drastically improves computation efficiency. Alternating Steepest Descent [51] also uses the same cost function as PowerFactorization but updates with steepest gradient descent, and the Low-rank Matrix Fitting algorithm [52] uses the undersampled matrix A as an initialisation but updates the unsampled locations with the matching locations in the current estimate of UV'. The AltMinComplete [53] splits the observed samples into subsets and solves the problem on a subset level. These alternating algorithms work with smaller matrices than other approaches and often avoid the need for an SVD computation - making them relatively efficient from a computation perspective. As such they are very popular, despite not being as robust as convex methods.

3.4.3.2 Limitations of Matrix Completion

Matrix completion requires the reconstruction to take place in the domain of the sampled matrix. Many of the methods mentioned above can be adapted for use with respect to an undersampled dataset d with the incorporation of an additional encoding matrix E applied to A, containing some sort of rank constraint (represented by $\| \ \|_*$, even if the nuclear norm is not strictly calculated).

$$\hat{A} = argmin_A \Big(\|d - EA\|_F^2 + \lambda \|A\|_* \Big) \tag{3.35}$$

This encoding matrix can incorporate non-Cartesian sampling trajectories and sensitivity map encoding for fMRI. As such, it may not be directly evident that these incoherence and sampling conditions required for matrix completion are fulfilled (e.g. in point 1, often the k-space sampling is far from random). As such, the robustness of reconstruction will have to be monitored closely when the limits of acceleration are pushed.

Many of the outlined matrix completion conditions are tough to guarantee in a practical setting, and so low-rank methods must be validated through constant testing. This is especially true where the matrix is only approximately low rank. The number of samples needed to define a low-rank matrix contains some uncertainty depending on the assumptions made [54, 55]. The DoF defined by equation 3.24 is a minimum sampling level, but non-Cartesian sampling may further render that limit imprecise. It does, however, indicate a likely fundamental limit regarding the minimum number of samples - with a sampling level below the DoF limit it is highly improbable that an accurate complete matrix can be reconstructed. Finding the intrinsic lowest rank [56, 57] of a matrix is also difficult, and so an estimation will have to be made.

3.4.4 Limitations of Low Rank approaches to fMRI

The low-rank assumption asserts that a low-rank representation of the fMRI dataset will capture the underlying functional information within a reconstruction. The enforced rank chosen during reconstruction must be large enough to contain some leeway, in order to allow for confounding factors arising from increase acceleration (e.g. motion, physiological noise, or increased serial correlation). However, the rank is fundamentally limited for a given acceleration factor in a dataset of a certain size. Equation 3.24 serves as a theoretical limit, although other factors (such as parallel imaging) render this limit imprecise. Petrov and Otazo [58, 59] both showed that the physiological noise is already represented in a low-rank reconstruction, and so can be dealt with in post-processing.

Increased serial correlation could result from sub-second TRs, and an increased fractional contribution from white noise can result from the reduced SNR per time frame (which manifests as a sharp peak in the autocorrelation at 0s lag). On a voxelwise level, this additional information can be modelled either as a higher order autoregression model or as a single order autoregression + white noise model [60], although both of these models assume that fluctuations from physiological noise have been accounted for [61]. However, it is more difficult to judge how autocorrelation jointly affects the spatial and temporal structures. Increased autocorrelation is known to reduce the effective degrees of freedom [62], a fact which would make the system more approximately low rank (see equation 3.24). Nevertheless, autocorrelation structures are known to be spatially heterogeneous in rsfMRI [62], and so a more full investigation into the relationship between rank and autocorrelation may be required for more definitive analysis.

Increased motion may introduce confounding factors, and extreme motion will violate the low-rank assumption. However, such strong motion will likely render the functional information irretrievable regardless of the method. Low/moderate motion (that is not removed prior to reconstruction) will introduce confounds, but these will be common confounds that can be represented by a couple of components. As the motion increases, the complexity required to represent the motion artefacts will also increase. This could drown out the functional information and also reorder components, potentially violating the low rank assumption.

To summarise: while various factors will introduce confounds, a low-rank reconstruction can still capture the functional information provided the enforced rank is high enough (and doesn't surpass the practical application of some theoretical upper limit). If the confounds are captured, then they will be removed post-reconstruction and prior to analysis, and have the same impact they would on any reconstruction. However, large confounds (particularly heavy motion) can drown out the functional information and must be monitored.

3.5 Dynamic Reconstruction Methods

In section 2.3.2.4, the importance of acceleration in fMRI was demonstrated. This section outlines the approaches used for acceleration of dynamic MRI, and specifically acceleration of fMRI where methods exist. In-plane acceleration of static MRI images (such as PI) do not incorporate any temporal information but can be used in tandem with methods that do.

3.5.1 Dynamic k-t Methods

Methods that jointly consider k-space and time are called k-t methods. By focusing on redundancies and/or structural information in k-t space, k-t methods have the potential for much greater degrees of acceleration than those methods limited to the gains achievable with coil geometry alone and have sprung up in many forms of dynamic MRI.

Keyhole methods were the first k-t acceleration approach [63, 64], and work by frequently resampling a central ‘keyhole’ band in k-space which provide the majority of contrast in an image, and carry an explicit assumption that outer k-space can be shared between multiple images over time. However, keyhole methods have a limited acceleratory capacity, and also produce a reconstructed dataset where the time frames are not fully independent, which violates frame-independence assumptions needed for statistical analysis. As an intuitive example, interpolating a whole extra image between each frame of a dataset would also increase the temporal resolution, but provide no new information (and indeed may be providing false information compared to the truth).

Keyhole methods are not the only approaches that rely on the temporal redundancy in k-t space acceleration. DIME [65] aims to model the time variation and only fit a set of parameters to that model rather than reconstruct a whole time series, assuming that the dynamic elements can be effectively modelled. k-t GRAPPA [66, 67] includes a time dimension to the GRAPPA kernel (section 2.2.2.2), but in doing so breaks the assumption of time-frame independence.

UNFOLD [68, 69, 70] obtains Fourier data in a grid which is sheared over time. This linear shift in k-t space produces additional phase shifts of all points in the sampling PSF except the central peak in x-f space. As can be seen for a linear undersampling example in from figure 2.3, the non-central points represent aliased copies, and so the aliased signal within a given voxel will then be modulated onto a higher frequency than the non-aliased signal. The aliased signal can then be removed with an anti-aliasing filter applied in x-f space after reconstruction. Conceptually, UNFOLD can be seen as similar to 2D CAIPI, but the sampling grid is sheared temporally rather than over k_z, and CAIPI seeks to maximise the distance between aliased copies whereas UNFOLD seeks to fully separate them. The limitation of UNFOLD depends on the dynamic FOV, as a larger

FOV would require a greater temporal shear in order to allow aliasing filters to separate out the copies [71].

k-t BLAST [72] uses the same lattice sampling pattern of UNFOLD, but does not require the sheer to fully separate the aliased copies. Instead, k-t BLAST also acquires an initial set of low spatial resolution training data from central k-space immediately before the main acquisition. The expected magnetisation covariance matrix from the low-resolution estimate is combined with the expected aliasing contribution in every voxel in order to filter the aliased information from the accelerated data. k-t SENSE [72] is k-t BLAST with coil sensitivity information also incorporated, and coil sensitivity profiles were also included in all subsequent extensions of k-t BLAST (e.g. k-t FOCUSS [73], which incorporated sparsity and is discussed in section 3.5.2; and k-t PCA [74] which additionally applied a decomposition while acquiring the low-resolution prior and final acquisition simultaneously, see section 3.5.3).

k-t PCA and k-t FOCUSS are not unique in incorporating underlying mathematical structures such as sparsity and low-rank information into the reconstruction. The ability to incorporate sparse and low-rank information into a dynamic reconstruction method led to great advancements in the field, as well as allowing more fMRI-specific approaches to reconstruction.

3.5.2 Compressed Sensing in BOLD fMRI

MRI datasets are known to be sparse in many domains [11, 75, 76]. However, different forms of MR imaging exhibit different forms of sparsity. Angiograms are sparse in the image domain [77], as the vessels take up a small fraction of the total brain volume. Cardiac MRI is sparse in temporal frequency (and can use wavelet transforms in image space with k-t SPARSE) [78], as the heartbeat strongly dictates as the main dynamic pattern. Less intuitive CS approaches also exist, e.g. breast images are typically sparse in a wavelet domain [15].

In fMRI, the BOLD signal comprises a small portion of the overall signal energy, meaning CS approaches must be handled with caution lest the compression remove the actual signal of interest. The following section outlines different approaches used for CS fMRI of the brain. Throughout this section, m is the 4D estimate of the dynamic image-space dataset, E is the encoding operator (which incorporates the Fourier transform and spatial coils), and y is the undersampled k-t space dataset. Ψ_{W_x} and Ψ_{W_t} are spatial and temporal wavelet transforms respectively. Similarly, TV_x and TV_t refers to total variation applied in space and time respectively.

The achievable acceleration factor (R) recorded by each paper will be noted in each case. It is worth stating that the criteria by which an achievable R is judged are very subjective. As such, these should be taken as indications of the approximate acceleration at which reasonable data fidelity is achieved, rather than as a hard and fast guide by which to rank the methods.

The method by Holland et al. [16] represents one of the most straightforward approaches to CS fMRI. They focused on promoting sparsity through Daubechies wavelet transforms on a spiral trajectory, with the sparsity being promoted in each image independently. Holland et al. found reductions in noise and improvements in mean z-score at R=4 (equation 3.36).

$$\hat{m}_{Holland} = argmin_m \Bigg(\|y - E\Psi_{Wx} m\|_2^2 + \lambda \|\Psi_W m\|_1 \Bigg) \tag{3.36}$$

This approach of Daubechies wavelet sparsity and spiral trajectories was continued by Jeromin et al. [17] albeit with an additional spatial TV term. When outlining reasonable values of the parameter weightings λ_1 and λ_2 (equation 3.37), they found different weightings for optimal tfMRI and rsfMRI. Additionally, they appeared to optimise in the domain of wavelet coefficients (c, where $\Psi_W c = m$), achieving high peak SNR (57.6) in spiral low-pass geometry at R=5. The inclusion of an additional TV term appears to allow for increased acceleration.

$$\hat{c}_{Jeromin} = argmin_c \Bigg(\|y - E\Psi_W c\|_2^2 + \lambda_1 \|TV_x(\Psi_W c)\|_1 + \lambda_2 \|c\|_1 \Bigg) \tag{3.37}$$

Other fMRI methods focussing on spatial sparsity also incorporate additional structural information. PICCS [79] demonstrates this by using the average spatial signal over time (the temporal mean) as a spatial prior m_{prior}. The PICCS cost function (see equation 3.38) used two separate spatial TV terms: the TV of the estimate x, and a TV term of difference between the estimate and the temporal mean (in addition to a data consistency term and an L1 stability term). A split Bregman approach was used, and acceleration factors of up to R=8 were suggested due to PICCS's success in low SNR scenarios.

$$\hat{m}_{PICCS} = argmin_x \Bigg(\|y - Em\|_2^2 + \gamma \|m\|_1 + \alpha \|\,TV_x(m - m_{prior})\|_1 + (1-\alpha)\|\,TV_x(m)\|_1 \Bigg) \tag{3.38}$$

Similar to PICCS, k-t FOCUSS [73] also attempts to incorporate additional structural information into the reconstruction. k-t FOCUSS is an expansion of the k-t BLAST/SENSE method covered in section 2.3.2.4. k-t FOCUSS begins with a low-resolution estimation of a sparse signal, which is then sparsified with pruning. Pruning ensures any coefficients which are turned to zero then fixed as zero, meaning the "solution" will only become sparser. The cost function is described in equation 3.39 and solves for the residual signal Δm, defined as the difference between the unknown signal m and the pruned low-resolution estimate signal m_0. F_t is used to represent the temporal Fourier transform, and W is a diagonal weighting matrix that is updated each cycle. k-t FOCUSS was demonstrated at acceleration factors of 6.5 when fully sampled reference frames bookended the acquisition [80].

$$\Delta\hat{m}_{FOCUSS} = argmin_{\Delta m}\Bigg(\|y - Em - EF_tW\Delta m\|_2^2 + \lambda\|\Delta m\|_2^2\Bigg), \quad m = m_0 + F_tW\Delta m \tag{3.39}$$

All of the methods so far have enforced transformed spatial sparsity of some form. However, temporal sparsity has also been explored. Double Temporal Sparsity-based Reconstruction (DTSR) [81] minimises temporal TV sparsity through the assumption that brain volumes are similar between frames, in addition to sparsity in the voxel temporal frequency domain. DTSR demonstrated radial retrospective undersampling at R=13, and an improvement in peak SNR at R=3.

$$\hat{m}_{DTR} = argmin_m\Bigg(\|y - Em\|_2^2 + \lambda_1\|F_tm\|_1 + \lambda_2\|TV_t(m)\|_1\Bigg) \tag{3.40}$$

Dual TRACER [82] is an adapted version of the TRACER [83] method that assumed minimal spatial change in consecutive frames with an L2 norm (temporal smoothing). While Dual TRACER is not technically a CS method, it is conceptually very similar. TRACER performed the execution of minimal spatial change by iterating forward and minimising the error between frames, meaning errors were liable to propagate. Dual TRACER executed the reconstruction twice: once forwards (equation 3.41), once backwards (equation 3.42), and subsequently averaged the two reconstructions. i is the iteration factor and n represents the frame. Dual TRACER was demonstrated at an acceleration factor of 20.

$$\hat{m}_{forward} = argmin_m\Bigg(\|y_n - Em_{n,i}\|_2^2 + \lambda\|m_{n,i} - m_{n-1,i}\|_2^2\Bigg) \tag{3.41}$$

$$\hat{m}_{backward} = argmin_m\Bigg(\|y_n - Em_{n,i}\|_2^2 + \lambda\|m_{n,i} - m_{n+1,i}\|_2^2\Bigg) \tag{3.42}$$

Dual TRACER also demonstrated that a Golden Ratio angle approach is particularly suited for temporal smoothing approaches, due to the non-uniform distribution of aliasing artefacts in separate frames (see figure 3.7). Golden Ratio angles never overlap, resulting in a unique trajectory or every frame in the dataset. The sampling for each frame has an associated unique PSF, and making the undersampling artefacts induced by those PSFs unique as well. These rapidly changing artefacts are exactly the type of signal which a temporal smoothing constraint will target. An fMRI signal at high TR should be comparatively smooth, and so will be more unaffected by a temporal smoothing constraint as long as the weighting is not too high.

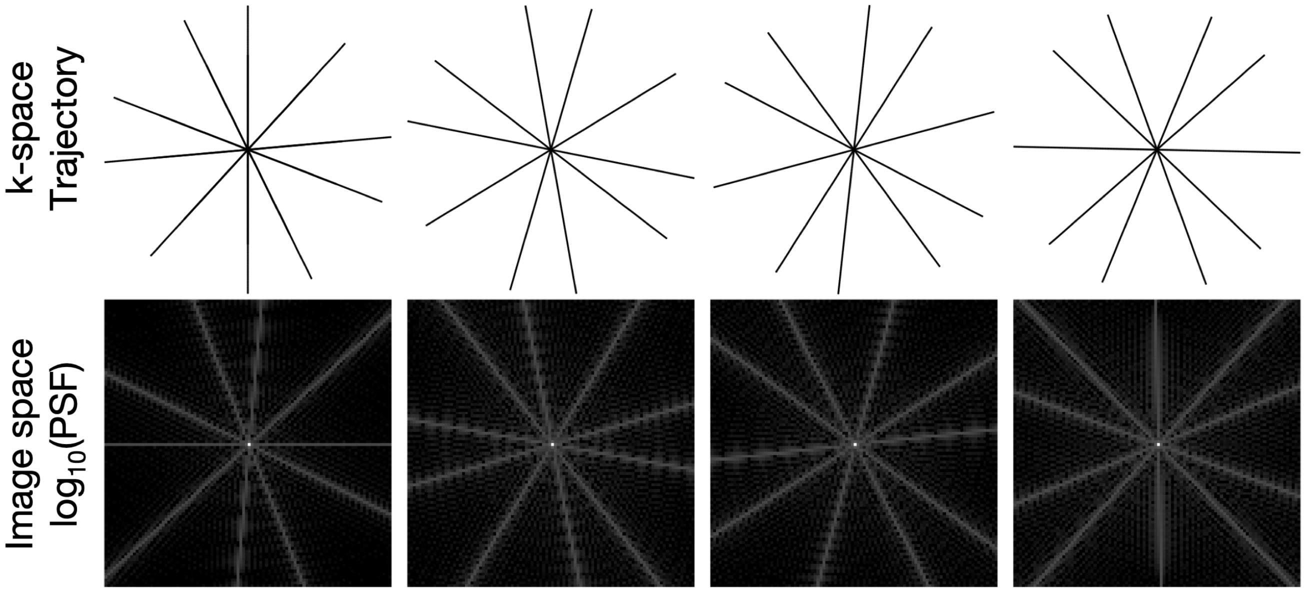


Figure 3.7: A demonstration of aliasing patterns in consecutive Golden Ratio angle frames. Each trajectory is formed of five Golden Ratio Angle spokes, and their corresponding PSF functions in image-space is shown on the bottom row (shown on a log scale for visual clarity).

Finally, a couple of methods incorporate both spatial and temporal sparsity constraints. The first such method is HSPARSE [84], which uses a modified version of k-t SPARSE. HSPARSE enforces in-frame spatial sparsity using either DCT or discrete wavelet transforms. The temporal sparsity is enforced either through discrete wavelet transforms/the Karhunen-Loeve transform for temporally complicated experiment designs such as event-related fMRI, or the DCT transform for block designs. HSPARSE compresses the temporal domain as a separate term (equation 3.43), and achieves reduction factors of roughly R=6.

$$m_{HS\hat{P}ARSE} = argmin_m\left(\|y - Em\|_2^2 + \lambda_1 \|\Psi_{Wt} m\|_1 + \lambda_2 \|\Psi_{Wx} m\|_1 \right) \tag{3.43}$$

As an alternate combined spatial/temporal sparsity approach, the Spatiotemporal Total Variation (TTV) approach applied to fMRI [79] used separate TV terms in space and time (equation 3.44). TTV was evaluated in the same paper as PICCS, with TTV being found to produce image degradation at acceleration factors that PICCS succeeded (R=8), but worked well at R=5.

$$m_{\hat{TTV}} = argmin_m\left(\|y - Em\|_2^2 + \alpha \| TV_x m\|_1 + (1-\alpha) \| TV_t m\|_1 \right) \tag{3.44}$$

There are many potential combinations of sparsity constraints and approaches that could be explored. It is worth noting that the reported acceleration factors in papers must obviously be read with nuance due to the different conditions under which they were obtained. With that in mind, the methods which focussed on a temporal form of sparsity - dual TRACER and DTSR - both contained the highest reported acceleration factors. Perhaps notable is that both temporal methods encouraged smoothness in the temporal domain by minimising the difference between frames in some manner. It should also be noted that those spatial sparsity methods which incorporated additional information (PICCS/k-t FOCUSS) appeared to achieve better acceleration than those methods which relied on more traditional constraints (Jeromin/Holland).

3.5.3 Low-Rank BOLD fMRI

Partially Separable Functions [85] is a low-rank approach in dynamic MRI that collects two datasets to estimate the different components - one with a high temporal resolution (and poor spatial), and one with a high spatial resolution (and poor temporal). An SVD can be used to estimate the temporal basis set, which is then fixed and projected onto the other dataset to estimate the spatial basis. This method will be referred to as k-t PSF (so as not to confuse the reader with the point spread function). k-t PSF was used in cardiac MRI [86] and breast MRI [87] before being demonstrated at R=5 in tfMRI in 2013 [88] (in combination with a sparsity constraint). k-t PSF will be explored more in chapters 5 and 6.

k-t PCA [74] is an extension to the k-t BLAST [72] method covered in section 2.3.2.4. k-t BLAST, similar to k-t PSF, can use a high temporal-resolution/low spatial-resolution acquisition in reconstruction (although k-t BLAST used the low spatial resolution prior's signal spatial covariance matrix for the purpose of dealiasing). The reconstruction problem in k-t BLAST is always underdetermined. The inclusion of sensitivity matrices allows k-t SENSE to be either under- or overdetermined (depending on the number of coils and the acceleration factor), but k-t PCA applies a PCA to reduce the dimensionality of the problem and ensure the problem can be overdetermined. Additionally, k-t PCA uses the temporal frequency components from the high temporal-resolution/low spatial-resolution acquisition portion of x-f space as the basis in the final reconstruction. The spatial components are then used to form the signal covariance matrix for the purpose of dealiasing. While k-t PCA has not been demonstrated in fMRI to our knowledge, its use of low-rank methods to improve a dynamic MRI dataset was considered sufficiently relevant to be worth including.

k-t FASTER (fMRI Accelerated in Space-Time via Truncation of Effective Rank) was the first known method to purely exploit the low-rank nature of fMRI [44], which can be understood as the combination of a few spatially coherent temporal processes (i.e. the activation maps which indicate distinct time-series occurring at certain voxels). fMRI often applies a dimensionality reduction [89] which explicitly enforces a low-rank representation of the system prior to analysis methods such as ICA. It should be noted that various noise sources (e.g. thermal noise, physiological noise, etc.) make the system only approximately low rank, although some of the coherent physiological noise can be estimated as low-rank processes. The rank constraint of k-t FASTER alone was enough to achieve notable acceleration in rsfMRI (R=4.3) but was simply combined with coil sensitivity information and a TURBINE trajectory for further acceleration in rsfMRI (R=5) [90]. Additional acceleration was achieved with the same set up in tfMRI (R=12.5), with the high signal variance of tfMRI allowing the dataset to be described with a lower rank than the rsfMRI approach. k-t FASTER uses an iterative hard thresholding algorithm with matrix shrinkage, and will be explored in much greater detail in chapter 4.

Both k-t PSF and k-t PCA estimate the spatial and temporal subproblems separately, using separate data subsets to solve each subproblem. Additionally, k-t PSF and k-t PCA both incorporate additional information: k-t PSF can use a sparsity constraint, and k-t PCA uses a signal covariance matrix. A version of k-t FASTER for tfMRI which incorporates a GLM constraint to help inform the temporal subspace has been demonstrated at R=16 [91]. The motivation of including GLM information in the reconstruction process mirrors the motivation of including a low-rank constraint to mirror the PCA/ICA subspace decomposition: the more directly the sampled information can reflect the final dataset which is used for analysis, the more efficient the use of sampled data is for image reconstruction.

k-t PSF and k-t FASTER will be tested more thoroughly in future chapters, although k-t PSF will estimate the temporal basis with an iterative-based framework (rather than SVD framework). A fuller comparison between the efficacy of these two methods will be carried out at that point.

3.5.4 Low Rank plus Sparse Reconstruction of BOLD fMRI

The relative brevity of the previous section should highlight how young the field of low-rank fMRI is. Surprisingly more developed is the field of combined low-rank and sparse (L+S) fMRI optimisation [36]. The lack of a clear sparse transform for fMRI in section 3.5.2 has prevented CS approaches from taking off in fMRI as they have in other MRI reconstruction problems. L+S approaches typically use a (very) low-rank approach approximation of the more static components of a dataset, in order to allow a sparse approach to fit the perturbations around the low-rank estimation [92]. This is not to say that low-rank problems cannot also represent the dynamic portion of the dataset, as was shown in section 3.5.3.

Incoherence is needed between the acquisition and both the low-rank component and the sparse component of the L+S reconstruction to allow removal of aliasing artefacts [93]. The low-rank and sparse components should also be incoherent with each other to allow separation of the domains, i.e. the low-rank components should not be sparse and the sparse components should not be low-rank [92]. Radial [94] and spiral [58] acquisitions are good sampling trajectories to encourage this incoherence. Exploratory analysis by Karker et al. [95] recently demonstrated reduced mean g-factor using Possion-PROPELLER - a novel sequence which combined PROPELLER-style blades and Poisson disc sampling to Golden Ratio angle rotation.

Various L+S approaches have been explored, many direct extensions to low-rank methods for fMRI. The aforementioned k-t PSF approach was investigated with sparsity on the spatial components at R=5 [88]. fLORA (functional LOw Rank Approximation) [96] is an extension to k-t PSF that estimates the temporal components from an SVD of the data, then estimates the spatial components with additional wavelet sparsity and component-wise regularisation. fLORA used a variable-density spiral acquisition on tfMRI, and was demonstrated in a retrospectively undersamped dataset at R=4 [97]. Both of these k-t PSF-based approaches only applied the sparsity to the spatial domain, directly estimating the temporal basis set with no additional constraints.

Three approaches used a soft singular value iterative threshold with soft sparsity thresholds. Otazo et al. first demonstrated the formulation in dynamic MRI [92], applying an iterative hard threshold to the sparse transform and an iterative hard threshold to the singular values. A data consistency term was applied to remove the residual, and then the process was iterated. In exploratory rsfMRI analysis, Otazo et al. then demonstrated that this L+S approach could be used for separating subsampled physiological noise in a manner comparable with retrospective methods (using an identity transform for sparsity and an EPI acquisition) [59]. Singh et al. [98] used a spiral acquisition and applied the soft iterative thresholds in the temporal Fourier domain to a block-design of retrospectively undersampled tfMRI. They showed results at R=3 with higher CNR than the pure CS or low-rank version of their approach [98]. Petrov et al. [58] expanded on this paper using a stack-of-spirals approach, and prospectively undersampled in the k_z-t domain. Petrov expanded on the block-design approach of Singh and also demonstrated their L+S approach in slow event-related designs, also using temporal Fourier sparsity but noting that it may not hold in fast event-related designs. They also increased the acceleration factor relative to Singh, displaying results at R=4.

The OptShrink method [99] (Optimal singular value Shrinkage denoising algorithm) was an approach introduced to improve soft singular value iterative threshold techniques. OptShrink uses a truncated SVD to provide a denoised estimate of the low-rank portion of the algorithm. Sparsity is still applied in the temporal Fourier domain. The authors demonstrated the reconstructions at R=6, and compared directly against soft singular value iterative thresholding, as well as CS with wavelet sparsity. tfMRI and rsfMRI datasets were retrospectively undersampled using uniform radial trajectories, and showed that OptShrink reduced the normalised mean square error compared to soft singular value iterative threshold techniques in acceleration factors up to R=6, while also improving z-scores.

PEAR [100] (an extension to k-t FASTER) used a fixed-rank approach with temporal Fourier sparsity. They observed that the low-rank component of other methods was very low rank, and that the important information tended to be focussed in the sparse domain. By forcing the rank to be slightly higher, they aimed to allow broader-band signals to be captured in the reconstruction. In retrospectively undersampled rsfMRI with radial trajectories at R=10, PEAR was compared with both soft singular value iterative threshold techniques and k-t FASTER reconstructions, showing slightly improved classification (judged through an ROC curve) of the ground-truth functional maps.

3.6 Conclusion

Regularisation in iterative reconstruction can guide the solution of an optimisation problem to take a specific form. By using prior knowledge about the mathematical structure of the dataset (such as sparsity or rank), the sampling limits set out by Nyquist-Shannon may be broken while still leading to high fidelity reconstructions.

As section 3.5 laid out, there are many attempts to exploit temporal structure in order to push the limits of acceleration. However, dynamic methods without sparsity/low-rank constraints have only demonstrated limited acceleratory capacity. CS methods, which have proven very effective in other fields, have been unable to find a naturally sparse domain in fMRI through which to apply compression.

In fMRI, the field of pure low-rank optimisation is still relatively ripe for innovation, and so appears to be one of the most promising fields to examine further. Various constrained methods (PICCS [79], k-t PSF [85], k-t PCA [74], k-t FOCUSS [101], low rank with GLM-based constraints [91]) have successfully incorporated addition structural information beyond the sparsity and rank constraints. In addition, the Dual TRACER approach (which encouraged temporal smoothness) was demonstrated at the highest acceleration factor of the CS approaches [82]. If the low-rank problem can be framed in such a way as to easily allow additional structural or smoothness constraints (e.g. on the subspaces), then there is potential for a novel approach to accelerated reconstruction.

3.A Norms

The mathematical family of norms are commonly used as the building blocks of a cost function. A norm is a strictly non-negative measure defining the size of a vector or matrix. Norms are typically used either to define the strength of a signal, or the size of an error. A norm is any function $p(\ )$ acting on any vector (a or b) which fulfils the following set of conditions:

1. $p(a+b) \leq p(a) + p(b)$ (Triangle Inequality)
2. $p(\lambda a) = |\lambda| p(a)$ (Absolutely Homogeneous, λ is any constant)
3. if $p(a) = 0$, a is the zero-vector

While the above is a strict set of definitions for a norm, many functions that fulfill only part of the conditions are still called norms for convenience (e.g. TV norms or L0-norms). Norms can take many forms, the most common of which is known either as the vector p-norm or the Lp norm:

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} \tag{3.45}$$

The choice of p in equation 3.45 defines different measures of the same underlying vector x. Any value of p $\geq$ 1 fulfils the conditions of a norm. Choosing 0 < p < 1 yields a quasinorm which only fulfils conditions 2 and 3 (an adapted version of condition 1 is needed: $p(a+b) \leq K(p(a) + p(b))$, with $0 < K < 1$). The larger p is, the more the norm of x will be characterized by x's higher magnitude elements.

The most commonly used Lp norm is the L2 norm (also known as the Euclidean norm), which is defined as the square root of the sum of the square of all elements. The L2 norm is also equivalent to the square root of an inner product, and can be used to calculate the energy of a signal. Large values dominate the characteristic L2 value of x.

$$\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2} \tag{3.46}$$

"Entrywise" matrix norms treat the $m \times n$ matrix A as a $(m \times n) \times 1$ vector (represented through $vec(A)$), and then apply the previously defined vector norm calculations. The Frobenius norm is the entrywise matrix equivalent of the L2 norm, and is the most common entrywise norm.

$$\|A\|_F = \|\text{vec}(A)\|_F = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{i,j}|^2\right)^{1/2} \tag{3.47}$$

Norms are very valuable functions in optimisation for producing a single value representation of the "size" of a vector or matrix. Different norms will yield different representations of the size of that vector/matrix. Optimisation will aim to shrink the size of a cost function, which is often defined by norms. The use of different norms will cause the underlying vector/matrix to shrink in different ways during the optimisation process.

Bibliography

[1] Rene Magritte and Harry Torczyner. *Magritte, ideas and images.* H. N. Abrams, 1977.

[2] Ekstrom Arne. How and when the fMRI BOLD signal relates to underlying neural activity: The danger in dissociation. *Brain Res Rev*, 62(2):233–244, 2013.

[3] Eliakim Hastings Moore. On the reciprocal of the general algebraic matrix. *Bulletin of the American Mathematical Society*, 26(9):394–395, 1920.

[4] Roger Penrose. A generalized inverse for matrices. *Mathematical Proceedings of the Cambridge Philosophical Society*, 51(3):406–413, 1955.

[5] Dan Kalman. Leveling with Lagrange: An Alternate View of Constrained Optimization. *Mathematics Magazine*, 82(3):186–196, 2009.

[6] Nasir Ahmed. How I came up with the discrete cosine transform. *Digital Signal Processing*, 1(1):4–5, 1991.

[7] David A. Huffman. A Method for the Construction of Minimum-Redundancy Codes. *Proceedings of the I.R.E.*, pages 1098–1101, 1952.

[8] David S. Taubman. *JPEG2000: Image Compression Fundamentals, Standards and Practice*, volume 11. 2002.

[9] Matt Mahoney. *Data Compression Explained.* 2010.

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

[11] Michael Lustig, David L. Donoho, and John M. Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007.

[12] Albert Cohen, Wolfgang Dahmen, Ronald Devore, Emmanuel J. Candès, and Terence Tao. Compressed Sensing and Best k-Term Approximation. *Journal of the American Mathematical Society*, 22(1):211–231, 2008.

[13] Emmanuel J. Candès and Terence Tao. Decoding by linear programming. *IEEE Transactions on Information Theory*, 51(12):4203–4215, 2005.

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

[15] Rachel W. Chan, Elizabeth A. Ramsay, Edward Y. Cheung, and Donald B. Plewes. The Influence of Radial Undersampling Schemes on Compressed Sensing Reconstruction in Breast MRI. *Magnetic Resonance in Medicine*, 67:363–377, 2012.

[16] Daniel J. Holland, C. Liu, Xiaowei Song, Erin L. Mazerolle, M. T. Stevens, Andy J. Sederman, Lynn F. Gladden, Ryan C. N. D'Arcy, Chris V. Bowen, and Steven D. Beyea. Compressed sensing reconstruction improves sensitivity of variable density spiral fMRI. *Magnetic Resonance in Medicine*, 70(6):1634–1643, 2013.

[17] Oliver Jeromin, Marios S. Pattichis, and Vince D. Calhoun. Optimal compressed sensing reconstructions of fMRI using 2D deterministic and stochastic sampling geometries. *BioMedical Engineering Online*, 11(25):1–36, 2012.

[18] Emmanuel J. Candès, Michael B. Wakin, and Stephen P. Boyd. Enhancing sparsity by reweighted L1 minimization. *Journal of Fourier Analysis and Applications*, 14(5-6):877–905, 2008.

[19] Scott Shaobing Chen, David L. Donoho, and Michael A. Saunders. Atomic Decomposition by Basis Pursuit. 43(1):129–159, 2001.

[20] Emmanuel J. Candès, Justin Romberg, and Terence Tao. Robust Uncertainty Principles : Exact Signal Frequency Information. *IEEE Transactions on Information Technology in Biomedicine : a Publication of the IEEE Engineering in Medicine and Biology Society*, 52(2):489–509, 2006.

[21] Michael Lustig. *Sparse MRI.* PhD thesis, 2008.

[22] Ingrid Daubechies, Michel Defrise, and Christine De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57(11):1413–1457, 2004.

[23] Amir Beck and Marc Teboulle. A Fast Iterative Shrinkage-Thresholding Algorithm. *Society for Industrial and Applied Mathematics Journal on Imaging Sciences*, 2(1):183–202, 2009.

[24] Ingrid Daubechies, Ronald Devore, Massimo Fornasier, and C. Sinan Güntürk. Iteratively reweighted least squares minimization for sparse recovery. *Communications on Pure and Applied Mathematics*, 63(1):1–38, 2010.

[25] Tom Goldstein and Stanley Osher. The Split Bregman Method for L1-Regularized Problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343, 2009.

[26] Benjamin Recht, Maryam Fazel, and Pablo A. Parrilo. Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization. *Society for Industrial and Applied Mathematics*, 52(3):471–501, 2007.

[27] George Mackiw. A Note on the Equality of the Column and Row Rank of a Matrix. *Mathematics Magazine*, 68(4):285, 1995.

[28] Martin Buehrer, Klaas P. Pruessmann, Peter Boesiger, and Sebastian Kozerke. Array compression for MRI with large coil arrays. *Magnetic Resonance in Medicine*, 57(6):1131–1139, 2007.

[29] Tao Zhang, John M. Pauly, Shreyas S. Vasanawala, and Michael Lustig. Coil compression for accelerated imaging with Cartesian sampling. *Magnetic Resonance in Medicine*, 69(2):571–582, 2013.

[30] Cornelius Lanczos. Matrix Calculus. In *Linear Differential Operators*, chapter 3, pages 100–162. Van Nostrand, 1961.

[31] Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, and Eero P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

[32] Harold Hotelling. Relations Between Two Sets of Variates. *Biometrika*, 28(3):321–377, 1936.

[33] Camillein D E Jordan. Essai sur la geometrie a n dimensions. *Bulletin de la Societe Mathematique de France*, 3:103–174, 1875.

[34] Andrew V. Knyazev and Merico E. Argentati. Principal Angles Between Subspaces in an A-Based Scalar Product: Algorithms and Estimates. *Society for Industrial and Applied Mathematics Journal on Scientific Computing*, 23(6):2008–2040, 2002.

[35] Ake Bjorck and Gene H Golub. Numerical Methods for Computing Angles Between Linear Subspaces Stable. *American Mathematical Society:Mathematics of Computation*, 27(123):579–594, 1973.

[36] Emmanuel J. Candès, Xiaodong Li, Yi Ma, and John Wright. Robust Principal Component Analysis ? pages 1–39, 2009.

[37] Chuangchuang Sun and Ran Dai. Rank-constrained optimization and its applications. *Automatica*, 82:128–136, 2017.

[38] Luong Trung Nguyen, Junhan Kim, and Byonghyo Shim. Low-Rank Matrix Completion: A Contemporary Survey. *IEEE*, 7:1–52, 2019.

[39] Maryam Fazel. Matrix Rank Minimization with Applications. *Dissertation*, 2002.

[40] Van Sy Mai, Dipankar Maity, Bhaskar Ramasubramanian, and Michael C. Rotkowitz. Convex Methods for Rank-Constrained Optimization Problems. *OJSA*, pages 123–130.

[41] Yehuda Koren. The BellKor Solution to the Netflix Grand Prize, 2009.

[42] Yehuda Koren. Collaborative filtering with temporal dynamics. *Communications of the ACM*, 53(4):447–456, 2009.

[43] Emmanuel J. Candès and Benjamin Recht. Exact Matrix Completion via Convex Optimization. *Foundations of Computational Mathematics*, 9(6):717–772, 2008.

[44] Mark Chiew, Stephen M. Smith, Peter J. Koopmans, Nadine N. Graedel, Thomas Blumensath, and Karla L. Miller. k-t FASTER: Acceleration of functional MRI data acquisition using low rank constraints. *Magnetic Resonance in Medicine*, 74(2):353–364, 2015.

[45] Guangcan Liu, Qingshan Liu, and Xiaotong Yuan. A New Theory for Matrix Completion. *Advances in Neural Information Processing Systems 30*, (Nips):785–794, 2017.

[46] Jian Feng Cai, Emmanuel J. Candès, and Zuowei Shen. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 20(4):1956–1982, 2010.

[47] Karthik Mohan and Maryam Fazel. Iterative reweighted algorithms for matrix rank minimization. *Journal of Machine Learning Research*, 13:3441–3473, 2012.

[48] Matan Gavish and David L. Donoho. The Optimal Hard Threshold for Singular Values is 4/\sqrt{3}. *IEEE Transactions on Information Theory*, 60(8):5040–5053, 2014.

[49] Jared Tanner and Ke Wei. Normalized iterative hard thresholding for matrix completion. *SIAM Journal on Scientific Computing*, 35(5), 2013.

[50] Justin P. Haldar and Diego Hernando. Rank-constrained solutions to linear matrix equations using PowerFactorization. *IEEE Signal Processing Letters*, 16(7):584–587, 2009.

[51] Jared Tanner and Ke Wei. Low rank matrix completion by alternating steepest descent methods. *Applied and Computational Harmonic Analysis*, 40(2):417–429, 2016.

[52] Zaiwen Wen, Wotao Yin, and Yin Zhang. Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm. *Mathematical Programming Computation*, 4(4):333–361, 2012.

[53] Prateek Jain, Praneeth Netrapalli, and Sujay Sanghavi. Low-rank Matrix Completion using Alternating Minimization. In *Proceedings of the 45th annual ACM symposium on Symposium on theory of computing*, pages 665–674, 2013.

[54] Emmanuel J. Candès and Terence Tao. The Power of Convex Relaxation : Near-Optimal Matrix Completion. *IEEE Transactions on Information Theory*, 56:2053–2080, 2009.

[55] Zhiqiang Xu. The minimal measurement number for low-rank matrix recovery. *Applied and Computational Harmonic Analysis*, 44(2):497–508, 2018.

[56] Alexander L. Chistov and Dima Yu Grigoriev. Complexity of quantifier elimination in the theory of algebraically closed fields. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 176 LNCS(April):17–31, 1984.

[57] Thomas P Minka. Automatic choice of dimensionality for PCA, 2000.

[58] Andrii Y. Petrov, Michael Herbst, and V. Andrew Stenger. Improving temporal resolution in fMRI using a 3D spiral acquisition and low rank plus sparse (L+S) reconstruction. *NeuroImage*, 157(August 2016):660–674, 2017.

[59] Ricardo Otazo, Alexandre Franco, Jingyun Chen, Charles Marmar, and Fernando Boada. Low-rank plus sparse (L+S) decomposition for separation of subsampled physiological noise in fMRI. *OHBM*, 45(4):1690, 2015.

[60] Jingyuan E. Chen, Jonathan R. Polimeni, Saskia Bollmann, and Gary H. Glover. On the analysis of rapidly sampled fMRI data. *NeuroImage*, 188(January):807–820, 2019.

[61] Saskia Bollmann, Alexander M. Puckett, Ross Cunnington, and Markus Barth. Serial correlations in single-subject fMRI with sub-second TR. *NeuroImage*, 166(October 2017):152–166, 2018.

[62] Soroosh Afyouni, Stephen M. Smith, and Thomas E. Nichols. Effective degrees of freedom of the Pearson's correlation coefficient under autocorrelation. *NeuroImage*, 199(May):609–625, 2019.

[63] Maxim Zaitsev, Karl Zilles, and N. Jon Shah. Shared k-space echo planar imaging with keyhole. *Magnetic Resonance in Medicine*, 45(1):109–117, jan 2001.

[64] Joop J. Van Vaals, Marijn E. Brummer, W. Thomas Dixon, Hans H. Tuithof, Hans Engels, Rendon C. Nelson, Brigid M. Gerety, Judith L. Chezmar, and Jacques A. Den Boer. "Keyhole" method for accelerating imaging of contrast agent uptake. *Journal of Magnetic Resonance Imaging*, 3(4):671–675, jul 1993.

[65] Zhi-Pei Liang, Hong Jiang, Christopher P. Hess, and Paul C. Lauterbur. Dynamic imaging by model estimation. *International Journal of Imaging Systems and Technology*, 8(6):551–557, 1997.

[66] Feng Huang, James Akao, Sathya Vijayakumar, George R. Duensing, and Mark Limkeman. K-t GRAPPA: A k-space implementation for dynamic MRI with high reduction factor. *Magnetic Resonance in Medicine*, 54(5):1172–1184, 2005.

[67] Zoran Stankovic, Jury Fink, Jeremy D. Collins, Edouard Semaan, Maximilian F. Russe, James C. Carr, Michael Markl, Mathias Langer, and Bernd Jung. K-t GRAPPA-accelerated 4D flow MRI of liver hemodynamics: influence of different acceleration factors on qualitative and quantitative assessment of blood flow. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 28(2):149–159, 2015.

[68] Bruno Madore, Gary H. Glover, and Norbert J. Pelc. Unaliasing by Fourier-encoding the overlaps using the temporal dimension (UNFOLD), applied to cardiac imaging and fMRI. *Magnetic Resonance in Medicine*, 42(5):813–828, nov 1999.

[69] Bruno Madore. Using UNFOLD to remove artifacts in parallel imaging and in partial-Fourier imaging. *Magnetic Resonance in Medicine*, 48(3):493–501, sep 2002.

[70] Jeffrey Tsao. On the UNFOLD method. *Magnetic Resonance in Medicine*, 47(1):202–207, jan 2002.

[71] Hsuan-Ming Huang and Yi-Yu Shih. Pushing CT and MR Imaging to the Molecular Level for Studying the "Omics": Current Challenges and Advancements. *BioMed Research International*, 2014:1–17, 2014.

[72] Jeffrey Tsao, Peter Boesiger, and Klaas P. Pruessmann. k-t BLAST and k-t SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations. *Magnetic Resonance in Medicine*, 50(5):1031–1042, nov 2003.

[73] Hong Jung, Jong Chul Ye, and Eung Yeop Kim. Improved k-t BLAST and k-t SENSE using FOCUSS. *Physics in Medicine and Biology*, 52(11):3201–3226, 2007.

[74] Henrik Pedersen, Sebastian Kozerke, Steffen Ringgaard, Kay Nehrke, and Yong Kim Won. k-t PCA: Temporally constrained k-t BLAST reconstruction using principal component analysis. *Magnetic Resonance in Medicine*, 62(3):706–716, 2009.

[75] Sairam Geethanath, Rashmi Reddy, Amaresha Shridhar Konar, Shaikh Imam, Rajagopalan Sundaresan, Ramesh Babu D. R., and Ramesh Venkatesan. Compressed Sensing MRI: A Review. *Critical reviews in biomedical engineering*, 41(3):183–204, 2013.

[76] Oren N. Jaspan, Roman Fleysher, and Michael L. Lipton. Compressed sensing MRI: A review of the clinical literature. *British Journal of Radiology*, 88(1056):1–12, 2015.

[77] S. Sophie Schauman, Mark Chiew, and Thomas W. Okell. Highly accelerated vessel-selective arterial spin labeling angiography using sparsity and smoothness constraints. *Magnetic Resonance in Medicine*, 83(June):892–905, mar 2019.

[78] Michael Lustig, Juan M. Santos, David L. Donoho, and John M. Pauly. k-t SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity. *Proceedings of the 14th Annual Meeting of the International Society for Magnetic Resonance in Medicine*, 50(5):2420, 2006.

[79] Cristina Chavarrías, Juan F. P. J. Abascal, Paula Montesinos, and Manuel Desco. Exploitation of temporal redundancy in compressed sensing reconstruction of fMRI studies with a prior-based algorithm (PICCS). *Medical Physics*, 42(7):3814–3821, 2015.

[80] Hong Jung and Jong Chul Ye. Performance evaluation of accelerated functional MRI acquisition using compressed sensing. *Proceedings - 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009*, (May):702–705, 2009.

[81] Priya Aggarwal and Anubha Gupta. Double temporal sparsity based accelerated reconstruction of compressively sensed resting-state fMRI. *Computers in Biology and Medicine*, 91:255–266, 2017.

[82] Xuesong Li, Xiaodong Ma, Lyu Li, Zhe Zhang, Xue Zhang, Yan Tong, Lihong Wang, Sen Song, and Hua Guo. Dual-TRACER: High resolution fMRI with constrained evolution reconstruction. *NeuroImage*, 164(February 2017):172–182, 2018.

[83] Bo Xu, Pascal Spincemaille, Gang Chen, Mukta Agrawal, Thanh D. Nguyen, Martin R. Prince, and Yi Wang. Fast 3D contrast enhanced MRI of the liver using temporal resolution acceleration with constrained evolution reconstruction. *Magnetic Resonance in Medicine*, 69(2):370–381, 2013.

[84] Zhongnan Fang, Nguyen Van Le, Man Kin Choy, and Jin Hyung Lee. High spatial resolution compressed sensing (HSPARSE) functional MRI. *Magnetic Resonance in Medicine*, 76(2):440–455, 2016.

[85] Zhi-Pei Liang. Spatiotemporal Imaging with Partially Separable Functions. *IEEE International Symposium on Biomedical Imaging*, 2:988–991, 2007.

[86] Bo Zhao, Justin P. Haldar, Cornelius Brinegar, and Zhi-Pei Liang. Low rank matrix recovery for real-time cardiac MRI. *IEEE International Symposium on Biomedical Imaging*, (May 2010):996–999, 2010.

[87] Justin P. Haldar and Zhi-Pei Liang. Spatiotemporal imaging with partially separable functions: A matrix recovery approach. *IEEE International Symposium on Biomedical Imaging*, (4):716–719, 2010.

[88] Fan Lam, Bo Zhao, Yinan Liu, Zhi-Pei Liang, Michael Weiner, and Norbert Schuff. Accelerated fMRI using Low-Rank Model and Sparsity Constraints. In *Proceedings of the International Society for Magnetic Resonance in Medicine 21*, volume 21, page 2620, 2013.

[89] Christian F. Beckmann and Stephen M. Smith. Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging. *IEEE Transactions on Medical Imaging*, 23(2):137–152, 2004.

[90] Mark Chiew, Nadine N. Graedel, Jennifer A. Mcnab, Stephen M. Smith, and Karla L. Miller. Accelerating functional MRI using fixed-rank approximations and radial-cartesian sampling. *Magnetic Resonance in Medicine*, 00:1–12, 2016.

[91] Mark Chiew, Nadine N. Graedel, and Karla L. Miller. Recovering task fMRI signals from highly under-sampled data with low-rank and temporal subspace constraints. *NeuroImage*, 174:97–110, 2018.

[92] Ricardo Otazo, Emmanuel J. Candès, and Daniel K. Sodickson. Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components. *Magnetic Resonance in Medicine*, 73(3):1125–1136, 2015.

[93] Venkat Chandrasekaran, Sujay Sanghavi, Pablo A. Parrilo, and Alan S. Willsky. Rank-sparsity incoherence for matrix decompositions. *IFAC Proceedings Volumes (IFAC-PapersOnline)*, 15(PART 1):1493–1498, 2009.

[94] Benjamin Trémoulhéac, Nikolaos Dikaios, David Atkinson, and Simon R. Arridge. Dynamic MR image reconstruction-separation from undersampled (k,t)-Space via low-rank plus sparse prior. *IEEE Transactions on Medical Imaging*, 33(8):1689–1701, 2014.

[95] Michelle Karker, Claire Lin, A Fessler, and Douglas C Noll. Evaluation of Sparse Sampling Approaches for 3D Functional MRI. In *ISMRM2019*, volume 60, page 2019, 2019.

[96] Hien M. Nguyen and Gary H. Glover. Field-corrected imaging for sparsely-sampled fMRI by exploiting low-rank spatiotemporal structure. *International Society for Magnetic Resonance in Medicine*, 22(c):1, 2014.

[97] Hien M. Nguyen and Gary H. Glover. Sparsely sampled functional magnetic resonance imaging using low-rank and sparsity constraints. *2014 IEEE 5th International Conference on Communications and Electronics, IEEE ICCE 2014*, (2):454–457, 2014.

[98] Vimal Singh, Ahmed H. Tewfik, and David B. Ress. Under-sampled functional MRI using low-rank plus sparse matrix decomposition. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 897–901, 2015.

[99] Priya Aggarwal, Parth Shrivastava, Tanay Kabra, and Anubha Gupta. Optshrink LR + S: accelerated fMRI reconstruction using non-convex optimal singular value shrinkage. *Brain Informatics*, 4(1):65–83, 2017.

[100] Lior Weizman, Karla L. Miller, Yonina C. Eldar, and Mark Chiew. PEAR: PEriodic and fixed Rank separation for fast fMRI. *Medical Physics*, 44(12):6166–6182, 2017.

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

4 | Alternating Minimisation Formulation of k-t FASTER

Contents

4.1 Introduction

k-t FASTER (fMRI Accelerated in Space-Time via Truncation of Effective Rank) is a reconstruction method that was designed to exploit the low-rank nature of fMRI. In the first k-t FASTER paper [1], an acceleration factor of R=4.27 was achieved in rsfMRI through the use of low-rank information alone, using a single (uniform) coil Cartesian sampling framework and solving with a matrix completion approach. A subsequent k-t FASTER paper [2] achieved an acceleration of R=12.5 by using a TURBINE [3] (hybrid radial-Cartesian) trajectory on tfMRI which incorporated coil spatial sensitivity information into the reconstruction [4]. Both papers used an iterative-hard threshold + matrix shrinkage (IHT+MS) approach to enforcing rank. The practicalities of thresholding and shrinkage will be covered in detail in this chapter.

There is a long history of using constraints to improve reconstructions in accelerated MR imaging [5]. There is also a history of combining different constraints together [6, 7, 8, 9]. By reformulating k-t FASTER as an Alternating Minimisation (AM) approach which decomposes the dataset into spatial and temporal component matrices, subspace-specific constraints can be incorporated alongside the low-rank constraint and the coil sensitivity information to help accelerate the reconstruction further. The ability to easily incorporate subspace-specific constraints is the main motivation for the AM approach, as it will also avoid the need for multiple potentially expensive SVD calculations (as SVDs get very computationally expensive as dataset size increases). It should be noted that it is possible to incorporate constraints into an SVD framework [10]).

This chapter will walk through the formulation of an AM approach to k-t FASTER. For the sake of clarity, AM k-t FASTER will refer to an alternating minimisation approach applied to k-t FASTER, and SVD k-t FASTER will refer to the original k-t FASTER approach using IHT+MS. Optimal reconstruction parameters, convergence criteria, and a variety of AM-based approaches will be investigated to deduce the ideal implementation of AM k-t FASTER for fMRI reconstruction. We will then demonstrate the applicability of AM k-t FASTER on Cartesian trajectories, before showing the general effectiveness on non-Cartesian trajectories, both in comparison with SVD k-t FASTER.

Throughout the chapter, A will refer to the dynamic image space optimisation variable as a full matrix (e.g. in SVD k-t FASTER). X and T will be used to represent the spatial and temporal domain respectively (where $A = XT'$), to represent the two matrices used as optimisation parameters in AM k-t FASTER. X and T will not be normalised nor necessarily be orthogonal, and so are distinct from the decomposed matrices produced by an SVD.

4.2 SVD k-t FASTER

Typical fMRI analyses represent data as specific time courses (either as brain activity, or as other coherent temporal processes e.g. physiological noise) with associated spatial maps showing which voxels have those time series. The creation of k-t FASTER as a low-rank approach to fMRI was motivated by the observation that PCA is often used to apply data compression to fMRI data, before ICA is applied as a denoising tool to generate a cleaner signal of brain activity [11]. Both decompositions make the spatial map vs. temporal process representation inherently low-rank. In a similar manner to how CS approaches seek to acquire a compressed signal at the point of "sensing" (rather than acquire a full signal and then compress), k-t FASTER aims to acquire a low-rank representation of the data without acquiring information that will later be discarded in the PCA. This maximises the sampling efficiency, as low-rank methods attempt to avoid acquiring data that provide no value to the final analysis (figure 4.1).

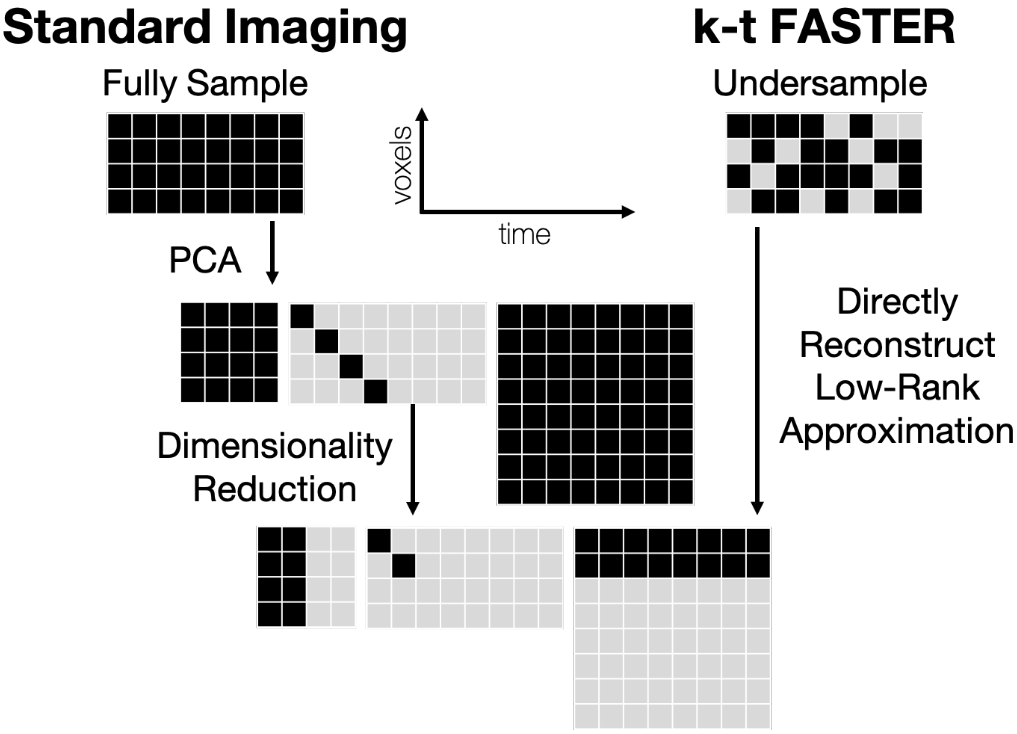


Figure 4.1: A demonstration of the increased sampling efficiency of k-t FASTER. The standard approach to fMRI reconstruction will often discard information through PCA-based dimensionality reduction. k-t FASTER directly approximates in the low-rank form of the desired final output representing a more efficient approach which does not require full initial sampling.

The first k-t FASTER paper showed that $> 80\%$ of the signal can be captured in the first 25 components of a space-time matrix in an average scan (see figure 4.2). It is this inherent low-rank nature of the space-time dataset which validates the use of low-rank approaches in fMRI.

The original Cartesian formulation of SVD k-t FASTER [1] and the updated non-Cartesian formulation [2] use slightly different algorithms. The Cartesian SVD k-t FASTER approach uses a low-rank matrix completion strategy [12] (see section 3.4.3.1) as a robust way to enforce a low-rank structure during reconstruction. Matrix d will represent the sampled data, and matrix A_i is the current estimate of the system in k-t space. Data consistency between the sampled data points in k-space and the current estimate is defined by an update step size ($0 \geq \mu \geq 1$) and a Boolean

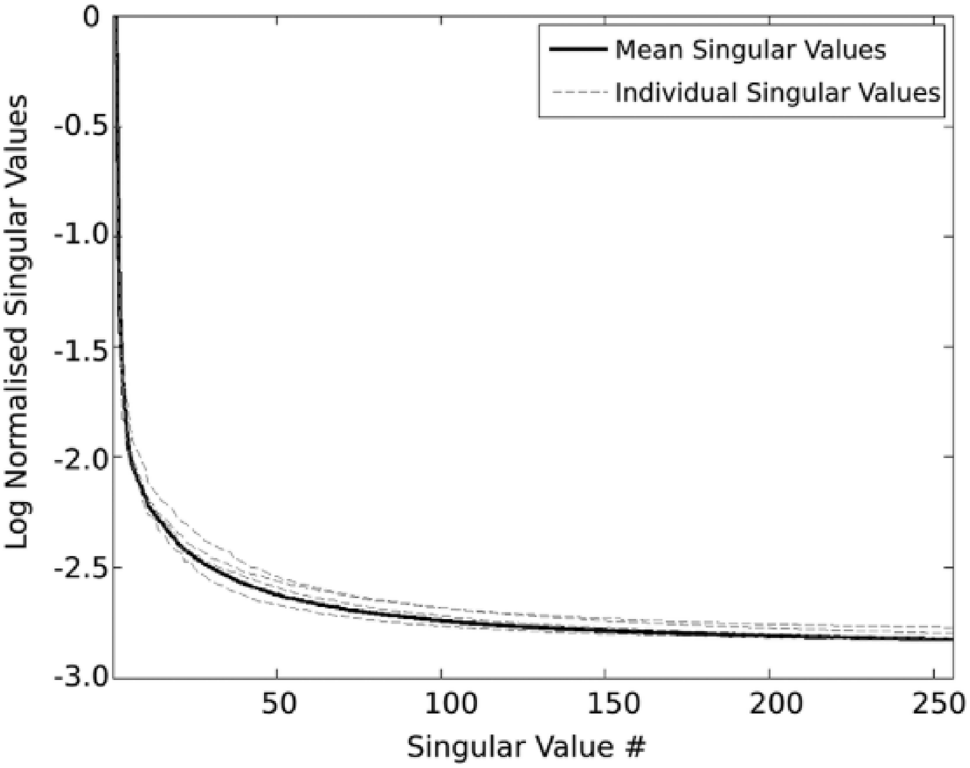


Figure 4.2: A normalised singular value curve produced from fMRI data with 512 time points from six different subjects (dotted: individual distributions, solid: mean distribution). Figure used with permission of Mark Chiew [1].

mask defining the location of the sampled points (Ω) which is applied through elementwise matrix multiplication ($.*$).

$$A_{i+1} = S\Big(A_i + \mu\big(\Omega .* (d - A_i)\big)\Big) \tag{4.1}$$

S is a shrinkage operator which can be implemented in a variety of ways. A hard thresholding approach enforces a fixed rank on the system by performing an SVD on the current estimate of the dataset and setting all components beyond an arbitrary number of singular values to zero:

$$S_r^{hard} = \begin{cases} \sigma_j = \sigma_j, & j \leq r \\ \sigma_j = 0, & j > r \end{cases} \tag{4.2}$$

Alternatively, a soft thresholding approach could apply a value-based threshold, setting all singular values less than a certain value τ to zero. This approach does not impose a fixed rank on the system, instead imposing a general low-rank nature with no strictly defined limit on the number of singular values. Additionally, the singular values which are not removed could still be shrunk by τ within the shrinkage operator.

$$S_\tau^{soft} = \begin{cases} \sigma_j = \sigma_j - \tau; & \sigma_j \geq \tau \\ \sigma_j = 0; & \sigma_j < \tau \end{cases} \tag{4.3}$$

SVD k-t FASTER uses a hybrid of the two shrinkage approaches called Iterative Hard Thresholding with Matrix Shrinkage (IHT+MS). IHT+MS imposes a hard upper limit on the number of singular values (r), but shrinks the remaining values by the multiple of the fixed shrinkage coefficient (c) and the $(r+1)^{\text{th}}$ singular value (the largest singular value that is cut by the hard threshold). This is a variant of the fixed point continuation algorithm [13], which also uses a hybrid approach. The floating shrinkage parameter is similar to a non-convex spectral penalty introduced by Lingala et al. [14]. This process is repeated for q iterations, or until a convergence criterion ϵ is reached.

$$S_{r,c} = \begin{cases} \sigma_j = \sigma_j - c\sigma_{r+1} & j \leq r \\ \sigma_j = 0 & j > r \end{cases} \tag{4.4}$$

The IHT+MS can be integrated into an overall matrix completion algorithm for k-t FASTER to produce pseudocode algorithms 1 and 2 for the Cartesian and non-Cartesian formulation respectively. In the Cartesian case, one single uniform coil is assumed. In the non-Cartesian [3] multicoil [4] framework, Ω can be replaced by E, which incorporates the NUFFT operator and the coil sensitivity information into the reconstruction.

Algorithm 1 Cartesian single-coil SVD k-t FASTER Pseudocode

```
Parameters
$d$: Undersampled k-t fMRI data
$\Omega$: A mask of 1s and 0s to indicate sampling locations
$A$: The low-rank matrix estimate of the dataset with potential maximum rank $n$
$r$: The maximum desired rank
$\mu$: The update step size
$c$: The shrinkage coefficient
$q$: Number of iterations
$\epsilon$: The convergence threshold
$\sigma_j$: The $j^{th}$ singular value in $\Sigma$

Algorithm
for $i = 1 : q$ do
    if $\|A_i - A_{i-1}\|_F < \epsilon$ then
        $U\Sigma V' = SVD\Big(A_i + \mu\big(\Omega. * (d - A_i)\big)\Big)$
        for $j = 1 : n$ do
            if $j \leq r$ then
                $\tilde{\sigma}_j = \sigma_j - c\sigma_{r+1}$
            else if $j > r$ then
                $\tilde{\sigma}_j = 0$
        $A_{i+1} = U\tilde{\Sigma}V'$
```

Algorithm 2 Non-Cartesian SVD k-t FASTER Pseudocode

```
Parameters
  $d$: Multicoil undersampled k-t fMRI data
  $E$: Non-Uniform Fast Fourier Transform (NUFFT) and multi-coil encoding operator
  $A$: The low-rank matrix estimate of the dataset with potential maximum rank $n$
  $r$: The maximum desired rank
  $\mu$: The update step size
  $c$: The shrinkage coefficient
  $q$: Number of iterations
  $\epsilon$: The convergence threshold
  $\sigma_j$: The $j^{th}$ singular value in $\Sigma$

Algorithm
for $i = 1 : q$ do
    if $\|A_i - A_{i-1}\|_F < \epsilon$ then
        $U\Sigma V' = SVD\Bigg(A_i + \mu E'(d - EA_i)\Bigg)$
        for $j = 1 : n$ do
            if $j \leq r$ then
                $\tilde{\sigma}_j = \sigma_j - c\sigma_{r+1}$
            else if $j > r$ then
                $\tilde{\sigma}_j = 0$
        $A_{i+1} = U\tilde{\Sigma}V'$
```

4.3 Formulating AM k-t FASTER

In SVD k-t FASTER, only matrix A is estimated. In AM k-t FASTER, two separate matrices (X and T) must be jointly estimated. X and T represent the spatial and temporal domain respectively. Data consistency in the AM optimisation problem can be formulated with respect to a transform operator E and the undersampled dataset d as follows:

$$\hat{X}, \hat{T} = argmin_{X,T}\left(\|EXT' - d\|_2^2\right) \tag{4.5}$$

To solve equation 4.5, one parameter is fixed while the other is updated to minimise $\|EXT' - d\|_2^2$, the general cost function. Then, that parameter is fixed and the other parameter is updated. This process of alternating the active variable in the cost function is the defining characteristic of alternating minimisation. Equation 4.5 can be broken down into two convex subproblems: the spatial subproblem which optimises for X, and the temporal subproblem which optimises for T.

The maximum rank of the reconstructed dataset in an AM implementation is encoded in the dimensionality of the decomposed matrices X and T, rather than through the singular value thresholding of section 4.2. X is $x \times r$ (x - number of voxels) and T is $t \times r$ (t - number of frames/time points). If $r < (x, t)$, then the maximum rank of the final reconstruction can be clearly stated as r through the following properties:

$$\begin{aligned} rank(X) &\leq \min(x, r) \\ rank(T) &\leq \min(t, r) \\ rank(XT') &\leq \min(rank(X), rank(T)) \end{aligned} \tag{4.6}$$

A visual breakdown of the matrix decomposition and enforcement can be seen in figure 4.3.

4.3.1 Different AM Implementations

The goal of an AM implementation is to identify the matrices X and T that achieve maximal data consistency with a target rank r. Two approaches were considered. The first approach was Incremented-Rank PowerFactorisation (IRPF) [15], which has previously been shown to be effective at recovering low-rank matrices from undersampled data. IRPF uses two nested loops. The outer loop increments from a rank-1 optimisation to a rank-r optimisation. The inner loop fully optimises X and T at the selected rank. Each outer loop is initialised based on optimisation from the previous iteration's inner loop. This approach is shown in pseudocode 3.

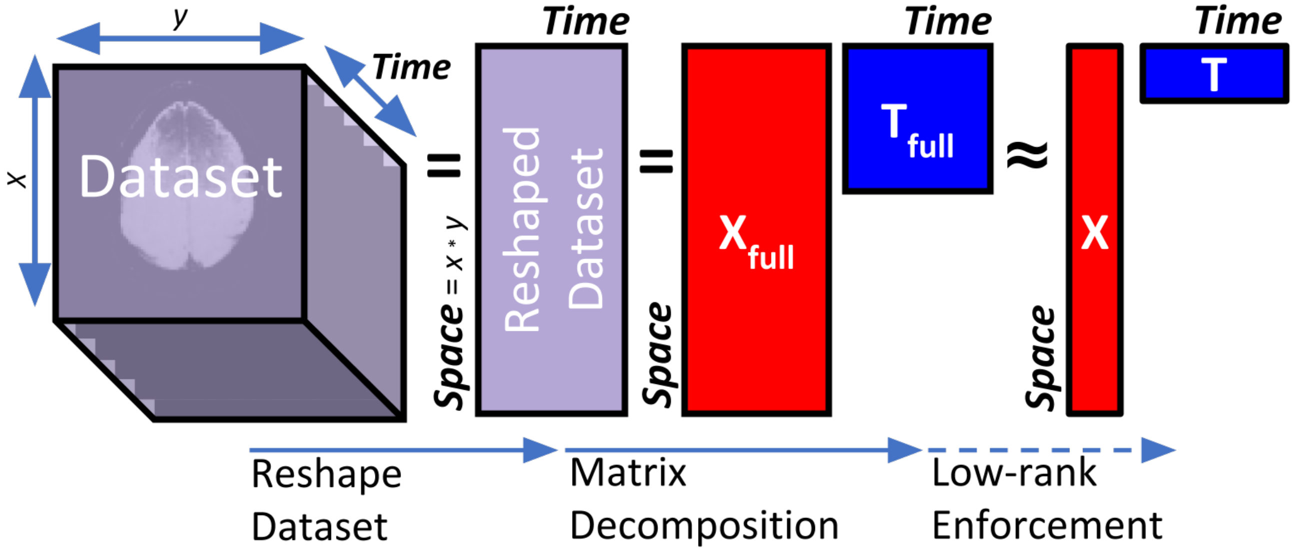


Figure 4.3: The matrix decomposition and low-rank enforcement of an fMRI dataset (a third spatial dimension could also be contained in the first dimension of X_{full}/X).

The second approach directly calculates a rank-r representation of the final dataset. This is equivalent to an IRPF approach which jumps straight to the final iteration of the outer loop, and so doesn't have the benefit of the IRPF initialisation in that loop from the rank-(r-1) optimisation. This approach is termed "All Components" (AC), and is shown in pseudocode 4. For both approaches, the issue of convergence and initialisation will be discussed in section 4.3.2. For now, χ represents the convergence criterion and ϵ represents the convergence threshold.

Both methods were tested on a retrospectively undersampled dataset and evaluated based on a) how well they represent the underlying truth, b) how aligned the subspaces are between the different methods, and c) length of time taken to converge.

4.3.2 Convergence

There are a number of different choices available when considering a convergence criterion. SVD k-t FASTER is optimised over a single parameter (matrix A), and uses the change in Frobenius norm of that parameter as the criterion.

$$\|A_i - A_{i-1}\|_F < \epsilon \tag{4.7}$$

The IRPF paper [15] defines their convergence criterion as the relative error between the transform of the estimate and the undersampled dataset. This is formulated as the Frobenius norm of the data consistency term normalised by the Frobenius norm of the undersampled dataset. It is

Algorithm 3 IRPF Pseudocode

```
Parameters
    $d$: Multicoil undersampled k-t fMRI data
    $E$: Non-Uniform Fast Fourier Transform (NUFFT) and multi-coil encoding operator
    $X$: The spatial component matrix
    $T$: The temporal component matrix
    $X_{1:j,i}$: The $i^{th}$ optimisation cycle of $X$ containing components $1 \rightarrow j$
    $r$: The maximum desired rank
    $\chi$: The convergence criterion function
    $CF_{j,i}$: The overall cost function for the $i^{th}$ cycle in the $j^{th}$ loop
    $q$: The maximum number of cycles

Algorithm
$initialise(X_{1,0}, T_{1,0})$
for $j = 1 : r$ do
    $rank(X) = rank(T) = j$
    $X_{1:j-1,0} = X_{1:j-1,q}$
    $T_{1:j-1,0} = T_{1:j-1,q}$
    for $i = 1 : q$ do
        while $\chi(X_i, T_i) < \epsilon$ do
            $X_{1:j,i+1} = argmin_{X_{1:j}} \|EX_{1:j}T'_{1:j,i} - d\|_2^2$
            $T_{1:j,i+1} = argmin_{T_{1:j}} \|EX_{1:j,i+1}T'_{1:j} - d\|_2^2$
            $CF_{j,i+1} = ||EX_{1:j,i+1}T'_{1:j,i+1} - d\|_2^2$
```

Algorithm 4 All Components Pseudocode

```
Parameters
    $d$: Multicoil undersampled k-t fMRI data
    $E$: Non-Uniform Fast Fourier Transform (NUFFT) and multi-coil encoding operator
    $X$: The spatial component matrix
    $T$: The temporal component matrix
    $X_{1:r,i}$: The $i^{th}$ optimisation cycle of $X$ containing all $r$ components
    $r$: The maximum desired rank
    $\chi$: The convergence criterion function
    $CF_i$: The overall cost function for the $i^{th}$ cycle
    $q$: The maximum number of cycles

Algorithm
$initialise(X_{1:r,0}, T_{1:r,0})$
for $j = r$ do
    $rank(X) = rank(T) = r$
    for $i = 1 : q$ do
        while $\chi(X_{1:r,i}, T_{1:r,i}) < \epsilon$ do
            $X_{1:r,i+1} = argmin_{X_{1:r}} \|EX_{1:r}T'_{1:r,i} - d\|_2^2$
            $T_{1:r,i+1} = argmin_{T_{1:r}} \|EX_{1:r,i+1}T'_{1:r} - d\|_2^2$
            $CF_{i+1} = ||EX_{1:r,i+1}T'_{1:r,i+1} - d\|_2^2$
```

calculated at the end of an optimisation cycle, after both the spatial subproblem and the temporal subproblem have been optimised.

$$\frac{\|EXT' - d\|_2}{\|d\|_2} < \epsilon \tag{4.8}$$

As a main motivation for the AM k-t FASTER approach is tied into the addition of extra constraints, a criterion is needed which can be easily extended beyond a cost function containing just a data consistency term. As such, the AM k-t FASTER convergence criterion will be defined as the absolute change in cost function and normalise by the updated cost function. It will be termed the normalised cost function gradient (equation 4.9); the algorithm will check to see if the criterion is satisfied after the temporal subproblem.

$$\frac{|CF_i - CF_{i-1}|}{CF_i} < \epsilon \tag{4.9}$$

The choice of ϵ will be determined empirically on a retrospective undersampled dataset as the value representing the point at which randomly initialised subspaces have converged (i.e. have a CCS score ≈ 1). It should be noted that the individual X and T subproblem cost functions will always decrease, as both the X and T subproblems are convex optimisation problems. However, there is no guarantee of the overall cost function decreasing between optimisation cycles once constraints are added.

4.3.3 Iterations

Alternately optimising two variables at once can be a lengthy procedure. Each subproblem optimisation must be long enough that X_i can change significantly, but not too long that time is wasted moving incrementally towards a solution that will change in the next cycle. For the purpose of clarity, one "cycle" will be defined as one optimisation of both the spatial and temporal subproblem, and one "iteration" will refer to the number of updates within a subproblem. The number of internal iterations will affect the total change in cost function during a cycle, and so will require a stricter ϵ in the convergence criterion.

4.3.4 Initialisation

The overall AM formulation of the low-rank problem is non-convex, despite containing two convex subproblems [15]. One result of this fact is that certain initialisations can lead to undesirable solutions. The most obvious example is that initialising T as zero-filled will allow an infinitely-varied expression of X to yield the exact same solution, since the cost functions are only defined by the product XT', and not (in this version without subspace constraints) by any standalone X and T terms.

The temporal mean is an image formed by concatenating the blades across all time frames into a single reconstruction, and is an intuitive choice for the first component of X as the mean spatial map is very likely to approximate closely to the most dominant spatial component (relatedly, the first temporal component will be roughly constant over time). The initialisation of other components is less clear. Some basic initialisations will be evaluated based on whether they form component matrices X and T which consistently converge to map the same underlying subspaces. The initialisations tested will be: fully randomised matrices, fully randomised and orthogonalised matrices, and fully randomised and orthogonalised matrices which contain the temporal mean as the first component. From this set of initialisations, we should be able to draw conclusions about the form of initialisations as a whole.

4.4 Methodology

All results in this chapter will use an acquired task fMRI dataset which has been retrospectively sampled. This dataset will be referred to as the "retrospective dataset". All datasets in this thesis will go through the same analysis, with any variations clearly stated where applicable. As such, this section will then be used to outline a standard pipeline for data processing that other chapters will mimic for the purpose of brevity.

4.4.1 Retrospective Dataset

The retrospective dataset consisted of a single slice (100×100) over 300 frames (TR_{frame}:1s) that was generated from five blocks of a 30s/30s on/off finger-tapping task. The original data was acquired as a full volume through a TURBINE acquisition with 20 blades/frame (TR_{blade}:50ms, TE_{blade}:30ms), and the single axial slice was chosen to contain a large amount of activation. The 32-channel coil dataset was compressed using an SVD to retain the 8 most dominant components for speed/memory purposes [16, 17]. The resolution was 2mm × 2mm × 2mm for each voxel. The dataset only contained magnitude information, with no added noise or phase variation.

4.4.2 Trajectory

Unless a Cartesian trajectory is specified, all data in the retrospective dataset were resampled from a ground truth using a Golden Ratio angle radial sequence which rotated a blade around the phase encoding axis at constant azimuthal increments of $\pi/\Phi \approx 111.25°$ [18] (mimicking a cross-section of the TURBINE [3] trajectory, see section 2.1.5.3). The angle was not reset at the start of each frame (e.g. each frame had a uniquely oriented set of blades).

4.4.3 Reconstruction Parameters

The reconstruction rank was chosen as 16 in all cases (a value used in recent literature for low-rank tfMRI [10]). A minimum residual algorithm (minres.m) [19] was executed in MATLAB, with a subproblem tolerance set at 1×10^{-15} (n.b. this tolerance was used for the separate optimisations of X and T and is distinct from ϵ, the overall convergence criterion). Each reconstruction used four parallel threads with an 8GB maximum of RAM. Reconstructions were computed using MATLAB R2019a [20]. Computation used the Oxford Biomedical Research Computing (BMRC) facility.

4.4.4 Processing Pipeline

Prior to reconstruction, the retrospective dataset was precleaned with motion correction tools and FIX in order to ensure a clean signal for all reconstructions. All task fMRI analysis was performed in FEAT [21] (part of the FSL package). No registration or spatial smoothing was used, but temporal high-pass filtering was enabled. An on/off block set-up was selected for the model window to generate the GLM. To account for residual autocorrelation, the resulting z-statistic maps were null-corrected using mixture modelling [22], which is part of the MELODIC functionality in FSL (see section 2.3.2.3 for more details).

4.4.5 Image Quality

Reconstruction image quality can be difficult to determine [23]. Incoherent artefacts are usually preferable to coherent artefacts, a factor which can be difficult to capture in a single value metric. Most common image quality metrics (e.g. root mean square error or structural similarity index) are based on signal power, and so are dominated by the first component in a component matrix (see figure 4.2). The low-rank reconstructions are attempting to directly reconstruct a reduced subspace representation of the functional information, and so the image quality metrics chosen must reflect that.

Canonical Correlations (CC) represent the alignment of two subspaces (see section 3.4.2). The CC between two viable matrices is an ordered r-element vector, with each subsequent element representing the cosine of the smallest remaining angle between the subspaces mapped by each matrix. The Canonical Correlation Score (CCS) is the normalised sum of those r elements, forming a $0 \rightarrow 1$ score. The CCS is used to compare the two subspaces mapped by two separate component matrices. The CCS is calculated between the reconstruction and the first r components of the decomposed retrospective ground truth unless otherwise stated. The spatial CCS will be referred to as "X CCS" and the temporal CCS as "T CCS". In order to provide an accurate comparison for radial trajectories, spatial subspaces will be windowed by a Tukey window of radius k_{max}. The CC vectors curves may also be plotted to show a visualisation of subspace alignment, with the correlation coefficients ordered by alignment (rather than by singular value magnitude).

Receiver Operator Characteristic (ROC) curves plot a False Positive Rate (FPR) against a True Positive Rate (TPR) (equations 4.10 and 4.11). Here, the ROC curves compare the z-statistic maps from reconstructions against the z-statistic map of a fully sampled reconstruction. This demonstrates the correct classification power of those reconstructed z-statistic maps. A threshold of z>3.1 was used to classify true positives in the retrospective dataset, the binary map of which is shown in figure 4.4 (any voxels in the reference dataset with a z-stat above this threshold are true positives and are labelled in green). The Area Under the Curve (AUC) of the ROC provides a simple comparison of many reconstructions, but the underlying z-statistic maps also provide valuable information as to the spatial location of false positives (which will be labelled in red) and false negatives (labelled in blue). Here we are strictly measuring how well the reconstructions are trying to strictly represent a ground truth, rather than the underlying activation. As such, all voxels which pass the threshold are considered as true positives, whereas a standard fMRI analysis might deem these as false positives because they lie outside of the sensorimotor cortex. All z-stat maps (aside from the truth) will be thresholded at the z-statistic corresponding to an FPR of 0.0015, as this was judged to be an approximate balance of allowing the reader to identify the location of likely false identification while not swarming the picture with too many negative points. The z-stat map is shown atop the mean brain image in the reconstruction.

The ROC AUC score is a standard metric for characterising an ROC curve, which evaluates the classification power of a model. However, there are some caveats that are needed when considering these ROC AUC values in fMRI. The nature of desirable z-statistic thresholds in fMRI lead to only the low FPR portion of the ROC curve being of practical interest. If a TPR of 1 is reached at a low FPR (as is the case with almost all ROC curves in this thesis), the ROC AUC score holds as a comparative metric. As there are only eighty-six true positive voxels in the ground truth (figure 4.4), failing to capture only one or two voxels until a high FPR can give a drastically decreased ROC AUC score, while not having any effect on the region of an ROC curve that reflects classification thresholds that are used for analysis. The two cases where a TPR of 1 is not reached at a low FPR are in IRPF reconstructions and Cartesian SVD k-t FASTER reconstructions, a fact which will also be explicitly stated in the text. The shape of the ROC curve in the relevant region and the z-stat parameter maps will still serve as a means of comparing estimated brain activity in these instances.

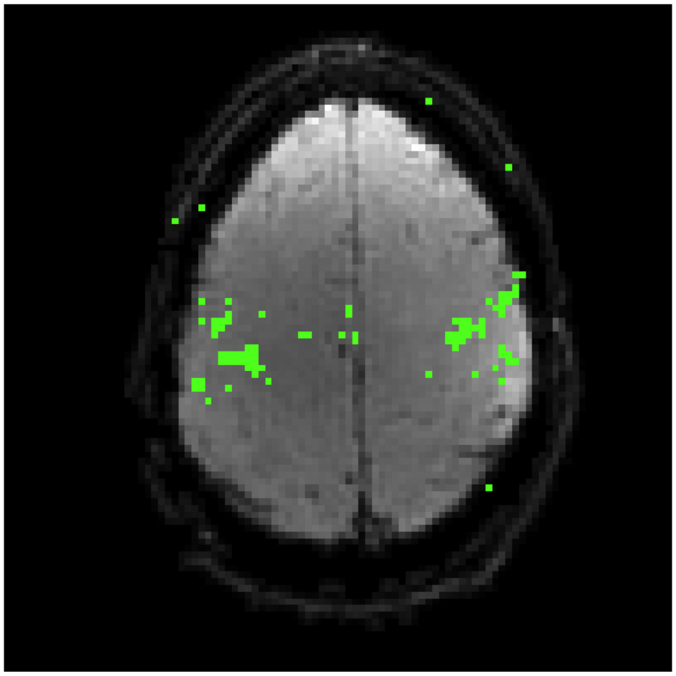


Figure 4.4: The true retrospective z-stat map. The pixels fulfilling a z-statistic significance >3.1 in the retrospective truth activation maps are shown in green. The background is the mean brain intensity of the truth.

$$FPR = \frac{FP}{FP + TN} \tag{4.10}$$

$$TPR = \frac{TP}{TP + FN} \tag{4.11}$$

It is very important to consider all of the relevant metrics concurrently (i.e. canonical correlation curves/scores, ROC curves/ROC AUC scores, quantitative visualisation of the z-statistic maps), as no metric alone is able to capture the full strength of the reconstruction. Combining the visualisations of these metrics ensures that reconstructions are comparable, although this means a lot of information has to be represented in a relatively small space. Where possible, the format of the figures will be kept consistent, with figure 4.5 providing a schematic for one of the most dense representations of reconstruction quality that will be used in this thesis.

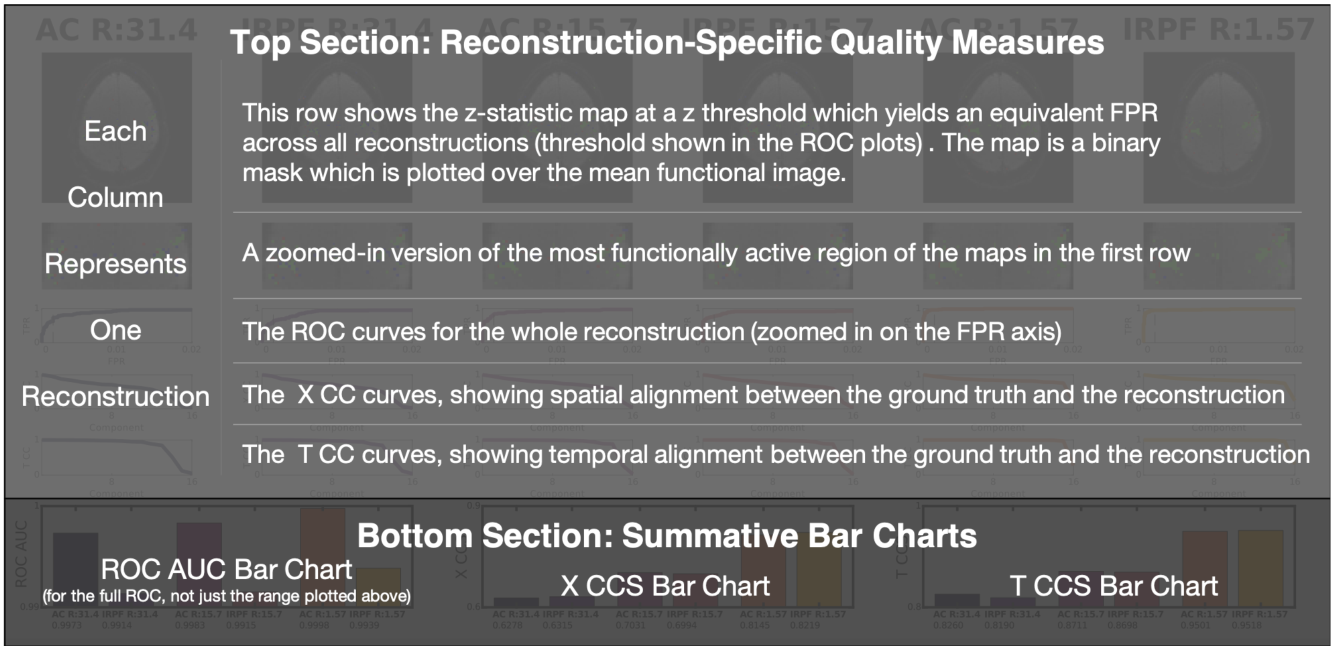


Figure 4.5: A comparison figure schematic. Each column represents a different method, with summative bar graphs on the bottom row. The z-stat maps (full and zoomed) are thresholded at the FPR shown with a dotted line on the ROC graph in row 3. On the z-stat maps, green pixels indicate true positives, red pixels indicate false positives, and blue pixels are false negatives. The full CC curves for space and time are shown in rows 4 and 5. Row 6 contains the ROC AUC, the X CCS score, and the T CCS score moving from left to right (the bar chart colour and line plots will correspond to a particular reconstruction). The bar chart labels will list the precise score shown in each plot.

In addition to image quality, the reconstruction time is also an important parameter that must be considered, especially in cases where differences in reconstruction quality are minimal. Reconstruction time will be displayed in a separate graph or table.

4.4.6 Colourmaps

Many standard colourmaps are deceiving to the eye, and lead to difficult-to-interpret results [24]. A good colourmap should be perceptually uniform, leading to natural conclusions based on perceived intensity. As an additional benefit, these colourmaps should also be colour-blind friendly, allowing the results to be visually accessible to a wide audience.

To that end, line graphs (and data grids) displaying more information than a single line used the inferno colourmap [25], in addition to distinct marker shapes where possible. Where figures contain a single plot with only one line, a black line is used.

4.5 Results & Discussion

The following subsection will work through a selection of graphs to demonstrate how the reconstruction parameters were chosen. This process was highly dimensional, with many factors that are not necessarily independent having to be jointly considered. In order to make a concise case for the set of parameters chosen for AM k-t FASTER, the figures will be limited to those that clearly show the effect of varying one parameter with other parameters fixed to their optimal setting. The varying parameter was optimised over a range of acceleration factors to demonstrate the effect of undersampling. Once the optimal AM k-t FASTER parameters have been selected, a comparison will be made with SVD k-t FASTER

4.5.1 Selecting AM parameters

IRPF vs AC Implementation Method

The first AM "parameter" that will be considered is the implementation method. The IRPF and AC methods were used for reconstructions at R = 31.42 (5 blades/frame), R=15.71 (10 blades/frame) and R = 1.57 (100 blades/frame). The R calculation represents the number of lines used relative to the total number of lines needed to fully sample k-space from a radial perspective (see equation 2.26). The acceleration factors used collectively represent: a high acceleration at which the reconstruction starts to break down, a reasonable acceleration at which most of the information is preserved, and a near fully sampled reconstruction.

Figure 4.6 shows a comparison between both reconstruction methods at different acceleration factors. The CCS scores record the alignment between the reconstructed component matrix and the ground truth. The CC curves display the ordered alignment of principal angles between the subspaces mapped by those matrices. The ROC curves and z-statistic maps represent the functional information in the comparison figure, with the ROC AUC bar chart representing the integral of the whole ROC curve. The z-statistic maps are shown at the threshold defined by the dotted line in the ROC curves. The zoom of the ROC graph may also lead to ROC AUC scores which can be easy to misinterpret. The TPR of IRPF methods doesn't reach 1 until FPR>0.8 for all R, whereas the TPR reaches 1 at an FPR<0.05 for all the AC methods (meaning the ROC AUC metric is unreliable in this instance). Despite this, a more qualitative assessment of the shape of the ROC curves and the z-stat map at a realistic threshold shows that the AC and IRPF reconstructions both approximately represent the same functional information for a given R over the range of practical z-stat thresholds.

Even more so than the ROC curve shape and z-stat map, the X/T CCS scores for the IRPF and AC reconstructions are nearly identical within each acceleration factor, meaning the reconstructed matrices map subspaces which are equally aligned with the underlying truth. This can be further seen in figure 4.7, which shows the alignment between the IRPF and AC reconstructions. The temporal subspace is fully aligned between both approaches, and the spatial subspaces are close to aligned (although the alignment decreases as R increases). Even though the spatial subspace is not fully aligned, the subspaces mapped by the spatial component matrices X_{AC} and X_{IRPF} are still more aligned with each other than with the underlying truth.

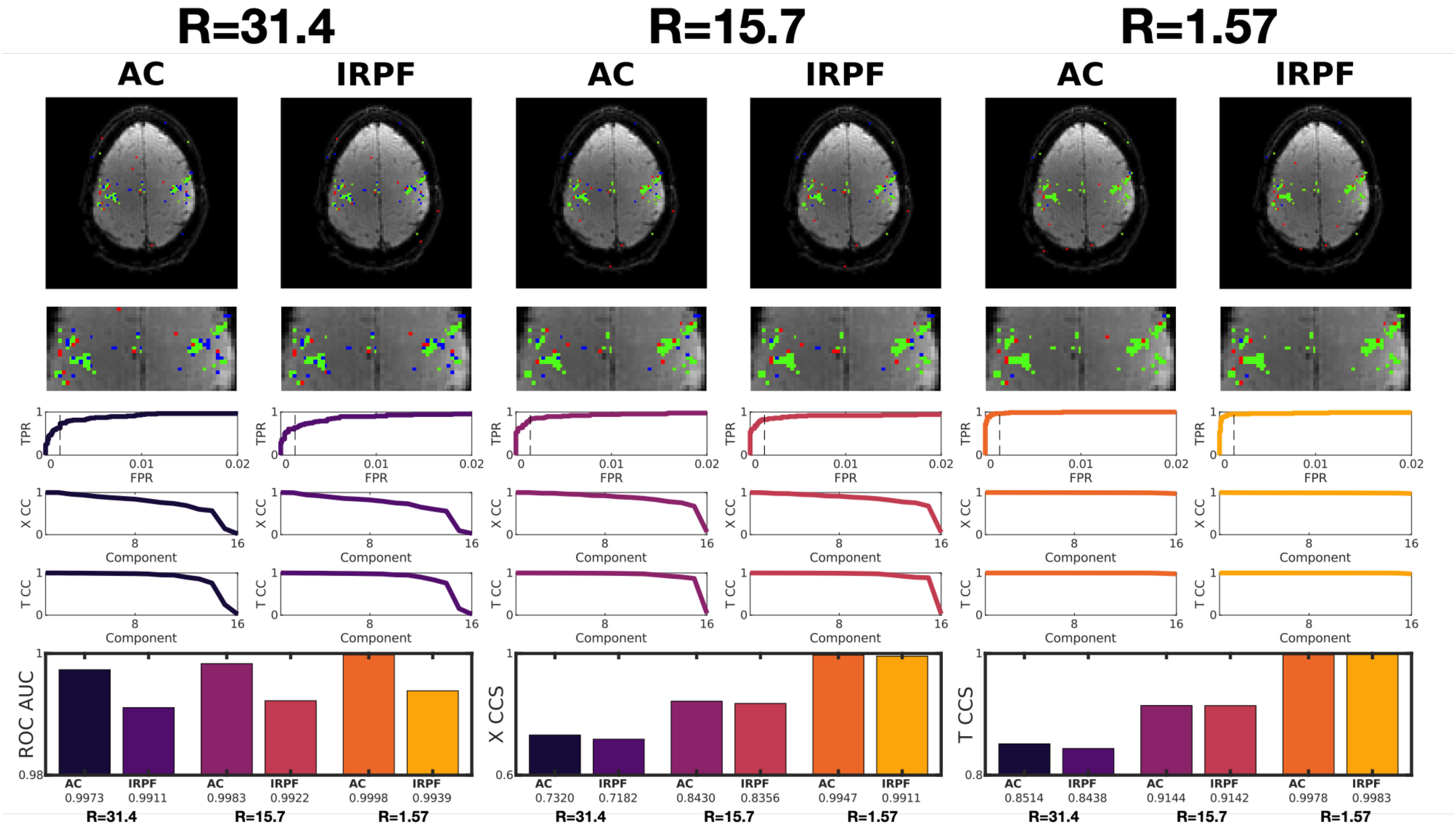


Figure 4.6: A comparison figure for the IRPF and AC approaches to low-rank reconstruction at a range of acceleration factors. Sample z-stat parameter maps, an ROC curve, and the CC curves are shown in a columnwise structure for each reconstruction. Summative bar graphs for all reconstructions are shown on the bottom row. Figure 4.5 contains a detailed schematic of the graph format.

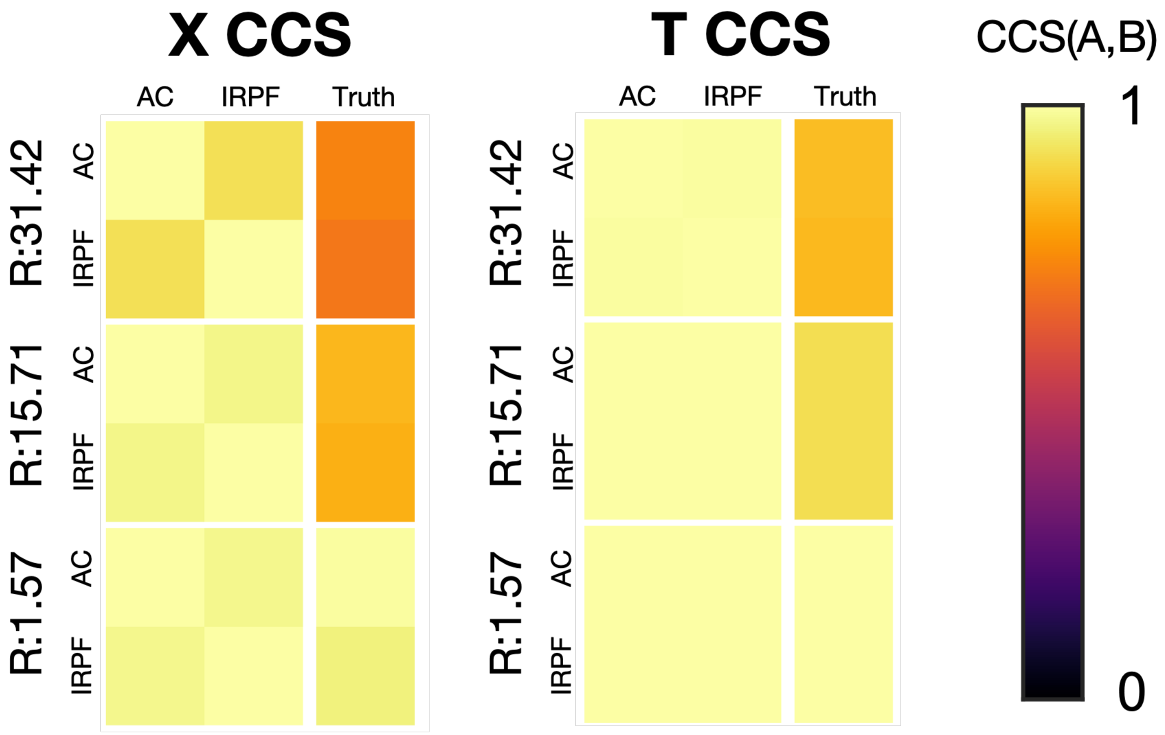


Figure 4.7: The Canonical Correlation Scores between different reconstructed subspaces for AC and IRPF reconstructions. A and B are placeholders for labelled matrices in the CCS(A,B) colourbar. All CCS scores shown are either calculated within an acceleration factor or compared against a ground truth. A CCS score of 1 represents total alignment, and 0 is zero alignment. Within each 2x2 grid, the diagonal values show perfect self-correlation, and the off-diagonal values both represent the alignment between IRPF and AC reconstructions at the specified acceleration factor.

The interpretation of figures 4.6 and 4.7 can be summarised as followed - either an AC or IRPF approach can be used to reconstruct the chosen retrospective dataset without any relative loss in reconstruction fidelity in choosing one method over the other, as both methods map roughly the same underlying subspace and capture approximately equal functional information over the range of relevant thresholds. If an advantage is to be given, then the AC method appears to slightly outperform the IRPF approach.

The more practical metric of calculation time is a more clear differentiator between AC and IRPF for use implementation purposes. Figure 4.8 shows the time taken to reach convergence for all methods, with IRPF taking dramatically longer than the AC formulation. This increased time to convergence is not unexpected, as the IRPF method must fully optimise for each component matrix at an incremented rank, whereas the AC method jumps straight to the optimisation at the desired final rank. When the decreased convergence time is combined with the either equal or slightly improved reconstruction performance of the AC approach, AC is left as the natural implementation of choice for AM k-t FASTER.

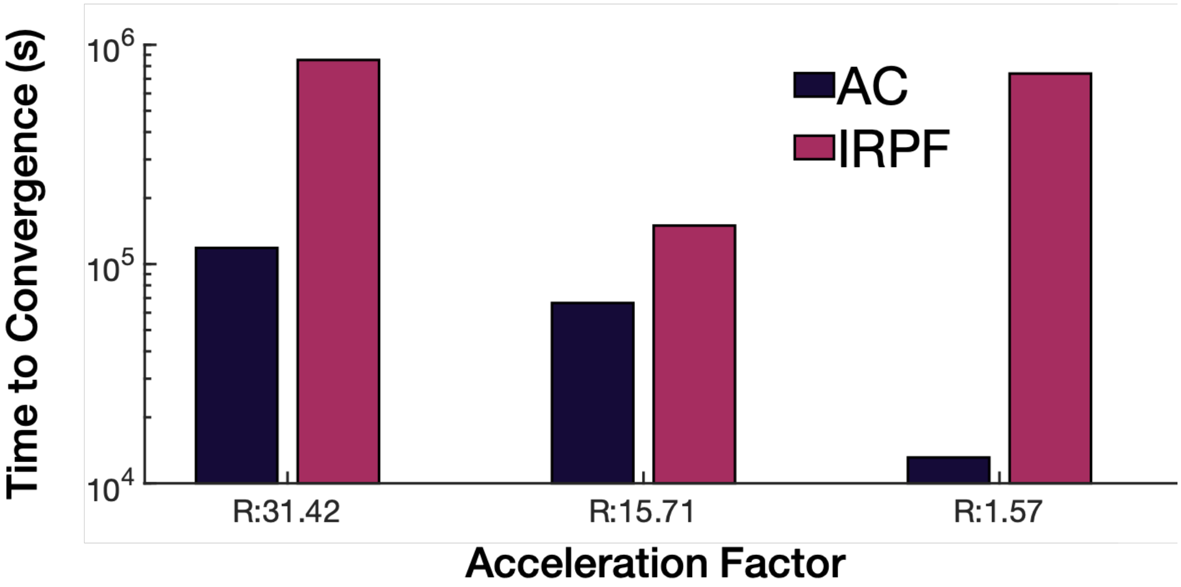


Figure 4.8: The time taken (in seconds) to converge for the different implementation methods at different acceleration factors. In all cases, IRPF took the longest time to converge.

Optimal Iteration Number

In order to select a number of internal iterations to use for each subproblem, it was found that the system converged faster when more iterations occurred in the spatial domain than the temporal domain. A value of 10× more iterations was chosen. With that relationship set, figure 4.9 shows the relationship between the number of iterations in the temporal subproblem, and the X CCS, T CCS, cost function, and time taken for three different acceleration factors over different cycle numbers. A value of 50 temporal subproblem iterations (and 500 spatial subproblem iterations) was chosen to represent a value at which the reconstructions would find a rough optimum each full cycle, in terms of CCS and cost. An increased number of iterations did not yield any additional benefit, and may have led to overfitting (e.g. more precisely fitting the current spatial estimate X_i to the imperfect current temporal estimate T_i and vice versa, which can be seen in the first few cycles of the 100 iteration results). A smaller number of iterations started to yield equivalent results over more cycles, and this could be an area in which the reconstruction speed is increased in future implementations. The iteration parameter optimisation was carried out with an AC formulation.

Convergence & Initialisation

For convergence, the normalised cost function gradient (equation 4.9) was used to determine the stopping point through some predetermined threshold ϵ. This section evaluates three different

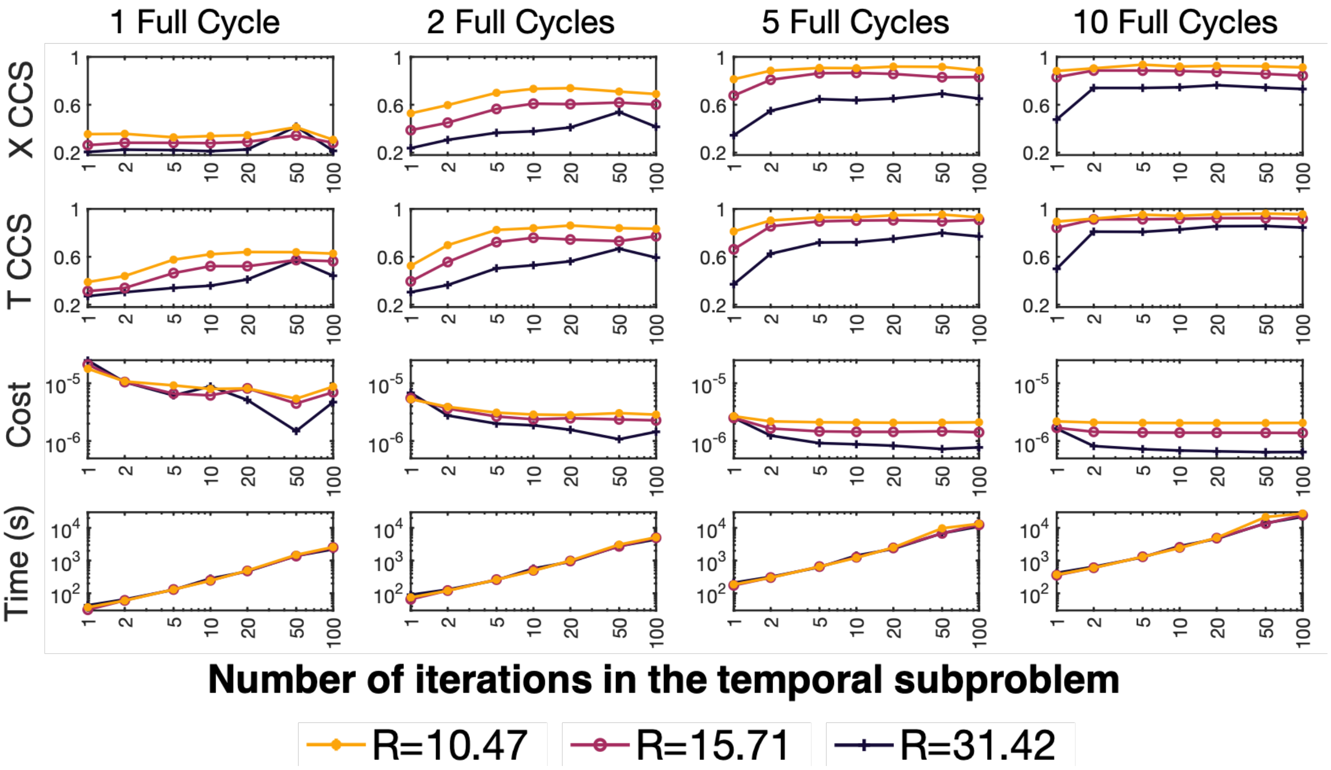


Figure 4.9: Various criteria (Row 1: X CCS, Row 2: T CCS, Row 3: Cost, Row 4: Time) are used to judge the reconstruction performance with varying iteration numbers in the subproblems. Three different acceleration factors are shown (R=31.42, 5 blades/frame; R=15.71, 10 blades/frame; R=10.47, 15 blades/frame) across a range of cycles (shown in each column). The number of iterations in the spatial subproblem was 10× higher.

ϵ thresholds for the initialisation approaches outlined in section 4.3.2. For each initialisation approach, five different instantiations were reconstructed at five different acceleration factors. The spatial CCS between each final reconstructed subspace were then compared, to determine how closely aligned the final reconstructions were despite the different initial starting points. A robust initialisation will lead to matrices that map the same subspace, with the ϵ parameter used to determine the degree of alignment that could sufficiently be labeled as converged.

The acceleration factors chosen were: R=31.42 (5 blades/frame), R=39.27 (4 blades/frame), R=52.36 (3 blades/frame), R=78.54 (2 blades/frame), and R=157.08 (1 blade/frame). The choice of highly unrealistic acceleration factors was made to determine the breaking point of acceleration for this clean retrospective dataset. As a general observation, the temporal subspaces were always more closely aligned than the spatial subspaces, while demonstrating similar distributions, and so the spatial CCS scores alone are considered representative of the worst-case scenario.

Figure 4.10 shows the effect of purely random initialisation, where the numbers in the initialised component matrices were generated by a normalised Gaussian distribution (MATLAB's randn.m function). Even at the lowest acceleration factor (R=15.71), the final matrices were only approximately aligned with the final reconstructions from other randomly initialised matrices, with 0.39 being the peak spatial CCS alignment for any reconstruction. While the reconstruction alignment improved with a more conservative conversion criterion and subsequent longer runtime, the improvement was notably slight.

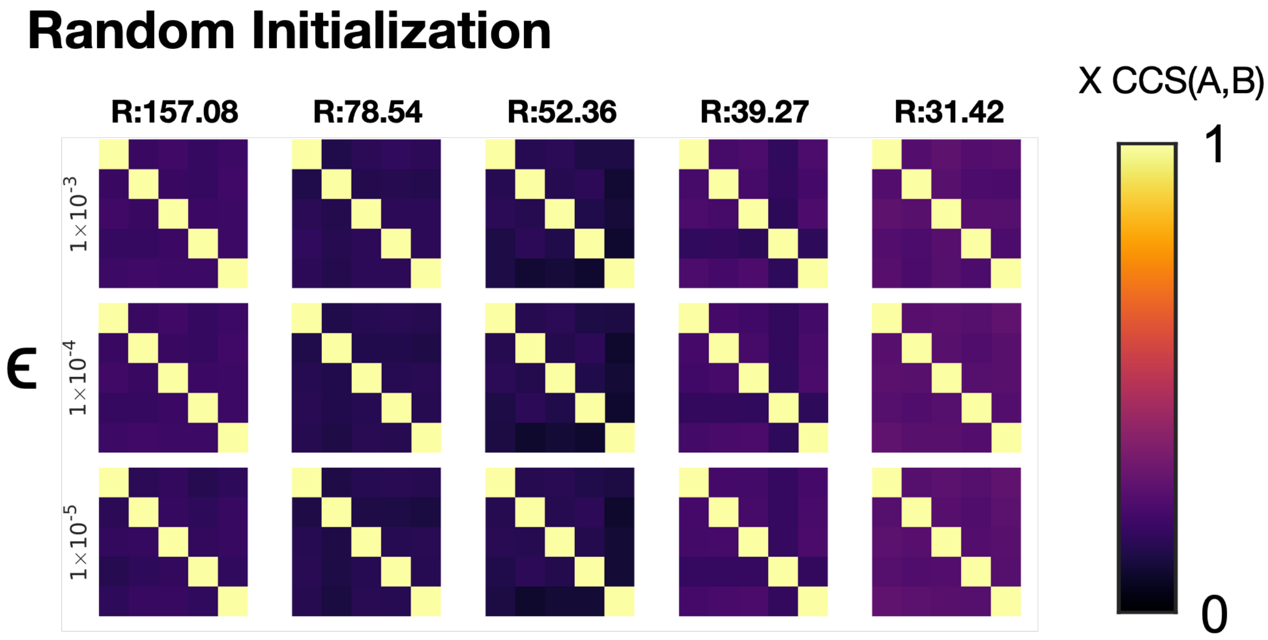


Figure 4.10: A demonstration of the final subspace reconstruction for different instantiations of random, normally-distributed numbers. Each mini grid compares five different reconstructions.

The issue of the final matrices being unaligned in figure 4.10 appears to be solved simply by orthogonalising the components during initialisation. Practically, this was tested by applying an orthogonalisation algorithm to randomised normally distributed entries, leading to the initialisation used for figure 4.11). Although the non-orthogonalised matrices initialised in figure 4.10 were full rank, the final solutions (also full rank) were noticeably less orthogonal (i.e. had higher inner products between components) than solutions from an initialisation with orthogonalised components. The conclusion can be drawn that any initialisation should have orthogonal components. Further testing confirmed that it is specifically the component matrix which is fixed first in the AM optimisation process that should contain orthogonal components. In this formulation, it is the T matrix that must contain orthogonal components. No other factor seems to have much impact on convergence. For example, an initialisation with the temporal mean is shown in supplementary figure 4.17, and shows a similar distribution to figure 4.11 despite having a first component which is already closely aligned to the ground truth. It is worth noting that the alignment is slightly higher in the high acceleration reconstructions with the temporal mean initialisation (R=157.08, R=78.54), but that the cut-off point (R=52.36, 3 blades/frame) remains the same.

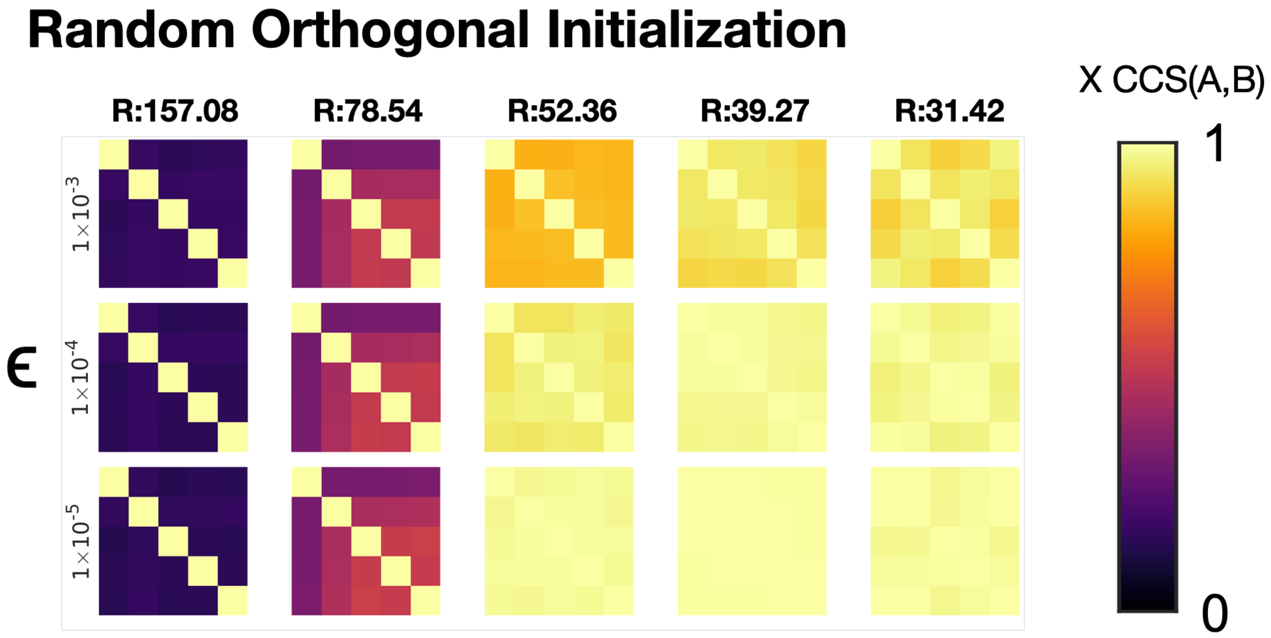


Figure 4.11: A demonstration of the final subspace reconstruction for different instantiations of random normally-distributed components which were orthogonalised before the reconstruction. Each mini grid compares five different reconstructions.

The aforementioned figure 4.11 can also be used to determine the maximum R and candidates for ϵ (ϵ is defined by equation 4.9). Full alignment of the final subspaces can be reached for approximately R≤50 in this particular retrospective dataset, regardless of the form of orthogonalised random initialisation. The minimum CCS for each ϵ value at R=52.36 were: 0.82 ($\epsilon = 1 \times 10^{-3}$), 0.93 ($\epsilon = 1 \times 10^{-4}$) and 0.98 ($\epsilon = 1 \times 10^{-5}$). R=52.36 appears to be the limit that a rank-16 coil-constrained radial-trajectory reconstruction will consistently cover the same subspace, with the reduced DoF in a low-rank matrix likely a limiting factor. This limit occurs at a slightly higher acceleration factor than might be expected from the reduction in DoF achievable from the rank constraint alone (equation 3.24). Reconstructing a 10000×300 matrix at rank 16 with a radial trajectory should lead to a hypothetical maximum acceleration of R=28.64, approximately 5.5 blades/frame (equation 2.26, $R = \frac{10000\times300\times\pi/2}{16\times(10000+300-16)}$). The addition of coils appears to increase the theoretical capacity for acceleration by a factor of roughly 2, which is a common acceleration factor in parallel imaging.

All reconstructions initialised with orthogonal components were able to produce similar graphs to figure 4.11. Using a temporal mean as the first component sped up reconstructions slightly, and so is recommended. A value of $\epsilon = 1 \times 10^{-5}$ will be used for all future reconstructions of the retrospective dataset, although ϵ may be varied depending on the dataset.

AM Parameter Summary

To summarise, the AM reconstruction method was found to be quicker than the IRPF method, while providing reconstructions of equivalent quality. A convergence threshold of $\epsilon = 1 \times 10^{-5}$ was found to guarantee a high level of subspace alignment from the created component matrices, provided the initial temporal component matrix had components that were orthogonal. Orthogonalised randomised inputs were demonstrated to show robustness, which does not discount the use of a more informed initialisation that also contains orthogonal components. An internal iteration number of 50 was chosen for the temporal subproblem and 500 for the spatial subproblem, although the iteration numbers seem to merely determine the speed of reconstruction if enough full cycles of the optimisation problem occur.

All of the above parameters were defined using the retrospective dataset with coil sensitivity information and a TURBINE trajectory. Because the temporal degrees of freedom do not increase with acceleration in this example, it is very possible that the optimal ratio between spatial iterations and temporal iterations could change in datasets of different dimensionality (although figure 4.9 showed that this factor is likely to affect the speed of convergence, but not the fact of convergence). Additionally, in a clean dataset with no chance of overfitting, it is possible that the convergence threshold of $\epsilon = 1 \times 10^{-5}$ may be subject to adjustment in other datasets too. Finally, it is worth noting that changes in dataset parameters (spatial resolution, number of time points, number of active voxels, SNR, etc.) could potentially change the optimal parameters with which to run the AM k-t FASTER formulation.

None of these reconstruction parameters take into account the effect of constraints, which will be covered in chapters 5 and 6. If a prior is also orthogonalised then the prior can be used as the subspace initialisation. These reconstruction parameters will be kept mostly identical for the constrained-subspace reconstructions, and any changes to these reconstruction parameters will be clearly stated.

4.5.2 AM k-t FASTER vs SVD k-t FASTER

Now that optimal parameters for AM k-t FASTER have been determined, a comparison between AM k-t FASTER and SVD k-t FASTER can be carried out. No significant difference is expected in terms of reconstruction performance between the two methods, although the form of the solutions are expected to differ. SVD k-t FASTER more strictly enforces orthogonality and the relative strength of components, whereas AM k-t FASTER contains no such restrictions other than a maximum rank of the solution. All tests were done on the retrospective dataset, with processing techniques outlined in section 4.4. SVD k-t FASTER used an IHT+MS approach with the recommended parameter set from [2]: a shrinkage coefficient of 0.5, a step size of 1, a convergence threshold of 1×10^{-4}, and the convergence criterion defined by equation 4.7. These parameters were originally derived heuristically for whole volume reconstruction, but are not expected to be affected by use in a single slice. The Cartesian single-coil IHT+MS formulation [1] follows algorithm 1, and the non-Cartesian coil-constrained IHT+MS formulation [2] follows algorithm 2.

4.5.2.1 Cartesian

The first test assumed a single uniform coil, and used a set of three Cartesian trajectories to ensure that AM k-t FASTER can match the Cartesian reconstruction power demonstrated by SVD k-t FASTER in the first paper [1]. Figure 4.12 shows a cross-section of the three trajectories across each slice for all frames. Figure 4.12a shows the moderately undersampled trajectory (R=5.0, 20 lines/frame), figure 4.12b is a slightly undersampled trajectory (R=3.1, 32 line/frame), and figure 4.12c is fully sampled (R=1.0, 100 lines/frame). The ten central lines were fully sampled for all trajectories, with the remaining lines pseudo-randomly distributed in each frame for the undersampled reconstructions. The pseudo-random sampling was chosen to establish a basic performance comparison, and is not practically used for fMRI acquisition.

Some reconstruction parameter changes were needed compared to the non-Cartesian multicoil reconstruction used in section 4.5.1, the justification for these changes is discussed with the other results. The Legendre polynomials were used as the initial temporal component set, while the spatial component matrix was zero-filled. In addition, 200 spatial iterations were used (rather than 500). Unlike in the previous examples (figures 4.9 and 4.11, coil-constrained radial trajectories), a specific orthogonal initialisation (in the form of the Legendre polynomials) and a specific iteration number were found to be required for convergence for the single-coil Cartesian case.

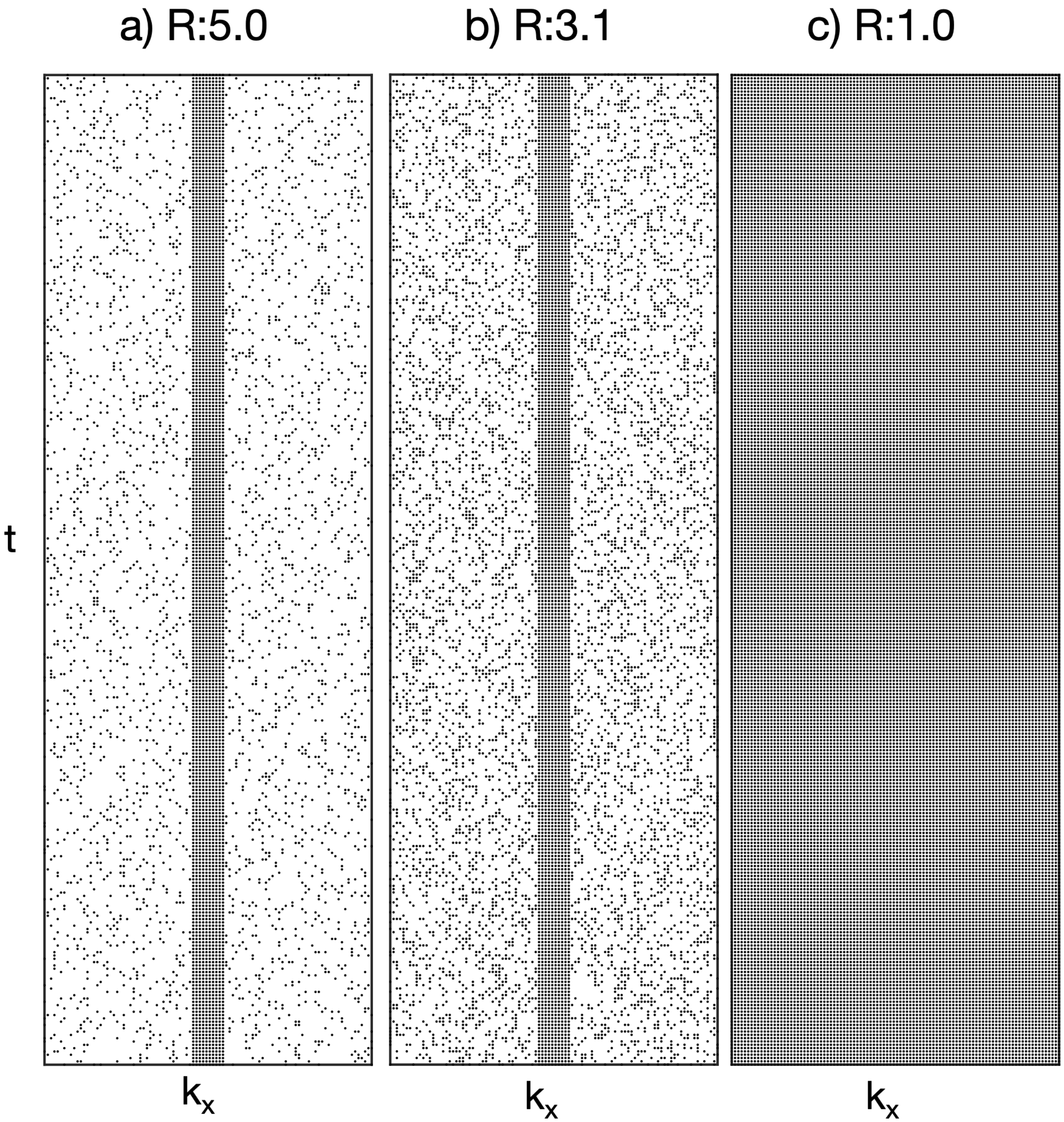


Figure 4.12: The three cartesian trajectories, where each point represents a line in k_y. Trajectory a) contains 20 lines/frame (including the 10 central lines, with 5 lines on each half of the remaining k-space), trajectory b) contains 32 lines/frame (including the 10 central lines and 11 lines on each half of the remaining k-space), trajectory c) is fully sampled and contains 100 lines/frame.

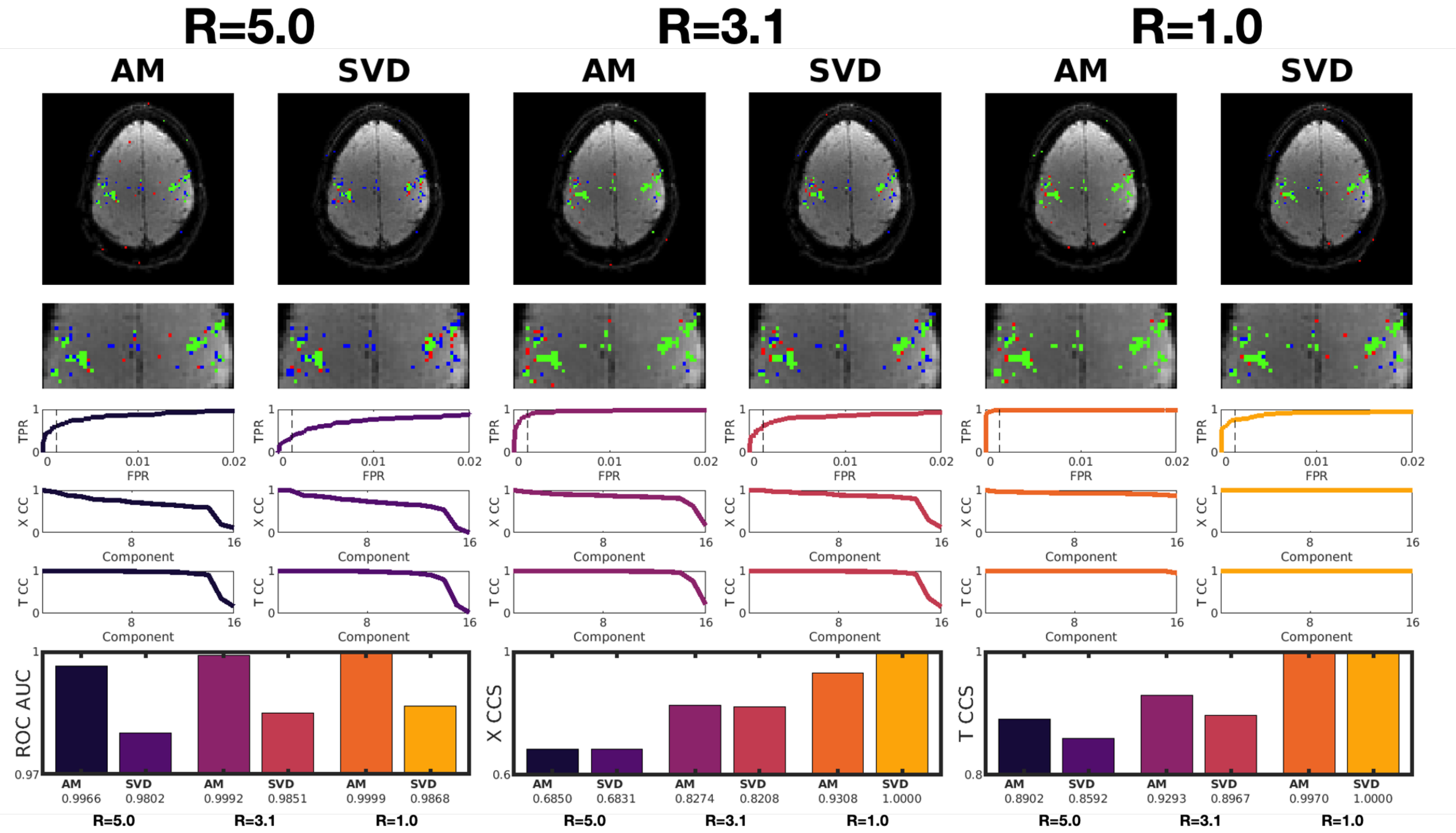


Figure 4.13: A comparison figure for AM and SVD k-t FASTER at three different acceleration factors for the Cartesian trajectories of figure 4.12. Sample z-stat parameter maps, an ROC curve, and the CC curves are shown in a columnwise structure for each reconstruction. Summative bar graphs for all reconstructions are shown on the bottom row. Figure 4.5 contains a detailed schematic of the graph format.

The results of reconstruction with Cartesian trajectories are shown in figure 4.13 and demonstrate that both AM and SVD k-t FASTER are capable of achieving accurate reconstruction at reasonable acceleration factors in this retrospective example. At R=1.0, SVD k-t FASTER has perfect CCS scores (i.e. the reconstructed component matrices are exactly aligned with the decomposed ground truth), but an imperfect ROC AUC score. The most likely factor for this is that the SVD k-t FASTER approach regulates the singular value distribution (via shrinkage relative to the largest removed value), a factor that would not affect the scale-independent CCS scores but which may affect the functional information. AM k-t FASTER does not impose any limitations on the singular value distribution, and so the distribution can more closely resemble the underlying truth (see figure 4.14). The SVD k-t FASTER singular value distribution is designed to exclude noisy components, which, in the case of fully sampled clean data, appears to negatively affect some aspects of reconstruction. In addition, while both nonlinear reconstruction approaches potentially limit the DoF available for robust conventional statistics by filtering out low-variance components, this factor may be exacerbated by SVD k-t FASTER's explicit shrinkage. Although mixture modelling helps this problem (e.g. by fitting a null distribution to the available residuals), no perfect solution for robust statistical inference of nonlinearly reconstructed data currently exists in fMRI. All SVD reconstructions only reach a TPR of 1 at FPR>0.75, meaning the ROC AUC does not hold as a comparative metric here. However, a qualitative look at the z-stat parameter maps shows that there is not a huge loss in reconstruction fidelity between the two methods at these acceleration factors, although there is still some noticeable improvement with the AM approach.

At increased acceleration both methods still achieve reasonable results, although AM k-t FASTER outperforms SVD k-t FASTER across all metrics in undersampled reconstructions. The T CCS scores are still very high for all reconstructions, indicating that the temporal information is mostly preserved in both methods. As R increases, the spatial alignment metrics indicate a worse performance, with the X CC curves of figure 4.13 indicating increased misalignment for nearly all components. In the R=3.1 reconstructions, the z-stat maps are still highly similar, with the bad ROC AUC score likely due to the aforementioned behaviour of the curve at high FPR. At R=5.0 the maps are noticeably worse for the SVD reconstruction, showing a failure to capture the signal in the supplementary motor area of the brain. This indicates that R=5.0 is likely close to a limit with this specific set-up in SVD k-t FASTER (and is above the reported value in the literature [1]), although AM k-t FASTER is still performing well. It should be noted that the AM k-t FASTER parameters were optimised on this dataset, which may also be a factor leading to the relative differences in performance.

The AM formulation was far more dependent on parameter selection in this instance than the SVD approach. In the specific Cartesian formulation for single-coil reconstruction, the AM approach was very dependent on the form of the initialisation (leading to the use of Legendre polynomials)

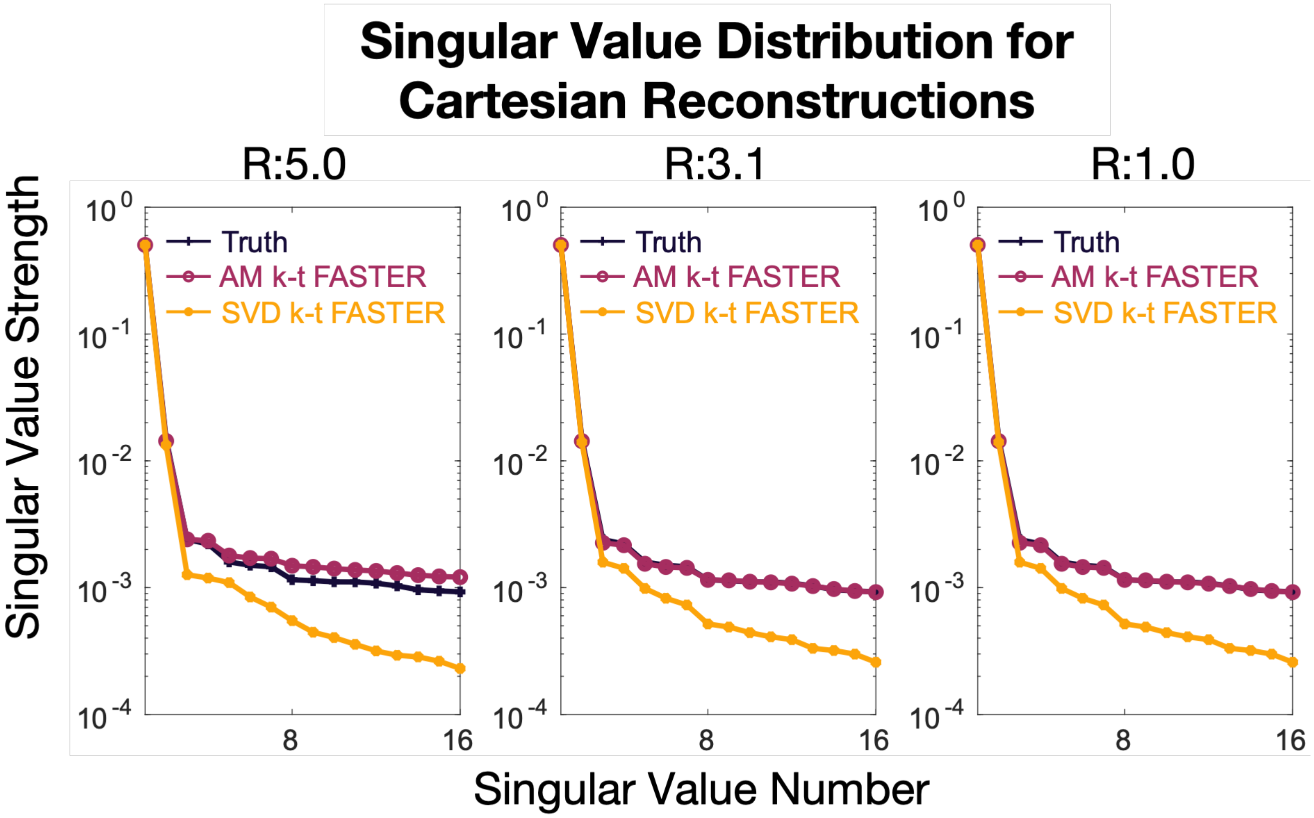


Figure 4.14: A comparison figure for the singular value distributions of the AM and SVD k-t FASTER at three different acceleration factors, along with the truncated distribution of the ground truth.

and the internal number of spatial iterations. A non-Cartesian single uniform coil reconstruction did not show this lack of robustness, nor did multicoil data sampled with a Cartesian trajectory. An additional test of Cartesian single-coil reconstructions with Tikhonov regularisation also proved robust to parameter selection. The general conclusion is thus - with any form of regularisation (or with a sampling trajectory which causes incoherent artefacts), the conclusions drawn above about parameter selection hold. The corollary to this conclusion is that AM k-t FASTER is not robust if no additional constraints are imposed. The specific set of parameters chosen for the Cartesian AM reconstructions were found experimentally, as prevention of over/underfitting had to be done through a number of calculations and a consistent starting point, rather than being a feature of the algorithm. All subsequent analysis in this thesis will use radial sampling, and much of it will incorporate coil sensitivity maps, satisfying the conditions for robust AM optimisation.

With the caveats about parameter selection aside, the general conclusion is clear - both the SVD and AM k-t FASTER approaches can work on Cartesian data. In addition, the AM k-t FASTER may have more potential at higher acceleration factors.

4.5.2.2 Non-Cartesian

Further to the Cartesian trajectories, some accelerated reconstructions were carried out for AM k-t FASTER and SVD k-t FASTER on a TURBINE reconstruction with Golden Ratio angle radial blades and multicoil acquisition. Acceleration factors of R=15.71, R=31.42 and R=52.36 were evaluated in figure 4.15. The reconstruction parameters found in section 4.5.1 were used for AM k-t FASTER, and the SVD k-t FASTER parameters were the same as in the Cartesian case.

The results of the non-Cartesian reconstructions can be seen in figure 4.15, and the spatial and temporal subspace similarities between reconstructions can be seen in figure 4.16. Collectively, the results show that the SVD and AM methods both produce similar low-rank solutions to the undersampled problem, but that those solutions take different forms.

Figure 4.15 compares the reconstructions against a ground truth. The CCS results collectively indicate that the AM approach forms component matrices which align closer to the subspace mapped by the ground truth. It should be noted that, while the ROC AUC scores for SVD k-t FASTER are broadly higher for the non-Cartesian coil-constrained reconstructions (figure 4.15) than for the Cartesian single-coil reconstructions at a lower acceleration factor (figure 4.13), the shape of the ROC curve at low FPR show a more accurate and subtle picture of the relative improvement. Despite this, the improved capacity for acceleration with coils and radial trajectories motivates the use of coil-constrained radial approaches for future reconstructions.

The comparison between the reconstructions is also worth noting. In figure 4.16, the alignment between the AM and SVD spatial subspaces within an acceleration factor is approximately equal to the alignment of each reconstructed spatial subspace with the ground truth - i.e. the reconstructed subspace not aligned with the truth is also not aligned with the other reconstruction. The temporal subspaces appear to be more aligned than the spatial subspaces in figure 4.16, but it is still far from a clear overlap. Contrast this graph with figure 4.7, in which the IRPF and AC approaches to AM optimisation were both more aligned with each other than the truth, because they mapped the exact same solution to the problem.

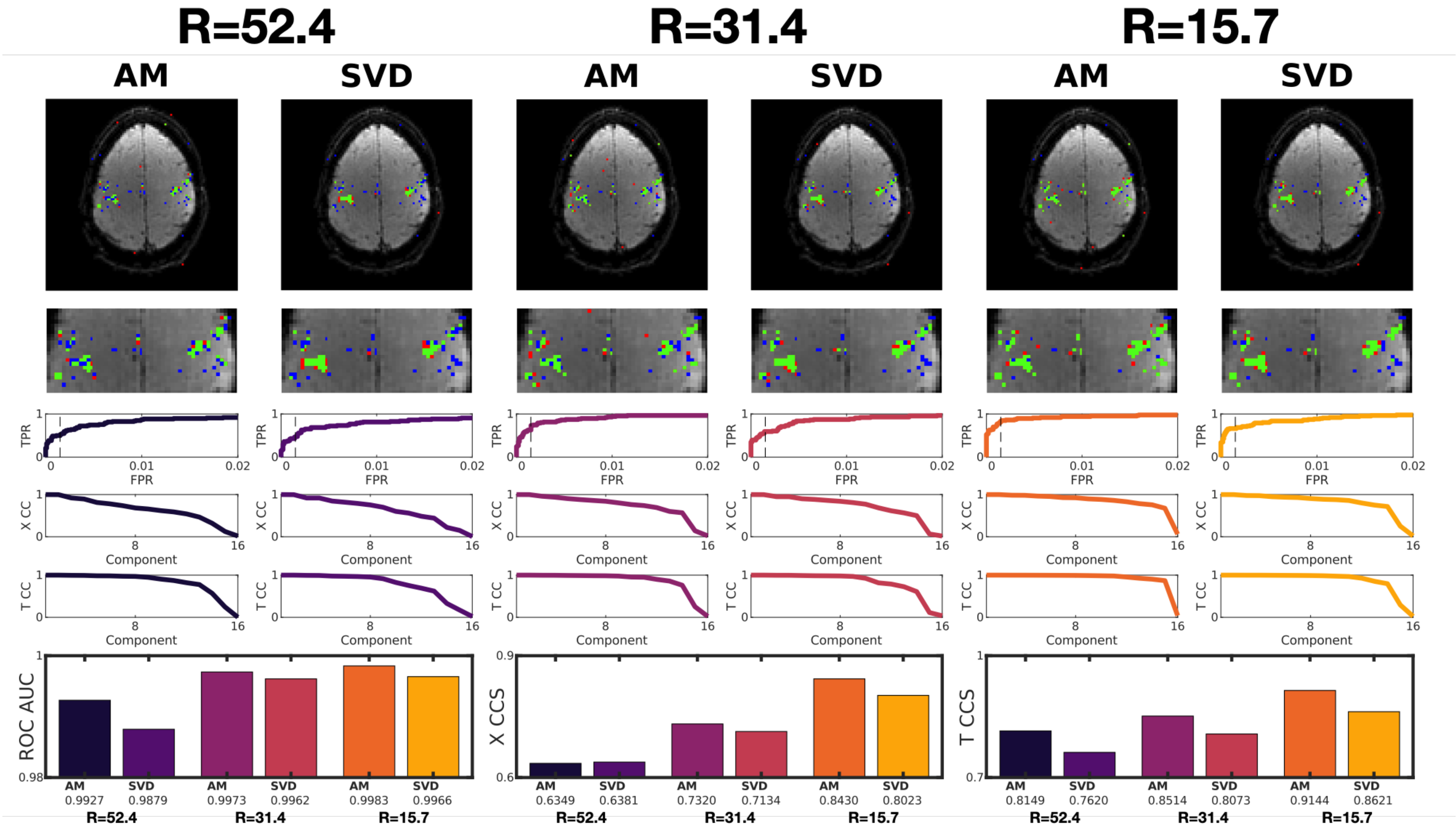


Figure 4.15: A comparison figure for AM and SVD k-t FASTER at 3 different acceleration methods for non-Cartesian TURBINE trajectories. Sample z-stat parameter maps, an ROC curve, and the CC curves are shown in a columnwise structure for each reconstruction. Summative bar graphs for all reconstructions are shown on the bottom row. Figure 4.5 contains a detailed schematic of the graph format.

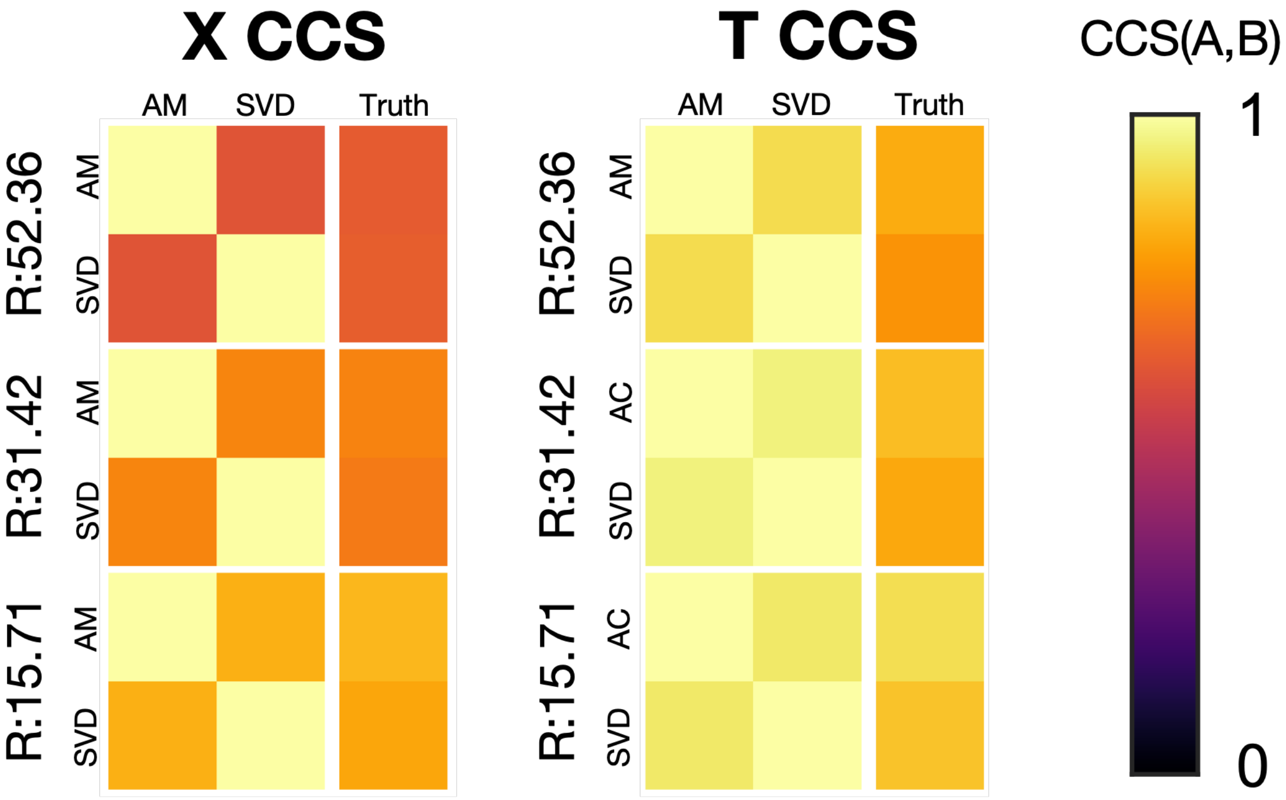


Figure 4.16: The Canonical Correlation Scores between different reconstructed subspaces of AM and SVD approaches to optimisation. As CCS(A,B)=CCS(B,A), the individual grids are symmetric along the downward diagonal. A CCS score of 1 occurs for CCS(A,A) to indicate total alignment. The right-hand column shows the CCS comparing the reconstructions and the first sixteen components of an SVD of the true retrospective dataset.

What must be taken into account is that these results were calculated on a single dataset as a sanity check to ensure that both implementations of k-t FASTER (as low-rank enforcements of fMRI data) were performing reconstructions with equal predictive power. As such, small variations in results between the two methods are likely not significant, especially as all of the AM parameters were all optimised on this dataset. AM and SVD both enforce the same low-rank constraint, but do it in very different ways, with SVD k-t FASTER having a very strict form for the singular values. Conversely, SVD k-t FASTER was quicker than AM k-t FASTER to reconstruct the dataset, achieving reconstruction times roughly 50× faster across high acceleration factors. However, the poor computational scaling of an SVD with dataset size is worth noting when considering the time penalty, especially considering that the retrospective dataset was a single slice with only 300 frames.

4.6 Conclusion

AM k-t FASTER achieves roughly consistent results to SVD k-t FASTER, if not slightly improved through the metrics chosen. However, the AM parameters were chosen for this dataset, and a similar SVD k-t FASTER parameter optimisation could improve that method's performance. Broadly though, any improvement from further development of AM k-t FASTER by incorporation of constraints should also demonstrate an improvement over SVD k-t FASTER.

The parameters in this chapter were chosen to guarantee convergence and prove the efficacy of the AM method, rather than necessarily for optimal computing time. With this dataset, the following parameters were found to be optimal: the AC implementation, $\epsilon = 1 \times 10^{-5}$ for convergence, orthogonalised components in initialisation, and 50 iterations per temporal subproblem.

The real advantage of AM k-t FASTER is the ease with which additional regularisation can be incorporated into the reconstruction. These parameters were selected by optimising AM k-t FASTER without the presence of any additional constraints, a factor which may lead to further time saving by reducing the necessary number of sub-problem iterations or the convergence criterion ϵ.

4.A Supplementary Figures

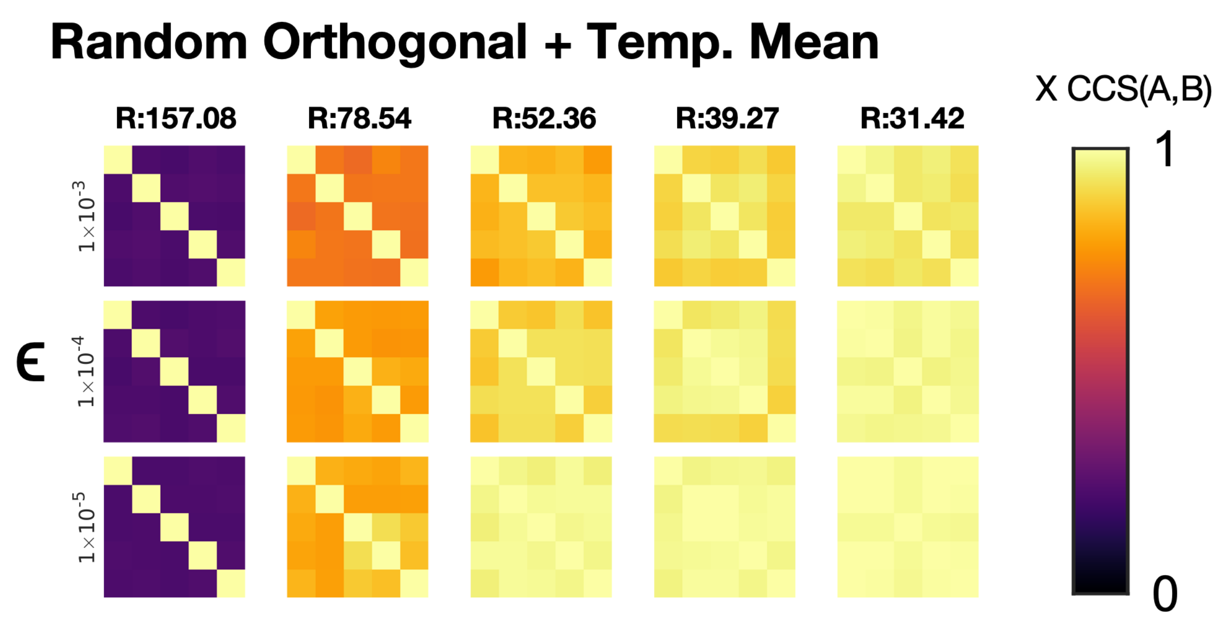


Figure 4.17: Subspace reconstruction for different instantiations of random orthogonalised normally-distributed components, with the temporal mean as the first spatial component. This graph shows that initialising with the temporal mean does not affect the robustness of the reconstruction.

Bibliography

[1] Mark Chiew, Stephen M. Smith, Peter J. Koopmans, Nadine N. Graedel, Thomas Blumensath, and Karla L. Miller. k-t FASTER: Acceleration of functional MRI data acquisition using low rank constraints. *Magnetic Resonance in Medicine*, 74(2):353–364, 2015.

[2] Mark Chiew, Nadine N. Graedel, Jennifer A. Mcnab, Stephen M. Smith, and Karla L. Miller. Accelerating functional MRI using fixed-rank approximations and radial-cartesian sampling. *Magnetic Resonance in Medicine*, 00:1–12, 2016.

[3] Nadine N. Graedel, Mark Chiew, Jennifer A. McNab, and Karla L. Miller. FMRI using a 3D radial-Cartesian trajectory: spatio-temporal tunability and artefact correction. In *Proceedings of the International Society for Magnetic Resonance in Medicine 23*, volume 23, page 1, 2015.

[4] Klaas P. Pruessmann, Markus Weiger, Peter Bornet, and Peter Boesiger. Advances in Sensitivity Encoding With Arbitrary k-Space Trajectories. *Magnetic Resonance in Medicine*, 651(46):638–651, 2001.

[5] Zhi-Pei Liang, Fernando E. Boada, R. Todd Constable, E. Mark Haacke, Paul C. Lauterbur, and Michael R. Smith. Constrained Reconstruction Methods in MR Imaging. *Reviews of Magnetic Resonance in Medicine*, 4(217):67–185, 1992.

[6] Ricardo Otazo, Emmanuel J. Candès, and Daniel K. Sodickson. Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components. *Magnetic Resonance in Medicine*, 73(3):1125–1136, 2015.

[7] Cristina Chavarrías, J. F. P. J. Abascal, P. Montesinos, and M. Desco. Exploitation of temporal redundancy in compressed sensing reconstruction of fMRI studies with a prior-based algorithm (PICCS). *Medical Physics*, 42(7):3814–3821, 2015.

[8] M. Murphy, Marcus T. Alley, J. Demmel, K. Keutzer, Shreyas S. Vasanawala, and Michael Lustig. Fast l-SPIRiT Compressed Sensing Parallel Imaging MRI: Scalable Parallel Implementation and Clinically Feasible Runtime. *IEEE Transactions on Medical Imaging*, 31(6):1250–1262, jun 2012.

[9] Hong Jung, Jong Chul Ye, and Eung Yeop Kim. Improved k-t BLAST and k-t SENSE using FOCUSS. *Physics in Medicine and Biology*, 52(11):3201–3226, 2007.

[10] Mark Chiew, Nadine N. Graedel, and Karla L. Miller. Recovering task fMRI signals from highly under-sampled data with low-rank and temporal subspace constraints. *NeuroImage*, 174:97–110, 2018.

[11] Martin J. McKeown, Tzyy Ping Jung, Scott Makeig, Greg G. Brown, Sandra S. Kindermann, Te Won Lee, and Terrence J. Sejnowski. Spatially independent activity patterns in functional MRI data during the Stroop color-naming task. *Proceedings of the National Academy of Sciences of the United States of America*, 95(3):803–810, 1998.

[12] Emmanuel J. Candès and Yaniv Plan. Matrix completion with noise. *Proceedings of the IEEE*, 98(6):925–936, 2010.

[13] Shiqian Ma, Donald Goldfarb, and Lifeng Chen. Fixed point and Bregman iterative methods for matrix rank minimization. *Mathematical Programming*, 128(1-2):321–353, jun 2011.

[14] Sajan Goud Lingala, Yue Hu, Edward Dibella, and Mathews Jacob. Accelerated Dynamic MRI Exploiting Sparsity and Low-Rank Structure: k-t SLR. *IEEE Transactions on Medical Imaging*, 30(5):1042–1054, 2011.

[15] Justin P. Haldar and Diego Hernando. Rank-constrained solutions to linear matrix equations using PowerFactorization. *IEEE Signal Processing Letters*, 16(7):584–587, 2009.

[16] Tao Zhang, John M. Pauly, Shreyas S. Vasanawala, and Michael Lustig. Coil compression for accelerated imaging with Cartesian sampling. *Magnetic Resonance in Medicine*, 69(2):571–582, 2013.

[17] Martin Buehrer, Klaas P. Pruessmann, Peter Boesiger, and Sebastian Kozerke. Array compression for MRI with large coil arrays. *Magnetic Resonance in Medicine*, 57(6):1131–1139, 2007.

[18] Stefanie Winkelmann, Tobias Schaeffter, Thomas Koehler, Holger Eggers, and Olaf Doessel. An optimal radial profile order based on the golden ratio for time-resolved MRI. *IEEE Transactions on Medical Imaging*, 26(1):68–76, 2007.

[19] C. C. Paige and M. A. Saunders. Solution of Sparse Indefinite Systems of Linear Equations. *SIAM Journal on Numerical Analysis*, 12(4):617–629, 1975.

[20] The Mathworks Inc. MATLAB R2019a, 2019.

[21] Stephen M. Smith, Mark Jenkinson, Mark W. Woolrich, Christian F. Beckmann, Timothy E. J. Behrens, Heidi Johansen-Berg, Peter R. Bannister, Marilena De Luca, Ivana Drobnjak, David E. Flitney, Rami K. Niazy, James Saunders, John Vickers, Yongyue Zhang, Nicola De Stefano, J. Michael Brady, and Paul M. Matthews. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23(SUPPL. 1):208–219, 2004.

[22] Christian F. Beckmann and Stephen M. Smith. Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging. *IEEE Transactions on Medical Imaging*, 23(2):137–152, 2004.

[23] Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, and Eero P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

[24] Yang Liu and Jeffrey Heer. Somewhere over the rainbow: An empirical assessment of quantitative colormaps. *Conference on Human Factors in Computing Systems - Proceedings*, April:1–12, 2018.

[25] Ander Biguri. Perceptually Uniform Colormaps, 2019.

5 | Exploring Constrained Low-Rank Reconstruction

Contents

5.1 Introduction

In chapter 4, k-t FASTER [1] (an algorithm designed for reconstructing fMRI images using low-rank constraints) was reformulated as an Alternating Minimisation (AM) algorithm. This new AM formulation of fMRI separated the spatial (X) and temporal (T) subspaces in the reconstruction, optimising over one subspace while fixing the other. A simple schematic of AM k-t FASTER is shown in figure 5.1.

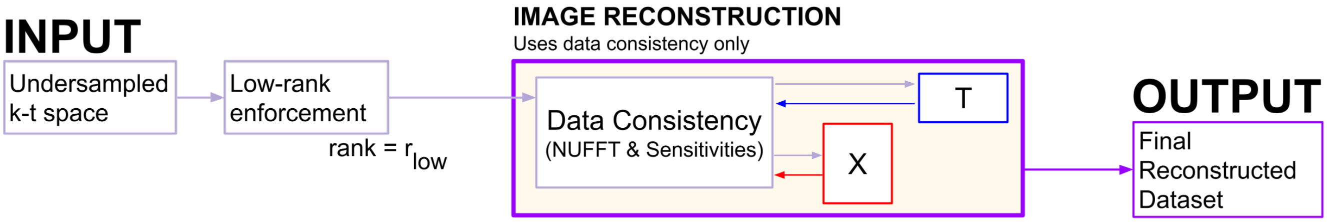


Figure 5.1: An unconstrained-subspace AM k-t FASTER schematic. A fixed low-rank reinforcement is applied in the dimensionality of X and T, and then X and T are calculated via the optimisation problem in equation 5.1, which comprises of a single term (data consistency) which uses updated versions of X when calculating T, and vice versa.

One major advantage of this new AM formulation is the ease with which subspace-specific constraints can be incorporated into a low-rank framework. k-t FASTER was already shown to be improved by the addition of coil sensitivity maps [2], and section 3.5 laid out many existing methods that exploit more than one type of constraint when reconstructing fMRI data. Within this chapter the following constraints will be considered:

- Tikhonov (minimise the energy within each component matrix)
- Low-Resolution Priors (LRP) (which exploit variable sampling density information)
- Temporal Subspace Smoothing (minimise the gradients within the temporal components)

The subspace constraints are incorporated into the reconstruction through regularisation. Alternate implementations of the Tikhonov and LRP constraints will also be explored, in order to better understand the impact of those constraints on each subspace. This chapter will walk through the motivation, formulation, and theory of each set of constraints, demonstrating the formulation through reconstruction on an artificial dataset, and noting any peculiarities encountered either during the reconstruction process or the analysis. First, however, the generically constrained cost function will be constructed in order to unify the notation.

To understand how the additional constraints are formulated, we will begin with the AM k-t FASTER reconstruction cost function with no subspace constraints.

$$\hat{X}, \hat{T} = argmin_{X,T}\Bigg(\|EXT' - d\|_2^2\Bigg) \tag{5.1}$$

E is a non-Cartesian multicoil encoding operator, X is the spatial component matrix, T is the temporal component matrix, and d is the multicoil data sampled in k-t space. The main advantage of the AM formulation is the ability to expand equation 5.1 to incorporate subspace-regularisation constraints. The constraints can be applied to either X and T, with each constraint requiring a separate weighting term. Initially, no assumption will be made about either the number or form of the subspace constraints, in order to demonstrate the full potential of the AM formulation developed in chapter 4. The generically-constrained AM k-t FASTER reconstruction can therefore be stated as follows:

$$argmin_{X,T}\Bigg(\|EXT' - d\|_2^2 + \sum_u \lambda_{x_u}\|\Psi_{x_u}(X - X_{prior_u})\|_{p_u}^{p_u} + \sum_v \lambda_{t_v}\|\Psi_{t_v}(T - T_{prior_v})\|_{p_v}^{p_v}\Bigg) \tag{5.2}$$

In equation 5.2, constraints are added to find more favourable solutions in the temporal and spatial domains. The operators Ψ_{x_u} and Ψ_{t_v} represent transforms into different domains. p_u and p_v represent the norms of the spatial and temporal constraints (L2 norms are typically squared and L1 norms are raised to the power of 1).

From this point on, any reconstruction with additional subspace constraints (i.e. with at least one non-zero value of λ_{x_u} or λ_{t_v}) will be referred to as “constrained-subspace” optimisation, and the AM formulation of k-t FASTER with no additional constraints (equation 5.1) will be referred to as “unconstrained-subspace” optimisation. Both unconstrained-subspace (all $\lambda = 0$) and constrained-subspace formulations contain fixed low-rank constraints enforced through the dimensionality of X and T, and can contain coil sensitivity information in operator E. If the term k-t FASTER is used for brevity, it will refer to the unconstrained-subspace formulation.

For this thesis, only a limited range of possible interpretations of equation 5.2 will be considered. Ψ_u will always be the identity matrix. Ψ_v will represent either the identity matrix, or the differential operator ∇ (see section 5.6). The differential operator on the temporal subspace may also work concurrently with terms using a $\Psi_v = I$ operator, and so it will be represented as a distinct constraint. Here, we will only consider L2 norms, although sparsity L1 norms could fit naturally into this framework to form either a CS or an L+S approach.

While equation 5.2 represents a wide range of possible constrained-subspace optimisations, a more useful general formulation going forward for this chapter will define the clear nomenclature for the constraints that will be covered in the following sections. The more specific formulation covering the range of constraints considered in the rest of this thesis is then:

$$\hat{X}, \hat{T} = argmin_{X,T}\Big(\|EXT' - d\|_2^2 + \lambda_X\|X - X_{prior}\|_2^2 + \lambda_T\|T - T_{prior}\|_2^2 + \lambda_{\nabla}\|\nabla T\|_2^2\Big) \tag{5.3}$$

Where λ_{∇} is the temporal smoothing-specific weighting. Different values of λ_X, λ_T, λ_{∇}, X_{prior}, and T_{prior} define the specific constrained-subspace approach that is being explored. For the constrained-subspace reconstructions, optimal λ values were found using both the AUC/CCS scores and a qualitative consideration of the z-statistic maps. Chapter 6 will walk through the process of finding optimal λ values. The results in this chapter will only focus on the optimal constrained-subspace reconstructions, in order to demonstrate proof-of-concept. The constrained-subspace reconstruction is laid out in pseudocode algorithm 5.

Algorithm 5 Constrained-Subspace Pseudocode

```
Parameters
  $d$: Multicoil undersampled k-t fMRI data
  $E$: Non-Uniform Fast Fourier Transform (NUFFT) and multi-coil encoding operator
  $\lambda_X$: Spatial regularisation weighting factor
  $\lambda_T$: Temporal regularisation weighting factor
  $\lambda_\nabla$: Temporal smoothing weighting factor

Constructing Constraints
if creating LRP constraints then
    $X_{prior0\ col\ 1}$ = Temporal mean (The average image over all frames)
    $X_{prior0\ cols\ 2:r}$ = 0
    $T_{prior0}$ = Random orthogonal rows

    $d_{win} = window(d)$
    while not converged do
        $X_{prior\ i+1} = argmin_{X_{prior}}\left(\|EX_{prior}T'_{priori} - d_{win}\|_2^2\right)$
        $T_{prior\ i+1} = argmin_{T_{prior}}\left(\|EX_{priori+1}T'_{prior} - d_{win}\|_2^2\right)$
    $X_0 = X_{prior}$
    $T_0 = T_{prior}$

else if not creating LRP constraints then
    $X_{prior} = 0$
    $T_{prior} = 0$

    $X_{0\ col\ 1}$ = Temporal mean (The average image over all frames)
    $X_{0\ cols\ 2:r}$ = 0
    $T_0$ = Random orthogonal rows

Final Reconstruction
while not converged do
    $X_{i+1} = argmin_X\left(\|EXT'_i - d\|_2^2 + \lambda_X\|X - X_{prior}\|_2^2\right)$
    $T_{i+1} = argmin_T\left(\|EX_{i+1}T' - d\|_2^2 + \lambda_T\|T - T_{prior}\|_2^2\right)$

Output
$D = XT'$: Final reconstructed x-t fMRI data
```

5.2 Unconstrained-Subspace Reconstruction

If λ_X, λ_T, and λ_∇ are all set to zero in equation 5.3, the system reduces to the unconstrained-subspace AM k-t FASTER formulation from chapter 4. Unconstrained-subspace reconstructions are only constrained by the dimensionality of X and T (enforcing a fixed rank on the system), and any coil sensitivity information in E (equation 5.1). The processing pipeline shown here to reconstruct the unconstrained-subspace AM k-t FASTER method in an artificial dataset will also be followed for all subsequent constrained-subspace formulations. A set of acceleration factors will be found at which a constrained-subspace method can demonstrate an improved reconstruction over the unconstrained-subspace approach.

5.2.1 FMRIB Dataset Methodology

As outlined at the end of chapter 4, the "All Components" method of reconstructing X and T was found to be the most efficient way to enforce a rank constraint in an AM framework (algorithm 4). The convergence criterion (equation 4.9) is defined by the normalised cost function gradient $|\frac{CF_i - CF_{i-1}}{CF_i}|$ dropping below 1×10^{-5} between optimisation cycles.

Reconstruction in this chapter will be tested on the "FMRIB" dataset for all constraint formulations. The FMRIB dataset is a highly artificial dataset composed of a brain phantom and five unique time-courses over five separate spatial patterns (the pattern representing the letters "F", "M", "R", "I", and "B" are shown in figure 5.2). The FMRIB dataset contains a single 64×64 slice over 300 frames, with no coil information. A unique noise pattern was added to form an SNR of 1000 in k-t space during reconstruction (equation 5.4). This level of noise was chosen to distinguish between the reconstruction power of different methods (n.b. this chapter aims to demonstrate proof of concept - realistic noise levels are investigated in chapter 6). The maximum rank was restricted to 16. Sampling was done with a radial Golden Ratio angle scheme. The SNR was defined in k-space as follows:

$$SNR = \frac{||Signal||_2}{||Noise||_2} \tag{5.4}$$

The preprocessing pipeline for FMRIB dataset is almost identical to the pipeline for the retrospective dataset in section 4.4, but the GLM for the FMRIB dataset takes the five distinct time courses as the basis function in $Y = \beta X$, rather than the convolution of the HRF with a stimulus pattern. As with chapter 4, three key main quantitative metrics will be considered: the spatial and the temporal Canonical Correlation Scores (X CCS and T CCS), and the AUC of the ROC

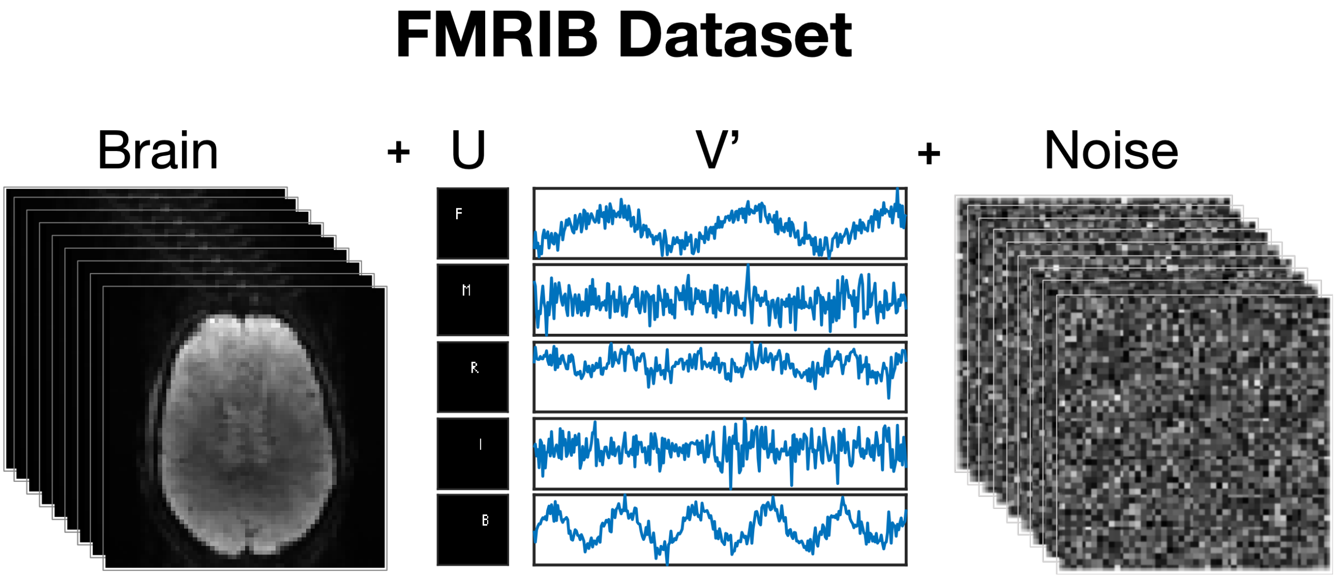


Figure 5.2: A representation of the "FMRIB" dataset. The final dataset can be considered as the combination of a set of spatial components (which mimic the shape of five letters) multiplied by a set of unique time courses, with a mean brain intensity and complex noise.

curve. In addition, the z-statistic map will be shown for qualitative comparison (overlain over the mean brain intensity), along with the ROC and canonical correlation curves.

Unlike with the retrospective dataset of chapter 4, the canonical correlation calculated against a ground truth for the artificial FMRIB dataset can only define a maximum of six principal components (figure 5.2, the static component plus the five distinct spatial patterns/time courses), despite the reconstructions being carried out at rank sixteen. The CCS for this dataset was then calculated by dividing the sum of the principal angle cosines by six to ensure the CCS scale still ranges from $0 \rightarrow 1$, a fact which will make smaller changes in the CCS more significant.

The z-statistic parameter map of the noiseless ground truth was thresholded at $z > 1.8$. The only reason a threshold was needed in this instance was due to the "F" component having a non-zero mean, leading to non-zero co-efficients for the static background. There is no practical difference between a threshold-defined truth and a manual selection of the voxels in the five spatial components - i.e. the letters F, M, R, I, and B - as a mask of significance. For the retrospective dataset, z-statistic parameter maps were shown at a z-statistic corresponding to an FPR value of 0.0015. To account for the smaller relative spatial dimensions of the FMRIB dataset, the reconstructed z-stat maps are shown at the z-statistic corresponding to an FPR value of 0.005, meaning that the plotted FMRIB dataset z-statistic map will always contain 20 false positives voxels. This FPR threshold is shown with a dotted vertical line on any ROC curves.

5.2.2 Unconstrained-Subspace Results & Discussion

Unconstrained-subspace reconstructions were formed at a range of acceleration factors from R=1.01 (100 blades/frame) to R=50.26 (2 blades/frame). The summative scores across the acceleration factors are shown in figure 5.3, with more in-depth comparison at certain key acceleration factors shown in figure 5.4

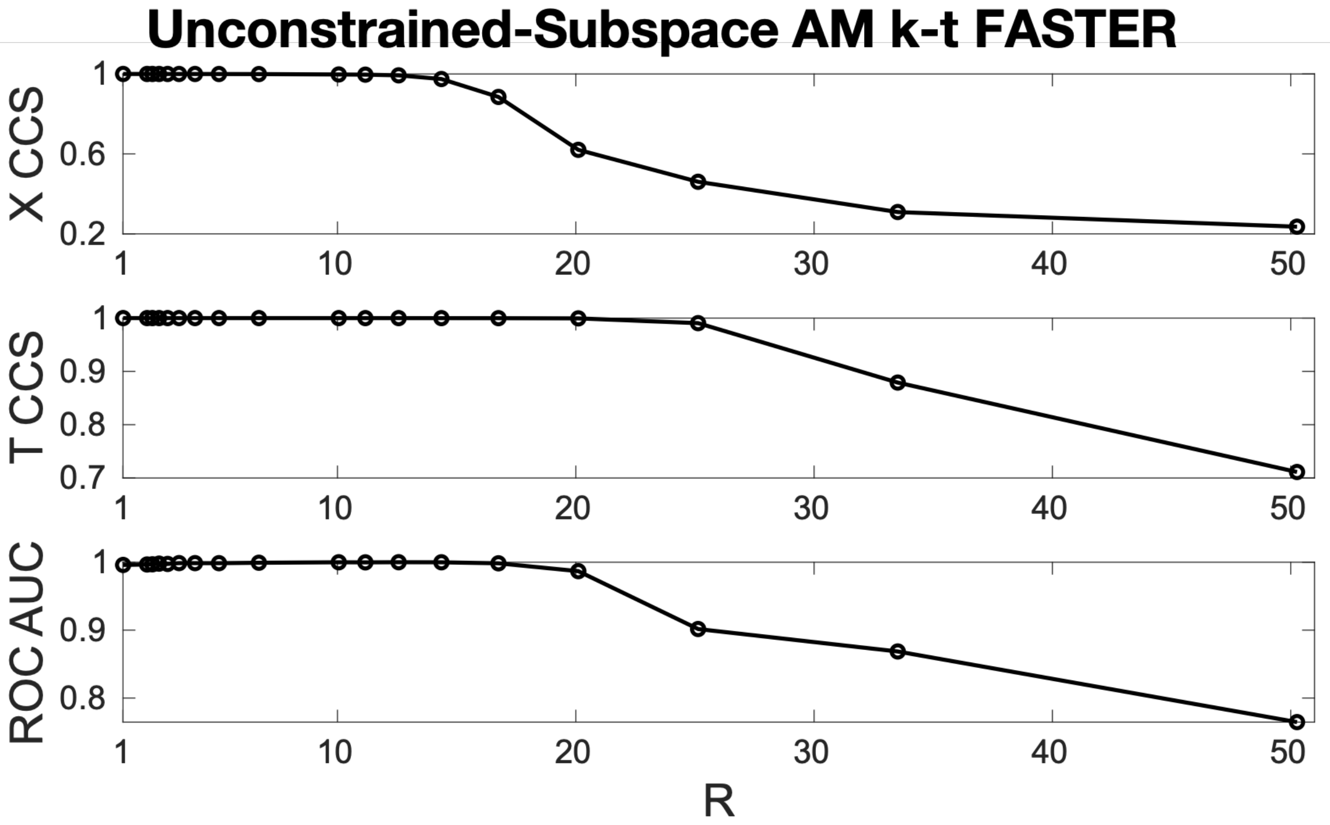


Figure 5.3: The X CCS, T CCS and ROC AUC results of a selection of acceleration factors from R=1.01 (100 blades/frame) to R=50.26 (2 blades/frame). The 9 most rightward points represent reconstructions at 10 → 2 blade(s)/frame.

In figure 5.3 the unconstrained-subspace k-t FASTER approach produces effective reconstructions at reasonably high acceleration factors (up to R≈15), but too strong an acceleration factor will cause the reconstruction to fail to capture the relevant functional information. The T CCS scores can be seen to be more robust, only failing to capture all six temporal components at R>20 (e.g. at 5 blades/frame or fewer). Therefore, reconstructions at acceleration factors of R=33.51, R=25.13 and R=20.11 (3, 4, and 5 blades/frame) will be considered in the constrained-subspace reconstructions. These factors represent a range where the previous method, pure k-t FASTER, fails and therefore is the range at which subspace constraints can provide improvement. The high R factors demonstrated here are achieved in a dataset with high SNR and low intrinsic rank, and don't necessarily represent the acceleration factors that would be achievable with in-vivo data.

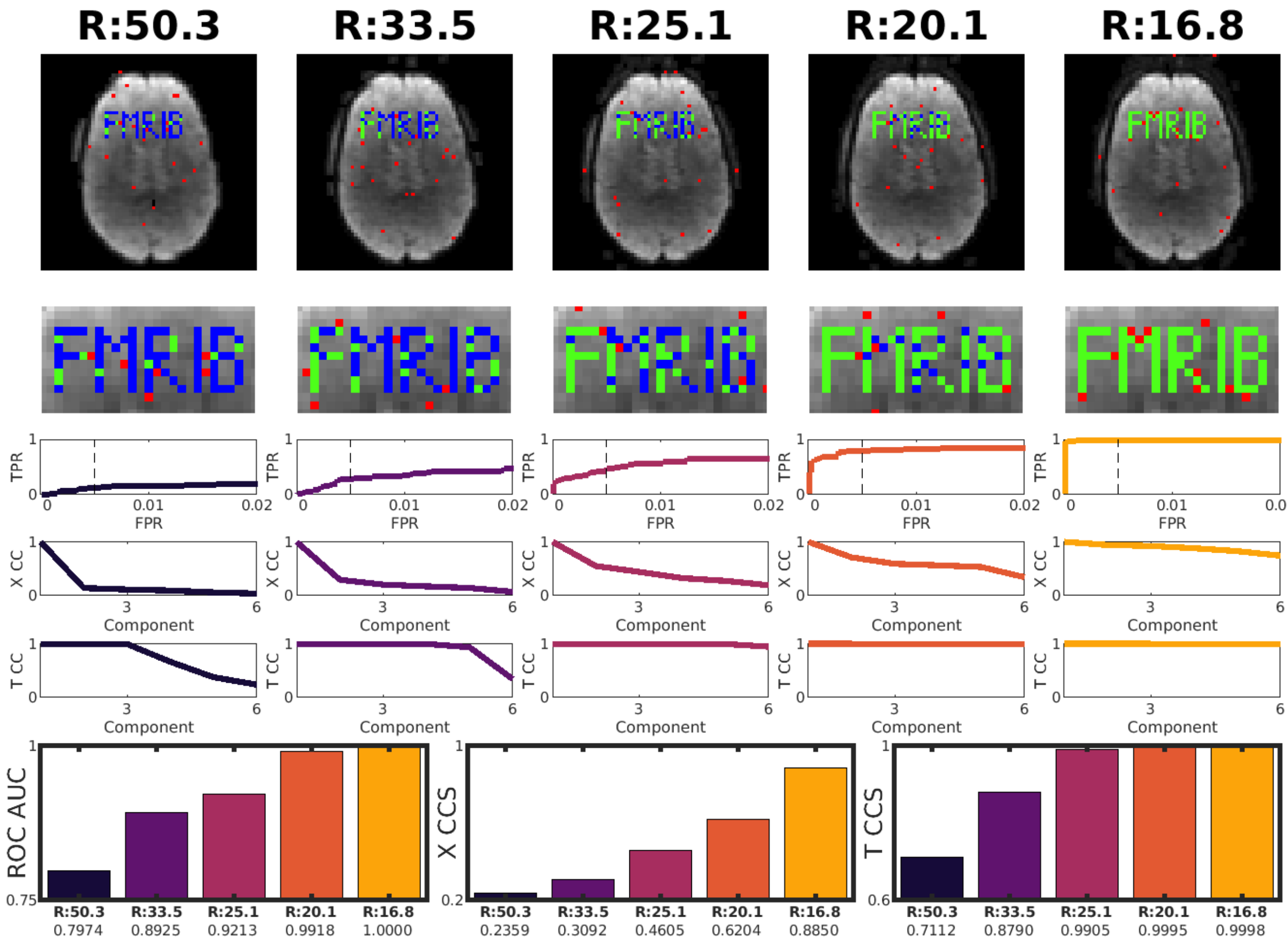


Figure 5.4: A comparison between unconstrained-subspace reconstructions of the FMRIB dataset at R=50.26 (2 blades/frame), R=33.51 (3 blades/frame), R=25.13 (4 blades/frame), R=20.11 (5 blades/frame), and R=16.76 (6 blades/frame). Sample z-stat parameter maps, an ROC curve, and the CC curves are shown in a columnwise structure for each reconstruction. The z-stat maps are shown at a threshold corresponding to an FP of 0.005 (the dotted lines on the ROC curves); they show true positives in green, false positives in red, and false negatives in blue. Summative bar graphs for all reconstructions are shown on the bottom row. Figure 4.5 contains a detailed schematic of the graph format.

Figure 5.4 shows a full comparison between these acceleration factors, as well as the immediately adjacent acceleration factors (i.e. between 2 → 6 blades/frame). It is worth noting that the mean brain intensity beneath the z-stat maps is fairly accurately reconstructed at all acceleration factors. The number of true positives displayed in the functional maps can be seen to decay with R, as do the X CCS and T CCS scores. At acceleration factors of $R \leq 16.76$, the X CCS, T CCS, and ROC AUC indicate practically perfect reconstructions of the FMRIB dataset. The z-statistic threshold was chosen to always display 20 false positives, but it can be seen from the ROC curve of R=16.76 that even better thresholding is possible in this instance.

5.2.3 Unconstrained-Subspace Summary

The unconstrained-subspace k-t FASTER approach does a good job in reconstructing the artificial FMRIB dataset, but at R>16.76 the reconstruction quality starts to decrease. It is in this decreased region of reconstruction quality that the constrained-subspace approaches must be shown to produce improved reconstructions.

5.3 Tikhonov Constraints

In this section, equation 5.3 will be adapted to consider non-zero λ_X and λ_T values, but in a case where X_{prior} and T_{prior} are set to be zero matrices. This set-up removes the priors, leaving pure Tikhonov regularisation terms. λ_∇ is set to zero. The formulation can be visualised either in equation 5.5 or in figure 5.5.

$$\hat{X}, \hat{T} = argmin_{X,T}\left(\|EXT' - d\|_2^2 + \lambda_X\|X\|_2^2 + \lambda_T\|T\|_2^2 \right) \tag{5.5}$$

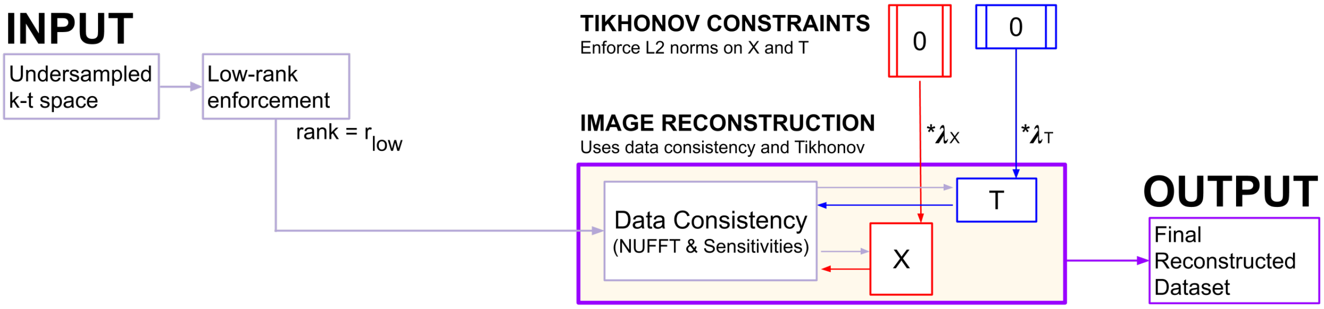


Figure 5.5: A schematic of Tikhonov constrained reconstruction. The 0 boxes represent $\|A - 0\| = \|A\|$ regularisation. X and T iteratively update until a convergence criterion is reached, and XT' is accepted as the final reconstructed dataset.

5.3.1 Tikhonov Theory

A common use of L2 regularisation in optimisation is to prevent overfitting in noisy data. Unconstrained-subspace AM optimisation has no constraint controlling the individual component matrices, only requiring that the product of the two component matrices is similar to the sampled data.

Tikhonov regularisation is the easiest and simplest way to constrain the two matrices during an AM optimisation to avoid an uneven energy distribution. It has the advantage of being both simple to implement and efficient to compute, while also requiring no prior dataset knowledge.

Tikhonov regularisation has previously been used in a low-rank context in collaborative filtering using matrix factorisation [3], which used Tikhonov regularisation on the two component matrices (X and T). The use of Tikhonov regularisation on component matrices in fMRI reconstruction is a novel step that could only be implemented concurrently with a data consistency term through an AM formulation.

5.3.2 Tikhonov Results & Discussion

Using values of $\lambda_X = \lambda_T = 1{\times}10^{-5}$ for the Tikhonov reconstruction and the processing pipeline from section 5.2.1, the reconstructions shown in figure 5.6 were generated. The Tikhonov-constrained results consistently outperform unconstrained k-t FASTER within an acceleration factor across all metrics except T CCS, wherein the difference between the two approaches T CCS was very small until R=33.51, where the Tikhonov result outperformed k-t FASTER by a large margin. The R=33.51 result for Tikhonov outperformed the R=20.11 k-t FASTER result for both X CCS and ROC AUC despite only containing 3 blades/frame. In fact, the ROC curve indicates that a much stricter threshold could have been used for all Tikhonov reconstructions without loss of useful information.

The small improvement of the R=33.51 Tikhonov-constrained ROC AUC score (ROC AUC:0.997) over the R=25.13 and R=20.11 Tikhonov ROC AUC scores (ROC AUC: 0.991 and 0.992) is not considered large enough to represent a tangible improvement - a visual inspection of the Tikhonov maps shows that there is only a difference of two/three true positives at the threshold chosen. When accounting for the existence of some underlying noise in the data, such a small change can be considered inconsequential. The slightly non-monotonic CCS scores across the acceleration factors for Tikhonov should also be noted. This non-monotonicity is likely due to almost all of the reconstructions operating in a near-perfect region, and so small changes in reconstruction quality due to factors such as the noise distribution can make a more highly accelerated reconstruction seem better than a less accelerated reconstruction.

Overall it can then be seen in this artificial FMRIB dataset that Tikhonov regularisation outperforms the unconstrained-subspace k-t FASTER reconstruction, in both quantitative and qualitative measures of reconstruction quality.

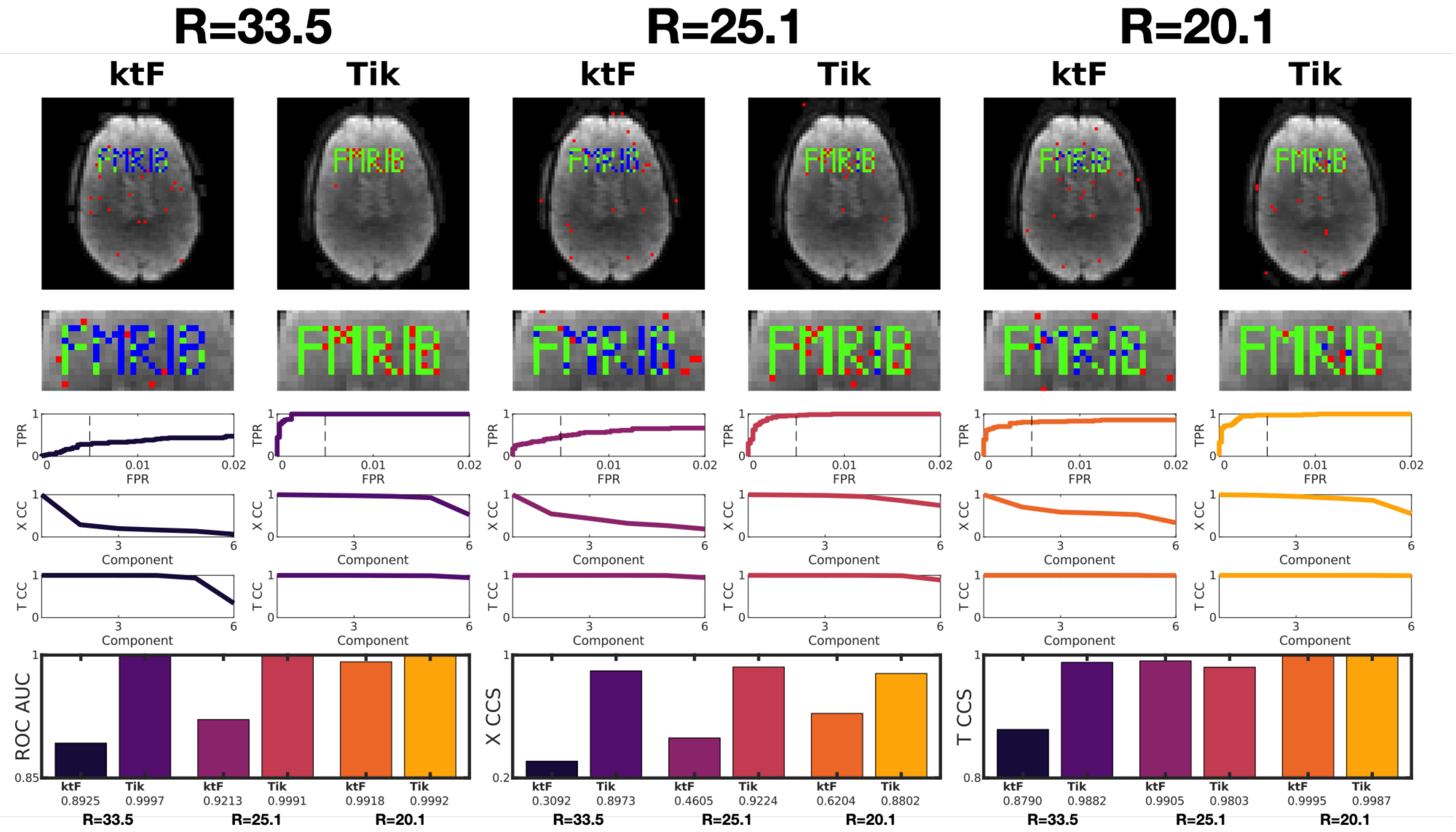


Figure 5.6: A comparison between unconstrained-subspace k-t FASTER (ktF) and Tikhonov-constrained (Tik) reconstructions of the FMRIB dataset at R=33.51 (3 blades/frame), R=25.13 (4 blades/frame), and R=20.11 (5 blades/frame). Sample z-stat parameter maps, an ROC curve, and the CC curves are shown in a columnwise structure for each reconstruction. Summative bar graphs for all reconstructions are shown on the bottom row. Figure 4.5 contains a detailed schematic of the graph format.

5.3.3 Tikhonov Conclusion

It is worth restating that, when compared to an unconstrained-subspace k-t FASTER approach, no extra structural information is added to a Tikhonov-constrained reconstruction beyond the inherent Tikhonov assumption that high energy should be penalised. As such, the improved reconstruction ability at high acceleration factors when compared with the unconstrained-subspace approach (see the R:33.51 results) make Tikhonov a very attractive constraint for future use.

5.4 Low-Resolution Prior Constraints

In the previous section, Tikhonov-constrained reconstructions imposed a low-energy form on X and T. In this section, low-resolution priors (LRP) are considered as a way to integrate known structural information into the k-t FASTER reconstruction process. Equation 5.3 will be adapted to consider non-zero λ_X and λ_T values, where X_{prior} and T_{prior} will be constructed from a low spatial-resolution version of the underlying dataset. This low spatial-resolution dataset is created through windowing, rather than through a separate acquisition (as is the case with some approaches). λ_∇ will be set to zero. Larger λ_X and λ_T values will result in stricter adherence to the LRP. k-t PSF [4] can be calculated using $\lambda_X = 0$ and $\lambda_T = \infty$ to represent a fully unconstrained spatial subspace and a completely fixed temporal subspace. As a reminder, the unconstrained-subspace k-t FASTER results from section 5.2 can be represented with $\lambda_X = 0$ and $\lambda_T = 0$.

$$\hat{X}, \hat{T} = argmin_{X,T}\left(\|EXT' - d\|_2^2 + \lambda_X \|X - X_{prior}\|_2^2 + \lambda_T \|T - T_{prior}\|_2^2 \right) \tag{5.6}$$

The term LRP will refer to the use of spatial and temporal priors that are driven from a low spatial-resolution reconstruction of the same data. Where the term LRP is ambiguous, “generated LRP” or “generated priors” will refer directly to the constraints, and “LRP-constrained reconstruction” will refer to a reconstruction that uses the generated LRP as constraints.

5.4.1 LRP Theory

In an accelerated radial k-space trajectory, there exists a central window of k-space which fulfils the Nyquist sampling criteria in the azimuthal direction due to the non-uniform sampling density. The k-space radius which satisfies Nyquist is $\frac{k_{max}}{R}$, where R is the acceleration factor defined by equation 2.26 and k_{max} is the maximum k-space extent. This relationship is shown in figure 5.7, and a schematic demonstrating how this windowing is used to generate an LRP can be seen in figure 5.8.

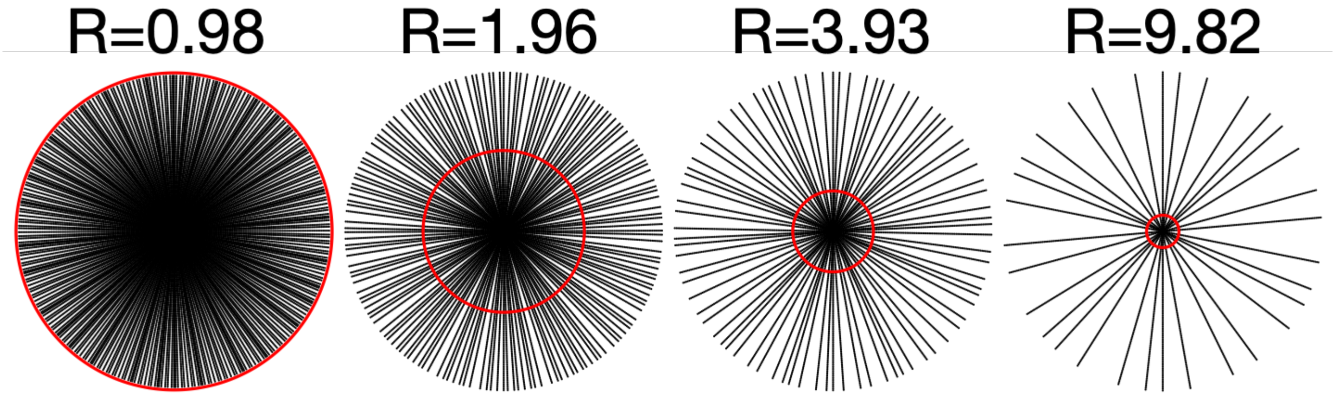


Figure 5.7: A demonstration of the changing radius which satisfies Nyquist in accelerated Golden Ratio angle radial data. The red circle demonstrates the radius at which the Nyquist criteria is fulfilled in uniform-angle radial trajectories, which is assumed to be similar for the Golden Ratio angle approach.

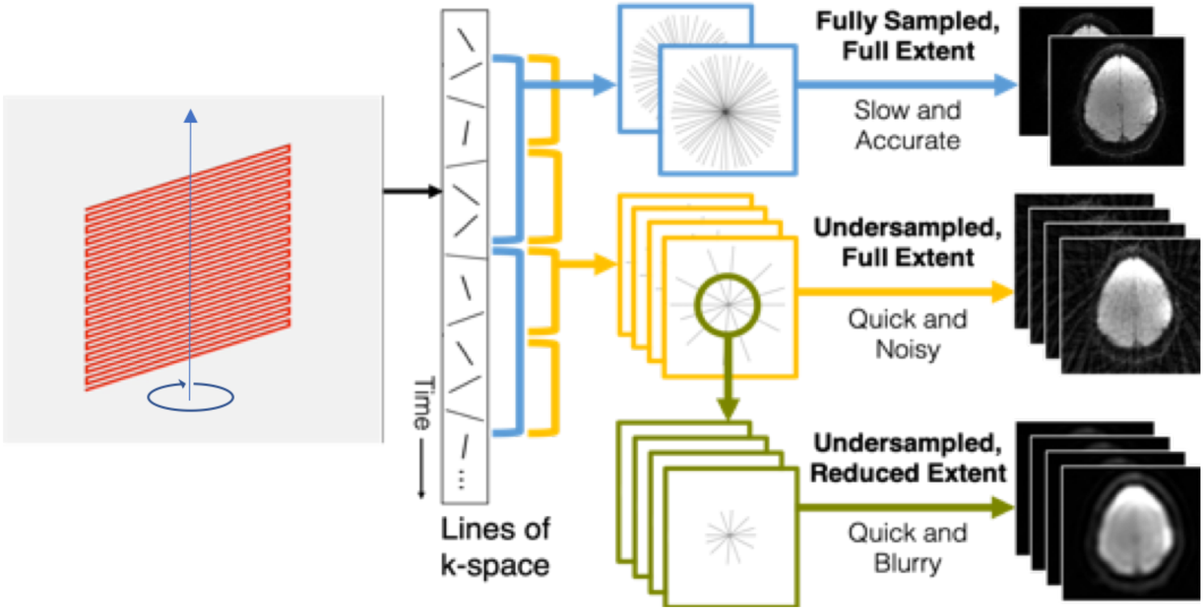


Figure 5.8: A demonstration of how the changing Nyquist radius can be used in accelerated Golden Ratio angle radial data to produce an unaliased LRP. An EPI plane (left) is rotated by $\approx 111.25°$ around the phase encode axis. These rotated planes can then be flexibly combined. If many planes are used (top, blue) then a clean image is easily generated, but at the cost of temporal resolution. If fewer planes are used (middle, yellow) then more images are generated per second, but with an increased number of artefacts. The central part of this undersampled k-space satisfies the Nyquist criterion, even if the full extent of the undersampled k-space does not. By windowing this central k-space (green, bottom), an accurate and high temporal resolution - albeit spatially smoothed - set of images can be created.

Windowed priors are produced from a reduced-extent k-space. If those windowed priors are incorporated into a final reconstruction, the central portion of k-space will be more greatly weighted in the final reconstruction. This will result in the central portion of k-space being more strictly enforced, while the outer portion k-space is guided more loosely during reconstruction. More strictly enforcing the outer k-space which violates Nyquist would lead to the appearance of artefacts. The importance of central k-space in dynamic MRI reconstruction has been previously exploited in low-rank methods. k-t PCA [5] and k-t PSF [4] both derive temporal components from low spatial resolution/high temporal resolution datasets. In both approaches the temporal basis from the low-spatial resolution dataset is strictly fixed during reconstruction, whereas the LRP-constrained k-t FASTER approach uses the central k-space information as a soft constraint on both subspaces.

As an aside, a weighted least-squares approach could also be taken to upweight central k-space ($||W(EXT' - d)||_2^2$, where W is an appropriate windowing function). Unlike the LRP approach, a weighted-least square approach would not require two separate reconstructions, which could dramatically reduce reconstruction time. However, Toeplitz embedding (section 2.A) could no longer be used, which would then add back to the reconstruction time. Additionally, separate k-space weighting could not be applied to the spatial and temporal subspaces (and indeed this approach could be done without alternating minimisation, by using the SVD k-t FASTER approach discussed in chapter 4). Overall, the weighted least-squares approach prioritises a simpler (and potentially quicker) formulation in exchange for reduced flexibility regarding the subspace weighting. The LRP-constrained approach will be used in this thesis to explore the benefits of such flexibility (and to allow easy comparison with the k-t PSF approach), but the weighted least-squares approach may be worth investigating in future studies.

5.4.2 Forming the LRP Constraints

When windowing the dataset to form priors, it is important to taper the edges of the window in a process known as apodisation. To that end, a Tukey window was used with α set to 0.4 for all windowing (α: the proportion of the window which is a sine lobe). Equation 5.7 was used to define the full-width half maximum (FWHM) for all windows (which begins the tapering just as Nyquist stops being fulfilled). This equation was chosen to represent the ideal theoretical radii for a given acceleration factor. Further information regarding the windowing parameters is included in section 5.A.

$$FWHM = \tfrac{\pi k_{max}}{2R} \tag{5.7}$$

The LRP is generated from the underlying data. This can be done by windowing the dataset with a Tukey window W with a FWHM defined by equation 5.7, and then following the standard reconstruction methods for unconstrained-subspace k-t FASTER. The cost function for the generated

LRP is formulated as follows:

$$\hat{X}_{prior}, \hat{T}_{prior} = argmin_{X_{prior},T_{prior}} \left(\|EX_{prior}T'_{prior} - Wd\|_2^2 \right) \tag{5.8}$$

5.4.3 Evaluating the Effect of Windowing

As an initial test of the theory behind LRP-constrained optimisation, a known ground truth was windowed and decomposed into X_{prior} and T_{prior}. While this approach will obviously not work in a realistic scenario with an unknown truth, it can be used as a gold standard to validate the assumption that, within a fully-sampled k-space window, an accurate reconstruction of the truth is possible.

5.4.3.1 Testing the LRP Radii

A set of priors at varying k-space extents were reconstructed at R=20.11 (5 blades/frame) and R=10.05 (10 blades/frame) to test the effects of windowing. Priors generated at the full k-space extent match the unconstrained-subspace k-t FASTER results from section 5.2. Figure 5.9 compares the alignment of the generated prior subspaces against both the unwindowed ground truth (top row) and the ground truth windowed to the same extent as each prior (bottom row).

The truth and the windowed truth have fully aligned temporal subspaces, but their spatial subspaces become more unaligned as windowing increases. The generated priors become less aligned with the true temporal subspace at small window radii, but are fully matched at high radii.

The spatial subspaces of the priors are only matched with the windowed ground truth in a middle range of radii. A peak spatial alignment is achieved between the windowed ground truth and the reconstructed priors at a radius of ≈0.4, although the peak is hard to determine for R=10.05. The superior reconstruction power of the R=10.05 reconstructions can be easily seen by the close alignment to the spatial and temporal subspaces at a wide range of radii.

The window radii used in this thesis were defined by equation 5.7, which was chosen so that the inner radius of the sinusoidal lobes would be placed just inside the Nyquist limit. This can be interpreted as tapering starting to occur at the red circle in figure 5.7. Equation 5.7 gave radii of $\text{FWHM}_{R=20.11}=0.08k_{max}$ and $\text{FWHM}_{R=10.05}=0.16k_{max}$ for the acceleration factors used here. The experimental results show that the reconstructed priors align most closely with the windowed ground truth at a higher radii than is predicted by the theoretical limit of equation 5.7. The

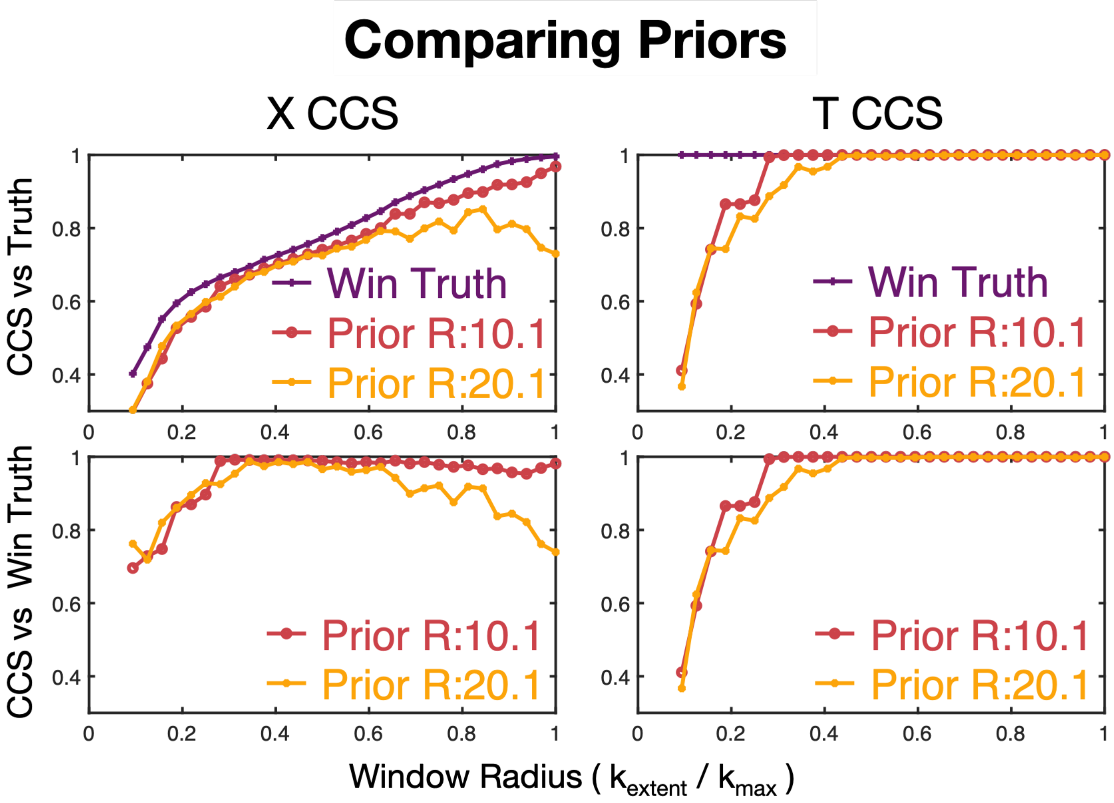


Figure 5.9: A comparison of the generated priors against both the windowed and unwindowed truth at R=20.11 (5 blades/frame) and R=10.05 (10 blades/frame). Where the truth is windowed, it is windowed to the same extent as the prior (indicated on the x-axis as the radius to the FWHM).

low-rank constraint combined with the incoherence of the artefacts may likely increase the radius at which a high-fidelity prior can be formed. Additionally, the small size of the artificial FMRIB dataset may cause overly harsh windowing of central k-space when using the theoretical radii in this dataset, in a way that would not be as negatively impactful for more realistic dataset sizes at the same acceleration factor. However, an exploration of priors with a wider radii than that defined by equation 5.7 may prove useful for future analysis.

5.4.4 LRP-Constrained Reconstruction

Once created, the LRPs are fed into the final reconstruction. The final reconstruction combines LRP regularisation on the subspaces with a data consistency term using the unwindowed dataset. The LRP regularisation terms are weighted by λ_X and λ_T parameters. This process is shown in figure 5.10.

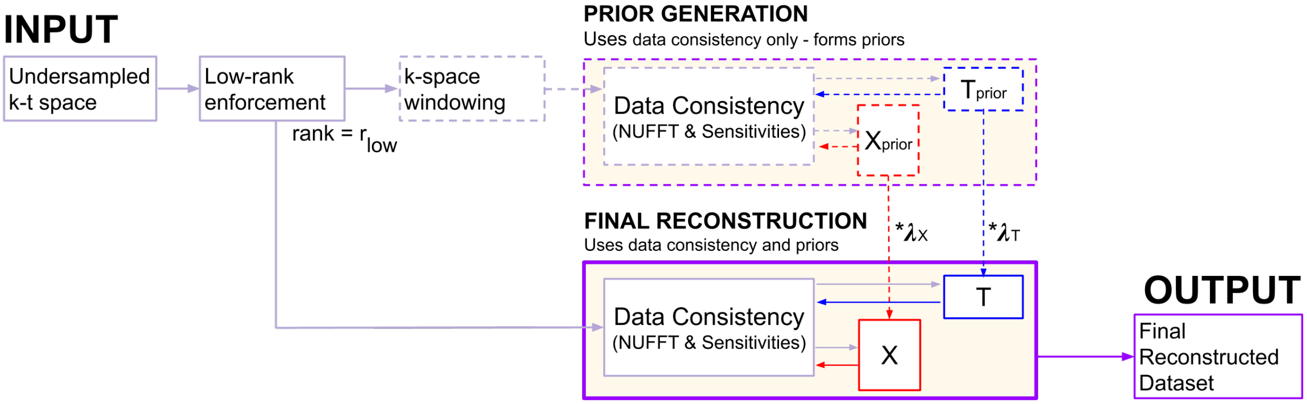


Figure 5.10: A schematic overview of a constrained-subspace reconstruction using priors. The undersampled dataset undergoes central k-space windowing and is reconstructed according to only the rank constraints and coil sensitivity information to create X_{prior} and T_{prior}. These priors are then fed as a regularisation constraint into the final reconstruction, producing the final dataset. The first term of the equation 5.6 is represented in the "Data Consistency" box, and the influence of the prior terms are shown via the downwards dotted arrows (weighted by λ_X and λ_T).

5.4.4.1 k-t PSF

k-t PSF [4] (partially separable functions) is a low-rank dynamic imaging approach which fits inside the framework of an LRP-constrained low-rank fMRI reconstruction, although it should be noted that this implementation differs from the approach used by Liang et al. (see section 3.5.3). k-t PSF reconstructs only the spatial coefficients against a predetermined temporal basis, a formulation which can be modelled by setting $\lambda_X = 0$ and $\lambda_T = \infty$ (see equation 5.9). This predetermined temporal basis set is formed from a reconstruction which must have a high temporal resolution, but can have a low spatial resolution. These conditions match the formulation of the generated LRPs described previously. Setting $\lambda_t = \infty$ fixes the T_{prior} constraint of the final reconstructed temporal component matrix. The spatial information garnered from the windowed reconstruction (X_{prior}) is discarded, and X is regenerated from the full unwindowed dataset. The k-t PSF schematic is shown in figure 5.11.

$$\begin{aligned} T &= T_{prior} \\ \hat{X} &= argmin_X \left(\|E(XT') - d\|_2^2 \right) \end{aligned} \tag{5.9}$$

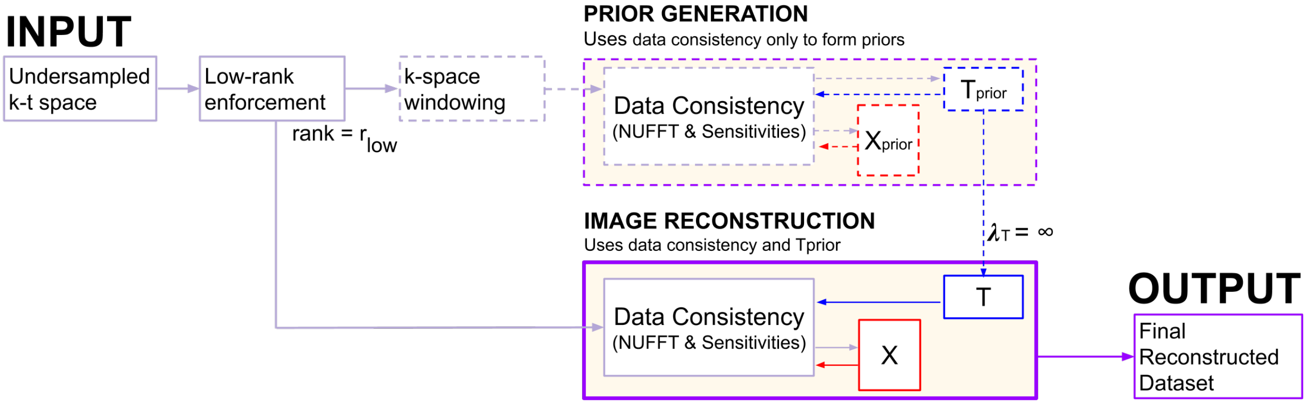


Figure 5.11: A k-t PSF schematic. Similar to the general LRP-constrained schematic, a set of LRP constraints are generated to create X_{prior} and T_{prior}. The temporal prior is then fed as a hard regularisation constraint into the final reconstruction, while X is reproduced through a more straightforward optimisation process.

5.4.5 LRP Results & Discussion

In figure 5.12, LRP reconstructions using $\lambda_X = \lambda_T = 1 \times 10^{-5}$ are compared against unconstrained subspace k-t FASTER reconstructions. The LRP-constrained results outperform unconstrained k-t FASTER in both X CCS and ROC AUC. In addition, the false negatives (red pixels) were highly localised around the region of interest in the LRP z-stat parameter maps when compared to the false negatives of the unconstrained-subspace k-t FASTER results. This indicates that the LRP-constrained results will be good at broadly identifying regions of interest, in addition to the improved overall reconstruction performance at this high acceleration factor. However, the worse T CCS scores of the LRP-constrained reconstructions should also be taken into consideration, as they indicate a poor temporal subspace reconstruction. Additionally, the reconstruction times are roughly twice as long for LRP-constrained reconstructions, due to the time required to generate the priors.

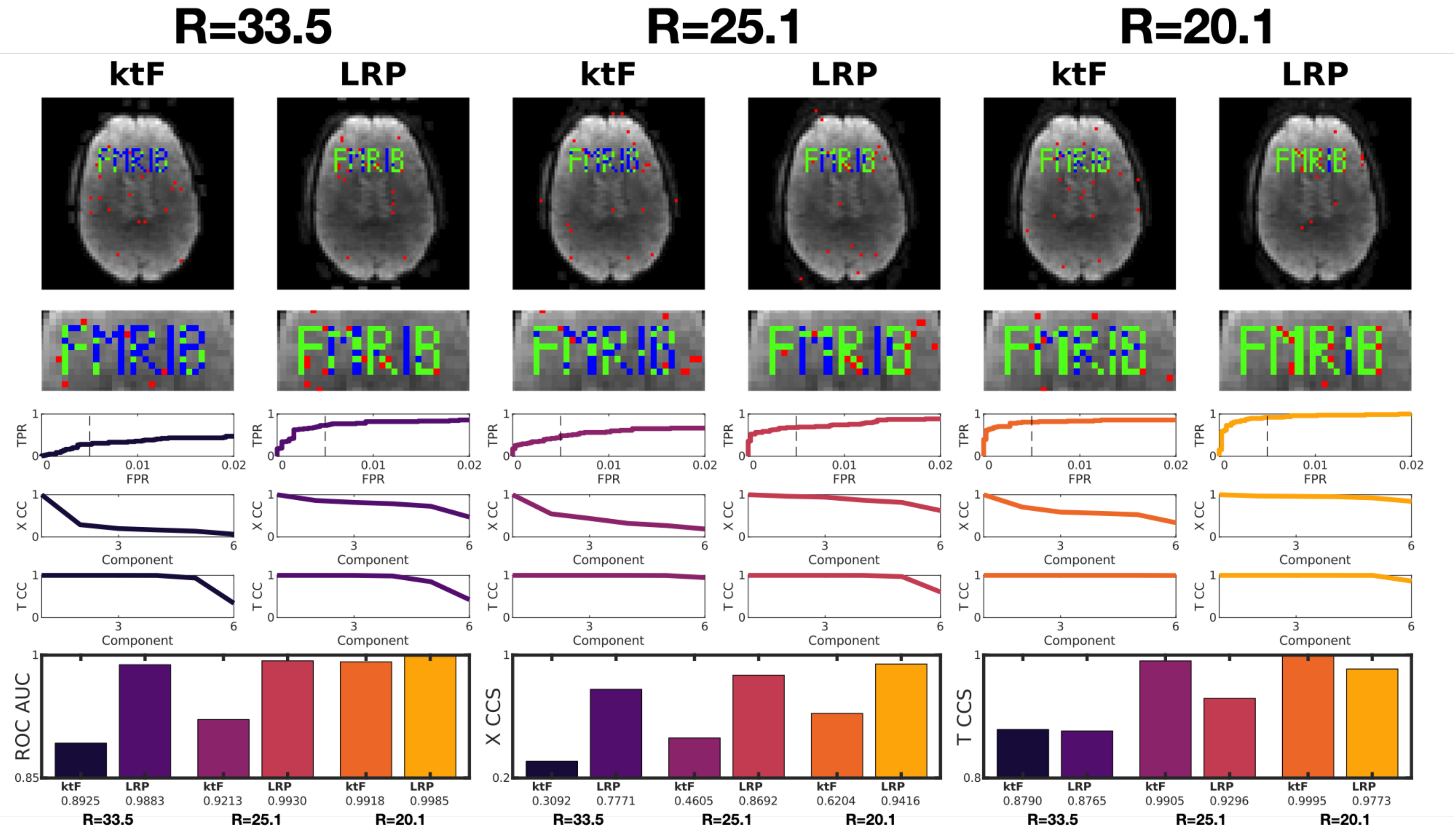


Figure 5.12: A comparison between unconstrained-subspace k-t FASTER (ktF) and LRP-constrained reconstructions of the artificial FMRIB dataset at R=33.51 (3 blades/frame), R=25.13 (4 blades/frame), and R=20.11 (5 blades/frame). Sample z-stat parameter maps, an ROC curve, and the CC curves are shown in a columnwise structure for each reconstruction. Summative bar graphs for all reconstructions are shown on the bottom row. Figure 4.5 contains a detailed schematic of the graph format.

Figure 5.13 shows a comparison of soft constrained-subspace reconstructions (Tikhonov-constraints and LRP-constraints) against unconstrained k-t FASTER reconstructions and the hard constraint of k-t PSF reconstructions. The overall improvement of the soft constrained-subspace results in X CCS and ROC AUC is noticeable over k-t FASTER, with the Tikhonov results performing best at high acceleration factors. The LRP-constrained reconstructions perform almost identically well until R=50.26, which with only two blades per frame appears to be beyond the realm in which the assumptions behind LRP-constrained reconstructions hold (although the failure of LRP in this extreme regime makes the high ROC AUC Tikhonov score very notable).

The poor performance of k-t PSF is also worth dwelling on. k-t PSF only achieves real parity with the other results at R≈5 (25 blades/frame). At R=12.57 (8 blades/frame) k-t PSF still contains approximately the same reconstruction power as the other methods, only failing to capture the "I" spatial component (the plateau in ROC score is consistently due to failing to identify the "I" component). The "I" component was the most prone to failure across all reconstruction methodologies, which may be due to "I" having the most relative power at high temporal frequencies, and so having the signal most likely to be lost in noise.

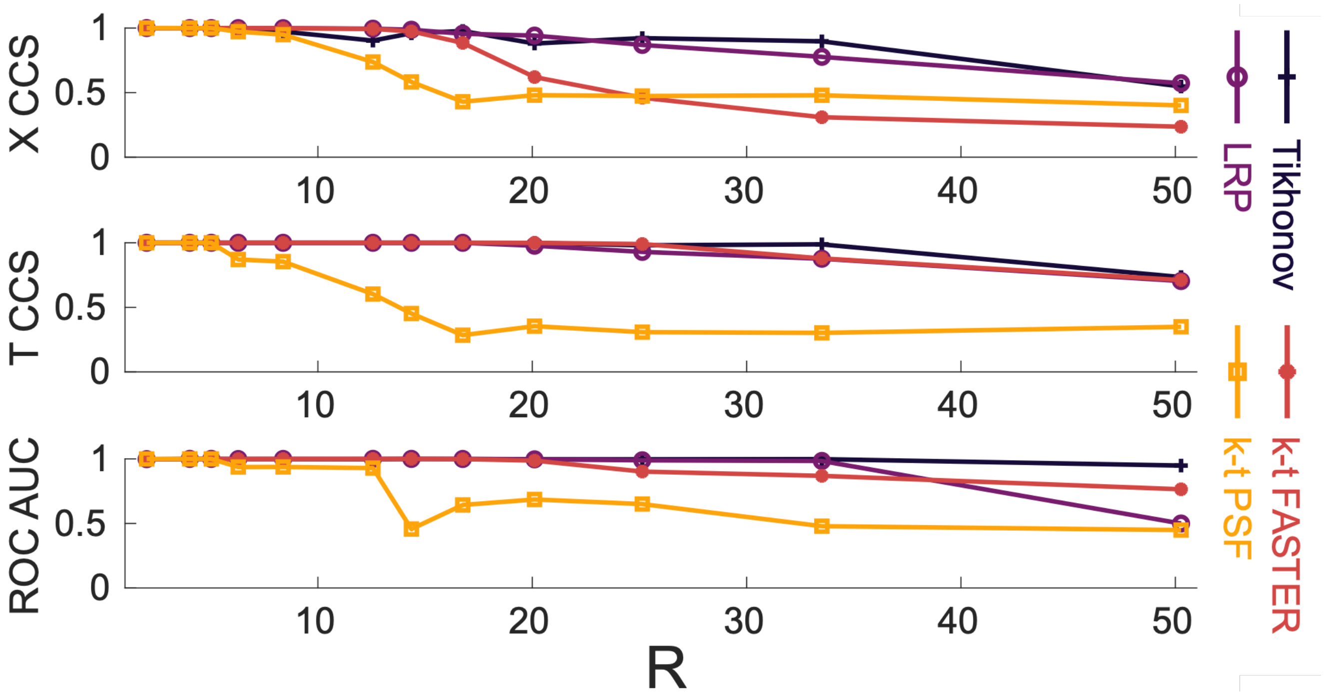


Figure 5.13: A comparison of the quantitative reconstruction scores for Tikhonov-constrained, LRP-constrained, k-t FASTER, and k-t PSF reconstructions at a range of acceleration factors.

R=14.36 (7 blades/frame) appears to represent a particularly bad reconstruction in figure 5.13. No clear reason could be found for this dip in ROC AUC. Figure 5.14 shows the R=14.36 results, which show the other three approaches all easily manage to successfully reconstruct the underlying functional information. k-t PSF successfully captures the "F" component and most of the "M", but the remaining components are completely absent from the z-stat image shown (and are not procured without many false positives obscuring the image). The shape of the T CC curve indicates that the remaining components are barely captured in the reconstruction.

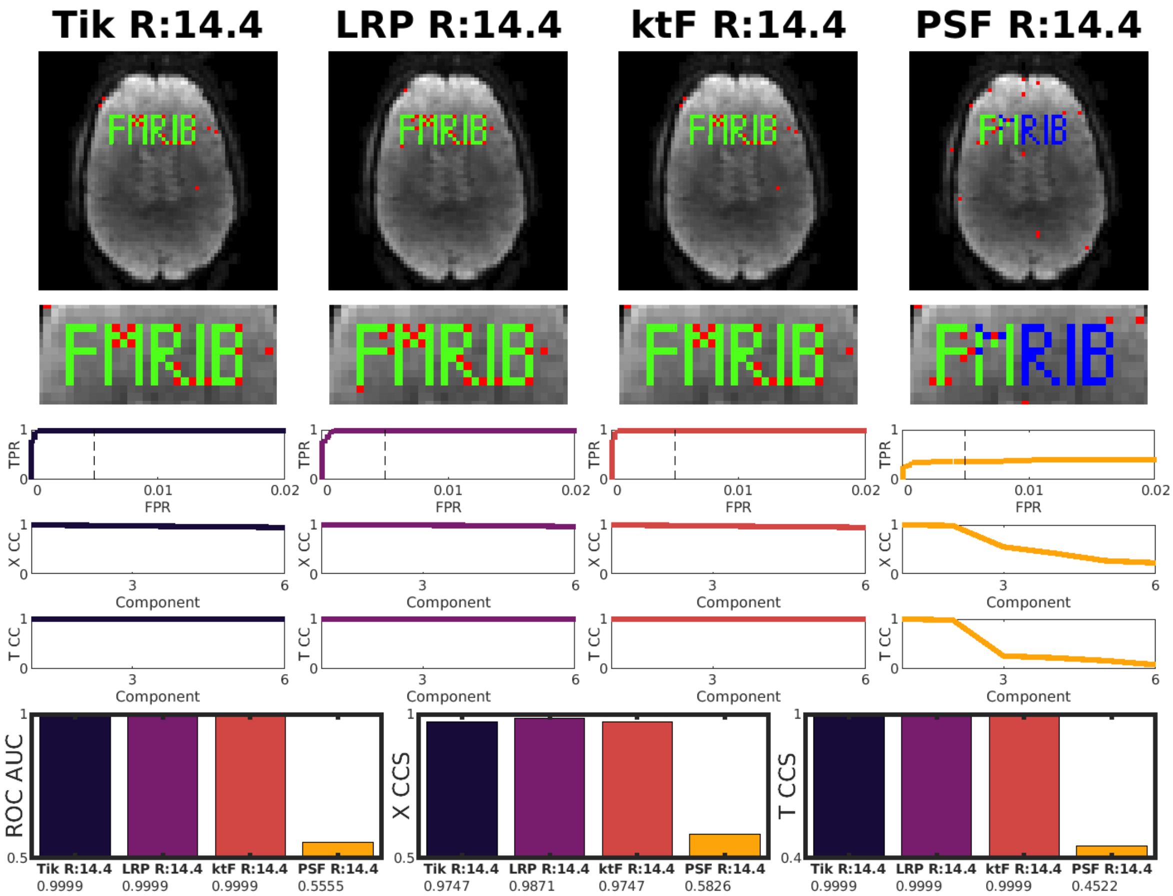


Figure 5.14: A comparison between Tikhonov-constrained (Tik), unconstrained-subspace k-t FASTER (ktF), LRP-constrained (LRP) and k-t PSF (PSF) reconstructions of the FMRIB dataset at R=14.36 (7 blades/frame). Sample z-stat parameter maps, an ROC curve, and the CC curves are shown in a columnwise structure for each reconstruction. Summative bar graphs for all reconstructions are shown on the bottom row. Figure 4.5 contains a detailed schematic of the graph format.

A closer look was taken at the temporal information for the R=20.11 reconstructions (5 blades/frame). The average signal over the spatial mask defined by all 5 maps in the ground truth is plotted in figure 5.15 (along with the noiseless ground truth, and the background signal at a random pixel). All k-t PSF time series contain a large amount of high temporal frequency content which is not present in the ground truth (see section 5.B.3 for more evidence behind this claim). As such, it is clear that the k-t PSF reconstruction was predestined for poor results with the fixed temporal subspace, with almost no component recognisable between the generated LRP and the ground truth. The presence of these high temporal frequency artefacts is somewhat unclear, although they are certainly linked to higher levels of windowing/acceleration.

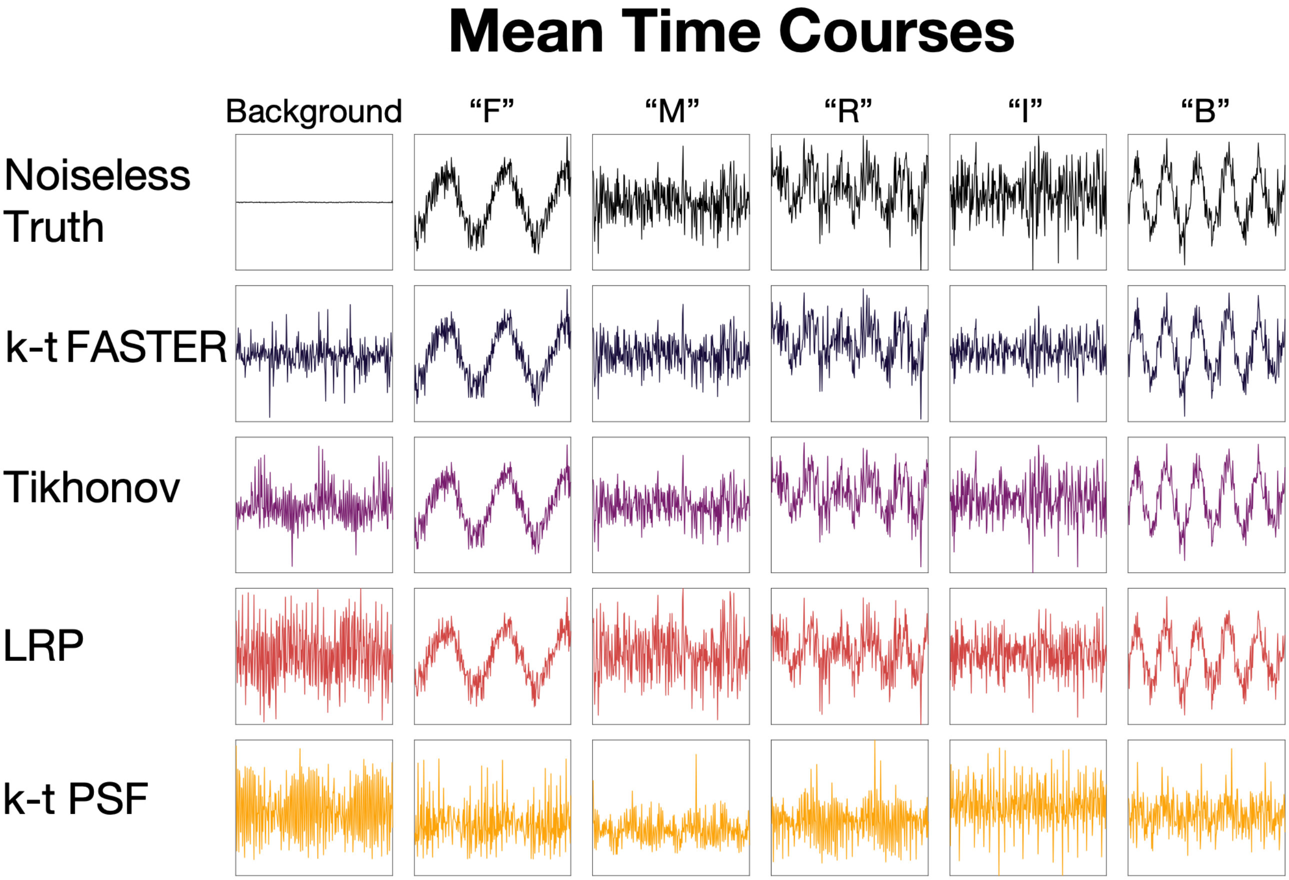


Figure 5.15: A comparison between the mean time courses over a random background pixel, and the five separate spatial component maps of figure 5.2. These time courses were taken from reconstructions at an acceleration factor of R=20.11. Each y-axis is normalised to show the shape of the temporal component.

5.4.6 LRP Conclusion

LRP-constrained reconstructions are clearly worth pursuing further, although the poor temporal subspace results in both the generated priors and final LRP-constrained reconstructions (and k-t PSF results) indicate that the temporal constraint may prove a limiting factor for this formulation. Further investigation of the LRP is needed to fully understand what the constraint is achieving. The high dimensionality of LRP-constrained reconstruction (window radius, λ_X, λ_T) makes the pursuit of an optimal LRP-constrained reconstruction somewhat challenging. Investigation of optimal λ_X and λ_T values will occur in chapter 6.

5.5 Note on Adapted LRP and Tikhonov Constraints

In order to investigate the results of the previous sections, three separate adaptations of the Tikhonov and LRP-constrained approaches were explored to gain further understanding about the nature of these constraints and the resulting reconstructions. While some insight was gained, the adapted constraints did not lead to any tangible improvement in reconstruction. As such, the formulation and analysis of these constraints can be found in appendix 5.B, and a short summary of the constraints and key findings will be listed here. The following terms are used to broadly distinguish between these constraints and the constraints of previous sections:

Pure Tikhonov/LRP The aforementioned Tikhonov and LRP-constrained reconstructions. The term "pure" is used to represent the formulations discussed in sections 5.3 and 5.4.

Adapted Tikhonov/LRP The general label for constrained-subspace reconstructions which incorporate Tikhonov and/or LRP information but do not follow the formulation of the pure approaches. All of the constraints in this section and section 5.B are adapted constraints.

The following terms will be used to describe specific different adapted methodologies:

Hybrid Constraints One subspace is constrained by Tikhonov and one subspace is constrained by LRP. The different combinations will be referred to as $X_{Tik}T'_{LRP}$ and $X_{LRP}T'_{Tik}$.

Tikhonov-Enhanced LRP (TE LRP) The priors are Tikhonov-constrained during generation. The final reconstruction is done with the same formulation as pure LRP-constrained reconstruction.

Null-Enforced LRP (NE LRP) The LRP enforcement is done via minimising the multiplication of the variable with the orthogonal complement of the prior (which maps the prior null space), rather than through minimising the euclidean distance between the variable and prior. This allows for use of a prior which rejects noise-like components.

The central finding of the adapted constraints section was the importance of the temporal subspace constraint in reconstruction. With the hybrid constraints, all reconstruction scores were found to be more strongly characterised by the temporal constraint than the spatial constraint (i.e. pure Tikhonov and $X_{LRP}T'_{Tik}$ reconstructions produced similar scores, and pure LRP and $X_{Tik}T'_{LRP}$ produced similar scores). Additional Tikhonov-enhancement in TE LRP did not aid the temporal subspace generation, and overweighting the Tikhonov constraints makes the components more non-orthogonal (i.e. have a larger inner product and become more correlated). The null-enforcement process for NE LRPs did provide the same regularisation as pure LRP reconstruction, but the temporal LRP was so bad that the rejection criteria would reject almost all components, leaving the null-enforcement process no information with which to improve the final result.

5.6 Temporal Subspace Smoothing

This section will investigate the effect of smoothing as a temporal subspace constraint using a non-zero λ_∇, while setting λ_X and λ_T to zero in equation 5.3. Implementing these λ conditions reduces the general equation 5.3 to the optimisation problem laid out in equation 5.10 and figure 5.16.

$$\hat{X}, \hat{T} = argmin_{X,T}\left(\|EXT' - d\|_2^2 + \lambda_\nabla \|\nabla T\|_2^2 \right) \tag{5.10}$$

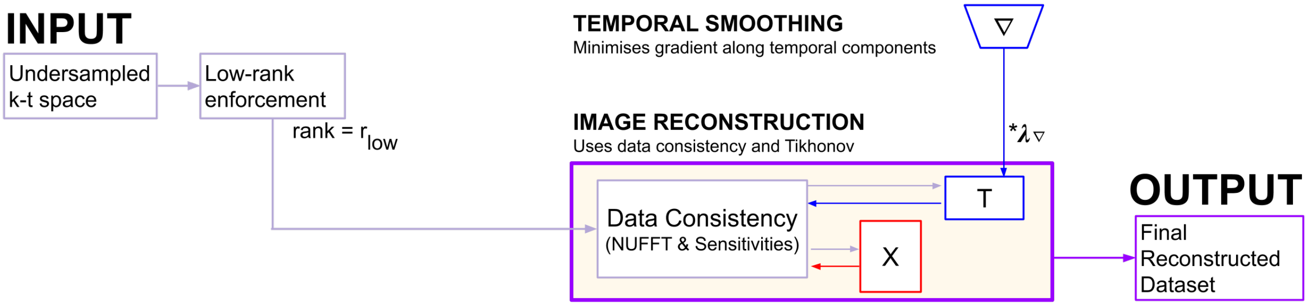


Figure 5.16: A schematic of the temporally smoothed reconstruction. The ∇ box represents a difference operator acting in the time direction along the $t \times r$ matrix T. X and T iteratively update until a convergence criterion is reached, and XT' is accepted as the final reconstructed dataset.

5.6.1 Smoothing Theory

In equation 5.10, ∇ is a difference operator which carries out the following underlying operation between each frame for all r components:

$$\nabla T_{i,1:r} = T_{i+1,1:r} - T_{i,1:r} \tag{5.11}$$

This ∇ operator can also be considered as a $(t\text{-}1) \times t$ matrix with constant diagonals. Matrix 5.12 shows ∇ for $t = 6$.

$$\nabla = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix} \tag{5.12}$$

As ∇ is a matrix operator, it can be differentiated in a cost function according to standard rules.

$$\frac{d\|\nabla T\|_2^2}{dT} = 2\nabla'\nabla T \tag{5.13}$$

The $\nabla'\nabla$ term creates a $t \times t$ matrix, which can be defined as follows for $t = 6$ (matrix 5.14).

$$\nabla'\nabla = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix} \tag{5.14}$$

This approach is not conceptually dissimilar to the other main non-sparsity approaches which enforce temporal smoothing in fMRI: Dual TRACER [6] and spatially adaptive temporal smoothing by Chiew et al. [7]. Chiew used the same finite-difference operator ∇ as shown in equation 5.11, but applied ∇ along the dataset time dimension (rather than the temporal subspace) for each voxel in SMS reconstructions. Dual TRACER enforced temporal smoothness through data consistency between frames (as covered in equations 3.41 and 3.42). As noted in section 3.5.2, Golden Ratio angle approaches are particularly suited to temporal smoothing approaches.

These approaches showed that temporal smoothing in fMRI does not remove the signal of interest (the BOLD signal). Section 3.5.3 showed that low-rank methods applied to fMRI will also preserve the signal of interest. As a temporal smoothing regularisation term is likely to preserve the relatively smooth BOLD signal (particularly at high acceleration) and reduce the magnitude of high frequency noise, it is not expected that the temporal smoothing term will cause the removal of the functional information through the low-rank enforcement.

5.6.2 Smoothing Results & Discussion

Figures 5.17 and 5.18 show the effect of a smoothing term at R=20.11 ($\lambda_\nabla = 1 \times 10^{-3}$) and at R=33.51 ($\lambda_\nabla = 1 \times 10^{-4}$). The temporally smoothed results were compared to other softly constrained reconstructions, which used $\lambda_X = \lambda_T = 1 \times 10^{-5}$ for all reconstructions. The temporally smoothed results outperformed k-t FASTER across all metrics, and is not shown. The smoothed reconstructions gave good T CCS scores across both acceleration factors, and at R=20.11 achieved the best ROC AUC score as well (although the improvement over Tikhonov is not large enough to draw concrete conclusions). However, while the T CCS smoothing score is still high at R=33.51, the X CCS score is now notably worse, and artefacts have begun to be introduced into the anatomical image. These artefacts originate from the first spatial component, which appears to have been particularly bad in this instance. These artefacts seem counterintuitive, and do not appear elsewhere in results with temporal smoothing. As such, the cause of these artefacts could not be deduced, and may simply be due to a poor reconstruction at high acceleration in a heavily artificial dataset.

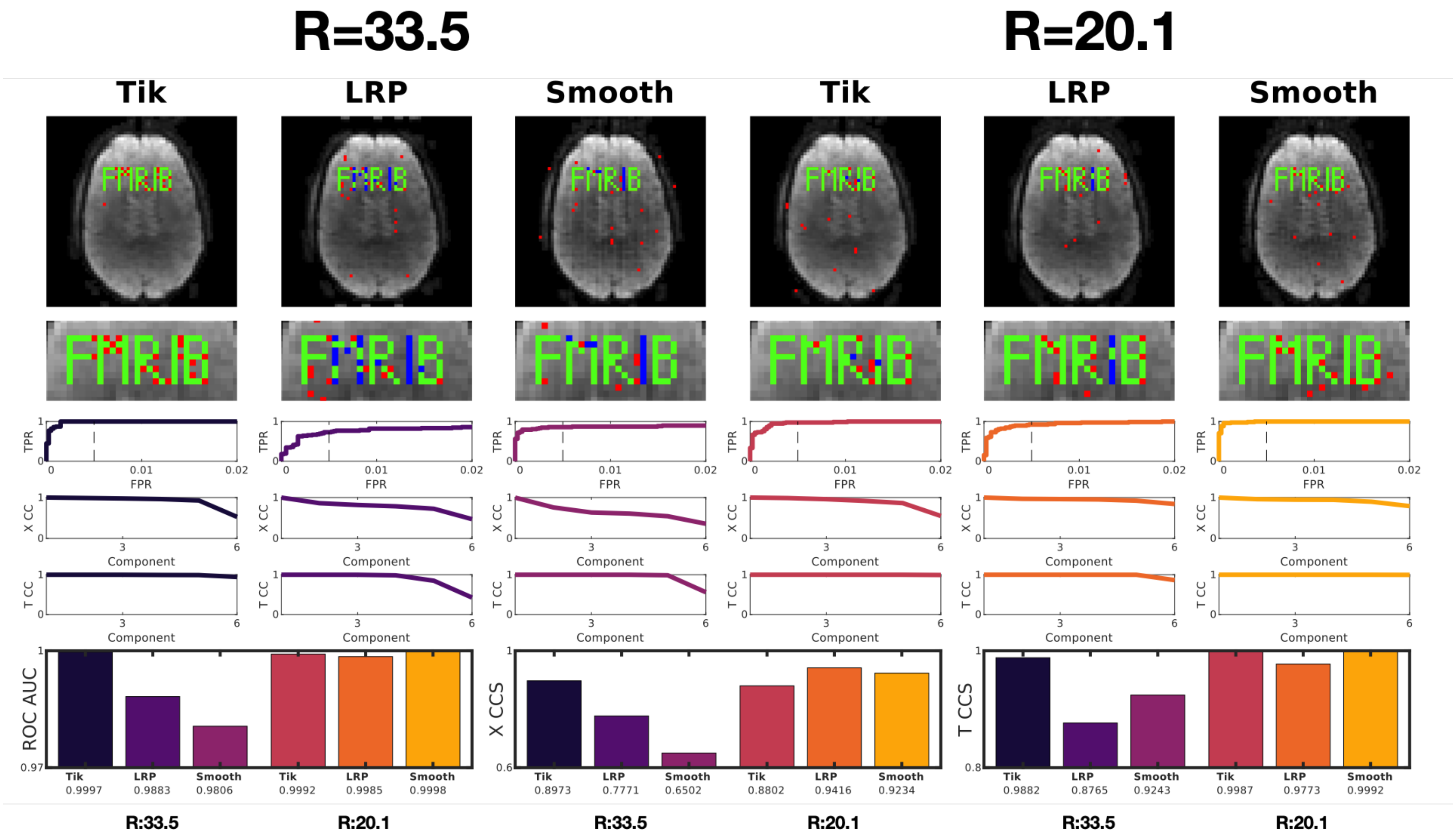


Figure 5.17: A comparison between the Tikhonov-constrained (Tik), LRP-constrained (LRP), and temporally smoothed (Smooth) reconstructions of the FMRIB dataset at R=33.51 (3 blades/frame) and R=20.11 (5 blades/frame). Sample z-stat parameter maps, an ROC curve, and the CC curves are shown in a columnwise structure for each reconstruction. Summative bar graphs for all reconstructions are shown on the bottom row. Figure 4.5 contains a detailed schematic of the graph format.

All five functional components were identified with relative ease. Figure 5.18 shows the mean temporal signal over each of the five spatial patterns and a randomly chosen background voxel in the brain. The temporal smoothing constraint is able to smooth the background noise, while preserving the spikes at high frequencies in the actual components. The LRP-constrained reconstruction has already been noted to contain a particularly large proportion of high temporal frequency information, which could indicate that the LRP constrained-reconstruction (or the generation of the LRP) would benefit from an additional smoothing constraint.

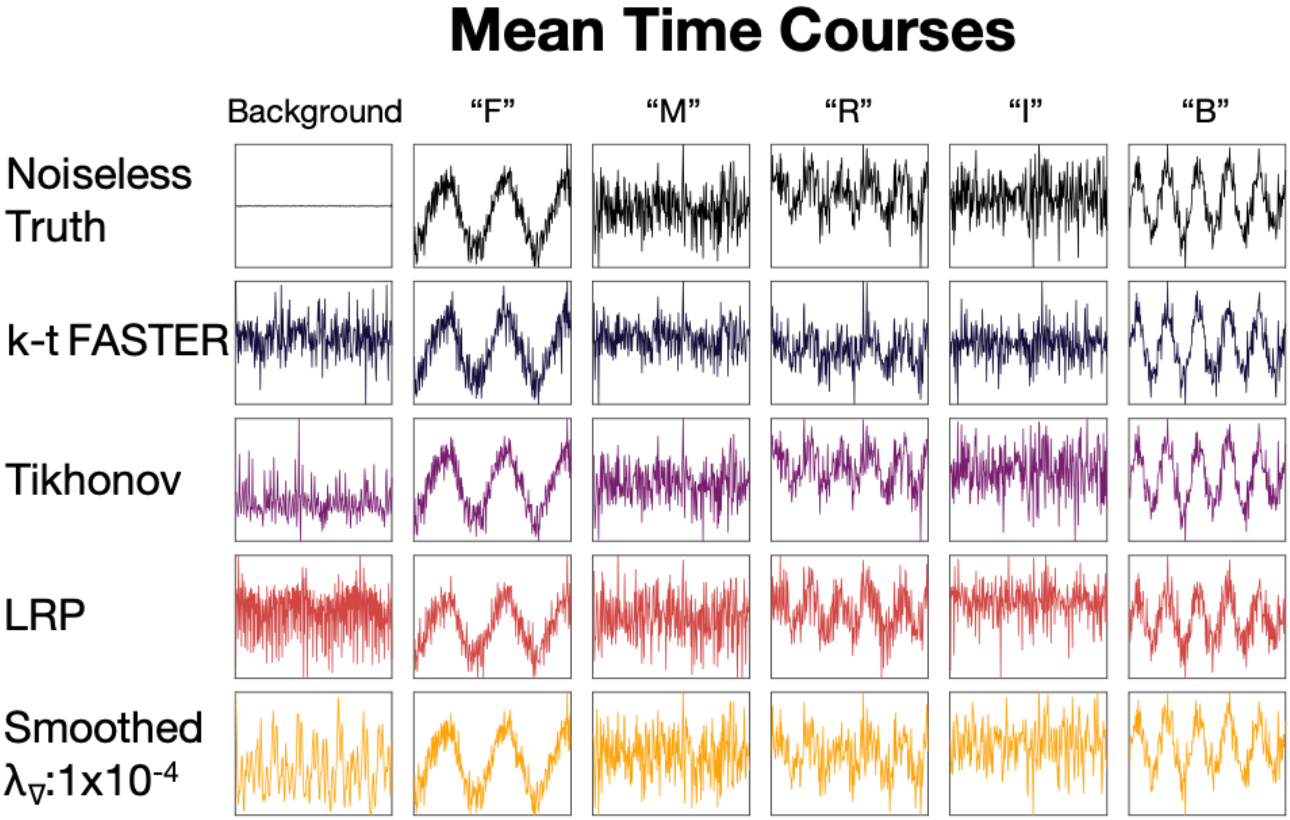


Figure 5.18: The mean time course over a particular spatial map for five different reconstructions: the truth (top row), a k-t FASTER reconstruction, (2nd row), Tikhonov/LRP-constrained reconstructions (3rd/4th row) and a smoothed reconstruction (bottom row). The columns indicate the spatial region over which the mean time course was calculated. The (non-truth) reconstructions had an acceleration factor of R=33.51. The y-axes are normalised.

5.6.3 Smoothing Conclusion

While including a smoothing term in the reconstruction did improve the unconstrained-subspace reconstruction, the effectiveness is dependent on the form of the underlying dataset signal. Testing temporal smoothing on a real dataset should be informative as to whether smoothing is worth including as a subspace constraint in fMRI data, or whether temporal subspace smoothing is just favourable to the basis functions used in the artificial FMRIB dataset's GLM.

5.7 Conclusion

The above chapter compared a variety of subspace constraints against the unconstrained-subspace approach of AM k-t FASTER in an artificial dataset. The Tikhonov, LRP, and smoothing constraints all outperformed the unconstrained subspace approach, but must all be validated further in realistic data before any concrete conclusions can be drawn.

5.A Windowing

This section describes in detail the parameters used to form the LRP. The importance of proper windowing is highlighted, and the fundamental mathematicals underlying the windows are stated.

5.A.1 Tukey

A Tukey 1D window (also known as a cosine-tapered window) is a composite window. A cosine lobe split in half and sandwiched around a central rectangular window. The primary design consideration in a Tukey window is α, a unit interval ($0 \leq \alpha \leq 1$) that defines the proportion of the window occupied by the sidelobes. $\alpha = 0$ results in a square window, α=1 is a pure cosine lobe (see figure 5.19).

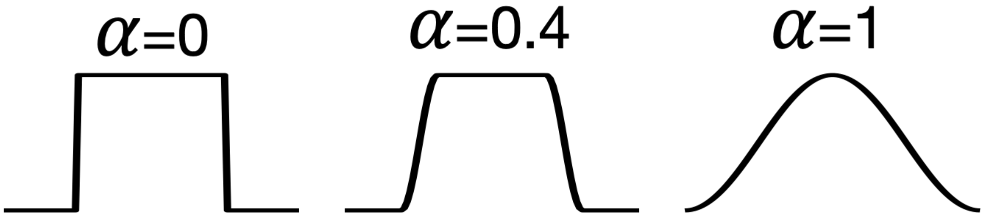


Figure 5.19: A demonstration of 1D Tukey windows. Selecting a larger α value will reduce the proportion of the window represented by the uniform rectangular window. All of these windows have the same FWHM.

A 1D discrete Tukey window w with N points is defined in equation 5.15. As Tukey windows are symmetric, $w_n = w_{(N+1)-n}$ in all cases. A discrete 2D window Cartesian window can be formed based on a Cartesian interpolation of a 1D window at each pixel radius.

$$
\begin{aligned}
w_n &= \frac{1}{2}\left(1 - cos\left(\frac{2n\pi}{N\alpha}\right)\right); && 1 \leq n < \frac{N\alpha}{2} \\
w_n &= 1; && \frac{N\alpha}{2} \leq n < \frac{N}{2}
\end{aligned}
\tag{5.15}
$$

5.A.2 Apodisation

It is important that the edge of any window is tapered, or else additional ringing will occur in the image. Figure 5.20, demonstrates that a hard window edge more uniformly spreads out the energy in the PSF, leading to ringing in the image domain. A window with a tapered edge produces less ringing and is therefore preferred.

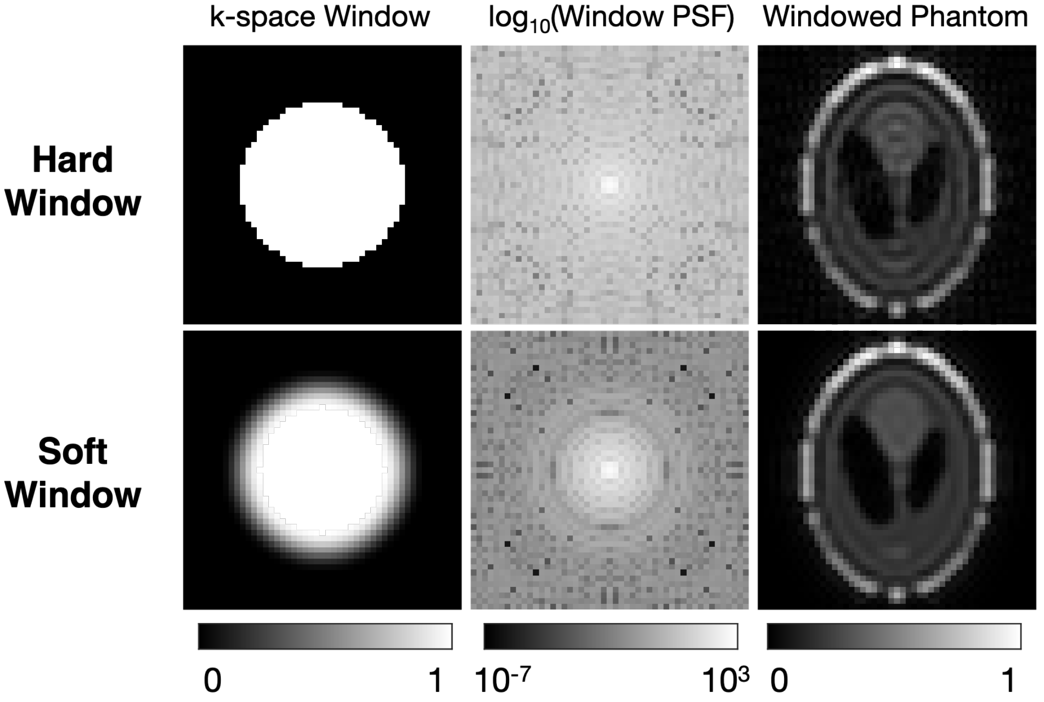


Figure 5.20: A demonstration of artefacts in windowing. Top row: a hard Tukey window with α=0. Bottom Row: a soft Tukey window with α=0.4.

5.B Adapted LRP and Tikhonov Constraints

Three separate adaptations of the Tikhonov-constrained approach (section 5.3) and LRP-constrained approach (section 5.4) were created in order to gain further understanding about the nature of those constraints. The terms which were first defined in the main text summary of section 5.5 are redefined here:

Pure Tikhonov/LRP The aforementioned Tikhonov and LRP-constrained reconstructions. The term "pure" is used to represent the formulations discussed in sections 5.3 and 5.4.

Adapted Tikhonov/LRP The general label for constrained-subspace reconstructions which incorporate Tikhonov and/or LRP information but do not follow the formulation of the pure approaches. All of the constraints in this section are adapted constraints.

The terms for the separate adapted constraints are also redefined:

Hybrid Constraints One subspace is constrained by Tikhonov and one subspace is constrained by LRP. The different combinations will be referred to as $X_{Tik}T'_{LRP}$ and $X_{LRP}T'_{Tik}$.

Tikhonov-Enhanced LRP (TE LRP) The priors are Tikhonov-constrained during generation. The final reconstruction is done with the same formulation as pure LRP-constrained reconstruction.

Null-Enforced LRP (NE LRP) The LRP enforcement is done via minimising the multiplication of the variable with the orthogonal complement of the prior (which maps the prior null space), rather than through minimising the euclidean distance between the variable and prior. This allows for use of a prior which rejects noise-like components.

5.B.1 Hybrid Constraints

The generic equation 5.3 permits the combination of different constraints within the same reconstruction. This section explores which constraint dominates in such a combined reconstruction, and whether this hybrid approach could outperform the pure approaches.

5.B.1.1 Spatial Tikhonov & Temporal LRP Constraints

The motivation behind combining spatial Tikhonov and temporal LRP constraints matches the motivation of the k-t PSF approach - namely that the temporal components from a densely sampled temporal domain are considered informative in a reconstruction, but the associated spatial information is not (unless the spatial domain is also densely sampled).

$$\hat{X}, \hat{T} = argmin_{X,T}\left(\|EXT' - d\|_2^2 + \lambda_X \|X\|_2^2 + \lambda_T \|T - T_{prior}\|_2^2 \right) \tag{5.16}$$

5.B.1.2 Spatial LRP & Temporal Tikhonov Constraints

The spatial domain is the source of aliasing when reconstructing an accelerated radial dataset, and so combining the spatial LRP constraint and the temporal Tikhonov constraint seems a logical step. The LRP-constrained optimisation aims to more greatly weight central k-space, a goal that can be theoretically achieved with just a spatial prior. However, a Tikhonov constraint could still usefully guide the temporal information towards a more favourable solution.

$$\hat{X}, \hat{T} = argmin_{X,T}\left(\|EXT' - d\|_2^2 + \lambda_X \|X - X_{prior}\|_2^2 + \lambda_T \|T\|_2^2 \right) \tag{5.17}$$

5.B.1.3 Hybrid Results & Discussion

R=20.11 was chosen to represent an acceleration factor at which both Tikhonov and LRP constraints perform equally, while improving upon the unconstrained-subspace results. Optimal values of $\lambda_X = \lambda_T = 1 \times 10^{-6}$ were used for both hybrid constraints. Values of $\lambda_X = \lambda_T = 1 \times 10^{-5}$ were used for the pure constraints.

Figure 5.21 shows that, between the two CCS scores, the T CCS scores more strictly correlate with the ROC AUC score. The temporal-subspace constraint then appears to be the most dominant factor in characterising the reconstruction quality. It should be noted that the presence of the spatial constraint does also make an impact, as having an unconstrained spatial subspace performed worse in all cases (see figures 6.3 and 6.4 in the next chapter). The LRP temporal constraint led to better X CCS scores and false positives more localised around the true activation area that the Tikhonov-constrained temporal subspace. However, the Tikhonov-constrained temporal subspace reconstructions produced almost perfect T CCS scores and better ROC curves (i.e. ROC curves with a sharper upper-left corner, particularly $X_{LRP}T'_{Tik}$).

The strong influence of the temporal constraint makes sense in the light of the time courses modelled in figure 5.15, and especially when considering the k-t PSF time course. The temporal components in the generated priors contain a lot of high-frequency content that is not present in the ground truth (or even in the k-t FASTER reconstruction). As such, it seems fair to conclude that the temporal constraint is the dominant underlying factor behind the ROC AUC scores seen in figure 5.21.

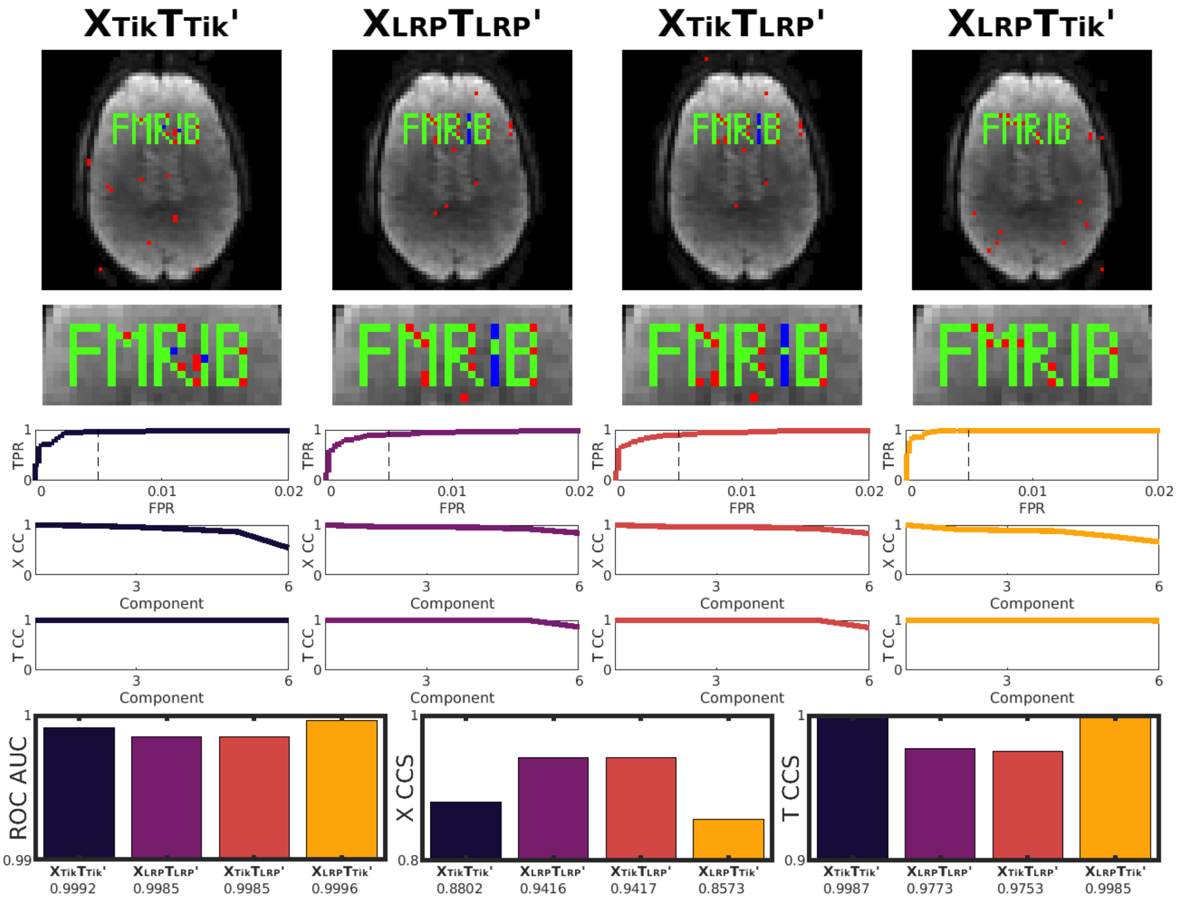


Figure 5.21: A comparison between different combinations of hybrid and pure Tikhonov/LRP approaches at R=20.11 (5 blades/frame). Sample z-stat parameter maps, an ROC curve, and the CC curves are shown in a columnwise structure for each reconstruction. Summative bar graphs for all reconstructions are shown on the bottom row. Figure 4.5 contains a detailed schematic of the graph format.

5.B.1.4 Hybrid Conclusion

Considering the artificial nature of this dataset, further testing is needed to evaluate whether the relation between temporal constraint and ROC curves hold in real data (which will be done in chapter 6). It is also worth noting that only the pure Tikhonov approach does not require an entire extra reconstruction (of the LRP), making the reconstruction time far shorter than the other approaches.

5.B.2 Tikhonov-Enhanced LRPs

In this section, the adapted constraint is formed by applying Tikhonov regularisation to the generation of the LRP. This is termed “Tikhonov-Enhanced LRP” (or TE LRP) to distinguish it from the pure LRP- and pure Tikhonov-constrained approaches.

The previous LRP-generation approach of section 5.4 was a windowed unconstrained-subspace reconstruction. A constrained-subspace reconstruction of the prior may also improve the final reconstruction. Section 5.3 demonstrated the benefit that adding Tikhonov constraints to a reconstruction can yield. A schematic is shown in figure 5.22 to help keep track of the different constraints, and the optimisation formulation used to generate TE LRPs is shown in equation 5.18.

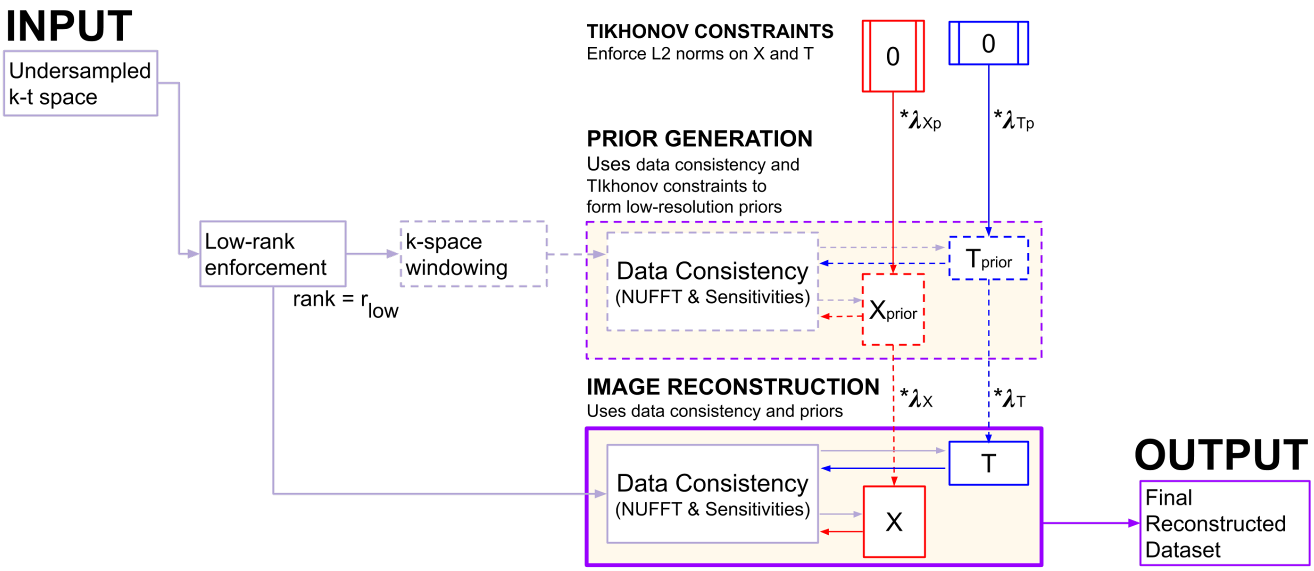


Figure 5.22: A schematic of LRP-constrained reconstruction, using Tikhonov-enhanced LRPs. The 0 boxes represent $\|A-0\| = \|A\|$ regularisation. X_{prior} and T_{prior} iteratively update until a convergence criterion is reached, and are then fed into the final reconstruction. Then, X and T iteratively update until a convergence criterion is reached, and XT' is accepted as the final reconstructed dataset.

$$\hat{X}_{prior}, \hat{T}_{prior} = argmin_{X_{prior},T_{prior}} \left(\|EX_{prior}T'_{prior} - Wd\|_2^2 + \lambda_{Xp}\|X_{prior}\|_2^2 + \lambda_{Tp}\|T_{prior}\|_2^2 \right) \tag{5.18}$$

5.B.2.1 TE LRP Results & Discussion

Figure 5.23 contrasts TE LRP-constrained reconstructions at R=25.13 and R=20.11 ($\lambda_{Xp} = \lambda_{Tp} = 1 \times 10^{-3}$) against pure Tikhonov and LRP-constrained reconstructions. Values of $\lambda_X = \lambda_T = 1 \times 10^{-5}$ were used for all final reconstructions. The T CCS scores of the TE LRP were far worse than the other approaches. The R=25.13 TE LRP result did improve the X CCS score and vastly improved the ROC AUC score (and ROC curve) over the R=25.13 pure LRP approach.

The TE LRP methods have even better localisation of errors around the active region than the LRP results, but the addition of the Tikhonov constraint to the generation of the LRP guided the final reconstruction to a more informative temporal subspace. Additionally, while the functional information in the R=25.13 TE LRP reconstruction was clearly better than the equivalent LRP-constrained reconstruction, the anatomical background of the z-stat parameter maps shows that there are likely some extra issues involved with the reconstruction that the FEAT analysis package struggled to parse (the RMSE for the TE LRP result was 10× higher than the other results at R=25.13).

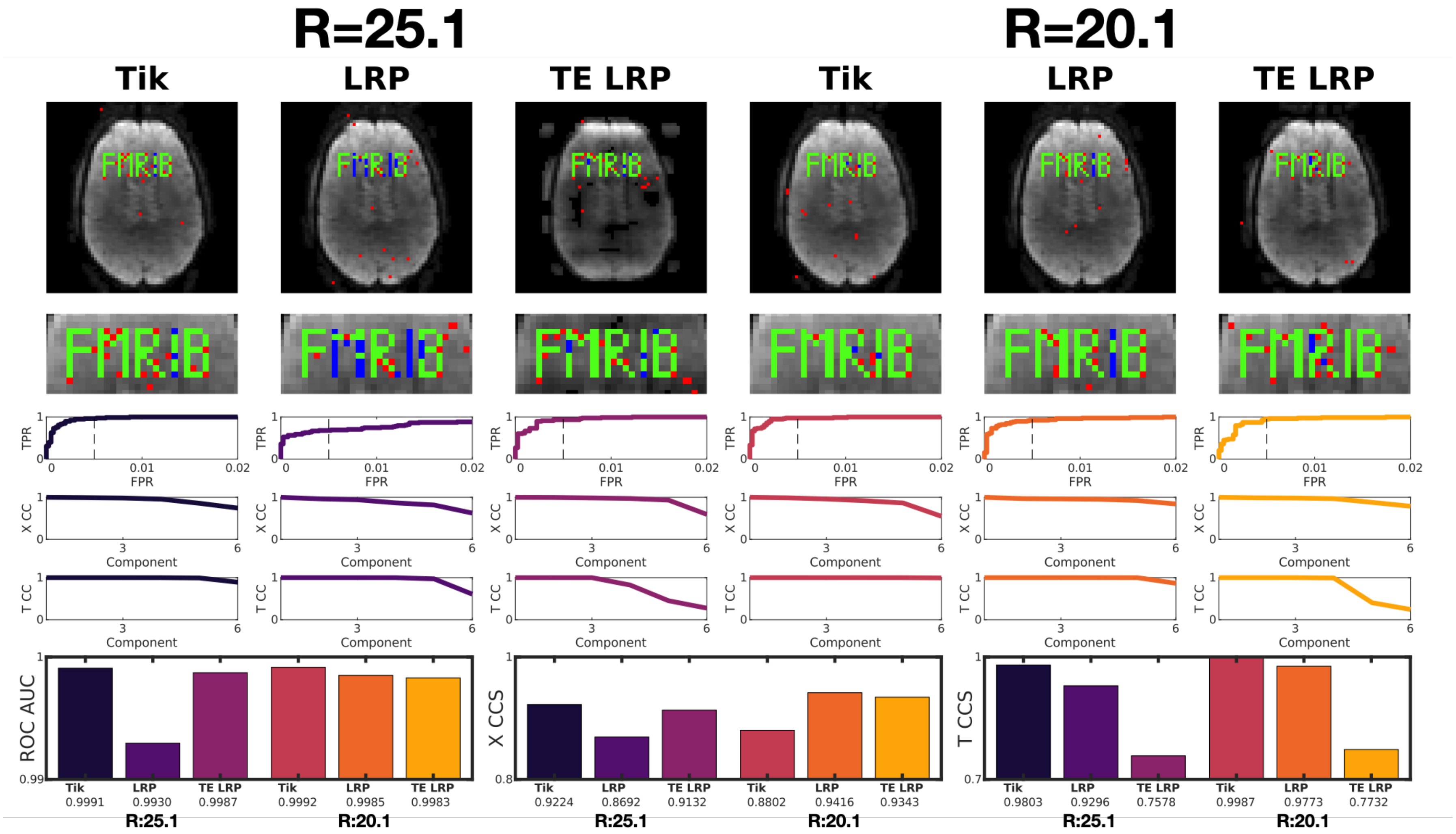


Figure 5.23: A comparison between Tikhonov/LRP-constrained reconstructions (at $\lambda_{Xp} = \lambda_{Tp} = 0$) and a TE LRP-constrained reconstruction (at $\lambda_{Xp} = \lambda_{Tp} = 1 \times 10^{-3}$) at R=25.13 (4 blades/frame) and R=20.11 (5 blades/frame). Sample z-stat parameter maps, an ROC curve, and the CC curves are shown in a columnwise structure for each reconstruction. Summative bar graphs for all reconstructions are shown on the bottom row. Figure 4.5 contains a detailed schematic of the graph format.

5.B.2.2 TE LRP Conclusion

While the reduction in temporal CCS in TE LRP-constrained reconstruction is unexpected, the improved ROC AUC score (with corresponding z-stat maps) and X CCS scores highlight TE LRPs as a method with some potential, if further understanding can be gained as to why the method falling short in some aspects.

5.B.3 Null-Enforced LRPs

For this section, an alternate way to implement the similarity terms of the form $\|A - A_{prior}\|_2^2$ will be formulated that allows for forms of A_{prior} that do not contain the same number of components as A.

5.B.3.1 NE LRP Theory

The rank subspace and null subspace of a matrix are fundamentally linked by the orthogonal complement (see section 3.4.2). An equation which forces similarity between a parameter A and the row space of a hypothetical matrix A_{prior} is also enforcing dissimilarity with the left null subspace $Null(A'_{prior})$. So far, the only term used to enforce similarity has been an L2 error norm.

$$\lambda_A \|A - A_{prior}\|_2^2 \tag{5.19}$$

An implicit restriction on the A_{prior} constraint is that it must be the same dimensionality as A, but what if this were not desirable? The poor temporal information resulting from LRP-constrained reconstructions has already been demonstrated in the artificial FMRIB dataset, but the more realistic retrospective fMRI dataset from section 4.4.1 produces the same result. This can be seen in the frequency content of the temporal components of an unwindowed k-t FASTER reconstruction (figure 5.24 at R=15.71, 10 blades/frame) when compared to the windowed k-t FASTER reconstruction of an LRP (figure 5.25, also at R=15.71). This means that the poor temporal information is likely not a result of the artificial FMRIB dataset time courses, but more likely a feature of the reconstruction.

In the unwindowed reconstruction, most of the temporal frequency information was concentrated in the central frequency bands (figure 5.24). In the generated LRP, the temporal frequency information was spread over a much greater range (figure 5.25), which might provide further insight into what was happening in the priors generated at a variety of window sizes (figure 5.9). More specifically, it can be seen that at least 40% of the energy was concentrated within a range of -0.1Hz to 0.1Hz for all components of the final reconstruction, but only four components in the LRP fulfilled this criterion (components fulfilling the criterion are shown in green, those failing are shown in red).

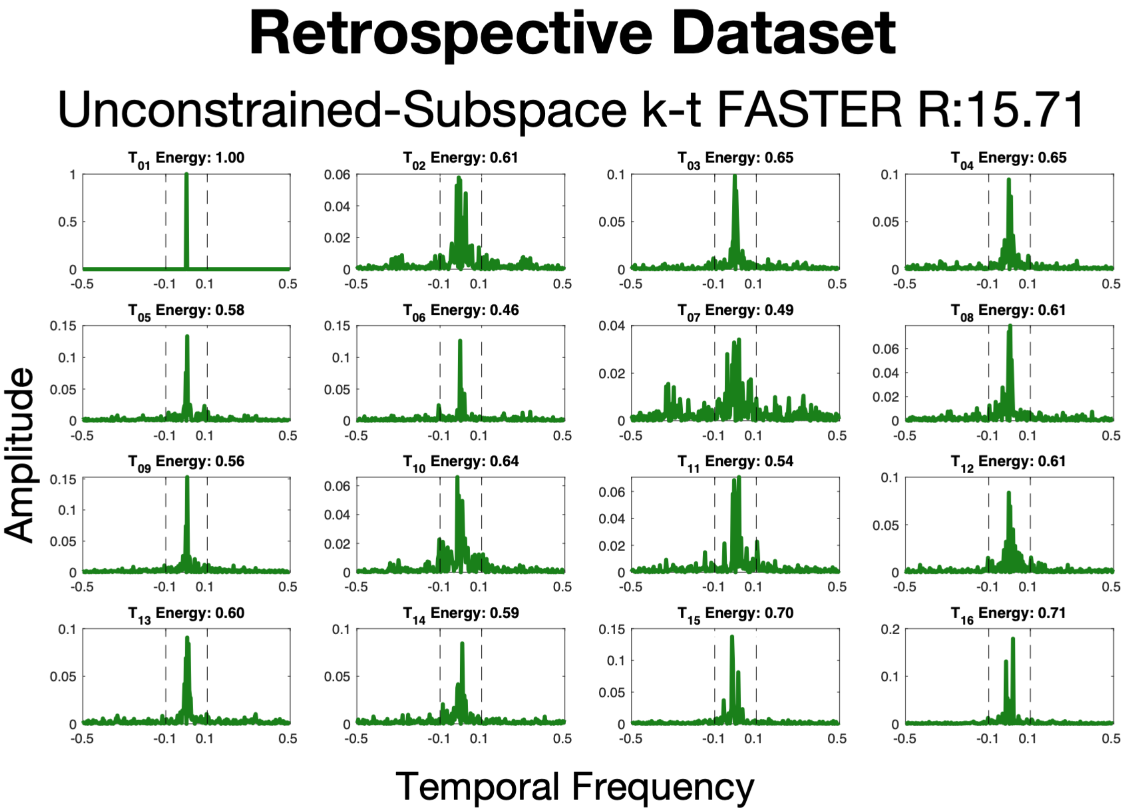


Figure 5.24: The normalised power spectrum in the components of an unconstrained-subspace k-t FASTER reconstruction at R=15.71. The retrospective dataset has a TR of 1s, meaning the maximum temporal frequency that can be represented is ±0.5Hz. Within each temporal component (T_{xx}), the proportion of energy contained within the bounds -0.1 Hz and 0.1 Hz is shown in the sub-figure title. Where T_{xx} >0.4, the curve is shown in green.

A prior-consistency term assumes that the information in the prior is worth incorporating as guidance for the final reconstruction. However, in figure 5.25, it can be seen that the energy in most components are composed of very high-frequency components, with temporal frequency curves are very spiked (unlike the smoother curves of the k-t FASTER reconstruction of figure 5.24). Even two of the components which “pass” the threshold test seem to contain some high spikes at high frequencies. In the components with dispersed energy, the similarity metric of equation 5.6 is enforcing similarity with an uninformative time-series. This is at best a useless enforcement and at worst a harmful one. Being able to remove undesirable components and enforce a comparison only with the remaining useful components would likely be a preferred outcome.

The k-t FASTER reconstruction shows that there are still at least 16 components of useful infor-

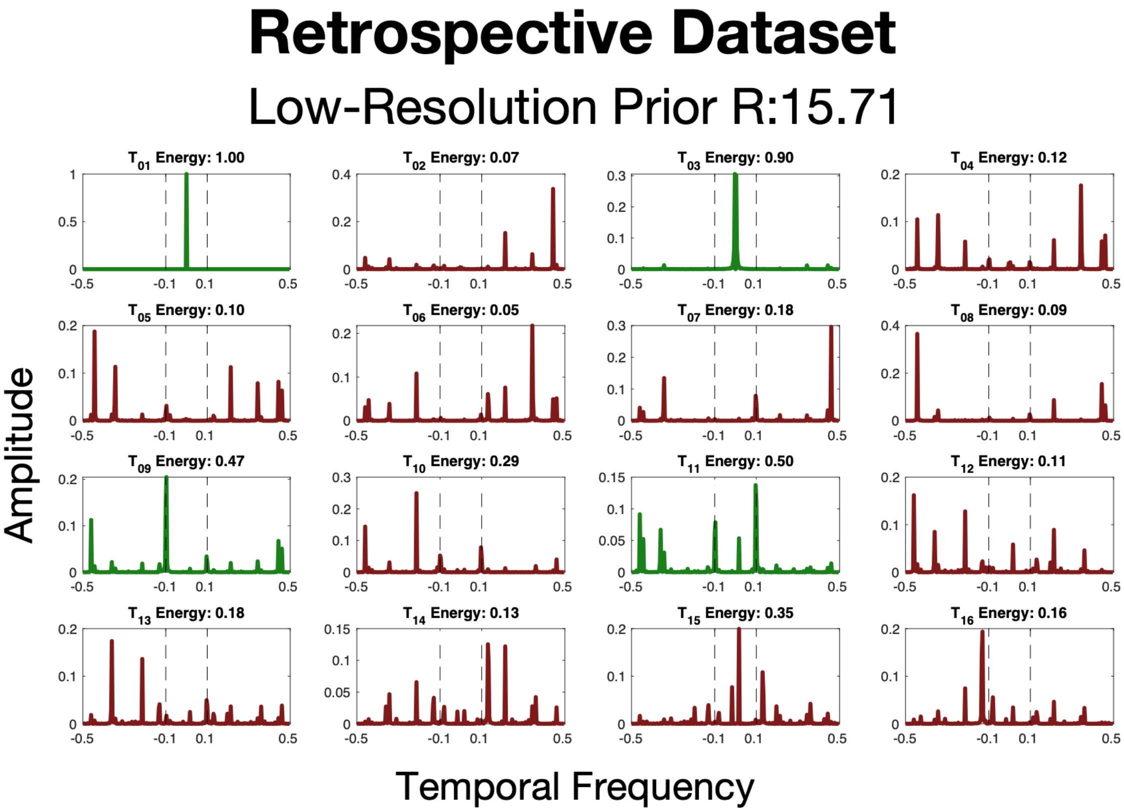


Figure 5.25: The normalised power spectrum in the components of a generated LRP at R=15.71 (windowed with a FWHM at $\frac{k_{max}}{10}$). The retrospective dataset has a TR of 1s, meaning the maximum temporal frequency that can be represented is ±0.5Hz. Within each temporal component (T_{xx}), the proportion of energy contained within the bounds -0.1 Hz and 0.1 Hz is shown in the sub-figure title. Where T_{xx} >0.4, the curve is shown in green.

mation that can be garnered from the final reconstruction. Trying to apply only the four useful temporal components in the LRP (using the $T_{xx} > 0.4$ metric) through a data consistency term will produce a dimensional mismatch. Given the lack of energy at low temporal frequencies in the components which fail the threshold, it is not certain that simply applying a low-pass frequency filter to the LRP temporal subspace would necessarily produce anything more informative.

An alternate way to enforce similarity with a reduced set of components of T_{prior} is by minimising the product of T and the orthogonal complement of T_{prior} (which is calculated as the null of T'_{prior}):

$$\lambda_T \|T_{prior}^{\perp} T\|_2^2 \tag{5.20}$$

Reformulating the LRP optimisation problem (equation 5.6) in the null-multiplication form specified by equation 5.20 produces the following formulation, will be termed the "Null-Enforced LRP" formulation (or NE LRP).

$$\hat{X}, \hat{T} = argmin_{X,T}\left(\|EXT' - d\|_2^2 + \lambda_X \|X_{prior}^{\perp} X\|_2^2 + \lambda_T \|T_{prior}^{\perp} T\|_2^2 \right) \tag{5.21}$$

5.B.3.2 NE LRP Methodology

Within the formulation of equation 5.21, there are two tests that should be carried out. The first test investigates whether the NE LRP formulation enforces similarity in instances where X_{prior} and T_{prior} are the same dimensionality as X and T (e.g. no components are removed). That formulation should then be compared against the similarity enforced in an $\|X - X_{prior}\|_2^2$ system (equation 5.6). A second test compares reconstructions using the null formulation with a full size prior and with a reduced prior. The evaluations are carried out concurrently using the artificial FMRIB dataset.

5.B.3.3 NE LRP Results & Discussion

Figure 5.26 shows the results of reconstructions at R=33.51 (3 blades/frame) and R=25.13 (4 blades/frame) for the LRP, the NE LRP using a reduced prior: N$_{RED}$(LRP), and the NE LRP using a full prior: N$_{FULL}$(LRP). In these highly accelerated regimes, only one component of the LRP was considered significant enough to include in the reduced NE LRP formulations. The full NE LRP performed equivalently to the typical LRP enforcement through all metrics, satisfying the first test by confirming that the null-enforcement approach does indeed work as an equivalent method through which to enforce the prior information.

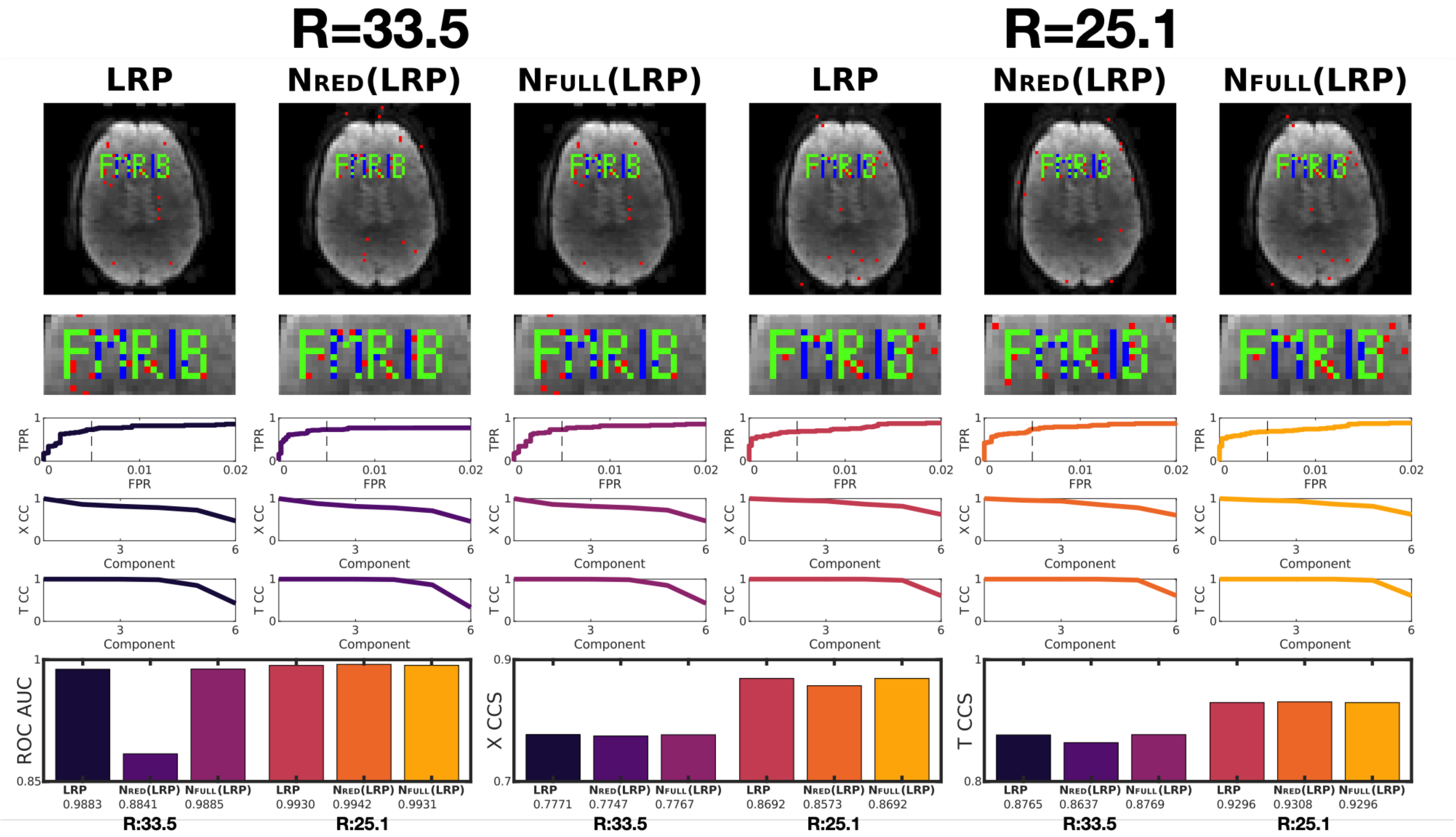


Figure 5.26: A comparison between the LRP-constrained reconstructions enforced through subtraction: LRP, null-enforced at reduced size: N_{RED}(LRP), and null-enforced at full size: N_{FULL}(LRP). Shown are reconstructions of the FMRIB dataset at R=33.51 (3 blades/frame) and R=25.13 (4 blades/frame). Sample z-stat parameter maps, an ROC curve, and the CC curves are shown in a columnwise structure for each reconstruction. Summative bar graphs for all reconstructions are shown on the bottom row. Figure 4.5 contains a detailed schematic of the graph format.

Compared to the LRP, The reduced NE LRP marginally improved the T CCS score and the ROC AUC for R=25.13 (and also for R=20.11, not shown), but dramatically reduced the ROC AUC score for R=33.51. The z-stat map at R=33.51 shows no noticeable quantitative difference, but the ROC curve indicates that the low score likely comes from a failure to identify the missing components beyond the z-stat threshold chosen. The ROC AUC score is even worse than the equivalent unconstrained-subspace k-t FASTER score at R=33.51, although the ROC curves in figure 5.26 indicate that the reconstruction is still improved in the range of relevant z-statistic thresholds in the reduced NE LRP case.

It is worth restating that only the first component in the LRP formulation passes the 40% energy threshold at all the above acceleration factors. The regime in which the NE LRP formulation would likely do best is in a low SNR situation where some components of the data are still accurately reconstructed. This regime does not appear to exist for the artificial FMRIB dataset, unless changes are made to the LRP formulation (e.g. a wider window size when generating the LRP).

5.B.4 Adapted Constraints Conclusion

In investigating the different forms of the subspace constraints, several interesting discoveries were made. The first discovery emphasises the importance of the temporal subspace constraint in AM reconstruction, as it was found that the hybrid constrained-subspace reconstructions were most strongly characterised by whichever constraint was acting on the temporal subspace.

Additionally, the poor LRP temporal subspace is not aided by additional Tikhonov-enhancement when generating the LRPs, and in fact overweighting the Tikhonov constraints makes the components more non-orthogonal (i.e. have a larger inner product and are more correlated).

Finally, the LRP-constrained reconstruction is not aided by null-enforcement of the LRP (whether reduced or otherwise). Only one component was deemed to have the energy focused in low temporal frequency at the acceleration factors where a constrained-subspace reconstruction would be used. As such, the LRP may need an alteration more fundamental than applying the components in a different manner. It is important to note that these conclusions are drawn from an artificial dataset, and not an fMRI dataset.

Bibliography

[1] Mark Chiew, Stephen M. Smith, Peter J. Koopmans, Nadine N. Graedel, Thomas Blumensath, and Karla L. Miller. k-t FASTER: Acceleration of functional MRI data acquisition using low

rank constraints. *Magnetic Resonance in Medicine*, 74(2):353–364, 2015.

[2] Mark Chiew, Nadine N. Graedel, Jennifer A. Mcnab, Stephen M. Smith, and Karla L. Miller. Accelerating functional MRI using fixed-rank approximations and radial-cartesian sampling. *Magnetic Resonance in Medicine*, 00:1–12, 2016.

[3] Yehuda Koren. Collaborative filtering with temporal dynamics. *Communications of the ACM*, 53(4):447–456, 2009.

[4] Zhi-Pei Liang. Spatiotemporal Imaging with Partially Separable Functions. *IEEE International Symposium on Biomedical Imaging*, 2:988–991, 2007.

[5] Henrik Pedersen, Sebastian Kozerke, Steffen Ringgaard, Kay Nehrke, and Yong Kim Won. k-t PCA: Temporally constrained k-t BLAST reconstruction using principal component analysis. *Magnetic Resonance in Medicine*, 62(3):706–716, 2009.

[6] Xuesong Li, Xiaodong Ma, Lyu Li, Zhe Zhang, Xue Zhang, Yan Tong, Lihong Wang, Sen Song, and Hua Guo. Dual-TRACER: High resolution fMRI with constrained evolution reconstruction. *NeuroImage*, 164(February 2017):172–182, 2018.

[7] Mark Chiew and Karla L. Miller. Improved statistical efficiency of simultaneous multi-slice fMRI by reconstruction with spatially adaptive temporal smoothing. *NeuroImage*, 203(August):116165, 2019.

6 | Implementing Constrained Low-Rank Reconstruction

Contents

In chapter 5, the theoretical basis behind a set of constrained-subspace approaches to low-rank reconstruction were explored in an artificial dataset. In this chapter, the same set of constraints will be implemented in an fMRI framework, including the addition of coil sensitivity maps.

6.1 Introduction

The AM framework proposed in chapter 4 was shown to produce equivalent (if not slightly improved) results on the retrospective dataset when compared to SVD k-t FASTER implemented with IHT+MS [1]. The main benefit of the AM method is the ease with which additional subspace-specific constraints can be incorporated into the reconstruction. Chapter 5 outlined a selection of such constraints and demonstrated their effective power to guide the optimisation algorithm towards favourable solutions.

These constraints will now be tested on more realistic datasets in order to evaluate their potential for adoption in actual fMRI data acquisition. In addition to reconstruction quality (evaluated with ROC AUC, X CCS, T CCS, and a visual assessment of parameter maps), other factors which may impact the practical implementation of an optimal constrained reconstruction will also be considered (e.g. reconstruction time, number of design parameters which are scan-dependent, how independent those design parameters are).

The Tikhonov-constrained reconstructions and LRP-constrained reconstructions will be considered together, as constraints which operate on both subspaces. The temporal smoothing constraints will be explored as an alternate spatiotemporal structure that can be exploited for acceleration, with particular focus on the effectiveness of temporal subspace smoothing in noisy and/or highly accelerated datasets.

6.2 Methodology

Unless otherwise specified the standard pipeline in section 4.4.1 will be followed, although key points will be re-outlined and any changes will be explicitly stated.

6.2.1 Datasets

The first dataset used in this chapter is the "retrospective dataset" previously outlined in section 4.4.1, although it will sometimes be referred to as the "clean retrospective dataset" to emphasise that it contains no noise. As a reminder, the clean retrospective dataset represented a single slice of fMRI data (100×100) over 300 frames (TR_{frame}:1s) and 8 compressed coils. The original dataset was generated from five blocks of a 30s/30s on/off finger-tapping task. Resampling was done through a TURBINE trajectory [2]. A z-statistic threshold of $z>3.1$ was used to determine the ground truth.

The high acceleration factor achieved with constrained-subspace reconstructions in the clean retrospective dataset (R=52.36, 3 blades/frame, figure 6.6) was higher than what would be considered reasonable in vivo. In order to address this, a “noisy retrospective dataset” was created. Complex Gaussian noise was applied in k-t space to the retrospective dataset at R=31.42 (five blades/frame). R=31.42 was chosen as the acceleration factor as it represents the undersampling factor at which the constrained-subspace approaches started clearly outperforming the unconstrained-subspace approach in the clean retrospective dataset (figure 6.5).

Using the SNR defined in equation 5.4, noise was added at three distinct magnitudes to generate high SNR data (SNR=100), medium SNR data (SNR=50) and low SNR data (SNR=20). These values are more representative of actual fMRI SNR values [3] than the artificial dataset SNR of 1000 used in chapter 5. At each SNR, five unique instantiations of complex noise were created (i.e. 15 different noisy undersampled datasets in total, which are referred to collectively as the noisy retrospective dataset). This additional Gaussian noise represents thermal fMRI noise, and merely represents a step towards more realistic data (other noise sources such as physiological noise and temporal autocorrelation are still not included). Evaluations of reconstruction quality in this dataset were done against the noiseless ground truth using the same methodology as the clean retrospective dataset.

The final dataset, which we will term the “prospective dataset”, refers to a short scan time (320s) experiment that collected task fMRI data across five 30s on/off epochs. The prospectively accelerated reconstructions used real TURBINE data [2] collected across eight slices (100×100), with the slices covering the primary motor cortex. The short experiment and the long experiment were carried out consecutively on the same subject, and used the same acquisition parameters/experimental set-up (TR_{blade} = 50ms, TE = 30ms, flip angle = 15°, BW = 1786 Hz/px).

The sampling was done prospectively and reconstructions were performed on the raw k-space, such that this dataset represents a real experimental setup (unlike the retrospectively re-sampled data). As there is no ground truth in real data, an experiment with a long scan time (640s) but otherwise identical experimental conditions (ten 30s on/off task epochs) was collected to provide a fully sampled comparison. An R=1.05 reconstruction of the long dataset contains enough temporal DoF to characterise the underlying functional signal and provide high-quality activation maps. This serves as a close approximation to the "ground truth" that the accelerated short dataset is attempting to represent. The different acceleration factors in the short prospective dataset (R=7.85, R=15.71, R=26.18) will lead to different temporal resolutions and temporal degrees of freedom, as well as affecting other statistical properties (such as physiological noise variance). The reconstruction details are listed in table 6.1.

The eight slices of the prospective dataset were reconstructed independently. For both the long

Dataset	**Blades**	**TR**(s)	**R**	$\frac{\textbf{Blades}}{\textbf{Frame}}$	**Frames**
Long	12,800	7.5s	1.05	150	85
Short	6,400	1.0s	7.85	20	320
Short	6,400	0.5s	15.71	10	640
Short	6,396	0.3s	26.18	6	1066

Table 6.1: The reconstruction details at different acceleration factors for the prospective dataset.

and the short dataset, z-statistic maps were created with FEAT [4], null-corrected using mixture modelling, and combined for volumetric modelling analysis. The mixture modelling will help account for the different confounding factors resulting from different temporal resolutions, allowing the analysis to focus on acceleration factor. A z-statistic threshold of z>4.8 was used on the reconstructed R=1.05 long prospective dataset to represent the truth in ROC curve comparisons (figure 6.1). As with the retrospective dataset, this z-stat threshold was chosen heuristically to match the expected statistical parameter map for a finger-tapping experiment. Short dataset slices were registered to the equivalent R=1.05 slice of the long dataset with FLIRT [5, 6, 7] (a linear registration tool) between the FEAT and mixture modelling processing steps. z-stat maps will be shown at a z threshold corresponding to a FPR of 0.0015 (the same threshold as the retrospective dataset). While it would be possible to construct the eight slices together to increase the shared temporal information and therefore improve reconstruction quality, the extra computational power to do so would greatly increase the reconstruction time.

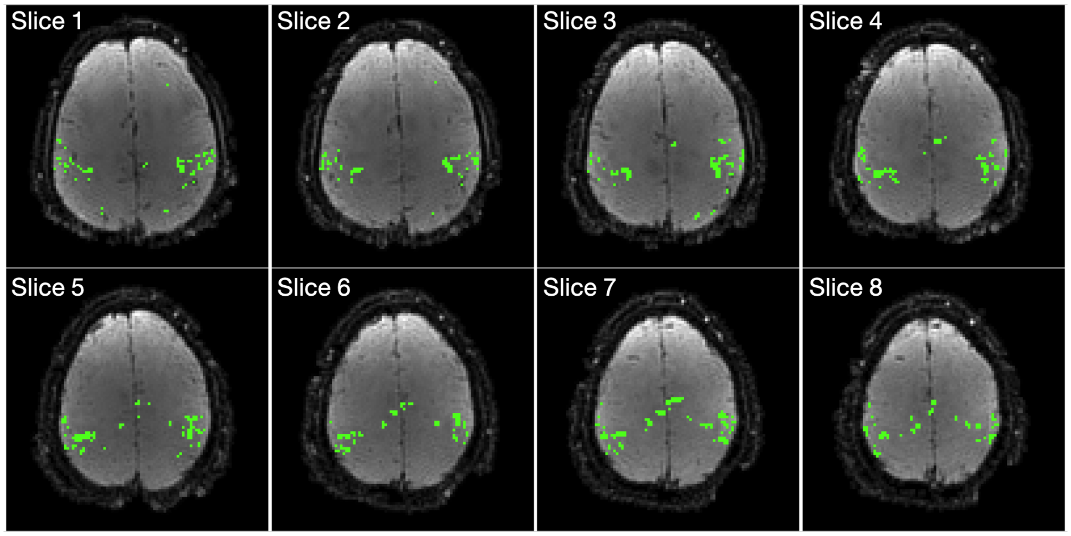


Figure 6.1: The true z-statistic parameter map for the eight slices of the long prospective dataset. The voxels with a z-statistic significance >4.8 are shown in green, and lain over the mean brain intensity image for each slice analysed.

6.2.2 Evaluating Constraints

Using an equivalent long dataset rather than a ground truth for validation will require some adjustments to the interpretation of the prospective results. Firstly, the ROC AUC scores are expected to be worse for an equivalent R in this prospective dataset, due to some inherent mismatch in the active functional areas that will naturally occur between scans. Without a ground truth the CCS is a less effective measure of reconstruction quality, particularly the T CCS as temporal interpolation is required to be able to match temporal components with different DoF. As such, CCSs will not be considered for the prospective dataset. Finally, a convergence criterion of $\epsilon = 1 \times 10^{-3}$ was found to be optimal for the prospective dataset (the robustness figures in section 4.5.1 can be used to justify that this value of ϵ will still yield a high level of robustness), although $\epsilon = 1 \times 10^{-5}$ was still used for the generation of the LRP. This lower convergence criterion was found to produce improved ROC AUC scores, which may be a result of overfitting occurring at the more precise criterion. It should be noted that the results which converged to $\epsilon = 1 \times 10^{-5}$ were still reasonable, and that the improved performance at $\epsilon = 1 \times 10^{-3}$ may be influenced by the comparison with the ground truth, rather than indicating a tangible reconstruction improvement.

With all methods, the value of the regularisation-weighting term λ was chosen to be optimal with respect to a combination of the X CCS, the T CCS, the ROC AUC, and a visual inspection of the activation to ensure a measured selection process (for the prospective dataset, only the ROC AUC and activation maps were used). A full set of λ grids will be shown in the retrospective dataset to demonstrate the effect of different λ pairings for the Tikhonov and LRP constraints (a 1D graph will suffice for temporal smoothing). Once the process for selecting optimal λs has been established, the optimal λ will be simply stated for a particular constraint/dataset/R combination, rather than restate the process and plot many large dimensionality graphs of similar appearance.

6.3 Tikhonov and LRP-Constrained Reconstructions

In this section, the effectiveness of Tikhonov-constrained reconstructions and LRP constrained reconstructions are evaluated in realistic data. The constraints will be compared against an unconstrained-subspace AM k-t FASTER reconstruction and an adapted k-t PSF reconstruction.

The reconstructions were first tested on the clean retrospective dataset at a variety of λ values, the summative results of which are shown on a grid. The grid search template is shown in figure 6.2 to demonstrate the range of λ covered, including boundary conditions at $\lambda = 0$ (zero influence from the constraint) and $\lambda = \infty$ (the solution is identical to the subspace constraint). In the Tikhonov λ grid, setting $\lambda = \infty$ will set the component matrix coefficients to zero. The special boundary case of ($\lambda_X = 0$, $\lambda_T = \infty$) in the LRP λ grid defines k-t PSF, and the special case of ($\lambda_X = 0$, $\lambda_T = 0$) defines an unconstrained-subspace k-t FASTER approach in both the Tikhonov and LRP grids.

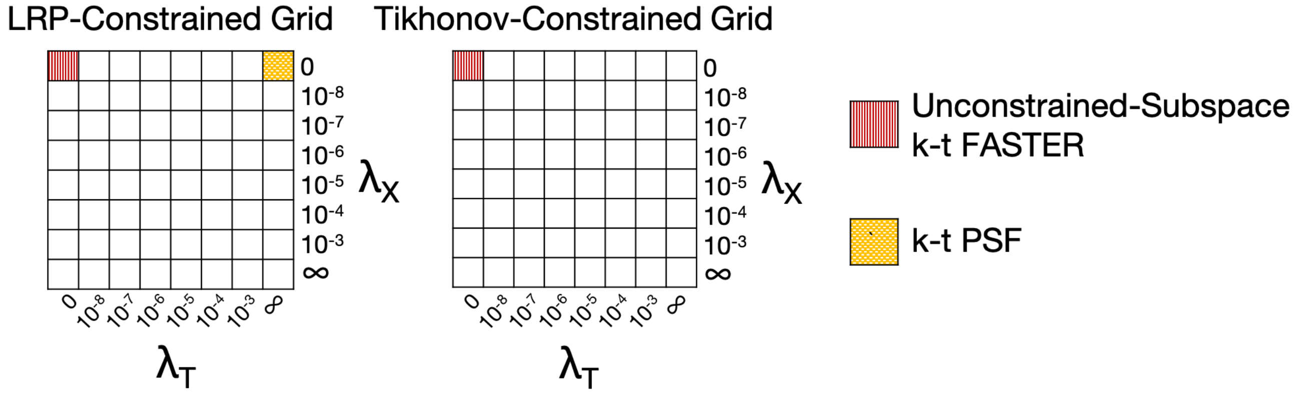


Figure 6.2: A λ grid template, demonstrating the location of the k-t FASTER result (top left corner) and the k-t PSF result (top right corner of LRP-grids). Each subplot will follow a logarithmic scale, except for the edges, which represent extreme values of λ. The left column/top row represents unconstrained subspaces, and the right column/bottom row represents reconstructions with hard subspace constraints.

6.3.1 Results

Figures 6.3 and 6.4 show the LRP-constrained and the Tikhonov-constrained CCS grids respectively at a range of acceleration factors. Critically, both grids show that improved reconstructions of the spatial subspaces can be found at λ pairings distinct from the λs which define k-t PSF and k-t FASTER. The LRP-constrained λ grid is defined by a distinct peak at a certain non-zero λ pairing, with the peak becoming more pronounced as R increases. The LRP grids are approximately defined by the product of two separable vectors - if a search in one dimension finds a peak in λ_X, that will be roughly the optimal λ_X value. These rough approximations are based on the search for optimal λ pairings across all datasets for the two constraints, and can reduce the LRP search from a 2D search to two 1D searches in an ideal setting, although the peak is hard to distinguish in the flatter parts of the grid.

The Tikhonov-constrained grid is defined by an approximate line of peak values for $\lambda_X\lambda_T = l$, where l is a constant. This indicates a one-dimensional search increasing values of l could suffice in finding the optimal Tikhonov λ pairing, reducing the number of unknowns that must be accounted for. An additional observation is that the Tikhonov solution became more low-rank at λ pairings which were too high, whereas using extreme λ values in the LRP grid did not lead to a reduction in rank. This is because overly high Tikhonov weighting will force the reconstruction into the most energetically favourable form, which purely represents the highest energy components. Similar

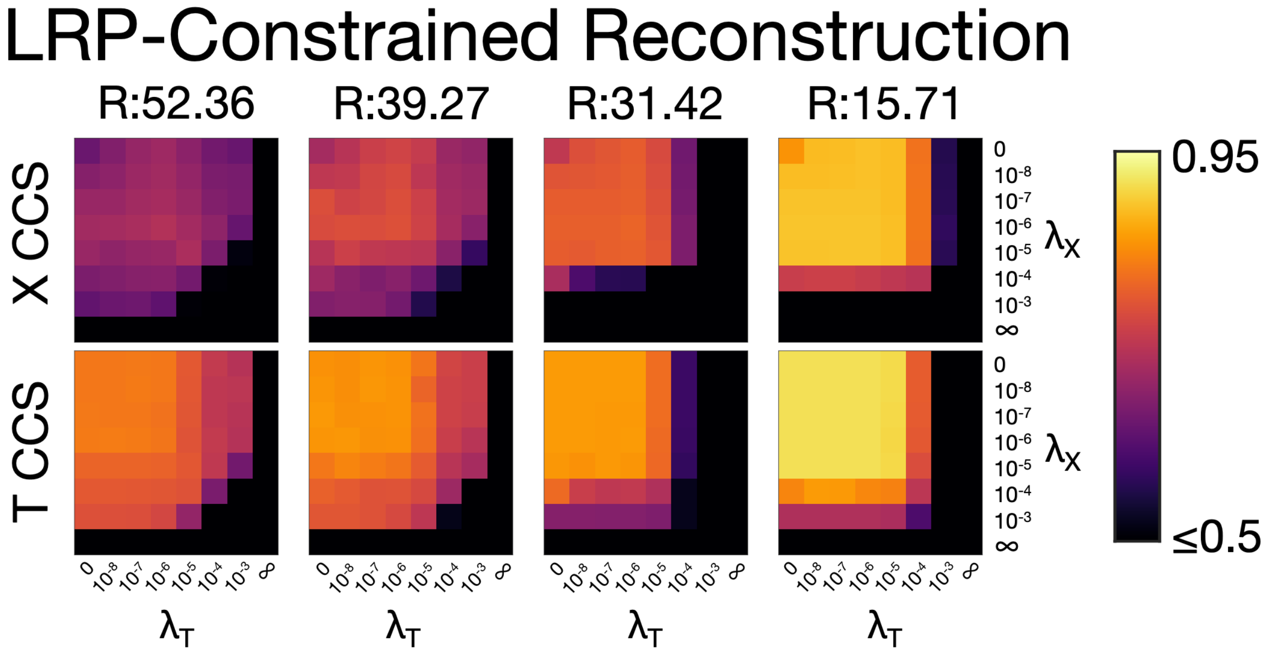


Figure 6.3: The spatial (top row) and temporal (bottom row) CCS scores for LRP-constrained reconstructions in the clean retrospective dataset across acceleration factors of R=52.36 (3 blades/frame), R=39.27 (4 blades/frame), R=31.42 (5 blades/frame) and 15.71 (10 blades/frame).

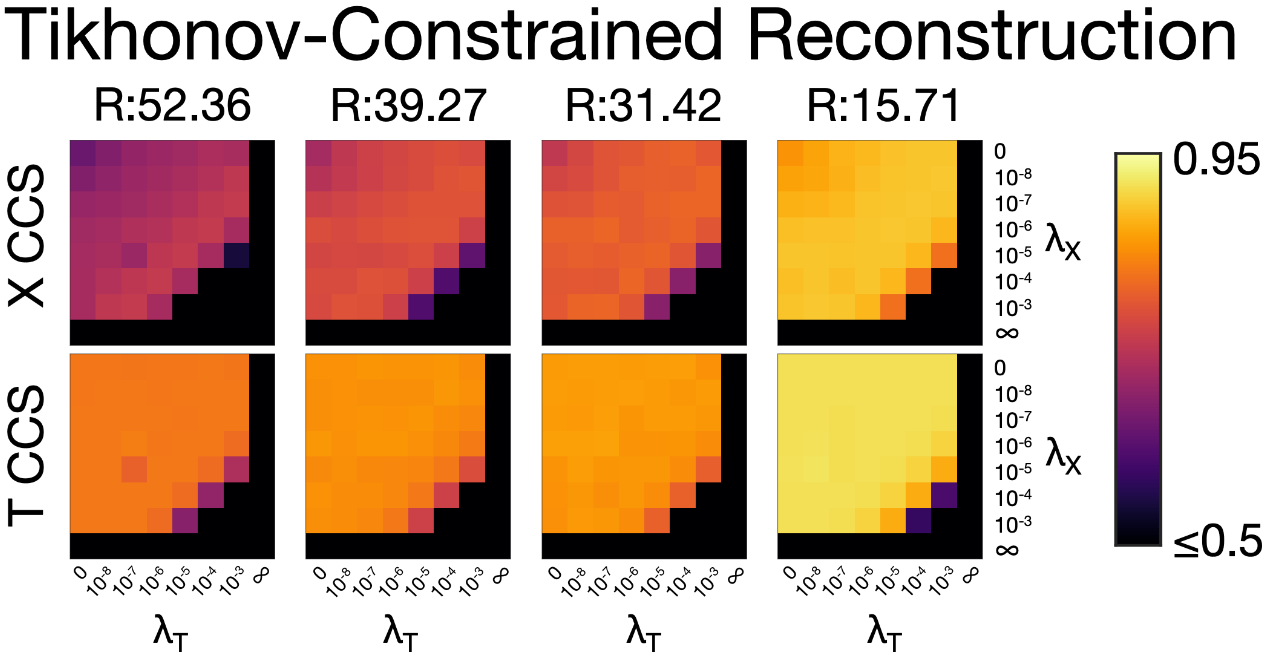


Figure 6.4: The spatial (top row) and temporal (bottom row) CCS scores for Tikhonov-constrained reconstructions in the clean retrospective dataset across acceleration factors of R=52.36 (3 blades/frame), R=39.27 (4 blades/frame), R=31.42 (5 blades/frame) and 15.71 (10 blades/frame).

to the LRP-constrained grid conclusions, this conclusion is a rough approximations based on the search for optimal λ pairings across all datasets.

An optimum λ pairing for the retrospective dataset was found to be approximately λ_X=λ_T=1×10^{-5} for both constraints across all R (except at high R in LRP). Figure 6.5 shows the CCS and ROC AUC scores across a wide range of R at λ_X=λ_T=1×10^{-5} for Tikhonov and LRP-constrained reconstructions, compared with k-t FASTER and k-t PSF. Under the AM formulation, k-t PSF performs poorly compared to the other methods (figure 6.5a). This erratic k-t PSF performance matches the results seen for the artificial dataset in section 5.4.5, with the poor temporal subspace reconstruction again a major factor in the low reconstruction scores.

The equivalent graph without the k-t PSF plot (figure 6.5b) shows that k-t FASTER and the constrained-subspace approaches perform similarly until R=15, when some improvements in temporal subspace and ROC curve can be seen in the constrained-subspace approaches. This roughly matches the expected acceleration from the literature [1]. Beyond R=31.42, Tikhonov-constrained reconstructions outperform LRP-constrained reconstructions. At R=39.27 and 52.36 the optimal λ pairings for the LRP approach were $\lambda_X = 1 \times 10^{-4}$ and $\lambda_T = 1 \times 10^{-6}$ respectively, which yielded scores for the LRP-constrained reconstruction which lay between the Tikhonov-constrained approaches and k-t FASTER. The temporal CCS is consistently worse for the LRP-constrained method when compared with k-t FASTER and Tikhonov-constrained reconstructions, reinforcing the aforementioned temporal subspace issue.

A more qualitative look at the optimal λ pairings for constrained reconstruction are shown in figure 6.6 at R=52.36 (3 blades/frame), figure 6.7 at R=39.27 (4 blades/frame), and figure 6.8 at R=31.42 (5 blades/frame).

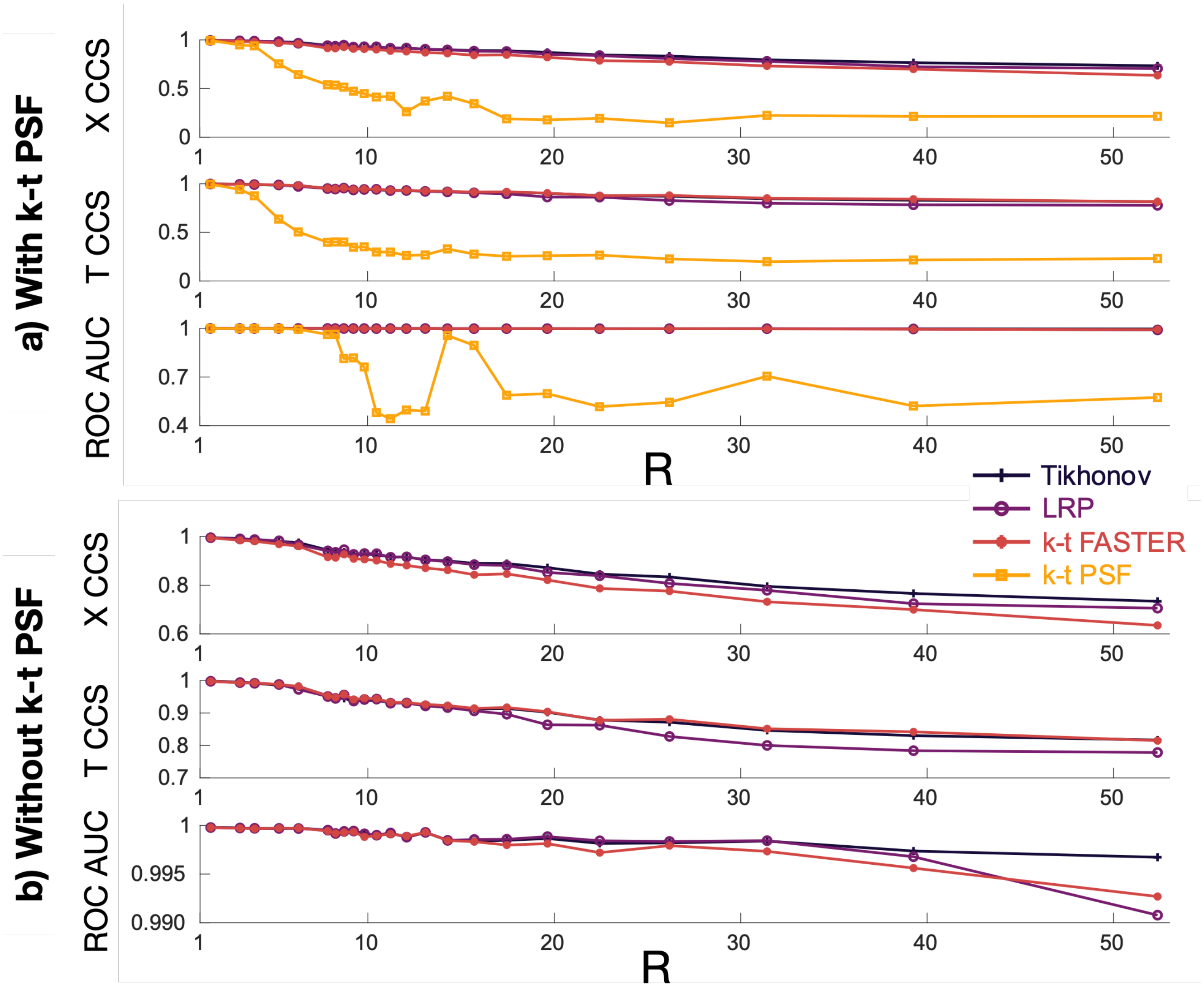


Figure 6.5: a) A line plot of Tikhonov-constrained reconstruction ($\lambda_X = 1 \times 10^{-5}$, $\lambda_T = 1 \times 10^{-5}$), LRP-constrained reconstruction ($\lambda_X = 1 \times 10^{-5}$, $\lambda_T = 1 \times 10^{-5}$), k-t FASTER and k-t PSF at different acceleration factors (R) for the Spatial CCS (top), Temporal CCS (middle) and AUC of the ROC curve (bottom) metrics. k-t PSF is not shown in the b) to allow a more focussed y-axis.

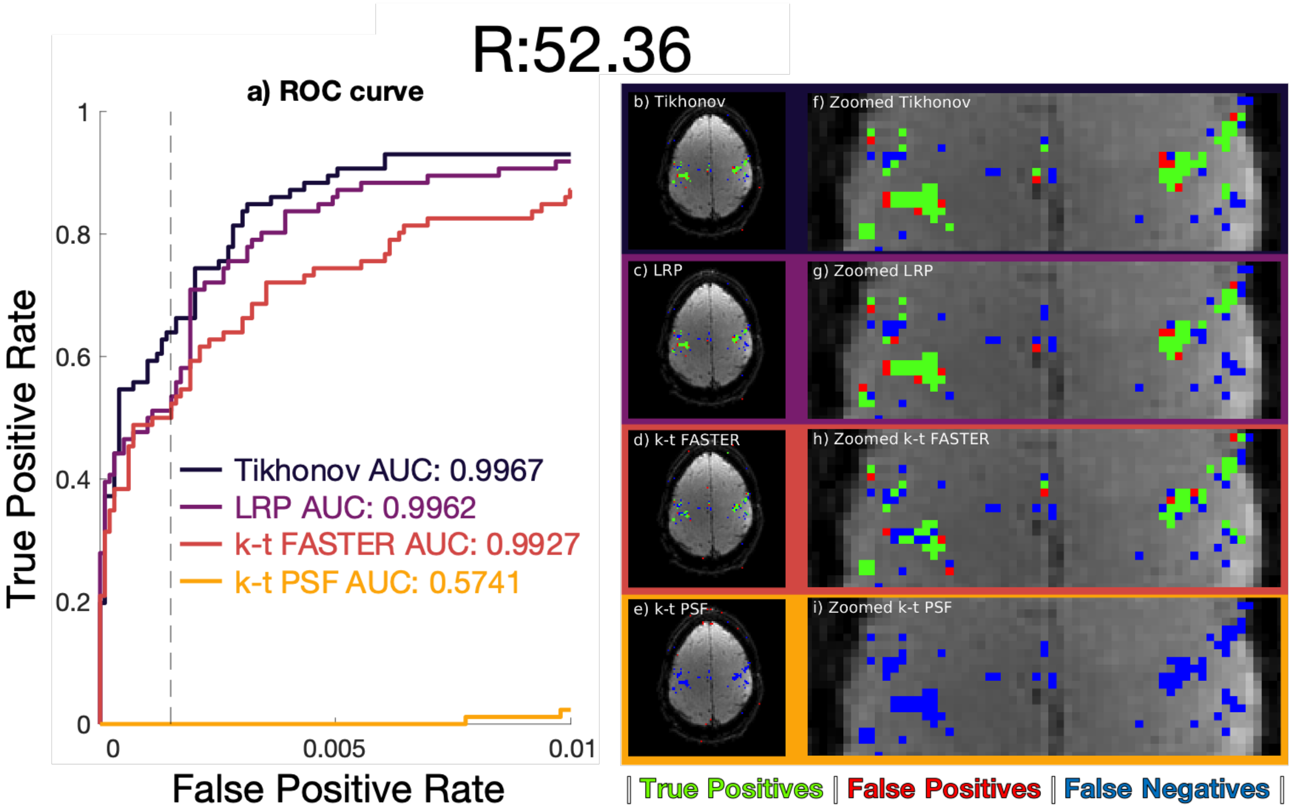


Figure 6.6: R=52.36 (3 blades/frame) retrospective dataset reconstructions. a) ROC curves, legend lists full curve AUCs. b)-e) Activation maps using a z-statistic corresponding to an FPR of 0.15%. f)-i) A medial zoom of the associated activation maps. b/f) Tikhonov: $\lambda_X = 1 \times 10^{-5}$, $\lambda_T = 1 \times 10^{-5}$, c/g) LRP: $\lambda_X = 1 \times 10^{-4}$, $\lambda_T = 1 \times 10^{-6}$, d/h) k-t FASTER, e/i) k-t PSF. Maps b)-i) use green true positive pixels, red false positives, and blue false negatives.

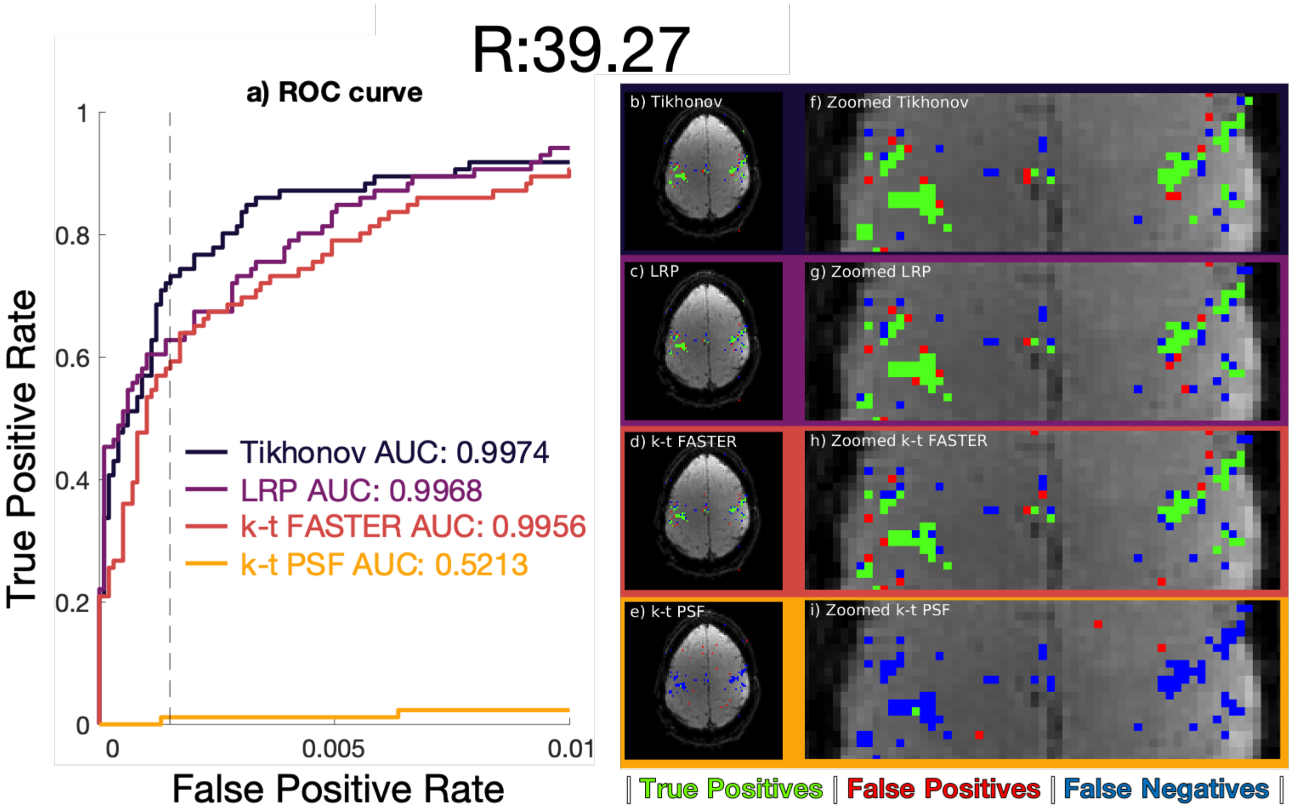


Figure 6.7: R=39.27 (4 blades/frame) retrospective dataset reconstructions. a) ROC curves, legend lists full curve AUCs. b)-e) Activation maps using a z-statistic corresponding to an FPR of 0.15%. f)-i) A medial zoom of the associated activation maps. b/f) Tikhonov: $\lambda_X = 1 \times 10^{-5}$, $\lambda_T = 1 \times 10^{-5}$, c/g) LRP: $\lambda_X = 1 \times 10^{-4}$, $\lambda_T = 1 \times 10^{-6}$, d/h) k-t FASTER, e/i) k-t PSF. Maps b)-i) use green true positive pixels, red false positives, and blue false negatives.

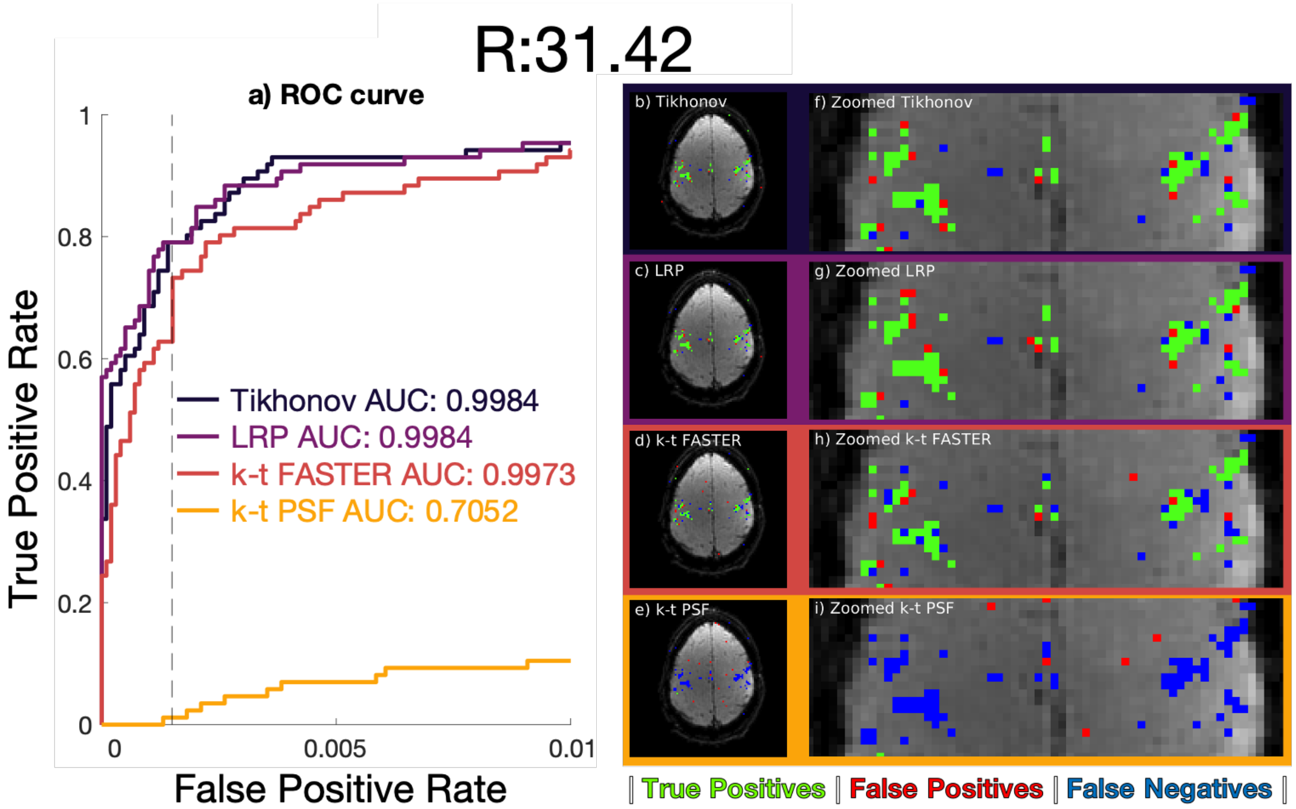


Figure 6.8: R=31.42 (5 blades/frame) retrospective dataset reconstructions. a) ROC curves, legend lists full curve AUC. b)-e) Activation maps using a z-statistic corresponding to an FPR of 0.15%. f)-i) A medial zoom of the associated activation maps. b/f) Tikhonov: $\lambda_X = 1 \times 10^{-5}$, $\lambda_T = 1 \times 10^{-5}$, c/g) LRP: $\lambda_X = 1 \times 10^{-5}$, $\lambda_T = 1 \times 10^{-5}$, d/h) k-t FASTER, e/i) k-t PSF. Maps b)-i) use green true positive pixels, red false positives, and blue false negatives.

Tikhonov, LRP and k-t FASTER reconstructions are observed to follow a broadly monotonic relationship between CCS scores and R (figure 6.5). While some improvement of Tikhonov-constrained reconstructions is noted over LRP-constrained reconstruction at high very R (figures 6.6 and 6.7, the improvement is slight and further testing was needed in a dataset with lower SNR to truly distinguish between the methods. This motivated the production of the noisy retrospective dataset (see section 6.2) at R=31.42, the highest R at which both constraints performed equivalently (figure 6.8). The optimal λ pairings in the noisy dataset were found to vary with SNR, but were optimal for both Tikhonov and LRP constraints (table 6.2).

Noise Level	SNR value	λ_X	λ_T
noiseless	∞	1×10^{-5}	1×10^{-5}
high SNR	100	1×10^{-4}	1×10^{-5}
medium SNR	50	1×10^{-4}	1×10^{-4}
low SNR	20	1×10^{-3}	1×10^{-4}

Table 6.2: The optimal λ pairings at different noise levels for the retrospective dataset at R=31.42.

The mean ROC AUCs of the noisy results are summarised in figure 6.9. As expected, all reconstructions captured a decreasing amount of functional information as SNR decreased. In t-tests performed between the different constraints within the three non-noiseless SNRs, all reconstructions within an acceleration factor were significantly separate ($p<0.05$) except Tikhonov vs LRP at high SNR and k-t FASTER vs k-t PSF at low SNR.

Figure 6.10 shows the reconstructed z-stat maps at one of the five noise instantiations at each SNR (the full range of z-stat maps can be seen in section 6.B). An ROC AUC of <0.9 was judged to be too poor to show any reasonable estimation of the activation area, and so a big variation below this value (such as in the k-t PSF results of figure 6.9) should not be considered overly meaningful. The value of 0.9 was derived from observing many z-statistic maps, and represents the point at which reliably localising activation becomes infeasible (e.g. the low SNR Tikhonov result (AUC: 0.8972) and the medium SNR k-t FASTER result (AUC: 0.8825) in figure 6.10).

Tikhonov-constrained reconstruction outperformed all other methods, identifying plausible activity even at the lowest SNR tested. LRP-constraints also performed well, with the medium SNR performing better than k-t FASTER's high SNR reconstructions. The unconstrained-subspace k-t FASTER approach appears highly susceptible to noise, with a roughly equivalent noiseless AUC score to the other methods at R=31.42 (figure 6.9) quickly decaying as SNR decreased. The more rigid k-t PSF approach failed to capture the activation even for a noiseless reconstruction at this acceleration. The soft approaches of Tikhonov- and LRP-constrained reconstructions appear to strike a balance in using constraints to guide the reconstruction towards a favourable solution in these artificially noisy examples.

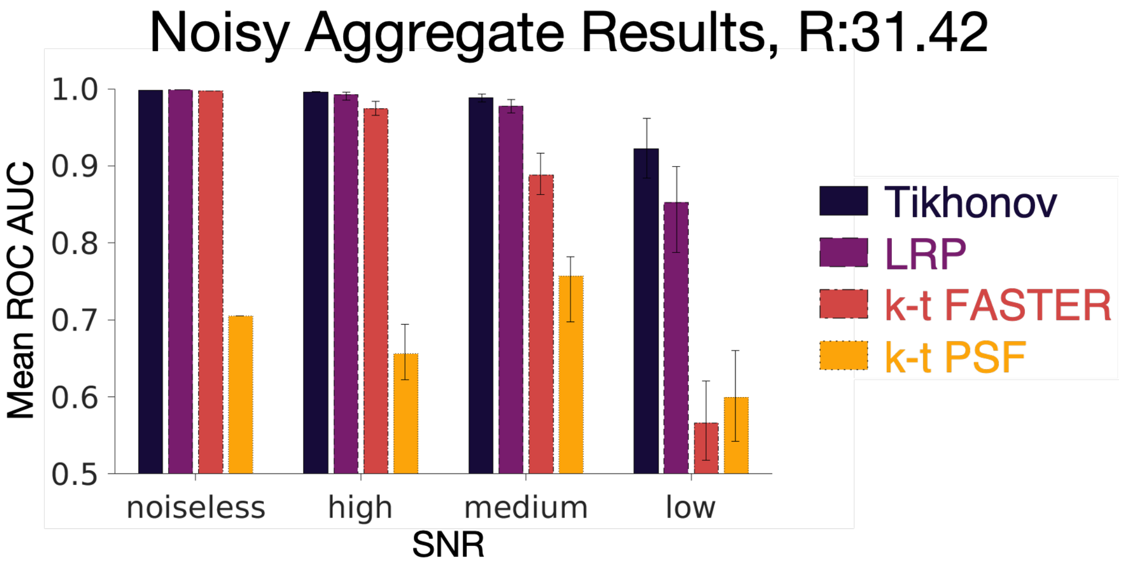


Figure 6.9: A comparison of mean ROC AUCs for noisy retrospective reconstructions. Each bar represents the mean AUC across the five different instantiations at a specific SNR (the noiseless results represent a single reconstruction). The error bars show the AUC value range.

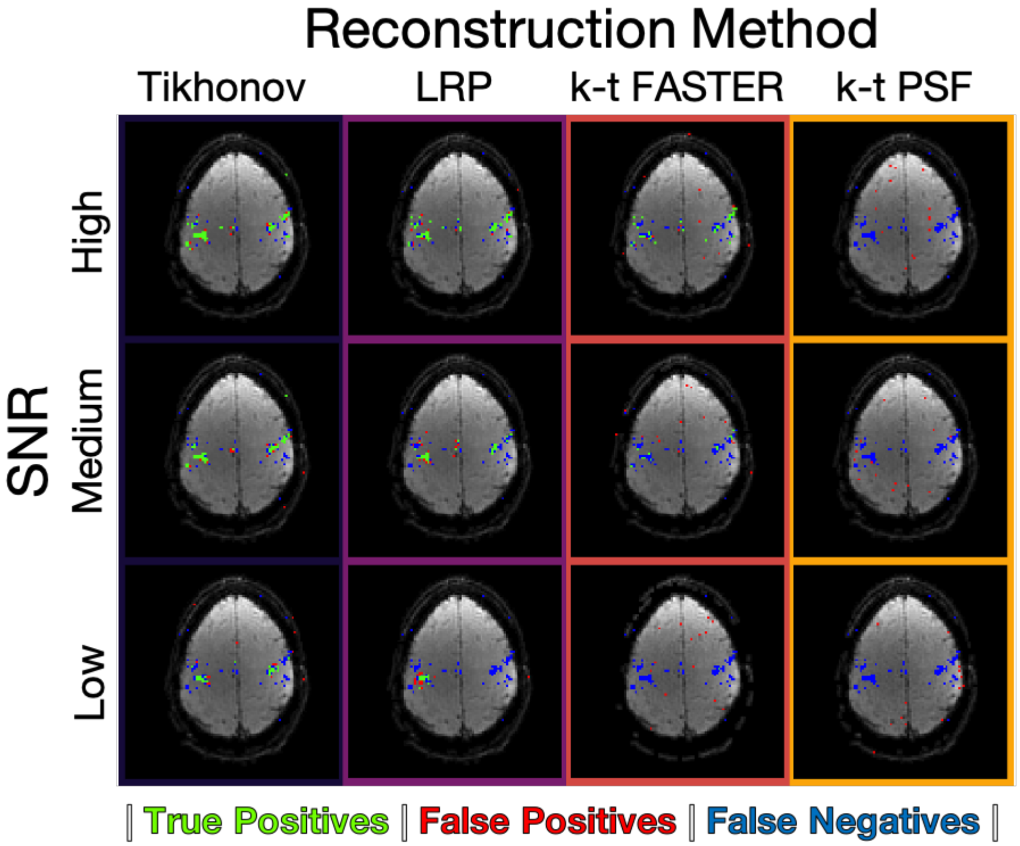


Figure 6.10: An example activation map at each noise value for each reconstruction method. Green pixels represent true positives, red pixels represent false positives, blue pixels represent false negatives. The z-statistics threshold yielded a false positive rate of 0.15%

Finally, a round of reconstructions were conducted on the prospective dataset. Due to the flexible nature of Golden Ratio angle methods such as TURBINE [8], different acceleration factors can be achieved with the same underlying data. The choice of acceleration factor trades off between more time frames (and a corresponding increase in the statistical DoF) and the loss of image fidelity resulting from spatial undersampling, with the overall data quality reflecting the balance of the two effects. The optimal λ pairings were chosen via grid search for each R and each constraint, and are stated along with the AUC results and the mean slice reconstruction time in table 6.3. The corresponding ROC curves are shown in figure 6.11. The λ distribution follows the same approximate shape in the prospective dataset as the retrospective dataset at each acceleration factor, although the LRP peak was broader along λ_T. This importance of correct weighting at increased R is demonstrated in figure 6.12, which uses the previous finding that equivalent Tikhonov reconstructions are defined by $\lambda_X \lambda_T = l$ to present a representative search in one dimension.

R	$\frac{Frames}{Blades}$	**Method**	λ_X	λ_T	**Mean Recon Time** (h)	**ROC AUC**
26.18	6	Tikhonov	1×10^{-2}	1×10^{-1}	11.6	0.9785
		LRP	1×10^{-3}	1×10^{-7}	(192.3+11.3) 203.6	0.9586
		k-t FASTER	0	0	13.0	0.9410
		k-t PSF	0	∞	(192.3+0.3) 192.6	0.4613
15.71	10	Tikhonov	1×10^{-1}	1×10^{-2}	6.3	0.9871
		LRP	1×10^{-1}	1×10^{-3}	(26.9+6.4) 33.3	0.9851
		k-t FASTER	0	0	5.8	0.9644
		k-t PSF	0	∞	(26.9+0.3) 27.2	0.9000
7.85	20	Tikhonov	1×10^{-1}	1×10^{-2}	2.9	0.9915
		LRP	1×10^{-1}	1×10^{-7}	(1.7+1.6) 3.3	0.9913
		k-t FASTER	0	0	1.4	0.9906
		k-t PSF	0	∞	(1.7+0.3) 2.0	0.9884

Table 6.3: The breakdown of parameters and scores for each approach at each acceleration factor for the prospective dataset. The mean recon time represents the mean reconstruction time between the separate slices. For k-t PSF and LRP-constrained reconstructions, the times in brackets indicate the split between the LRP generation and the final reconstruction.

At the lower acceleration factor (R=7.85, TR=1s) the constrained-subspace approaches and the unconstrained-subspace k-t FASTER approach produce equivalent reconstructions, and k-t PSF also produces a reasonable representation of the functional information. At the middle acceleration factor (R=15.71, TR=0.5s), both constrained-subspace approaches outperform k-t FASTER, but are indistinguishable from each other (see figure 6.14). These results contrast with the retrospective dataset at R=15.71, where the constrained-subspace approaches only just began to outperform the ROC AUC score of k-t FASTER (figure 6.5b). At R=26.18 (TR=0.3s), the Tikhonov-constrained reconstructions still provide good localisation of the functional information (see figure

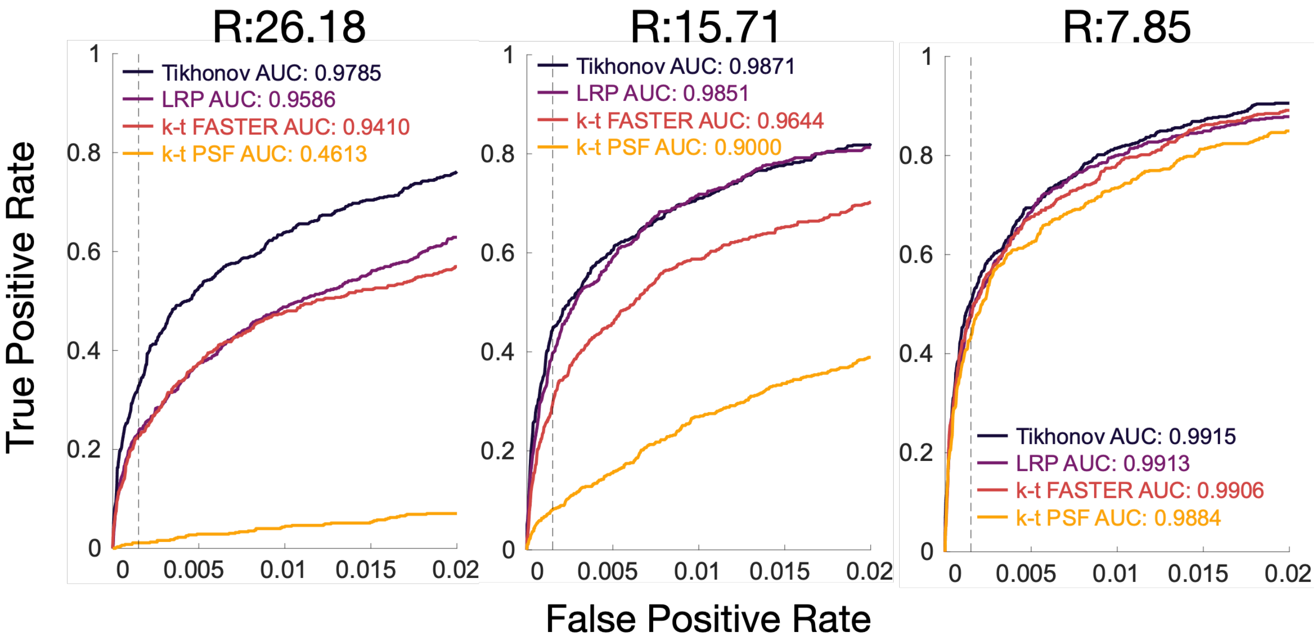


Figure 6.11: Volumetric ROC curves comparison between Tikhonov, LRP, k-t FASTER and k-t PSF reconstructions. The ROC curves were calculated over eight slices of the prospective dataset for R=26.18, R=15.71, and R=7.85. The z-statistic parameter map of the comparative long dataset was thresholded at z>4.8. The FPR is shown on the x-axis up to 0.02, and the AUC is quoted for the full curve. The corresponding z-stat maps for the different acceleration factors are in figures 6.13 (R:26.18), 6.14 (R:15.71), and 6.15 (7.85).

6.13), whereas the LRP-constrained approach is now equivalent to a k-t FASTER reconstruction at the 0.15% FPR threshold plotted (albeit with a higher overall ROC AUC score). k-t PSF performs noticeably better in the prospective dataset than in the retrospective dataset, but still poorer than the other approaches. The increased temporal resolution and/or temporal degrees of freedom may be responsible for this outcome. The full set of z-stat maps are shown in section 6.B.

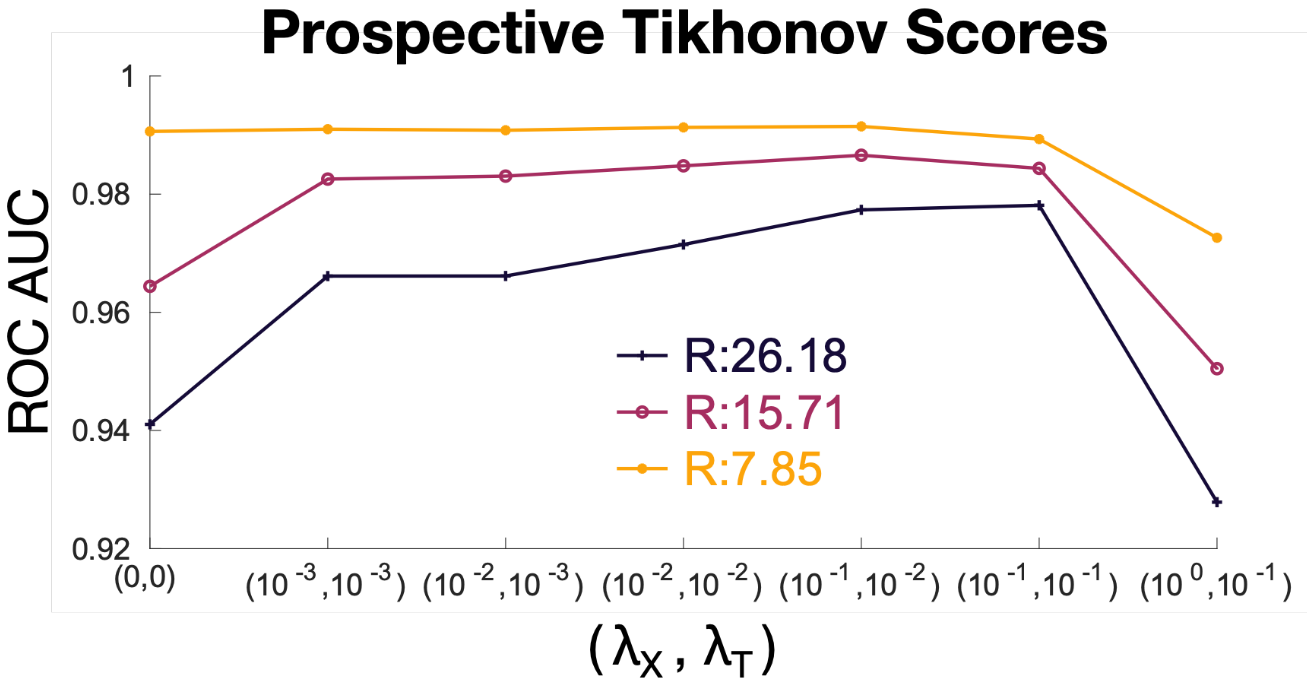


Figure 6.12: A comparison of different Tikhonov weightings at three different acceleration factors in the prospective dataset, with the x-axis showing a logarithmically increasing values of $\lambda_X \times \lambda_T$.

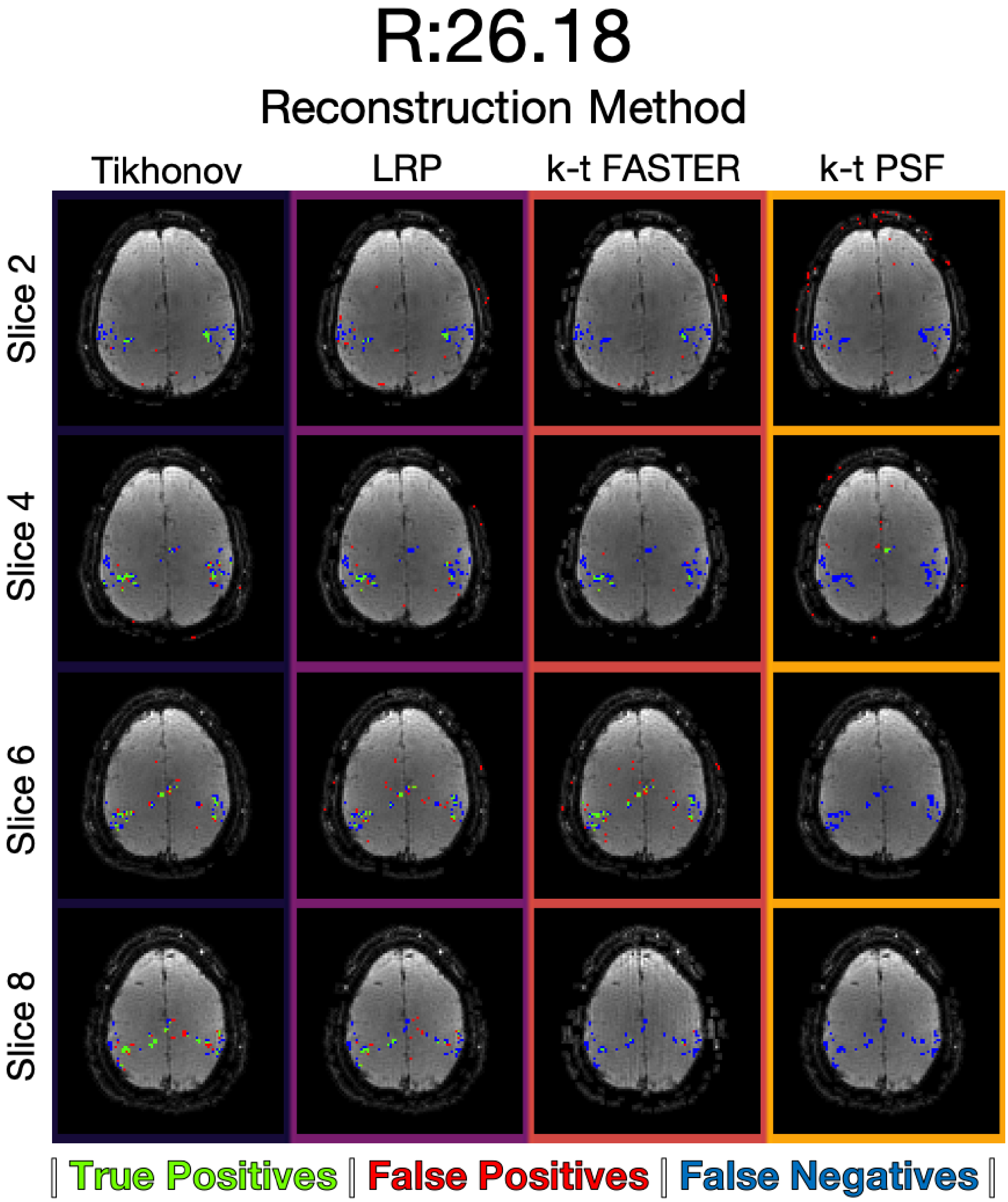


Figure 6.13: The activation maps for every second slice of the reconstruction at R=26.18 (TR=0.3s), at a threshold defined by a volumetric 0.15% FPR shown in figure 6.11. The activation maps of all slices can be found in Supplementary Figure 6.31.

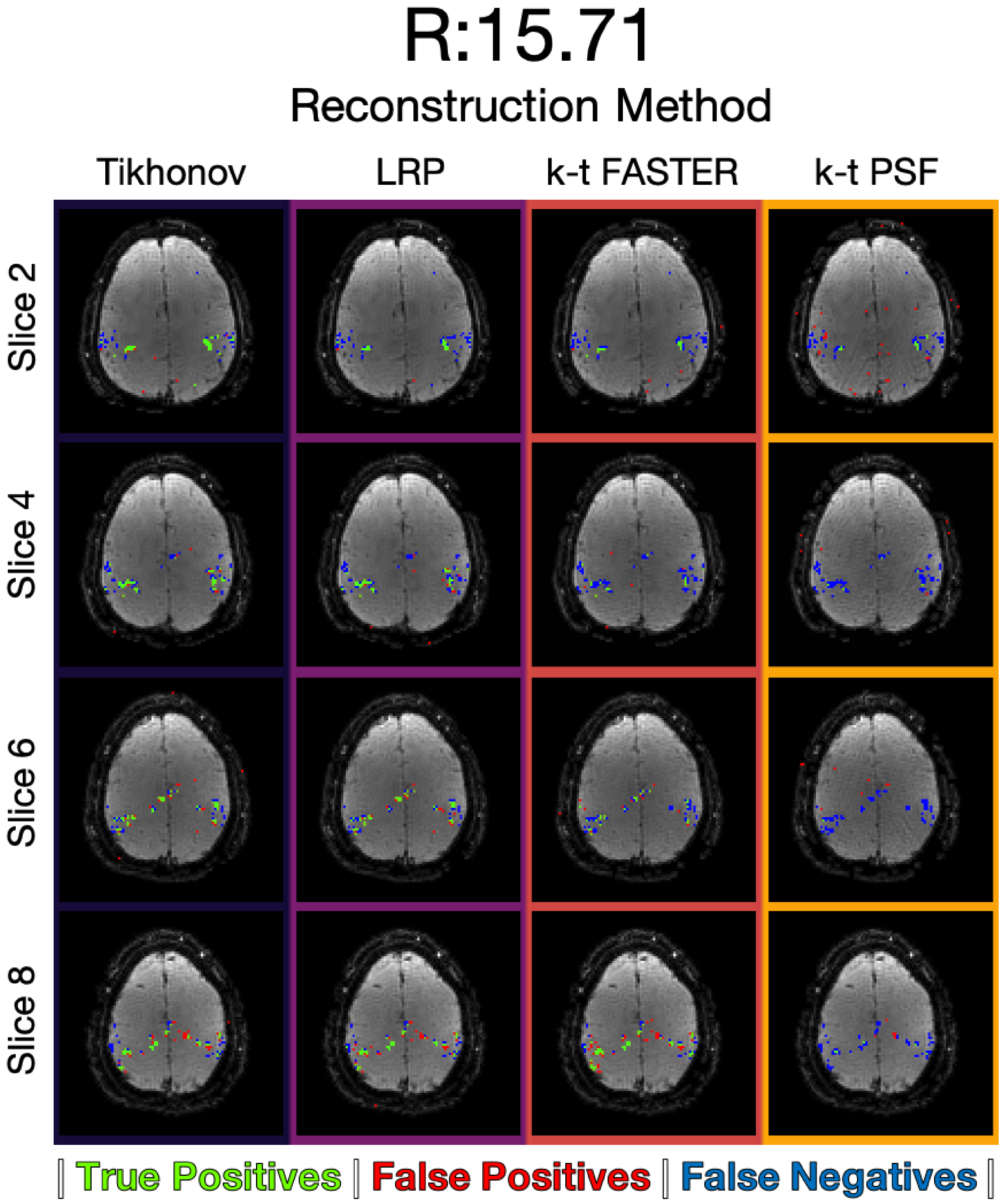


Figure 6.14: The activation maps for every second slice of the reconstruction at R=15.71 (TR=0.5s), at a threshold defined by a volumetric 0.15% FPR shown in figure 6.11. The activation maps of all slices can be found in Supplementary Figure 6.32.

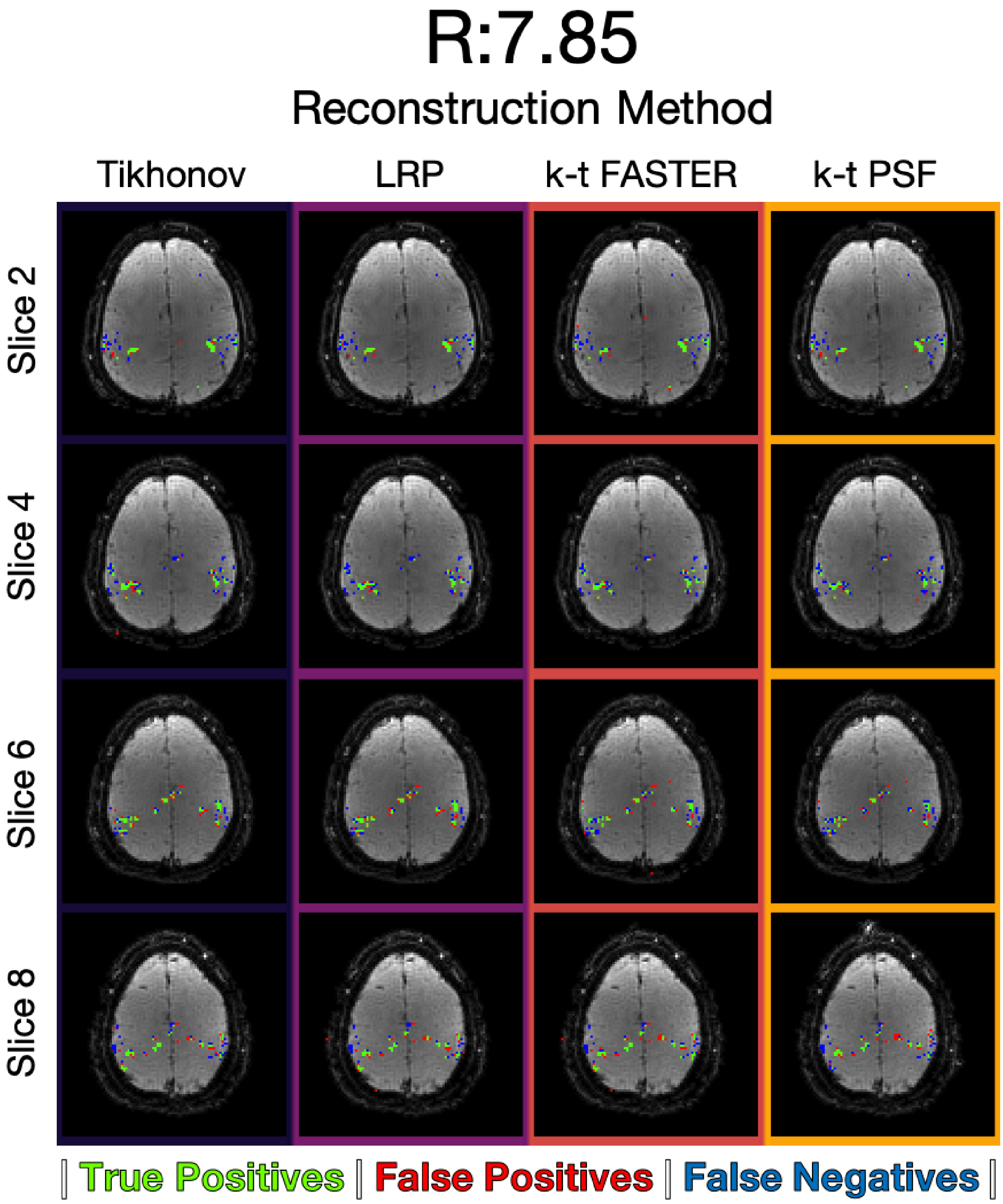


Figure 6.15: The activation maps for every second slice of the reconstruction at R=7.85 (TR=1s), at a threshold defined by a volumetric 0.15% FPR shown in figure 6.11. The activation maps of all slices can be found in Supplementary Figure 6.33.

6.3.2 Discussion

Across the different realistic datasets a clear trend emerges: the addition of soft subspace-constraints to the AM k-t FASTER formulation produces improved CCS and ROC AUC scores at high acceleration/low SNR where fixed subspace constraints (k-t PSF) or unconstrained-subspace reconstructions begin to break down.

The Tikhonov-constrained reconstructions proved the more effective of the two softly-constrained methods, producing high fidelity reconstructions even when the system was pushed to the acceleration factors or SNR levels where other methods began to fail (e.g. the prospective R=26.18 results, or the low SNR retrospective noisy results). LRP-constrained reconstructions recorded similar reconstruction scores to the Tikhonov-constrained approach outside of these extreme cases. The LRP-constrained activation maps at low SNR showed reduced fidelity (figure 6.30), but the false positives were localised close to the area of interest. By comparison, at lower SNR k-t FASTER produced false positives more broadly distributed across the FOV, not localised to voxels adjacent to true positive activations.

At high acceleration, the fixed temporal constraint of k-t PSF forced the temporal information to be equivalent to an LRP temporal subspace. Particularly in the retrospective dataset, this approach produced poor reconstructions (see figure 6.5a), achieving parity with the other methods at roughly R=3-5. A wider window size ($\frac{k_{max}}{4}$) at R=31.42 for the noiseless retrospective dataset was found to produce an improved k-t PSF result (ROC AUC = 0.9329) which was still worse than the other methods. This poor performance of the temporal subspace in the retrospective dataset then makes more notable the relative improvement seen in the prospective dataset (achieving good results at R=7.85, higher than the values reported in the literature [9]).

The optimal λ value varied between datasets, and was dependent on SNR (in the noisy retrospective dataset) and loosely with R (in the prospective dataset). It is clear that a soft constraint can help guide the dataset to improved reconstruction scores, but it is unclear yet if there is a concrete heuristic for deriving the optimal λ parameters. However, the distributions of the reconstruction scores with respect to the λ parameter were found to be independent of the dataset, although the sharpness of the peak of the distribution within a dataset did change with acceleration factor. In the ideal case, the LRP distribution was described by the product of two separable vectors, and so two 1D searches along each dimension would find an optimal solution. The optimal Tikhonov λ pairing could be found with a 1D search normal to $\lambda_X \lambda_T = l$.

6.3.3 Tikhonov and LRP Summary

Adding soft constraints to k-t reconstructions has been shown to improve low-rank reconstructions of highly-undersampled task fMRI data at high acceleration levels. We demonstrate that Tikhonov and LRP constraints provide higher fidelity reconstructions than existing k-t FASTER and k-t PSF methods. At R>10-15 (roughly TR<0.8-0.5s), the Tikhonov constraints appear to provide improved reconstruction with similar computation cost and only one additional design parameter required. While LRP constraints also demonstrated improvement over k-t FASTER results, they were not as robust as Tikhonov constraints, while also requiring a longer reconstruction time and more parameters to optimise.

6.4 Note on Adapted LRP and Tikhonov Constraints in Realistic Data

As with section 5.5, the "adapted" Tikhonov and LRP constraints were explored (here on realistic data) in order to better understand the nature of the "pure" Tikhonov and LRP reconstructions, and the effects caused by varying λ values were explored. These results mostly verified the conclusions of the previous chapter. Any results which expanded on the previous findings have been included in section 6.A, and a short summary is given here. While an exploration was made with TE LRP constraints (which applied Tikhonov constraints to the generation of the prior), the results were not found to add anything to the discussion and so are not included here.

The hybrid formulation used different constraints for each subspace ($X_{LRP}T'_{Tik}$ and $X_{Tik}T'_{LRP}$), and confirmed that the temporal subspace constraint was the more important factor in characterising the reconstruction. The $X_{LRP}T'_{Tik}$ did outperform the pure LRP approach in R=26.18 prospective reconstructions, and produced an ROC curve roughly equivalent to a pure Tikhonov approach.

The NE LRP approach rejected poor components of the LRP before reconstruction. In prospective data, the null-enforcement of the reduced LRP improved upon the pure LRP reconstruction, although not quite enough to make NE LRP reconstructions equivalent to pure Tikhonov regularisation.

If the priors can be improved, then both the NE LRP and the hybrid approach may prove to be a more effective way to integrate LRP constraints (or other priors) into a final reconstruction. However, without evidence that either adapted approach can improve upon pure Tikhonov methods, the increased reconstruction times required to generate the priors limits the desirability of both approaches when compared to a pure Tikhonov reconstruction.

6.5 Temporal Smoothing

The application of temporal subspace smoothing (with a constraint specifically on the temporal subspace) is a novel approach to acceleration of fMRI, although as mentioned in section 5.6 it is conceptually similar to other temporal smoothing approaches. The motivation and formulation of temporal subspace smoothing were introduced in section 5.6, where temporal smoothing was shown to be a promising approach when applied to the artificial dataset. The reconstruction is governed by the following optimisation problem (where ∇ is the differential operator, and λ_∇ is the weighting parameter):

$$(\hat{X}, \hat{T}) = argmin_{X,T}\left(||EXT' - d||_2^2 + \lambda_\nabla ||\nabla T||_2^2 \right) \tag{6.1}$$

The aim of the temporal smoothing term is to exploit a different underlying structure of the temporal subspace than the Tikhonov or LRP constraints, and could help regularize the final reconstruction (or the generation of a temporal prior) in highly accelerated regimes [10]. Previous results in this thesis have emphasised the importance of an informed temporal subspace, further motivating the investigation of incorporating additional temporal information into the reconstruction where appropriate.

6.5.1 Results & Discussion

A range of reconstructions at varying values of λ_∇ were analysed in the clean retrospective dataset at R=31.42 (5 blades/frame), R=15.71 (10 blades/frame) and R=7.85 (20 blades/frame). The reconstruction scores are shown in figure 6.16, and broadly show that no clear improvement was achieved over the unconstrained-subspace k-t FASTER approach in this dataset. The best-case scenario for temporal subspace smoothing at R=31.42 can be visualised in figure 6.17, emphasising that no advantage was seen.

There are two features worth exploring in the line plot of figure 6.16, which are the results at the edge of the main plateau in all plots ($\lambda_\nabla = 1 \times 10^{-4}$) and the result at the start of the second plateau in the CCS plots ($\lambda_\nabla = 1 \times 10^{-1}$). The first six temporal components of the $\lambda_\nabla = 1 \times 10^{-4}$ and $\lambda_\nabla = 1 \times 10^{-1}$ reconstructions are shown in figure 6.18, along with the k-t FASTER result and the decomposed ground truth. The $\lambda_\nabla = 1 \times 10^{-4}$ result clearly has reduced temporal frequency content when compared to the unconstrained temporal subspace (much more similar to the first four components of the truth), and yet does not produce a notably improved reconstruction. The $\lambda_\nabla = 1 \times 10^{-1}$ result characterises the second plateau, which occurs far above

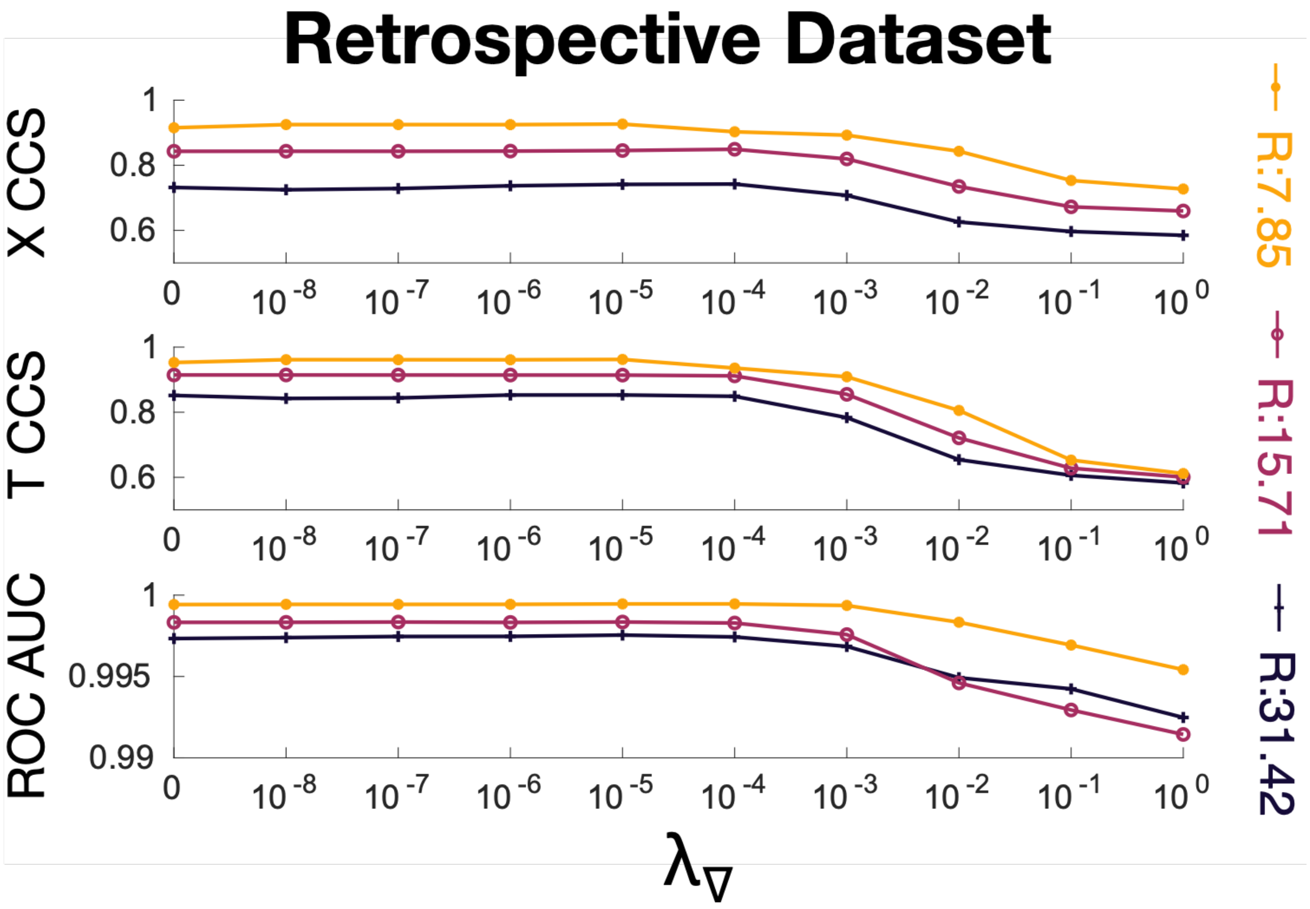


Figure 6.16: A comparison of different temporal subspace smoothing strengths at three different acceleration factors in the clean retrospective dataset.

a CCS score of zero (the CCS plateau continues to at least $\lambda_\nabla = 1 \times 10^1$, but the ROC AUC score dropped to 0.9351 for R=31.42, which made other ROC AUC information indiscernable if the x-axis was expanded to accommodate it). This secondary non-zero plateau is an interesting curiosity, as it implies that general alignment with the majority of the subspace can be found in a heavily smoothed dataset, which could be useful information for the forming of a temporal prior. However, this general alignment of the subspaces came at the cost of representation of the functional information, meaning it would likely not be useful for fMRI.

Chiew et al. [10] observed that in multiband acceleration, temporal smoothing was found to only increase GLM efficiency when considering voxels with high noise amplification (i.e. with high g-factors). In voxels with no noise amplification (i.e. similar to our clean retrospective dataset), they observed a similar smoothing parameter curve for geometric efficiency in PI to the curve that we plotted for CCS and ROC AUC in figure 6.16. Therefore, the next step was to verify whether temporal subspace smoothing in a low-rank framework would also be proven more effective in the presence of noise (especially as temporal subspace smoothing had proven effective in the more

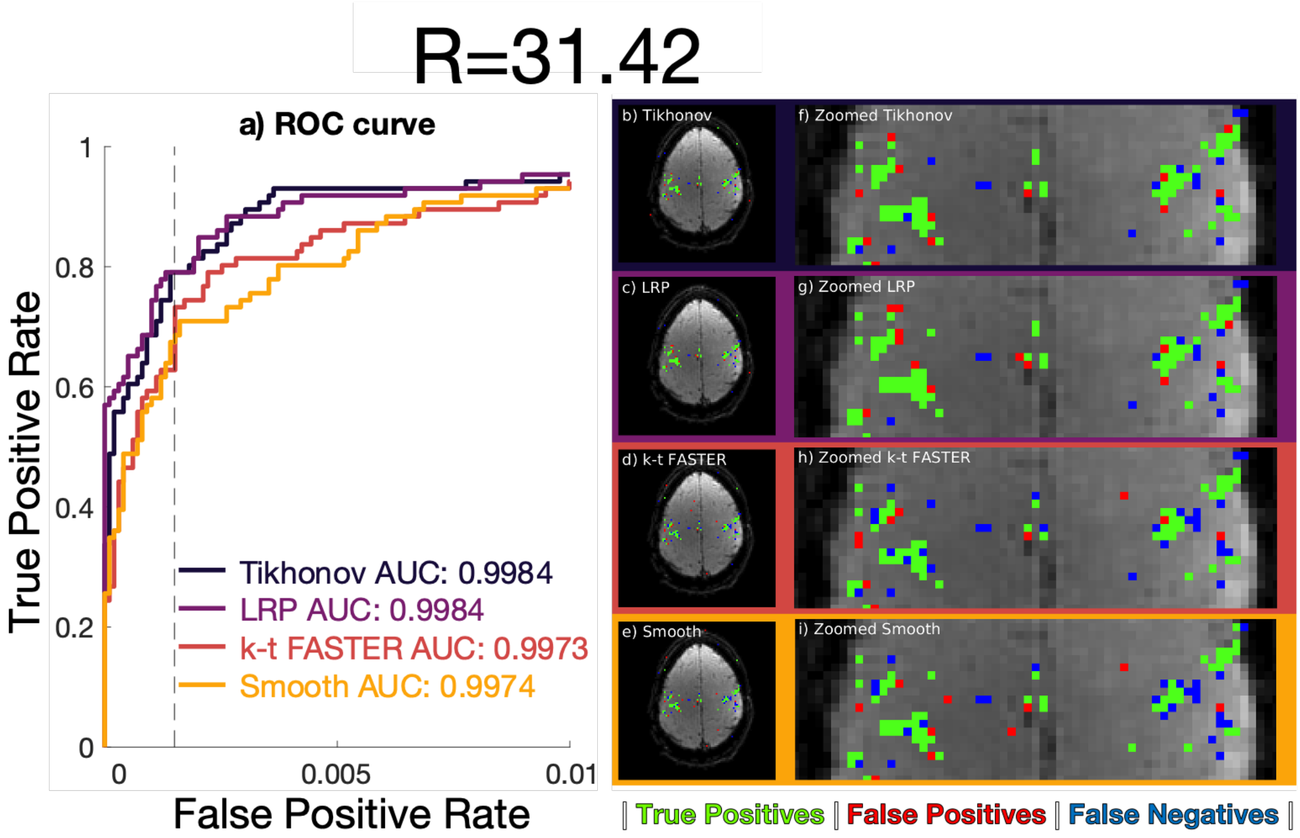


Figure 6.17: R=31.42 (5 blades/frame) retrospective dataset comparison. a) ROC curves, legend lists full curve AUCs. b)-e) Activation maps using a z-statistic corresponding to an FPR of 0.15%. f)-i) A medial zoom of the associated activation maps. b/f) Tikhonov: $\lambda_X = 1 \times 10^{-5}$, $\lambda_T = 1 \times 10^{-5}$, c/g) LRP: $\lambda_X = 1 \times 10^{-5}$, $\lambda_T = 1 \times 10^{-5}$, d/h) k-t FASTER, e/i) temporally smoothed: $\lambda_\nabla = 1 \times 10^{-4}$. Maps b)-i) use green true positive pixels, red false positives, and blue false negatives.

noise-laden artificial dataset of chapter 5).

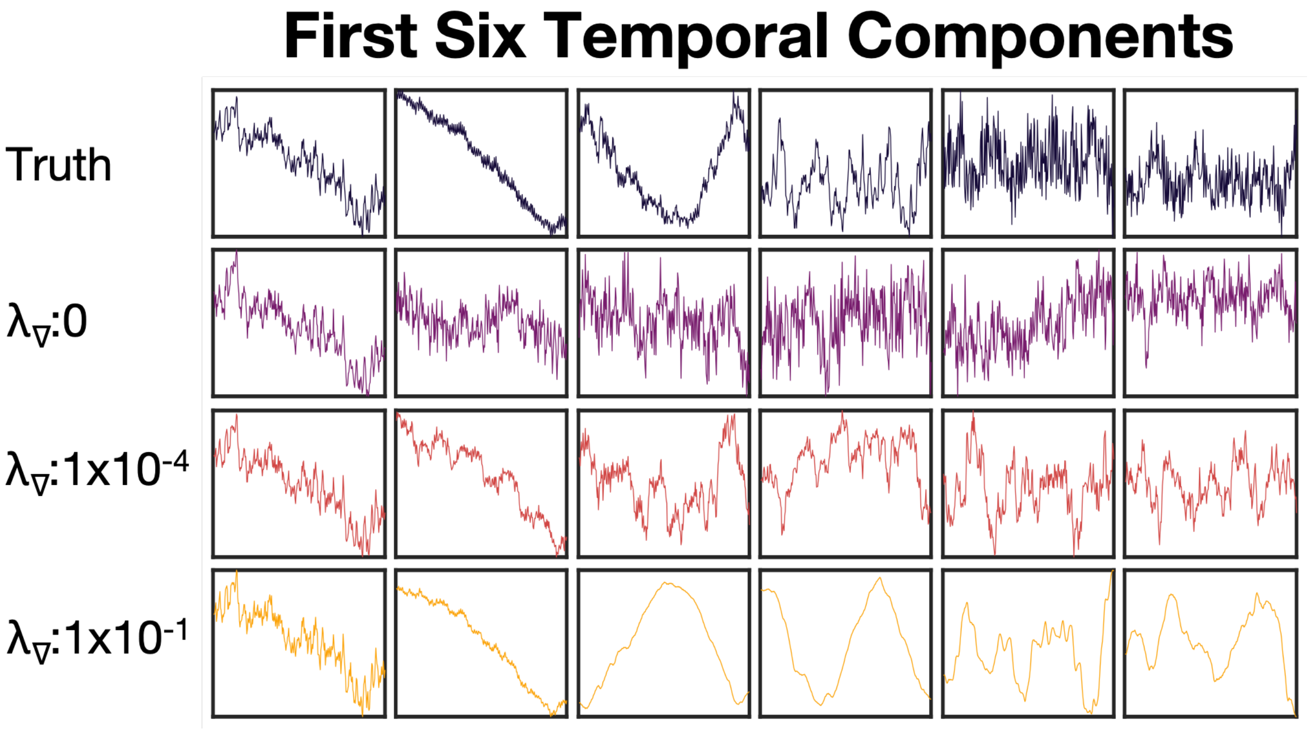


Figure 6.18: The first six components of the truth (top), an unsmoothed reconstruction (2nd row), a lightly smoothed reconstruction (third row) and a heavily smoothed reconstruction (bottom). These were taken from reconstructions at an acceleration factor of R=31.42, and are ordered by relative strength. Each y-axis is normalised to clearly show the shape of the temporal component.

The noisy retrospective datasets at high, medium, and low SNR (corresponding to SNRs of 100, 50, and 20 - see section 6.2) were reconstructed over five instantiations at a range of smoothing factors, and are shown in figure 6.19. The most substantial improvement in CCS scores occurred in the low SNR dataset, where the rightward CCS plateau of figure 6.16 was achievable regardless of the noise level.

The optimal noisy smoothing results occurred at approximately $\lambda_\nabla = 1 \times 10^{-4}$ for all acceleration factors, although all of the results up until $\lambda_\nabla = 1 \times 10^{-1}$ seem to lie within the range of possible optimal ROC AUC values. This finding contrasts with the noisy Tikhonov/LRP results, where the optimal (λ_X, λ_T) pairing clearly changed with SNR value. In all cases, the optimal temporal subspace smoothing showed a moderate improvement over the unconstrained-subspace approach (at $\lambda_\nabla = 0$), but none of the averaged smoothing results improved upon the Tikhonov- or LRP-constrained averaged results, a fact highlighted in figure 6.20. The smoothing results at non-noiseless SNR are significantly separate ($p<0.05$) from all other results except LRP-constrained reconstructions at low SNR.

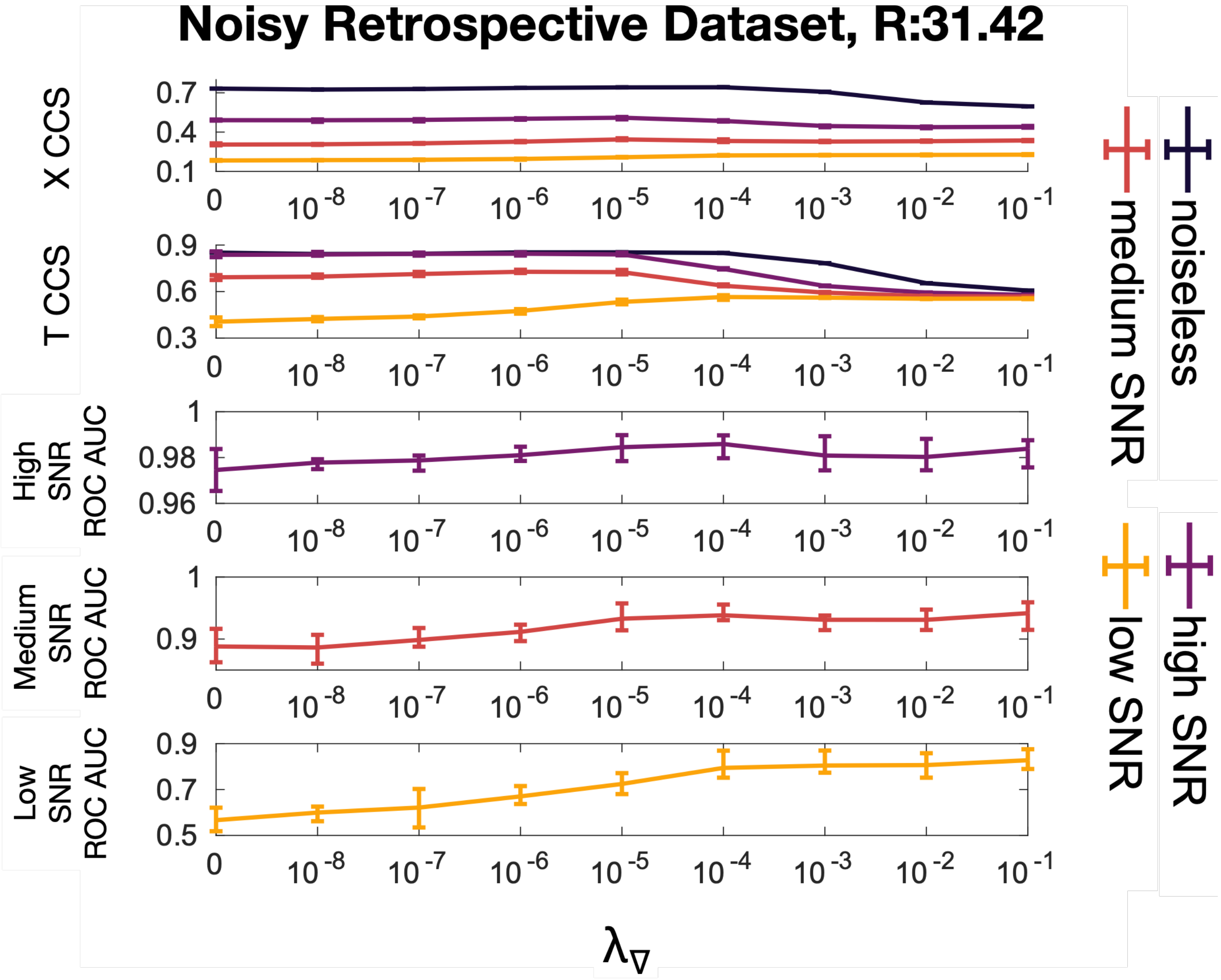


Figure 6.19: A comparison of different smoothing strengths for high SNR, medium SNR, and low SNR (and noiseless) reconstructions, with each line showing the mean score across five instantiations (for the noiseless reconstruction, only one reconstruction was carried out at each point). At each smoothing factor, the range of scores amongst the instantiations is shown as error bars.

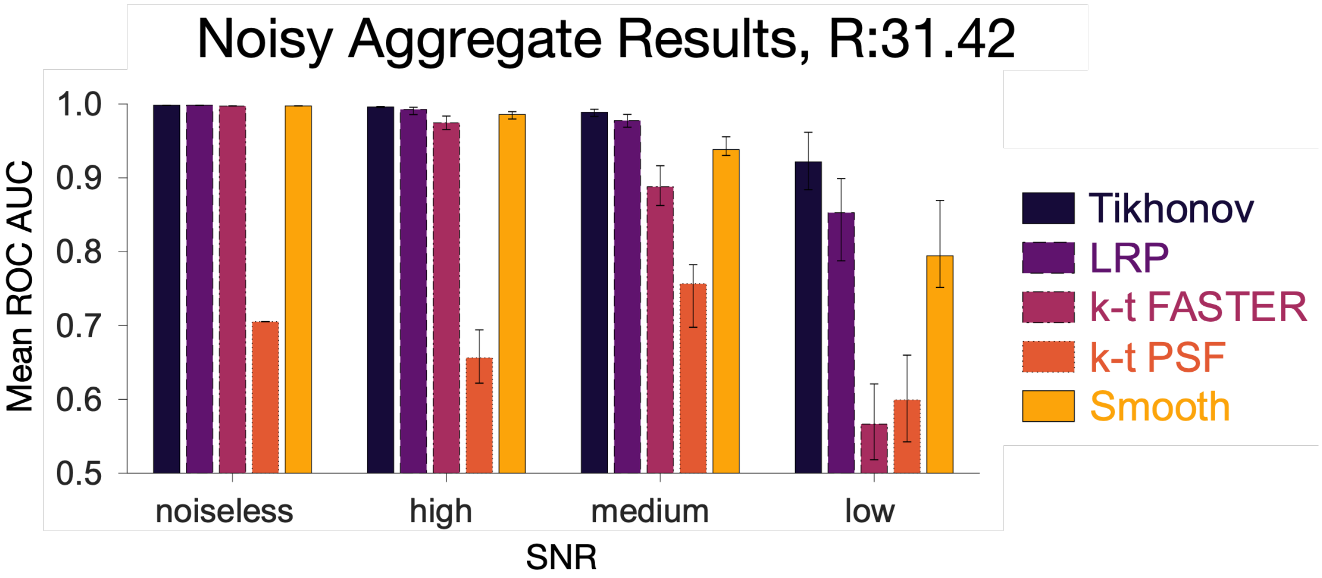


Figure 6.20: A comparison of mean ROC AUCs for noisy retrospective reconstructions. Each bar represents the mean AUC across the five different instantiations at a specific SNR (the noiseless results represent a single reconstruction). The error bars show the AUC value range. This figure is an extension of figure 6.9, but containing the optimal smoothing results.

The prospective dataset is perhaps the most suited to temporal smoothing, given the increased temporal resolution that arises with acceleration. As the time dimension is more rapidly sampled, the capacity to express unwanted high temporal frequencies beyond the expected signal increases, and so temporal smoothing has more potential to inform the reconstructed temporal subspace.

The optimal λ_∇ increased as R increased (figure 6.21), which makes intuitive sense due to the increased temporal degrees of freedom (this finding is notable for the fact that optimal λ_∇ did not change in the noisy retrospective dataset with decreased SNR). The optimal λ_∇ values were: $\lambda_{\nabla R=26.18}$=1×10^2, $\lambda_{\nabla R=15.71}$=1×10^1, and $\lambda_{\nabla R=7.85}$=1×10^{-3}.

Optimal results in the prospective dataset (figure 6.22) showed no improvement over k-t FASTER in the lower acceleration factor (R=7.85, 20 blades/frame, TR=1s), which coincidentally had the same TR_{frame} (1s) and therefore temporal resolution as the retrospective dataset. As the acceleration factor (and the temporal resolution) increased, the temporal subspace smoothing showed increasingly impressive reconstruction scores, relative to both k-t FASTER and the Tikhonov/LRP-constrained approaches.

The smoothed results retained approximate parity with the other soft constraints at R=7.85, achieved the highest ROC AUC score at R=15.71 (but not by a clear margin over the other soft constraints), and outperformed all other approaches at R=26.18 (6 blades/frame, TR=0.3s, $\lambda_\nabla = 1 \times 10^2$, figure 6.23). In fact, within the slices, the false positives in the temporally smoothed reconstruction mostly lie outside the brain, and so the reconstruction could be further improved with intelligent masking. These R=26.18 smoothing results are very notable, because these are the only results which show the smoothing constraint as the definitively best constraint in terms of reconstruction scores in realistic data. While the increased importance of a smoothing constraint at high acceleration is not unexpected, the level of improvement over other results occurred to a degree that no previous result would indicate.

The shape of the prospective λ_∇ curve (figure 6.21) is the same as the clean retrospective curve for R=7.85, but takes on a very different shape as acceleration increases. The existence of a middle plateau implies a benefit to light smoothing, but only at heavy smoothing ($\lambda_\nabla > 1 \times 10^{-1}$) do the reconstructions improve dramatically in quality. Similar to the Tikhonov results in figure 6.12, the importance of selecting optimal an λ value increases with R. There is also an increased reconstruction time of the temporally smoothed reconstructions when compared with Tikhonov reconstruction at high acceleration factors (R=26.18: t_{Tik} = 11.6 hours vs t_{smooth} = 29.6 hours; R=15.71, t_{Tik} 6.3 hours vs t_{smooth} = 11.2 hours; R=7.85, t_{Tik} 2.9 hours vs t_{smooth} 1.4 hours).

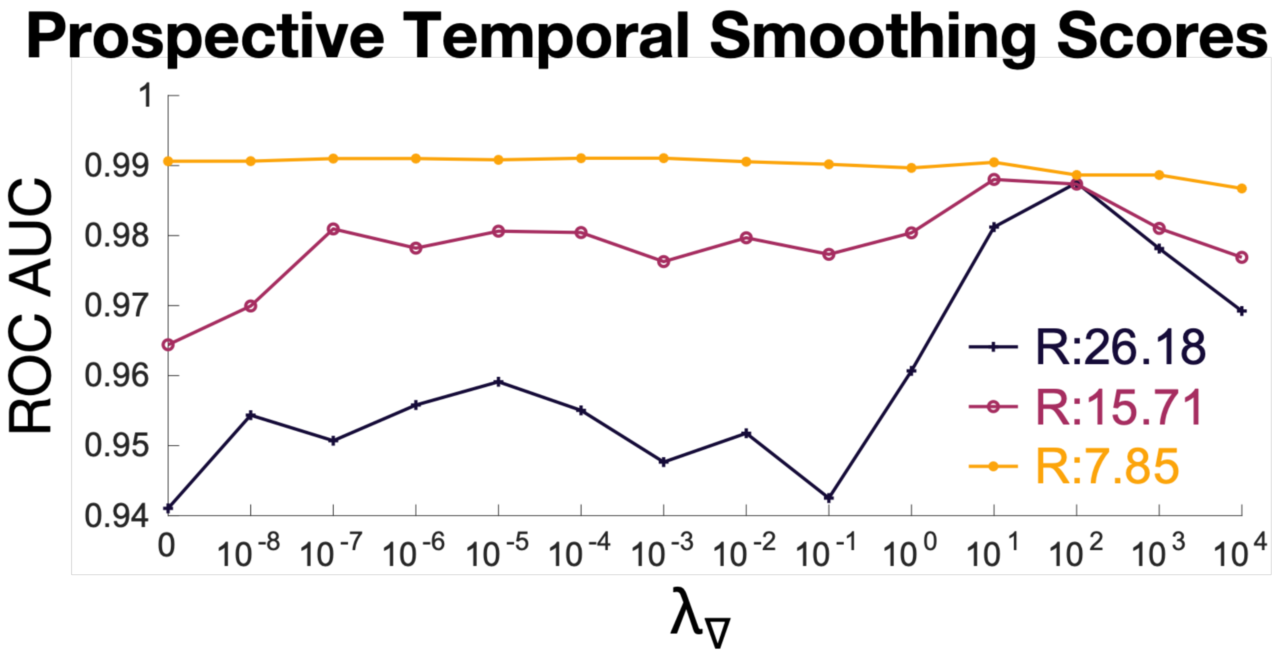


Figure 6.21: An ROC AUC comparison of different smoothing strengths at three different acceleration factors in the prospective dataset.

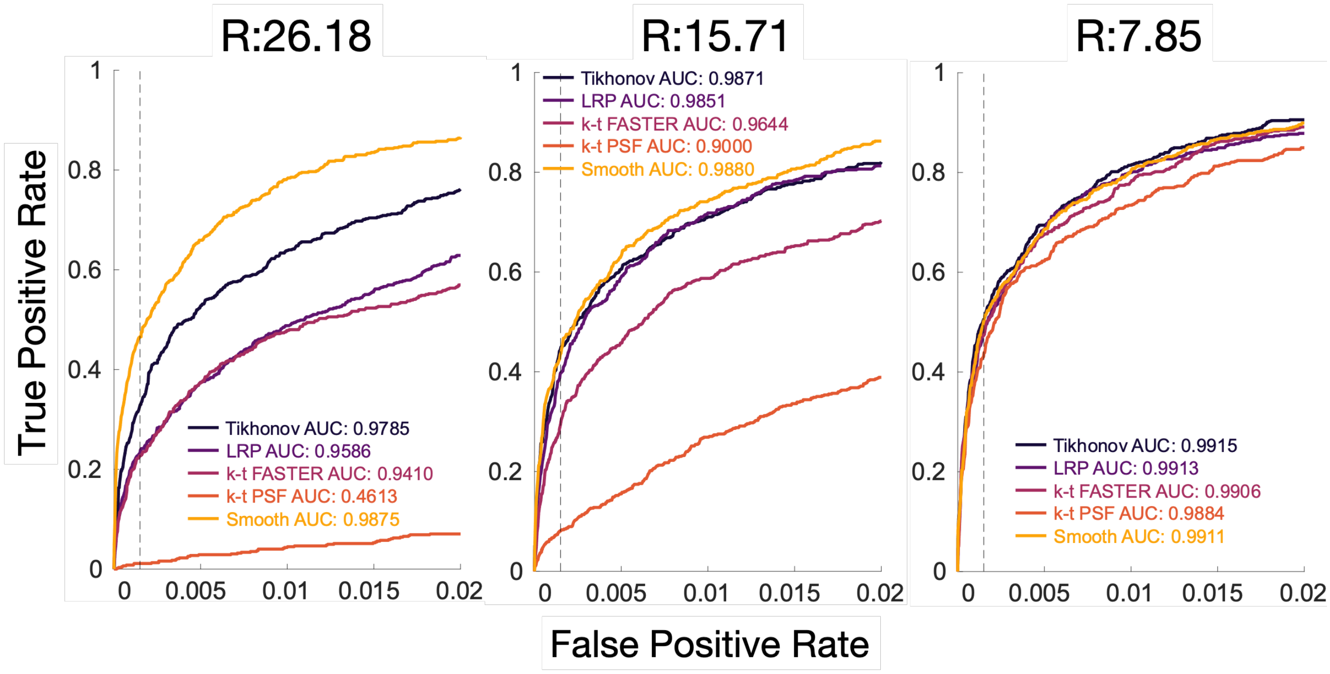


Figure 6.22: An ROC comparison of the optimal results for all different pure constraints at three different acceleration factors in the prospective dataset.

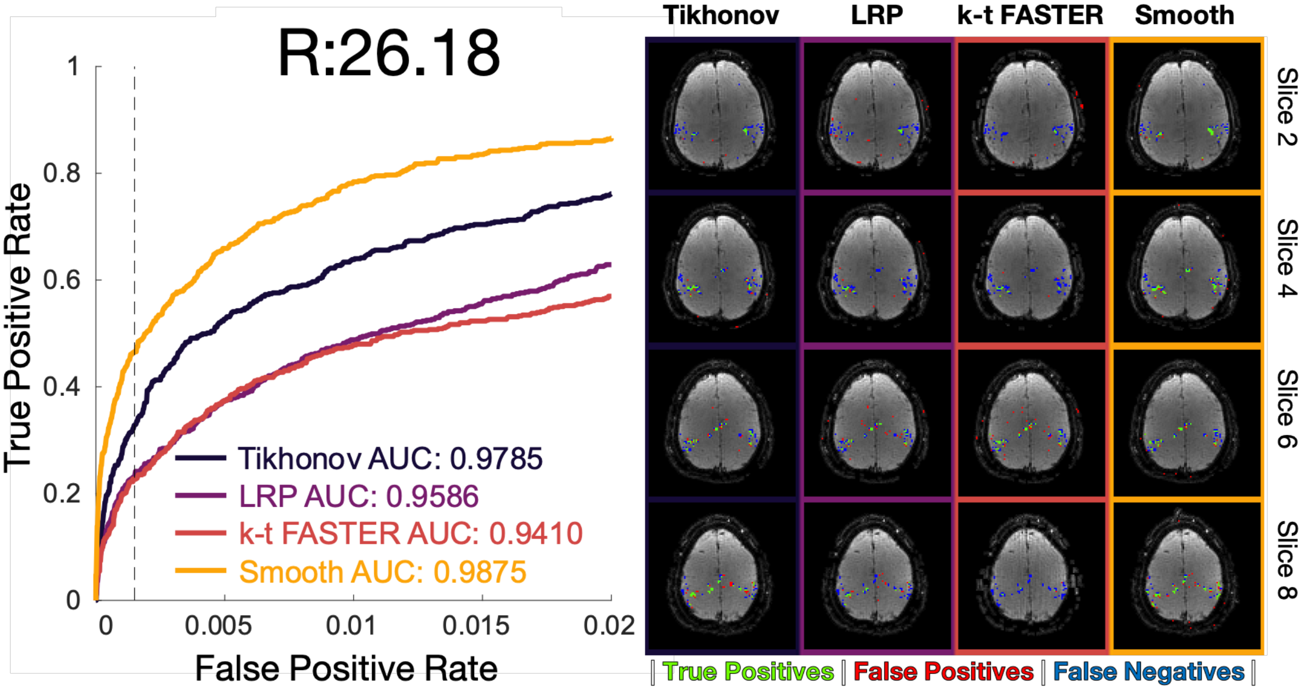


Figure 6.23: A comparison between Tikhonov, LRP, k-t FASTER and temporally smoothed reconstructions (at optimal λ values) in the prospective dataset at R=26.18. A subsection of activation maps at a z-statistic corresponding to an FPR of 0.0015. The complete set of activation maps can be found in figure 6.31.

6.5.2 Smoothing Conclusion

The smoothing prospective results achieved at high acceleration are very impressive, far more impressive than would be expected given the relatively poor results in the retrospective test datasets. The increased temporal degrees of freedom in high acceleration appears to benefit greatly from increased temporal subspace smoothing (possibly due to the less uniform and more rapidly changing PSF - see figure 3.7). These results are drawn for a slowly varying task, and as such additional testing would be needed to verify if smoothing would be as beneficial for rsfMRI, or for identifying and removing rapidly changing physiological noise. However, the results shown here are already highly promising and certainly warrant further future investigation, especially if a measure can be found to decrease the reconstruction times without sacrificing reconstruction quality.

6.6 Conclusion

All of the reconstructions here can be considered in the context of the general constrained framework provided in Chapter 5, including the previously proposed approaches of k-t FASTER and k-t PSF. While the constrained-subspace AM k-t FASTER framework absorbs ideas and expands upon the theoretical basis of these two methods, the Tikhonov- and LRP-constrained approaches exhibited a clear advantage for highly accelerated reconstructions (at R≥15.71 in the retrospective dataset, and at an acceleration factor between R=7.85 (TR=1s) and R=15.71 (TR=0.5s) in the prospective dataset). The temporal subspace smoothing constraint only showed clear benefit at high acceleration in the prospective dataset, also beginning to show improvement over the unconstrained method between R=7.85 (TR=1s) and R=15.71 (TR=0.5s).

The LRP-constrained reconstructions hold the implicit assumption that any reasonable trajectory with a greater sampling density in central k-space will be improved with a prior to more strictly enforce that central weighting (although this assumption may be somewhat contradicted by the poor temporal subspace generated by the LRP). The temporally smoothed result holds best in methods with a unique sampling trajectory in each frame, especially when the underlying paradigm is slow. The Tikhonov-constrained reconstructions have no explicit limiting assumption. As such, further testing on prospective datasets with different paradigms and conditions might yield different conclusions regarding the most favourable constraints (e.g. in rsfMRI, a temporal smoothing constraint may prove less beneficial than the more general Tikhonov constraint).

While LRP-constrained approaches benefited from various innovations (e.g. null-space enforcement after thresholding components), Tikhonov results proved more consistently high quality between the two, even in extreme cases of acceleration or noise. Additionally, Tikhonov-constrained reconstructions were the quickest to reconstruct and required an optimal λ pairing found through a 1-D parameter search only - reducing the dimensionality of the design constraints. However, the smoothing result (also requiring only a 1-D parameter search) demonstrated further improved performance at very high acceleration factors (R=26.18, TR=0.3), showing that there is potential for informed constraints to allow drastically improved results fMRI.

6.A Adapted Tikhonov/LRP Constraints

In section 5.B, various different adapted approaches were explored in order to better understand the impact of the Tikhonov and LRP-constrained approaches on reconstructed subspaces using the artificial dataset. These different approaches were collectively described as “adapted” constraints (contrasting to the “pure” constraints). Similarly adapted reconstructions were carried out for the clean retrospective dataset and on a limited search of the prospective dataset. While most of these results simply reinforce the conclusions of section 5.B, a small subsection of results are included to demonstrate their application in more realistic data. A summary of these findings can be found in the main text in section 6.4. The TE LRP constrained reconstructions in realistic data were not found to produce any interesting results, and are not included here.

6.A.1 Hybrid Constraints

In section 5.B.1, the hybrid reconstruction approaches that mixed Tikhonov and LRP constraints were found to be heavily characterised by whichever constraint was applied to the temporal subspace. This finding held in the more realistic datasets, and in particular can be seen in the shape of the ROC curves in the prospective results (figure 6.24). At R=7.85 and R=15.71, the reconstruction results are broadly similar (with the pure methods slightly outperforming hybrid methods). At R=26.18 both reconstructions which use the temporal LRP constraint - pure LRP and the hybrid $X_{Tik}T'_{LRP}$ method - begin to deteriorate, whereas both reconstructions which use the temporal Tikhonov constraint did better. Figure 6.24 emphasises that, even in real data, the temporal subspace is a defining factor in characterising the functional activity at high acceleration.

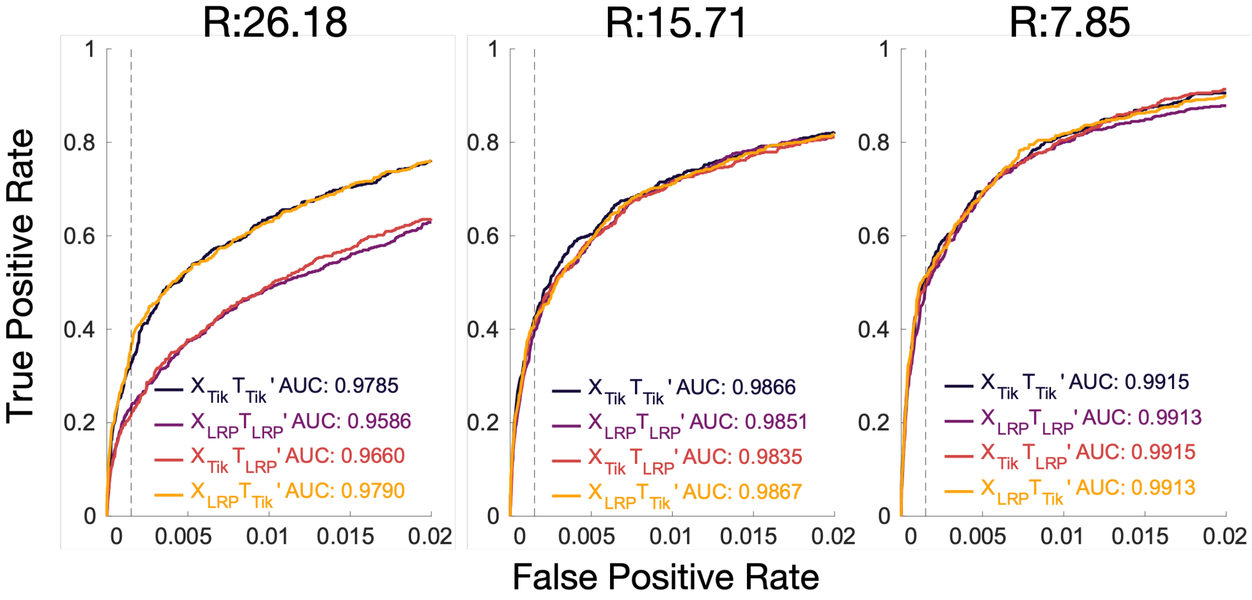


Figure 6.24: Volumetric ROC curve comparison for hybrid application of the Tikhonov and LRP constraints. The ROC curves were calculated over eight slices of the prospective dataset for R=26.18, R=15.71, and R=7.85. The z-statistic parameter map of the comparative long dataset was thresholded at z>4.8. The FPR is shown on the x-axis up to 0.02, and the AUC is quoted for the full curve.

Another interesting observation occurred when generating the hybrid λ CCS grids for the retrospective dataset (figure 6.25 shows the R=31.42 reconstructions). The hybrid grids contained a distribution which was a mix of the two pure distributions, but contained a flatter peak in the direction that was characterised by the Tikhonov-constrained subspace. The optimum in these hybrid grids was usually found in a region with a higher Tikhonov weighting than LRP weighting. Given the generally improved reconstruction of Tikhonov-constrained reconstruction in this dataset, the larger Tikhonov weighting should not be seen as a surprise.

In conclusion, the hybrid approaches highlight the importance of the temporal subspace constraint

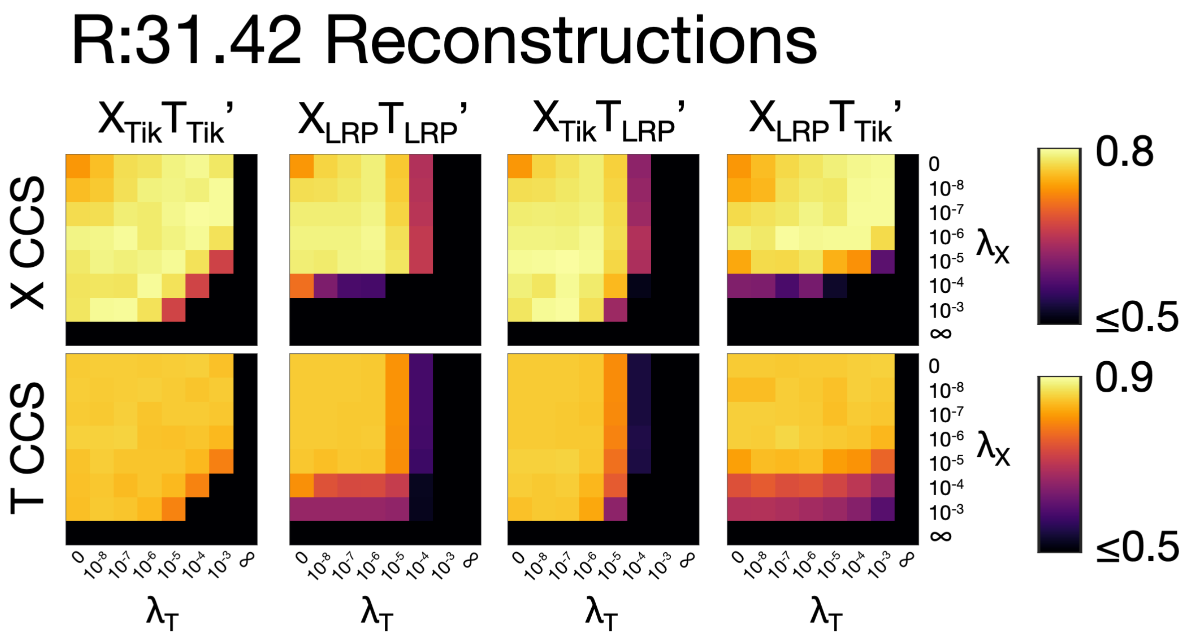


Figure 6.25: The CCS scores for Tikhonov, LRP, and hybrid-constrained reconstructions against a ground truth in the clean retrospective dataset at R=31.42 (5 blades/frame). The top row represents the spatial CCS scores, and the bottom row are the temporal CCS scores. Where any Tikhonov constraint is fixed ($\lambda = \infty$), the subspace is identically zero. The X CCS and T CCS grids are shown at different scales to better highlight the distribution. An extended grid guide is shown in figure 6.2.

in characterising the reconstruction. This logic can be used to make the following deduction: at high acceleration factors with increased windowing for the LRP, it is the temporal subspace constraint which drives the reduced reconstruction quality between the pure Tikhonov- and pure LRP-constrained approaches. It should be noted that this observation does not imply that there are no inaccuracies in the spatial subspace constraint, merely that the quality of the spatial subspace constraint is less impactful to the reconstruction process.

6.A.2 Null-Enforced LRPs

In section 5.B.3, NE LRP-constrained reconstructions accounted for the bad information in the LRP temporal prior by rejecting components that didn't satisfy an energy threshold in the temporal frequency domain (rejecting components with ≤40% energy between 0.1Hz and 1Hz).

At high acceleration factors/windowing in the artificial FMRIB dataset, this led to the rejection of all but the most dominant component. In the clean retrospective dataset at R=31.42, the generated LRP again only had one non-rejected component and so will not be covered in great detail. NE LRP λ CCS grids were formed for the retrospective dataset, and found to follow a similar shape to the pure LRP grid in figure 6.3. A reduced NE LRP reconstruction in the retrospective dataset with optimal λ improved upon the X CCS and the T CCS score of optimal pure LRP-constrained reconstructions, and produced an equivalent ROC AUC score.

An exploratory search with NE LRP was carried out in the prospective dataset (the criteria for rejection/acceptance was applied on a slice-wise basis). Within the different acceleration factors R=7.85 LRP contained an average of eleven non-rejected components across the eight slices (± 1 component), R=15.71 LRP contained four or five non-rejected components, and all R=26.18 LRPs had only two non-rejected components.

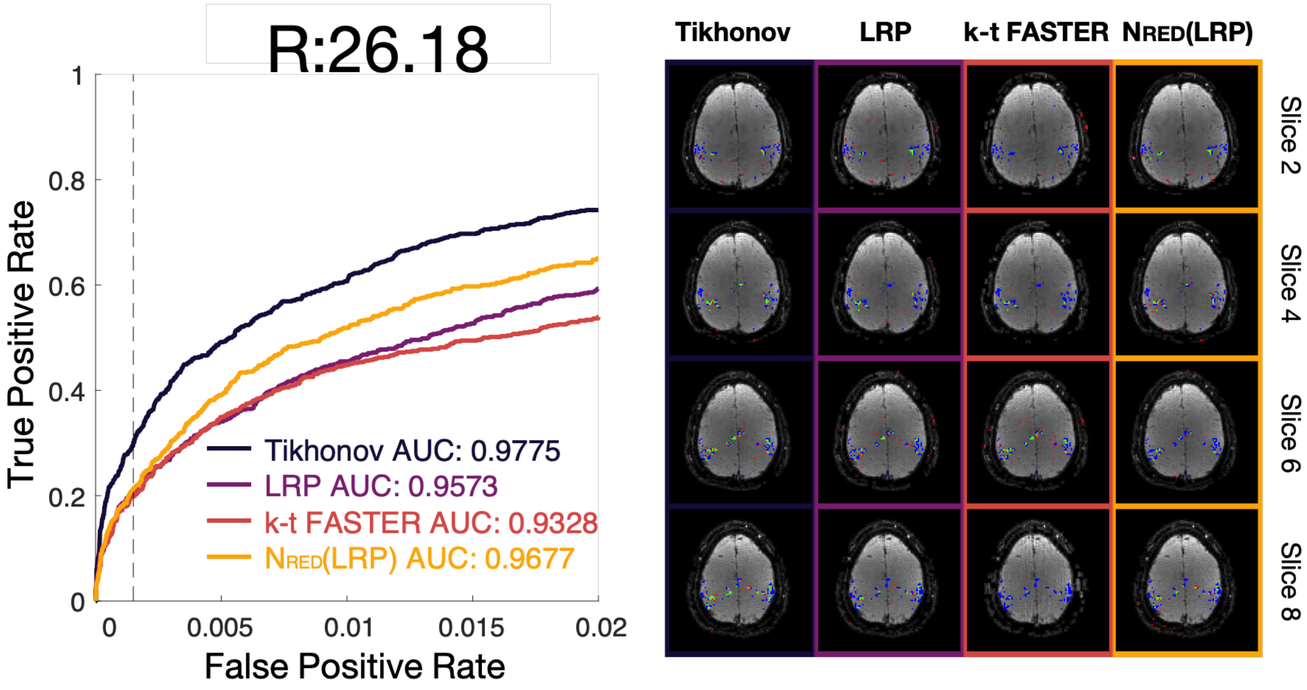


Figure 6.26: A comparison between Tikhonov, LRP, k-t FASTER and reduced NE LRP reconstructions (at optimal λ values) in the prospective dataset at R=26.18. Left: The volumetric ROC curves over eight slices. The false positive rate is shown on the x-axis up to 0.02, and the AUC is quoted for the full curve. Right: A subsection of activation maps at a z-statistic corresponding to an FPR of 0.0015.

The optimal λ results of the NE LRP R=26.18 ($\lambda_X = 1 \times 10^{-2}, \lambda_T = 1 \times 10^{0}$) reconstruction are shown in figure 6.26 with previously plotted approaches. At the other acceleration factors (R=15.71 and R=7.85), NE LRP produced equivalent results with the other methods. At R=26.18, the NE LRP approach outperformed the pure LRP-constrained approach in terms of reconstruction score. However, a qualitative look at the ROC curve shows that there is little statistical improvement in the classification for the low FPR region in which an experiment would hope to operate, and a qualitative look at the z-stat maps show that this acceleration factor is beyond the point at which reliable analysis can occur in NE LRP reconstructions.

Overall, the null-enforcement of the reduced LRP does appear to improve the pure LRP reconstruction. The energy threshold was derived heuristically, and so there is potential for some improved reconstruction quality with investigations of slightly different thresholds, or with the reduction done by an alternate metric. However, the Tikhonov-constrained results outperform NE LRP at high acceleration factors. While null-enforcement is an effective adaptation to a constraint with a mixture of poor and good components, it is likely only worth revisiting if an approach to LRP generation can be found which produces better temporal components at high acceleration.

6.B Supplementary Figures

In this chapter, a lot of data and graphs were generated. In many graphs in this chapter, only a subsection of the results are shown (e.g. one of the noise instantiations in figure 6.10, every alternate slice in figure 6.13). In order to demonstrate that the data selected in the chapter are representative and not cherry-picked, more complete graphs are shown here. The optimal Tikhonov/LRP/k-t FASTER/k-t PSF/smoothing reconstructions are shown in each graph. These graphs do not seek to draw any new conclusions, merely validate that the summative graphs shown in the above chapter are representative of the full dataset, and to present a conclusive picture of all constraints.

From the Tikhonov and LRP constrained noisy retrospective results (section 6.3.1), only the mean ROC curve (figure 6.9) and the activation maps from one of the five instatiations of the noisy retrospective dataset (figure 6.10) were shown. The ROC curve for all results across all instantiations are shown here in figure 6.27, along with the activation maps for the high SNR results (figure 6.28), medium SNR results (figure 6.29), and low SNR results (figure 6.30), along with the optimal smoothing results ($\lambda_X = 0$, $\lambda_T = 0$, $\lambda_\nabla = 1 \times 10^{-4}$ for all SNR).

For the prospective dataset, only alternate slices of the eight-slice dataset were shown. The results with subspace smoothing are also shown here alongside Tikhonov, LRP, k-t FASTER, and k-t PSF. The full set of slices are shown for R=26.18 (figure 6.31), R=15.71 (figure 6.32), and R=7.85 (figure 6.33).

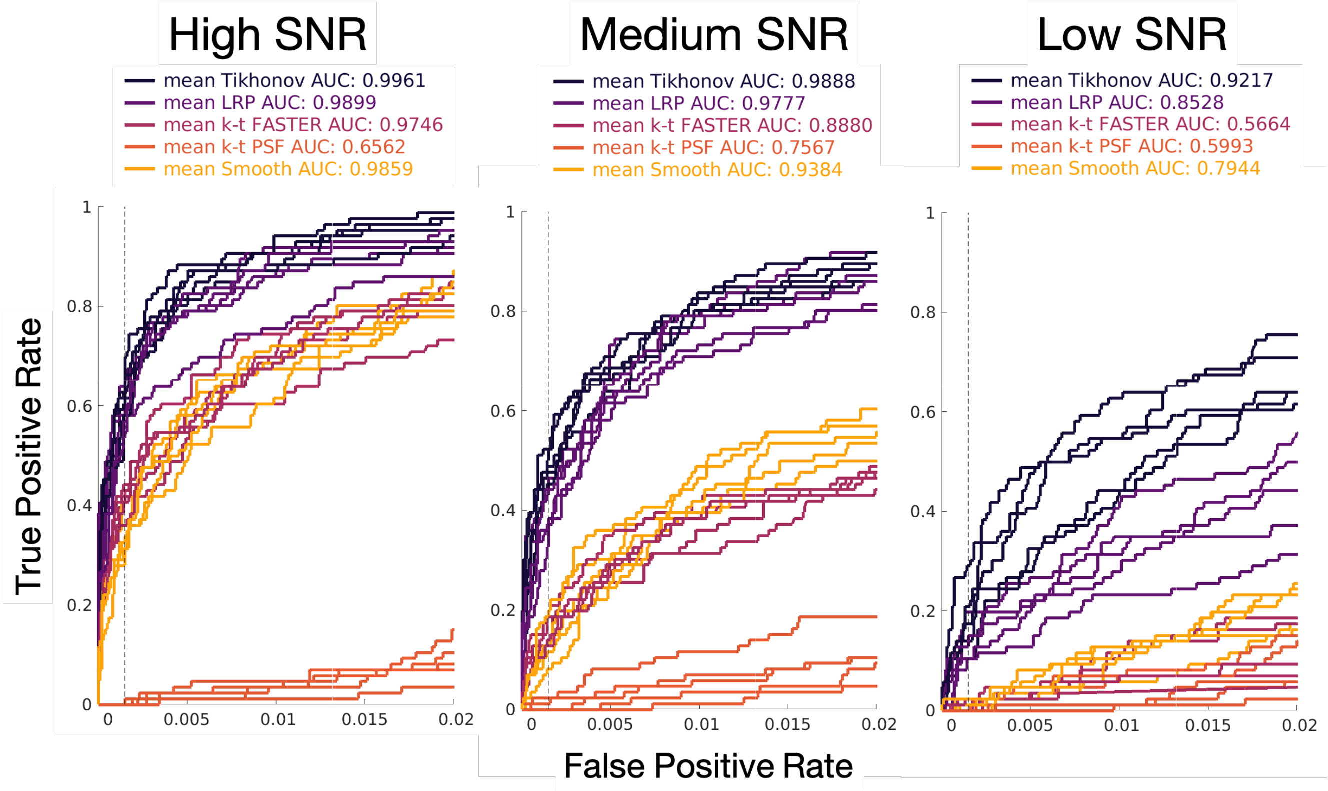


Figure 6.27: The individual ROC curves for the five different instantiations of the noisy retrospective dataset for each noise level and reconstruction method. Figure 6.9 showed the mean ROC curves. In this figure, the variation within each constraint is more apparent, although the general relationship between the methods remains the same. While there is a fairly clear distinction between k-t PSF and k-t FASTER, and between k-t FASTER and the other two methods at each SNR, the Tikhonov and LRP curves separate clearly at low SNR.

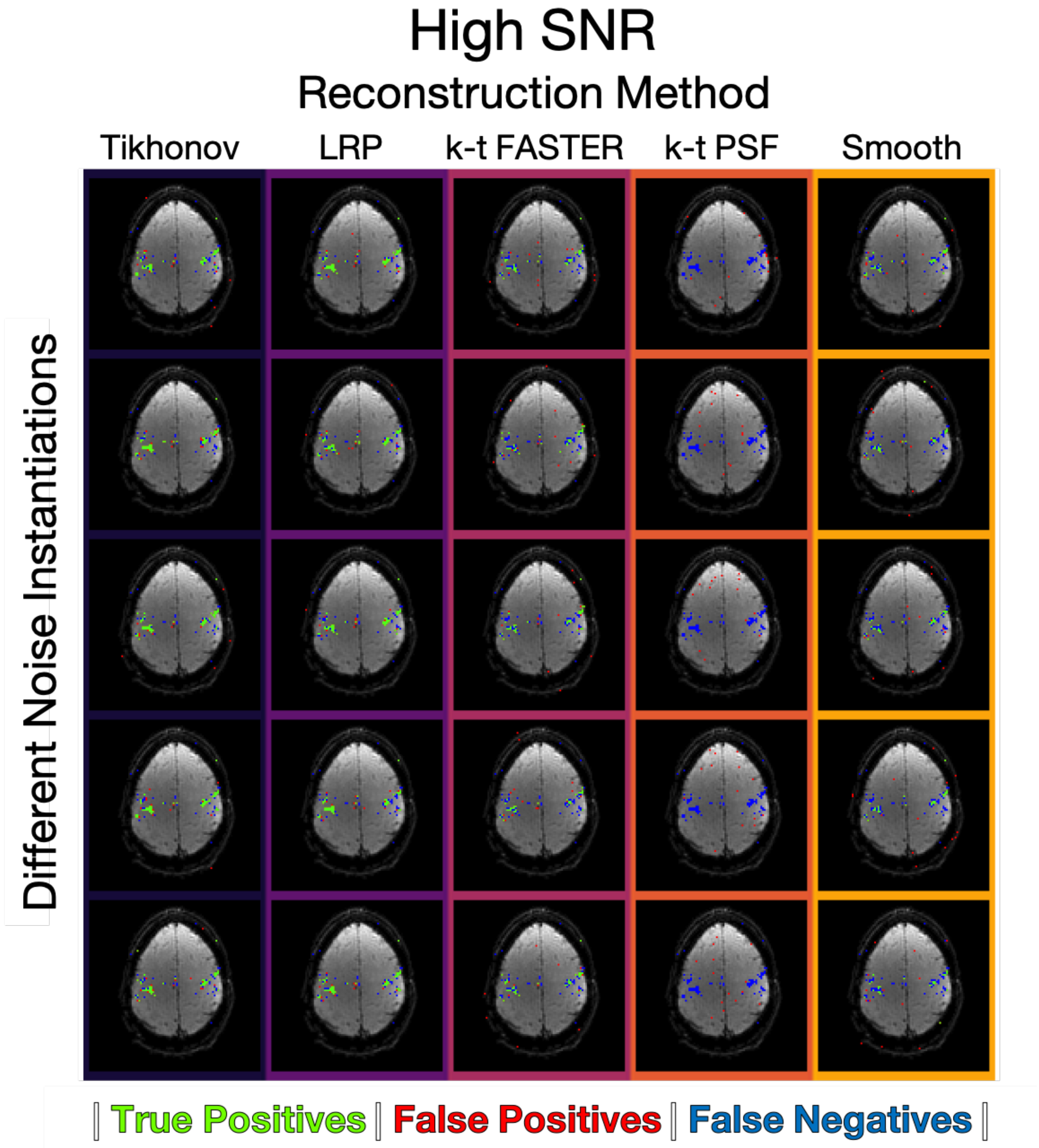


Figure 6.28: The five activation maps at high SNR (with random noise) for each reconstruction method. Green pixels represent true positives, red pixels represent false positives, blue pixels represent false negatives. The z-statistics threshold yielded a false positive rate of 0.15%

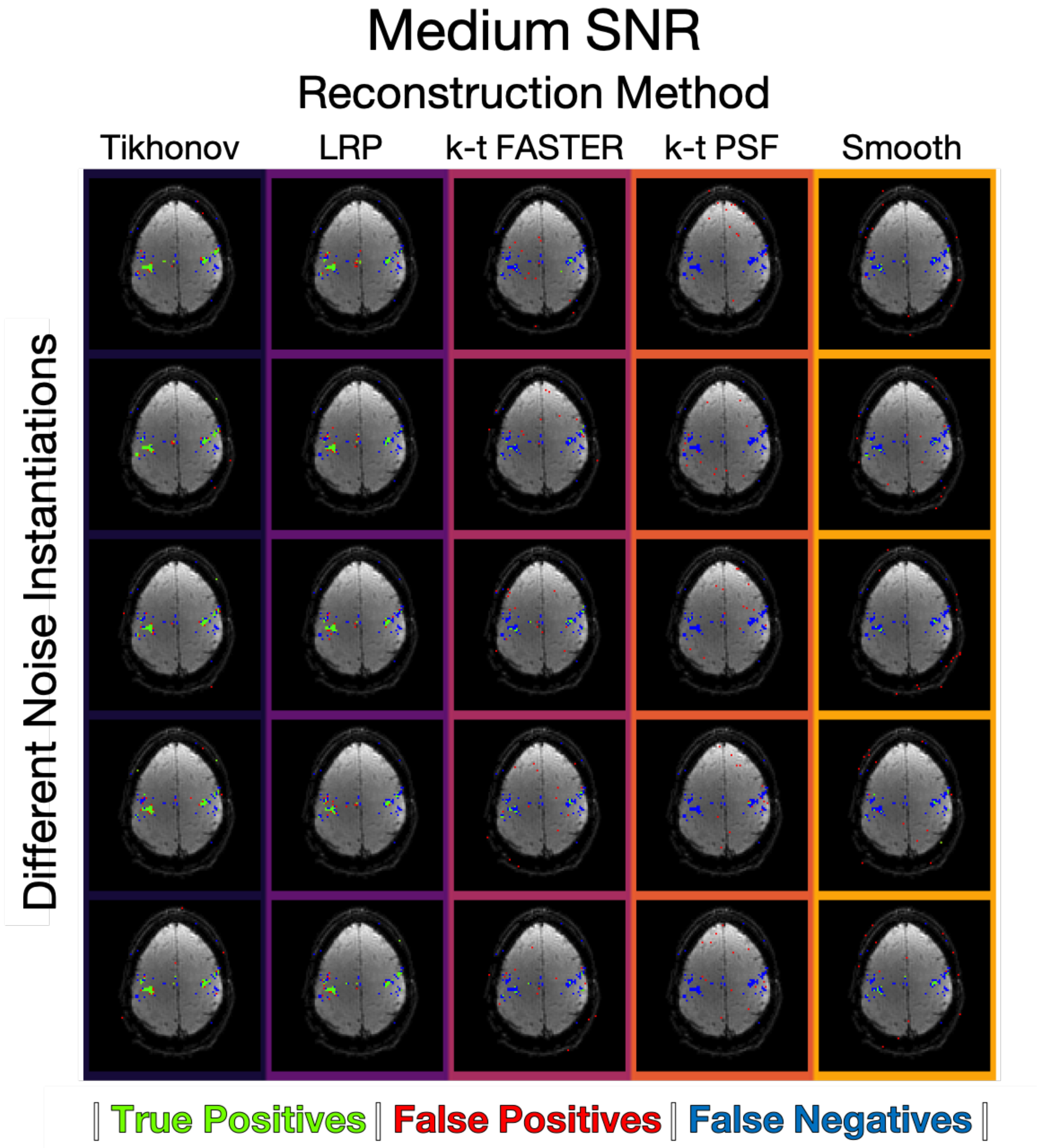


Figure 6.29: The five activation maps at medium SNR (with random noise) for each reconstruction method. Green pixels represent true positives, red pixels represent false positives, blue pixels represent false negatives. The z-statistics threshold yielded a false positive rate of 0.15%

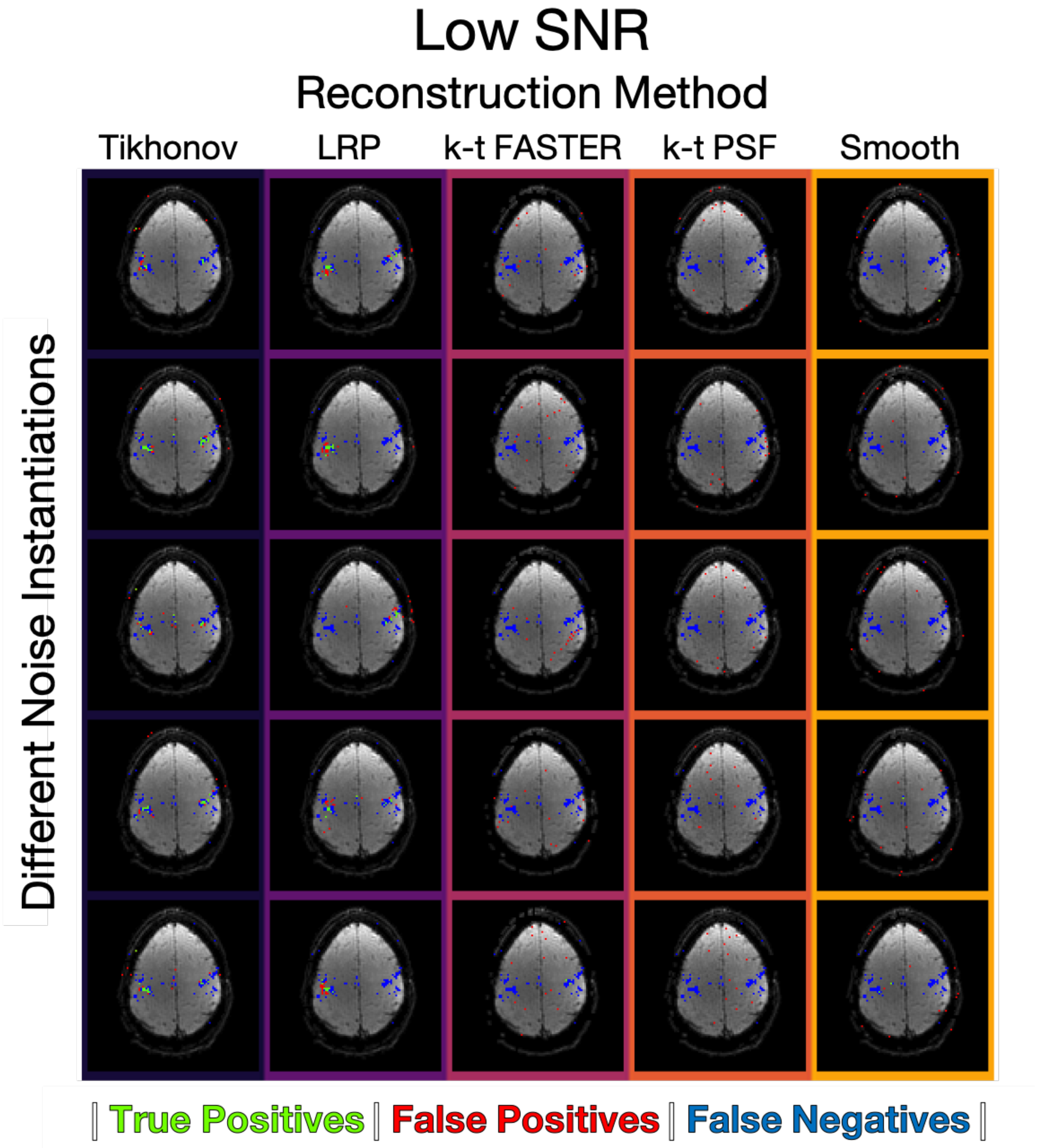


Figure 6.30: The five activation maps at low SNR (with random noise) for each reconstruction method. Green pixels represent true positives, red pixels represent false positives, blue pixels represent false negatives. The z-statistics threshold yielded a false positive rate of 0.15%

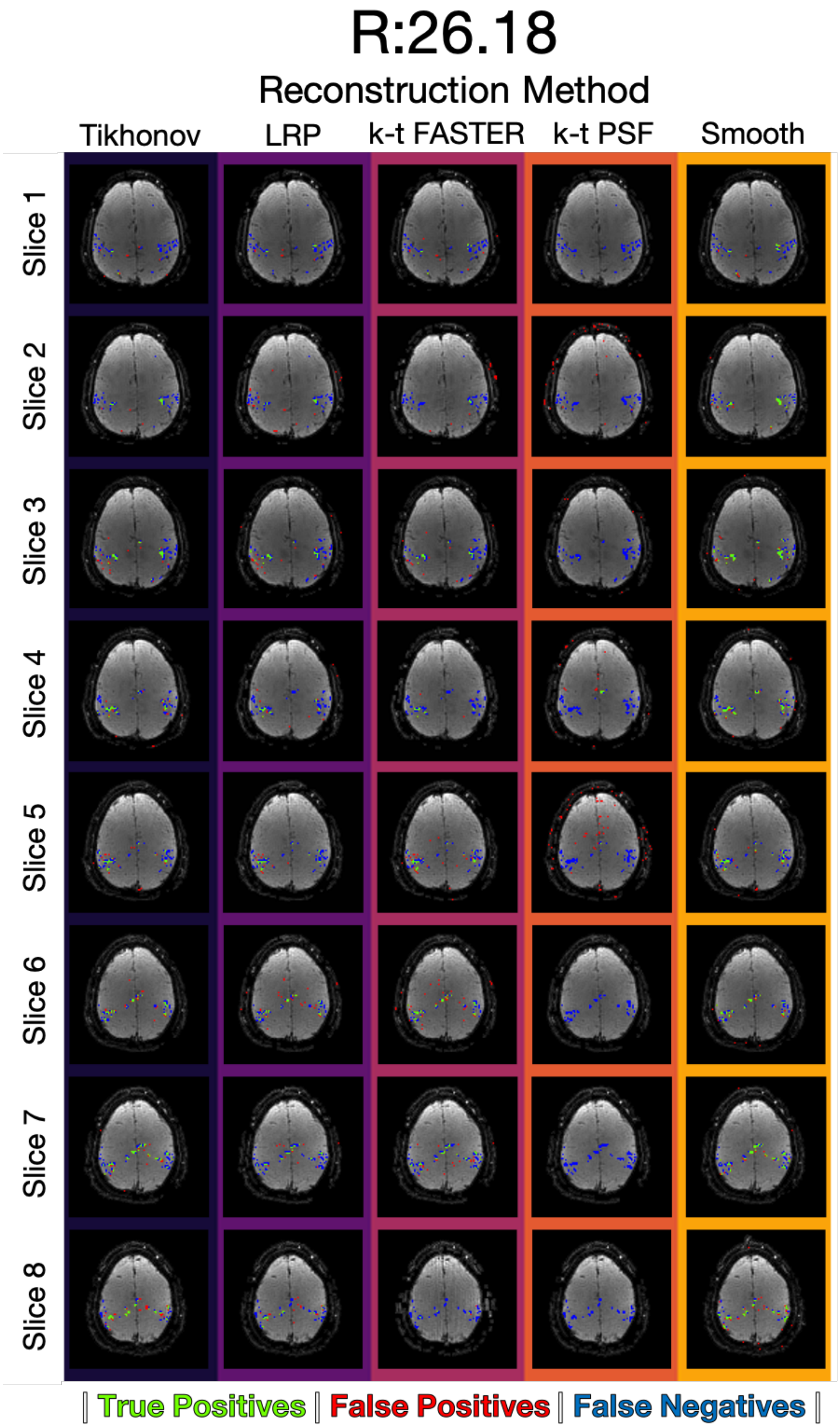


Figure 6.31: The activation maps for every slice of the reconstruction at R=26.18, at a threshold defined by a volumetric 0.15% FPR shown in figure 6.11.

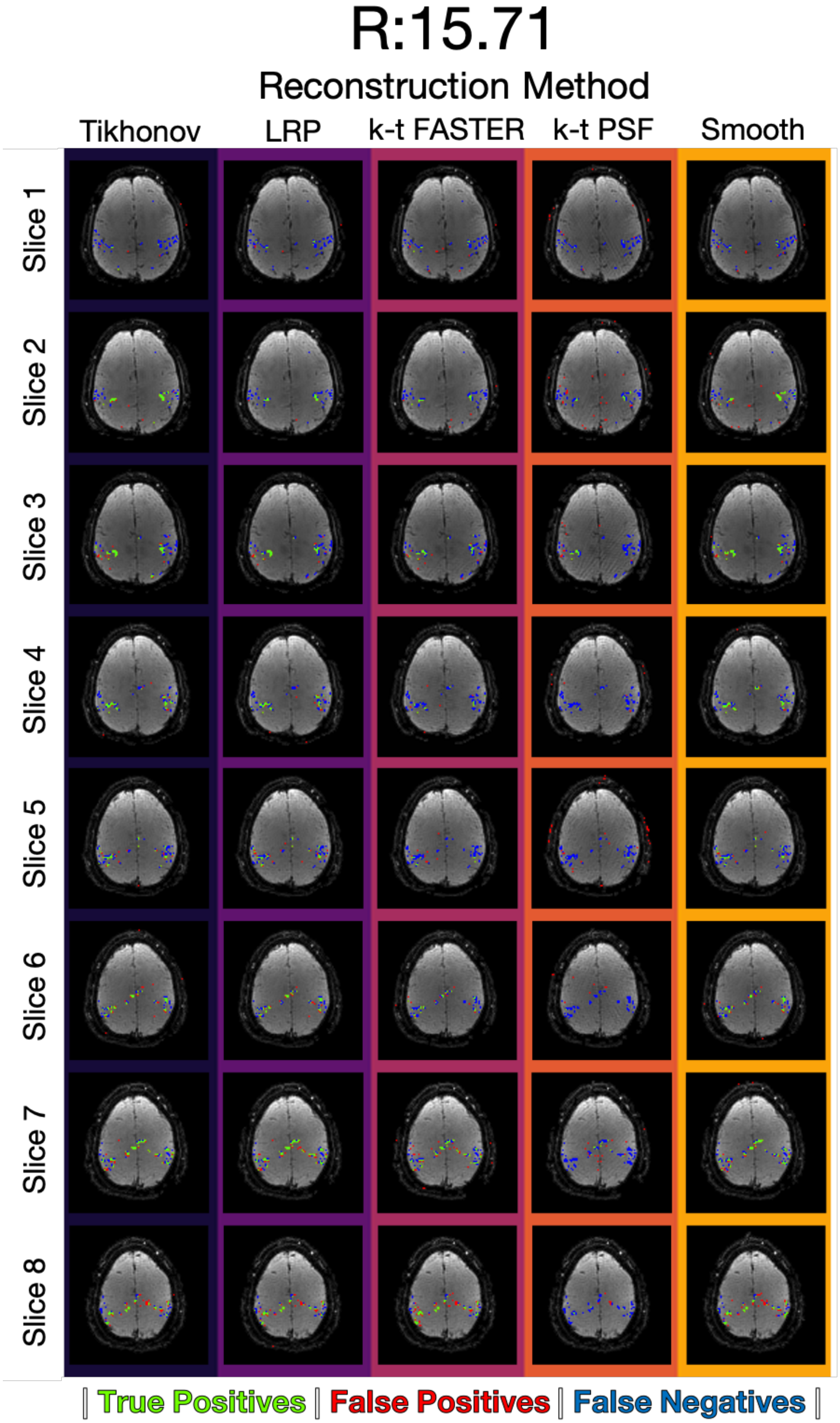


Figure 6.32: The activation maps for every slice of the reconstruction at R=15.71, at a threshold defined by a volumetric 0.15% FPR shown in figure 6.11.

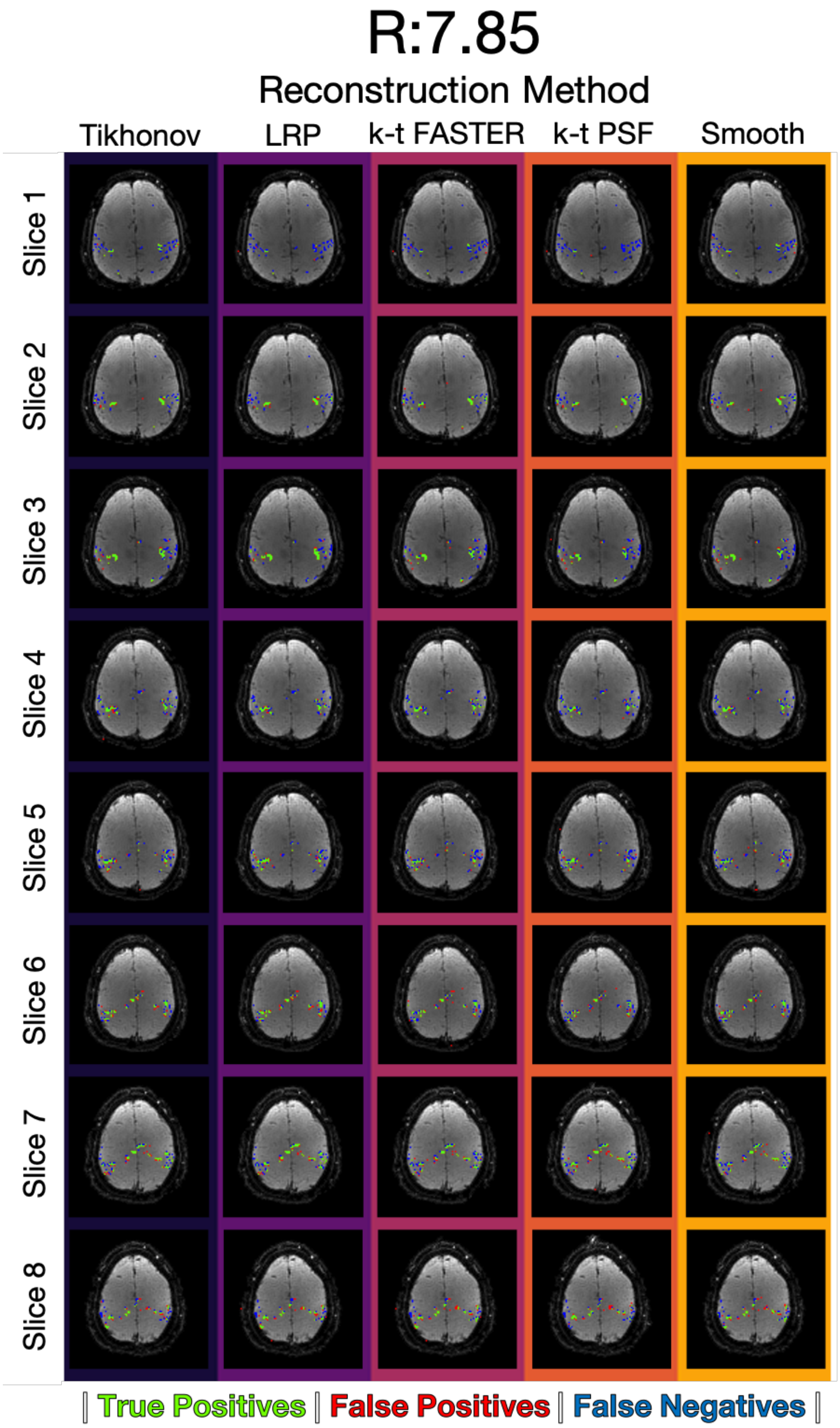


Figure 6.33: The activation maps for every slice of the reconstruction at R=7.85, at a threshold defined by a volumetric 0.15% FPR shown in figure 6.11.

Bibliography

[1] Mark Chiew, Nadine N. Graedel, Jennifer A. Mcnab, Stephen M. Smith, and Karla L. Miller. Accelerating functional MRI using fixed-rank approximations and radial-cartesian sampling. *Magnetic Resonance in Medicine*, 00:1–12, 2016.

[2] Nadine N. Graedel. *Three-Dimensional Hybrid Radial Cartesian Echo Planar Imaging for Functional MRI*. PhD thesis, 2016.

[3] Marijke Welvaert and Yves Rosseel. On the definition of signal-to-noise ratio and contrast-to-noise ratio for fMRI data. *PLoS ONE*, 8(11):1010, 2013.

[4] Stephen M. Smith, Mark Jenkinson, Mark W. Woolrich, Christian F. Beckmann, Timothy E. J. Behrens, Heidi Johansen-Berg, Peter R. Bannister, Marilena De Luca, Ivana Drobnjak, David E. Flitney, Rami K. Niazy, James Saunders, John Vickers, Yongyue Zhang, Nicola De Stefano, J. Michael Brady, and Paul M. Matthews. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23(SUPPL. 1):208–219, 2004.

[5] Mark Jenkinson and Stephen M. Smith. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5(2):143–156, 2001.

[6] Mark Jenkinson, Peter R. Bannister, J. Michael Brady, and Stephen M. Smith. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage*, 17(2):825–841, 2002.

[7] Douglas N. Greve and Bruce Fischl. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, 48(1):63–72, 2009.

[8] Nadine N. Graedel, Mark Chiew, Jennifer A. McNab, and Karla L. Miller. FMRI using a 3D radial-Cartesian trajectory: spatio-temporal tunability and artefact correction. In *Proceedings of the International Society for Magnetic Resonance in Medicine 23*, volume 23, page 1, 2015.

[9] Fan Lam, Bo Zhao, Yinan Liu, Zhi-Pei Liang, Michael Weiner, and Norbert Schuff. Accelerated fMRI using Low-Rank Model and Sparsity Constraints. In *Proceedings of the International Society for Magnetic Resonance in Medicine 21*, volume 21, page 2620, 2013.

[10] Mark Chiew and Karla L. Miller. Improved statistical efficiency of simultaneous multi-slice fMRI by reconstruction with spatially adaptive temporal smoothing. *NeuroImage*, 203(August):1–14, dec 2019.

7 | Conclusion

Contents

In this thesis, an alternating minimisation approach to low-rank fMRI reconstruction with incorporated coil sensitivity information was formulated, and Tikhonov constraints, low-resolution priors, and temporal subspace smoothing constraints were tested.

7.1 Main Conclusions

The alternating minimisation approach was shown to perform equivalently to the existing k-t FASTER implementation of iterative hard thresholding and matrix shrinkage, when no additional subspace constraints were included. Alternating minimisation takes a longer time to optimise, and is less robust for Cartesian trajectories that do not incorporate any additional constraints into the reconstruction. As some form of constraint is nearly always included in modern reconstructions, the long reconstruction time is the main limiting factor.

The constrained-subspace approaches were then all shown (with appropriate weighting factors) to improve upon the unconstrained-subspace implementation of alternating minimisation k-t FASTER in realistic fMRI data, as well as an artificial example dataset.

Tikhonov constrained-subspaces were proven to improve the acceleration over a wide range of acceleration factors. In the prospectively sampled dataset, the temporal subspace smoothing was also shown to produce very high quality results (measured by successful classification in an ROC curve). However, in the retrospectively sampled dataset with added noise, the temporal subspace smoothing did worse than both Tikhonov and LRP-constrained subspaces. As such, further testing is needed into the types of dataset that temporal subspace smoothing can improve the reconstruction at high acceleration. However it is worth restating the specific findings of this thesis: at the highest acceleration tested in the prospective dataset (R=26.18 with a volumetric TR of 0.3s) the smoothing constraint yielded the most high-fidelity reconstructions.

The above conclusions were all derived using a qualitative analysis of the images produced, as well as a set of quantitative reconstruction metrics: canonical correlation scores - which measured the alignment between the subspace mapped by component matrices, and ROC curves - which measured the classification of activation characterised by the z-statistic maps. These metrics are different from conventional image quality metrics (RMSE or SSIM) as they are tailored to favour reconstructions that provide exactly the required information for fMRI analysis. Collectively, the qualitative and quantitative metrics reveal that very high acceleration factors are possible with these constrained-subspace low-rank approaches.

7.2 Future Work

While the constrained-subspace low-rank approaches did improve on the unconstrained-subspace low-rank approaches for fMRI reconstruction, there are many factors that must be finalised before they can be adopted in practice. The first and simplest step is further testing of the constrained-subspace approaches on additional datasets. This would allow more insight into optimal λ values, as well as validating the results found here. A larger λ_X was required in the prospective dataset for both the LRP and Tikhonov constraints compared to the retrospective dataset, and a larger λ_∇ was needed for the smoothing constraint. The higher optimal λ at high acceleration may indicate the presence of increased noise as it did in the retrospective case (i.e. as the propensity for artefacts increases with increased noise, the amount of required regularisation also increases).

In addition to parameter selection, further dataset testing would allow more precise characterisation of the response of the system to different confounding factors, especially testing the robustness of the reconstructions in the presence of motion. Motion can violate the low-rank assumptions within fMRI, and so adequate motion-correction is required. However, a major challenge with rank inflation due to motion is that this effect cannot be corrected after reconstruction using conventional time-series registration. The data collected for this study was performed on healthy volunteers with very little apparent motion, although the TURBINE k-space trajectory enables motion correction using low spatial resolution navigators [1]. One solution could involve joint estimation of the subspaces and motion parameters, on the supposition that a motion-free portion of reconstruction would have the lowest rank/nuclear norm.

The slow joint-optimisation of two subspaces in alternating minimisation is made worse in the case of the LRP-constrained reconstruction, which must carry out the join optimisation twice (once for the priors, and once for the final reconstruction). A focus of future work must be to speed up the overall reconstruction regardless of constraint choice (as seen by the long reconstruction time of table 6.3). There is a strong possibility the convergence criterion ϵ could be more relaxed in constrained-subspace optimisations. In addition, a smaller number of internal iterations in each subproblem (see figure 4.9) could allow for a quicker computation.

Alternatively, there may be more effective approaches to exploit the spatiotemporal information. Regarding minor alterations to the formation of the LRP, it's possible that small adjustments to the k-space window (adjusting the window smoothness or increasing the FWHM) might provide more optimal results. In addition, all priors were reconstructed using the same acceleration factor as the final reconstruction, whereas small changes could prove effective. For example, the spatial prior could be created at a lower R (and so allowing a larger k-space window), while the temporal prior could be separately created at the same R as the final dataset but with a larger window. Non-Tukey windows may provide more efficient windowing, and spatial priors based on functional parcellations could be used to directly relate to the underlying functional networks [2].

An additional constraint that might be considered makes use of the null-enforcement approaches. One of the main benefits of acceleration is the ability to increase the sampling frequency to a point that physiological noise may be discerned. Rather than separating such noise in pre-processing, forcing dissimilarity with known physiological temporal frequencies may allow an increase in sampling efficiency, in a similar manner to the original motivation behind low-rank methods.

The main desirable future focus is likely to be a combined investigation of the smoothing and Tikhonov constraints, however. As simple, single weighting constraints which produced very impressive results at high acceleration factors in prospective data, they represent a path forward to greatly accelerated fMRI reconstruction with a low-rank formulation.

7.3 Final Remarks

The field of fMRI is multifaceted, and is filled with many subtleties. The work presented here is my best attempt at expanding this broad field in some small way. Beyond this point in time, I have no doubt that the field will expand and evolve further, hopefully swallowing the concepts in this thesis as small building blocks for a greater structure. I do not know whether the next great innovation will be so simple that I am amazed that I never thought of it, or so complex that I feel I could never understand it. I have been blessed enough to work with some fantastic researchers, young scientists, and mentors in this field. And so I also know that, whatever the problems of the future of the field may be, there are no finer set of people to try to find the solutions.

Bibliography

[1] Nadine N. Graedel, Jennifer A. McNab, Mark Chiew, and Karla L. Miller. Motion Correction for Functional MRI With Three-Dimensional Hybrid Radial-Cartesian EPI. *Magnetic Resonance in Medicine*, 78:527–540, 2017.

[2] Eric C. Wong. Direct imaging of functional networks. *Brain Connectivity*, 4(7):481–487, sep 2014.